# Social Media Sentiment of Hydrogen Fuel Cell Vehicles in China: Evidence from Artificial Intelligence Algorithms

**Tuo Ye**
School of Public Administration
South China University of Technology, Guangzhou, China
Email: payetuoscut.edu.cn

**Songyu Zhao**
**(Corresponding Author)**
School of Public Administration
South China University of Technology, Guangzhou, China
Email:202210188089@mail.scut.edu.cn

**Chi Keung Marco Lau**
Teesside University International Business School, Teesside University, UK
Email: c.lau@tees.ac.uk

**Frankie Chau**
Durham University Business School, Durham University, UK
Email: h.c.f.chau@durham.ac.uk

**Abstract**

This paper uses social media data to find influential topics from public perceptions of hydrogen energy vehicles. The paper combines a modified K-means algorithm and the Latent Dirichlet Allocation method with sentiment analysis to analyze the clusters, a prerequisite for understanding popular topics and their sentiments in a big data pool. The paper focuses on the posts on Hydrogen Fuel Cell Vehicles, which originated from the Bilibili platform (Chinese social media platform), with 42,063 comments related to Hydrogen Fuel Cell Vehicles. The paper observes 10 clusters with different topics and sentiments. It is also found that the Chinese public perception of hydrogen energy vehicles is mainly neutral, especially regarding the future development of hydrogen energy and the market of hydrogen energy vehicles. The positive stakeholders are optimistic about hydrogen energy. Finally, this paper provides the findings' technological, operational, and strategic development implications.

**Keywords:** hydrogen fuel cell vehicles; K-means algorithm; Latent Dirichlet Allocation method; social media sentiment; public perception

## 1. Introduction

In recent years, China has made significant progress in its energy transition, with the share of renewable energy in energy consumption rising sharply and the energy consumption structure becoming cleaner and lower carbon. According to the Renewable Energy Policy Network for the 21st Century (REN21) (2023), China is the global leader in the renewable energy industry, leading the world in hydropower, photovoltaic, and wind power. It has been leading the development of renewable energy (REN21, 2023). In addition, China has the highest amount of investment in renewable energy, accounting for about 32% of global investment, followed by Europe (25%) and the United States (19%). With rapid economic development, China's demand for energy is growing. Since 2010, China has become the world's largest Energy-consuming country, surpassing the United States and accounting for more than 27% of total global energy consumption (Zhang et al., 2017). In addition, economic growth has also improved the lives of Chinese citizens, boosting China's energy demand. China is currently the world's largest producer and consumer of coal, which dominates energy production and consumption. However, using coal generates a lot of pollution, so the Chinese government has been committed to energy transformation, reducing the dependence on coal and improving the cleanliness of the energy structure. As the global climate becomes severe, all countries are accelerating their energy transition and reducing their reliance on fossil energy. China is no exception, and while actively developing renewable energy, it is also exploring overseas energy markets to ensure energy security, such as the Belt and Road Initiative, which provides a significant opportunity for China to explore overseas energy markets (Nie et al., 2021).

The development of renewable energy has a significant impact on a country's economy and environmental protection. On the one hand, alternative energy sources are needed as natural resources continue to be depleted. On the other hand, economic development and the natural environment must reach a certain balance for the economy (Zhang et al., 2016). In this context, the development and application of renewable energy sources such as photovoltaics, hydrogen energy, and hydropower have become the focus of technological development (Buttner et al., 2017). Hydrogen energy is a clean, efficient, and stable secondary energy source essential in energy transition. In recent years, countries have been actively developing hydrogen energy technology and promoting its industrialization, which may help realize the energy transition. At the same time, it also has broad application prospects, such as power generation, transportation, industry, and other fields, providing new options for energy transition. It can also help balance the power system and reduce costs. In addition, buffered by hydrogen storage facilities equipped by downstream industries such as natural gas infrastructure and hydrogen supply chain, renewable energy generation is volatile and complex to meet the steady demand of the power system (Noussan et al., 2020).

Hydrogen can be an energy storage medium to help power systems balance supply and demand. Specifically, hydrogen is produced by utilizing renewable power, and the remaining renewable power is converted into hydrogen. Based on the need of the power system, hydrogen energy is released to generate electricity to supplement the power. Therefore, it can help the power system consume renewable electricity and improve the stability of the power system (Marchenko et al., 2015). Among them, hydrogen energy is mainly applied to automobiles, and hydrogen fuel cell vehicles have the advantages of long-distance driving and fast refuelling, so they have a broad application prospect in large-scale transportation. Still, in the medium and long term, they may face the challenge of using natural gas or biogas as the energy source for vehicles (Aminudin et al., 2023). However, in recent years, the production cost of hydrogen buses has been significantly reduced. Shortly, the price will be further reduced as the number of hydrogen vehicles increases. In addition, hydrogen trucks are in the development stage. In recent years, they are expected to be deployed on a large scale in China.

Therefore, this study will obtain the public discussion on hydrogen energy and take the hydrogen energy vehicle market as an example to understand the public perception of new energy vehicles. The data is collected from Blibili (similar to Youtube), as it is the second pop-up video website and the most prominent communication centre for young people in mainland China, combined with the Latent Dirichlet Allocation (LDA) model and K-means clustering method, and SnowNLP for sentiment analysis. We get 10 themes with sentiments and give relevant suggestions from technology, operation, and strategy.

Our study contributes to the existing research in several ways. First, as a representative emerging economy, China differs from developed countries regarding energy consumption and industrial structure. This paper remains debatable regarding the perceived factors of energy transition and how these relate to each other. Second, this paper constructs an AI-based algorithmic model and mines comments on social media, which can provide good suggestions for energy transition in developing countries and is more objective and rational than traditional research. Third, this paper applies qualitative comparative analysis to compare LDA and K-means methods to find the differences, making our findings more comprehensive and providing an essential theoretical basis for government, business decision-making, and related organizations.

The rest of the paper is structured as follows. Section 2 reviews the previous literature on hydrogen vehicle consumption. Section 3 presents an algorithm for artificial intelligence that considers the applicability of the data. In Section 4, we describe the data collected for this study and the analytical model for the research and present the analysis results of the public perception of hydrogen energy vehicles. Section 5 is based on the results; a sentiment analysis compares the results between LDA and K-means, followed by a comparative analysis of their potential impact. Finally, Section 6 concludes with the implications.

## 2. Literature Review

### *2.1 Energy Transformation*

The energy transition is a crucial way to achieve sustainable human development. It cannot wait any longer, as it has already penetrated all aspects of our lives (Kovač et al., 2021). To achieve carbon neutrality, we must find renewable energy sources as alternatives to fossil fuels. Burning conventional fossil fuels produces large amounts of carbon emissions and is the leading cause of climate change. Renewable energy is a clean and sustainable source of energy that can help us reduce carbon emissions and combat climate change (Zhao, 2022). China's rapid economic growth has driven rapid growth in energy demand. Over the past 45 years, China's energy consumption has increased by 847%, far exceeding that of developed economies. This has put tremendous pressure on energy supply and environmental sustainability. China's industrial output has grown nearly 20 times between 1978 and 2022, driving the growth in energy demand. The urbanization rate rose from 17.9% to 64.7% between 1978 and 2022, driving growth in energy demand. Residential consumption increased nearly 30-fold between 1978 and 2022, driving growth in energy demand. Global energy consumption will continue to grow over the next two decades. This is mainly due to rapid economic and population growth in developing countries. Natural gas, oil, and non-fossil energy sources will be the main drivers of future growth. To cope with the increase in energy demand, we need to shift from a traditional fossil-fuel-based energy mix to a renewable energy-based one. Renewable energy is a clean and sustainable source of energy that can help us reduce carbon emissions and combat climate change.

We are currently going through a phase of the energy transition, but that does not mean we are starting from scratch. Significant progress has been made in developing renewable energy sources. However, the global economy is still highly dependent on fossil fuels. We are now in the accelerated phase of the energy transition. The cost of renewable energy has fallen, and the technology has matured, so the deployment of renewable energy is accelerating. However, the global economy is still highly dependent on fossil fuels. Therefore the energy transition still requires a long process (Barbir, 2009). Hydrogen is a clean and sustainable energy source essential to the energy transition. This has long been recognized by the scientific community and the International Energy Agency (IEA) (Elam et al., 2003). Hydrogen has great potential as a zero-carbon emission energy source that can help us combat climate change. However, hydrogen technology is still in the early stages of development, and several challenges need to be overcome, but hydrogen remains a worthwhile venture. As technology advances, the cost of hydrogen technology will come down, storage and transportation difficulties will be resolved, and safety concerns will improve. Hydrogen plays an increasingly important role in the energy transition, helping us to reduce carbon emissions and combat climate change (Hetland & Mulder, 2007). For this reason, hydrogen today occupies an essential place in energy storage (McPherson et al., 2018).

### *2.2 Hydrogen Fuel Cell Vehicles Consumption*

The success of hydrogen technology as an essential component of the energy transition depends not only on factors( such as the efficiency of the technology and the cost of production) but also on its eventual acceptance by the public (O'Garra et al., 2005). The technology must be simple enough to be understood and utilized by the people to enable hydrogen to participate in the energy transition system (Mourato et al., 2004). For instance, the public does not know the principles of the internal combustion engine, but they drive to work regularly. At the same time, research focusing only on corporations, entrepreneurs, environmentally conscious groups, industrial and scientific associations, etc., will not build a hydrogen energy society. Educating the public and the recognizability of hydrogen energy should play an important role (Hienuki et al. 2019). Therefore, the automotive industry is the best gateway to reach the public and the best option to enhance the adoption of hydrogen energy.

Most studies have focused on electric vehicles or HFCVs. When HFCVs were in their early stages, several studies conducted questionnaires to analyze consumer perceptions of HFCVs, focusing mainly on the U.S. and European markets. For example, Lopez Jaramillo et al. and Hardman et al. used surveys to investigate consumer perceptions of HFCVs in the United States (Hardman & Tal, 2018; Lopez Jaramillo et al., 2019). The study shows that consumers have very similar levels of preference for HFCVs and Internal Combustion Engine Vehicles (ICEVs). This suggests that consumers have positive expectations about the environmental friendliness and safety of HFCVs. In addition, compared to EVs, HFCVs are positively rated regarding charging time, driving distance, fuel cost, and fuel type, corresponding to the vehicle's attributes. Hardman & Tal (2018) reported that consumers use HFCVs with decreasing accessibility to EV charging stations. Byun identified consumer preferences for cars and predicted the dynamic market share of eco-friendly vehicles by utilizing data from a hybrid logit model and a discrete choice experiment conducted on Korean consumers (Byun et al., 2018). In most studies, the fuel type that has attracted the most attention in the eco-friendly vehicle category is predominantly electric vehicles (Byun et al., 2018; Choi et al., 2018; Lee et al., 2021). However, considering the trends and prospects of the automotive market, academic studies are gradually emphasizing HFCVs to predict a more realistic market share soon. However, less research has been conducted on developing the Chinese market. Therefore, based on the above, this study will examine China's public perception of hydrogen-fueled vehicles (HFCVs). Unlike previous research methods, this paper employs a new artificial intelligence algorithm to understand the main characteristics of the hydrogen-fueled vehicle market.

## 3. Methodology

From the perspective of public demand, we design a model based on LDA and K-means algorithm, and through Snownlp sentiment analysis, effectively mine the adequate information in user comments, strengthen the information exchange with users, and sort out the internal mechanism of rapid iteration of new products. The public concern for hydrogen energy vehicles constructed in this paper is shown in Figure 1.

### *3.1 LDA Topic Modeling*

### *3.1.1 LDA Topic Extraction*

Extracting topics from comment text using the LDA model, which is a probabilistic topic model. As a three-layer Bayesian probabilistic model, it has a three-layer structure of "word, topic, document." LDA topic model considers that a document has several topics and a topic is embodied by several words. Using LDA to obtain the distribution of text topic information, the paper is a random combination of potential issues with a certain probability; each topic is a combination of words with a particular possibility. The basic calculation process of extracting topics from frequency text using LDA is provided in Eq. (1). In the "document-word" matrix, φ denotes the "topic-word" matrix; θ denotes the "document-topic" matrix, and the matrices φ and θ are obtained by unsupervised learning (Jelodar et al., 2019).

$$P(\theta, t, w, \varphi|\alpha|) = \prod_{n=1}^{N} P(\theta|\alpha)P(t_n|\theta)\, P(\omega_n|\theta)P(\varphi|\beta) \qquad (1)$$

Where α is the document parameter; β is the vocabulary parameter; θ is the document-topic matrix sampled from β; φ is the topic-vocabulary matrix sampled from β; t is the topic tested from θ; N is the total number of words in the document collection.

For the topic model, this paper utilizes the perplexity degree to determine the optimal number of topics, K, to ensure the effectiveness of topic extraction. Perplexity is a standard metric for evaluating the performance of language models. The lower the perplexity, the better the clustering effect and the higher the topic fit. The formula for calculating the perplexity degree is shown in Eq. (2).

$$Perplexity(T) = exp\left\{-\frac{\Sigma_{d=1}^{M}\log p\ (w_d)}{\sum_{d=1}^{M} N_d}\right\} \qquad (2)$$

where T is the number of required test document sets, $w_d$ is the sequence of words, and $N_d$ the number of words in d.

### *3.1.2 Feature Word Extraction*

According to Eq. (1), the "document-topic" matrix φ is obtained, as shown in Eq. (3). There are m topics and n words in the matrix, each column is the distribution probability of each word in m topics, and each row is the distribution probability of each issue in n terms.

$$\begin{bmatrix} & w_1 & \cdots & w_n \\ t_1 & P(w_1+t_1) & \cdots & P(w_n+t_1) \\ \vdots & \vdots & & \vdots \\ t_s & P(w_1+t_m) & \cdots & P(w_n+t_m) \end{bmatrix} \quad (3)$$

According to Eq. (1), the "topic-vocabulary" matrix θ is obtained, as shown in Eq. (4). There are s documents and m topics in the matrix, each column is the distribution probability of each subject in s documents, and each row is the distribution probability of each paper in m topics.

$$\begin{bmatrix} & t_1 & \cdots & t_1 \\ d_1 & P(t_1+d_1) & \cdots & P(t_m+d_1) \\ \vdots & \vdots & & \vdots \\ d_s & P(t_1+d_s) & \cdots & P(t_m+d_s) \end{bmatrix} \quad (4)$$

To find the high-value topic keywords and better describe the content of the topic, Eq. (5) is taken to calculate the weights of m words. The most prominent keyword is selected as the feature word of the issue. Where w is the weight value of vocabulary wi for the topic and document set {d1,d2,...,dm}; $P(w_i|t_j)$ is the distribution probability of vocabulary $w_i$ in topic $t_j$.

$\sum_{d\in\{d_1,d_2,\cdots,d_m\}} c(w_i|d)$)is a summation of the frequency of occurrence of the vocabulary word $w_i$ in the set of maximal distributional probability texts $\{d_1, d_2,\ldots, d_m\}$of the topic $t_j$.

$$w = P(w_i|t_j) \sum_{d\in\{d_1,d_2,\cdots,d_m\}} c(w_i|d) \quad (5)$$

***3.2 K-means***

The k-means algorithm is an unsupervised clustering algorithm. it can be a document within the file automatically divided into several clusters (Cluster). The paper's content in each cluster has a more remarkable similarity, the range of the groups between the similarity of the smaller. The k-means algorithm is a dynamic clustering algorithm. The algorithm uses iterative operations to test and adjust the classification of data samples. Due to the K-means algorithm, the initial clustering centre selection is not appropriate to cause the clustering results to have a considerable error. In response to the problem, scholars proposed the K-mean + + algorithm to the K-means algorithm to improve the selection of the initial point of the location of the algorithm. Algorithmic steps are as follows: the initial point of the clustering centre selection is inappropriate to cause the clustering results to have a considerable error.

In step 1, a point is randomly selected from the data set as the initial clustering centre $c_1$.

Step 2: Calculate the shortest distance between each sample point and the existing clustering centre, denoted by D(x), and subsequently calculate the probability that each sample is selected as the next clustering centre; finally, like the next clustering centre according to the roulette wheel method, see Eq. (6); and

$$\frac{D(x)^2}{\sum_{x \in X} D(x)^2} \quad (6)$$

Step 3: Repeat step 2 until K clustering centres are selected.

Step 4: For each sample $x_i$ in the dataset, calculate its distance from the K clustering centres and classify it into the corresponding clustering centres with the smallest spaces.

Step 5: For each category $c_i$, recalculate its clustering centre. See Eq. (7);

$$C_i = \frac{1}{|c_i|} \sum_{x \in c_i} x; \quad (7)$$

In step 6, steps 4 and 5 were repeated until the position of the clustering centres did not change anymore.

### *3.3 Sentiment Analysis*

#### *3.3.1 SnowNLP*

Regarding text granularity, sentiment classification can realize document-level and sentence-level analysis at different levels of coarseness and fineness. Sentence-based sentiment classification is to identify the sentiment tendency of sentences with finer granularity but with overall significance (Chen et al., 2018).

The sentiment intensity of each emotion category is calculated through the extracted features, and the emotion category with the most incredible intensity is taken as the emotion category of a single microblog text. In this paper, we choose the emotion vocabulary ontology library of the Dalian University of Technology as the emotion lexicon, which categorizes the emotions into "joy," "good," "anger," "sadness, "Fear," "Evil,," and "Surprise" are divided into 7 categories and 21 subcategories, and the initial intensity of emotion words is set as 1, 3, 5, 7, and 9. The initial emotional intensity of the emotion words was set to 1, 3, 5, 7, and 9, with 9 indicating the maximum intensity and 1 the minimum intensity. Each word corresponds to a polarity under each category of emotion. 0 represents neutral, 1 represents positive, 2 represents negative, and 3 represents positive and negative gender. Eq. (8) for the sentiment value of a word is.

$$s(w) = v(w)p(w) \quad (8)$$

In Eq. (1), *s(w)* denotes the sentiment value of the lexicon, *v（w）* denotes the sentiment intensity of the lexicon, and *p(w)* indicates the sentiment polarity of the lexicon. Since negative words and degree adverbs affect the results of sentiment discrimination, based on the reference to the existing negative word lexicon and degree adverb lexicon, this study supplements the negative word lexicon and degree adverb lexicon, which are suitable for analyzing the corpus. According to the Chinese language convention, when there are an odd number of negations before and after an emotion word, the polarity of the emotion word is deflected.

In addition, to improve the accuracy of sentiment feature discrimination, the degree adverbs are added and manually discriminated to form the degree adverb lexicon with the weight value of the influence on the sentiment words, see Eq. (9).

$$WordSentiment = v(w) * v(adv) \quad （9）$$

Where v(adv) is the weight of the degree adverb, v(w) is the intensity value of emotion words. For each text, the intensity values of the corresponding emotion words after modification are calculated for each type of emotion, summed up according to the category, and the emotion category with the most considerable intensity is taken as the final emotion classification of the text, to portray the emotion evolution law of public opinion reversal events under the perspective of time dimensions (Wu et al., 2023).

***3.3.2 Sentiment Value Calculation***

The sentiment value calculation of SnowNLP is based on the theory of plain Bayesian classification algorithm, which belongs to the generative classification algorithm, and the calculation formula (10) is as follows.

$$P(category \,|\, word) = \frac{p(word \backslash category)P(category)}{P(word)} \qquad (10)$$

Before calculation, due to the vulnerability of plain Bayesian classification to the classification task, it is necessary to retrain the sentiment analysis model by importing positive and negative samples based on the microblog corpus data. In this paper, we choose the dataset of 6000 positive and negative comments with completed sentiment annotation for retraining, and the sentiment value can be calculated after the model is trained. In data processing, we convert the sentiment interval from [0,1] to [-0.5,0.5] so that the closer the sentiment value of a single microblog text is to -0.5, the more damaging it is. The closer the sentiment value is to 0.5, the more favourable it is. With the help of the sentiment value obtained by SnowNLP, we portray the sentiment evolution law of the public opinion reversal events from the spatial perspective to reflect the details of the relevant regions more accurately.

## 4. Data Acquisition and Preparation

### *4.1 Data Acquisition*

In this paper, we mainly use Python software to obtain the public's view on the consumption of hydrogen energy vehicles in B station and take the following steps for data collection. Firstly, using Python as the scripting language and MySQL as the database, we crawl the videos related to keywords such as "hydrogen+automobile," "hydrogen energy+automobile," "hydrogen fuel+automobile," and so on, and obtain a total of 840 video information (including title, release time, publisher, clicks, etc.). ", "hydrogen fuel + car," and other keywords-related videos were crawled, and a total of 840 video information (including title, release time, publisher, clicks, etc.) were obtained, whose release time was mainly concentrated between January 1, 2021, and August 20, 2023 (including a small number of 2018-2020 videos). Then, manual denoising was performed for the video information obtained in the previous step (removing duplicates and videos that were obviously irrelevant to the topic), and 712 video information was finally accepted. Finally, according to the above results, Octopus Collector V8.5.7 was applied to capture online comments and real-time comments in the videos. 14905 online comments and 27130 real-time comments were obtained, totalling 42,035 comments.

### *4.2 Data Preprocessing*

It is necessary to preprocess the acquired raw data to improve the reliability of data analysis. In the first step, the data cleaning function embedded in the Baidu Flying Paddle EasyDL platform is utilized to delete duplicated online comments and real-time comments and remove meaningless characters such as nicknames, emoticons, URLs, etc. The second step is to use Excel to convert all the data from traditional Chinese to simplified Chinese. In the second step, the Chinese Traditional-Simplified Conversion function with Excel was used to convert all the data from Traditional to Simplified. In the third step, the two authors used manual judgment to eliminate the content unrelated to this chapter, and any elimination needs to obtain the agreement of the two authors. In the fourth step, using the RE library in Python, the review texts with longer textual content were sliced, i.e., divided into several phrases with punctuation intervals, and finally, we obtained 88,956 short readers.

After the data preprocessing, it is necessary to apply the lexicon software to lexicalize the acquired text data again. Since the content of the collected data is basically all in Chinese, this paper uses a Chinese lexical package named Jieba to perform lexical processing on the data in a Python environment. Firstly, two lexicons closely related to the automobile industry downloaded from the Sogou Thesaurus are added to the customized word list, such as "automobile industry-specific vocabulary" and "automobile professional vocabulary," etc., to ensure the credibility of the subsequent analysis results. Secondly, in the elimination of deactivated words, considering that the clustering effect of the complete deactivated word list, which integrates multiple deactivated word lists, is significantly improved compared with the clustering result of a single deactivated word list, the experiment comprehensively uses the deactivated word list of Harbin Institute of Technology, the deactivated word list of Robotics Intelligence Laboratory of Sichuan University, and Baidu's deactivated word list for filtering the text, and deletes the words "me," "you," "he," "the" and other words that appear frequently but have no practical significance, as well as false words, pronouns, prepositions and so on. The result is a more accurate segmentation result.

**5. Empirical Results**

More and more people choose to find information online, especially with the emergence of online video sites and cultural communities, which facilitate the public's search for information and access to knowledge. Therefore, analyzing the discussion about hydrogen energy vehicle consumption in online video websites is necessary. Bilibili (now called B station) is an online video website founded in 2009. which has a more mature mechanism of online comment and real-time comment function, and in this way, realizes a strong interactive connection between creators and users. This feature can help to better understand the public's view on hydrogen energy vehicle consumption. The analysis process of this paper is as follows: In the first step, we use the LDA algorithm to mine the hot topics that the public is concerned about in the hydrogen energy vehicle consumption field. In the second step, we add the K-means algorithm to supplement the topic mining. Finally, we conduct sentiment analysis for the content of topic mining to analyze the public's perception of different topics.

***5.1 Exploratory Analysis***

After completing the Chinese word segmentation work in this paper, this study visually highlights the keywords that appear more frequently in the dataset by generating a word cloud map (see Figure 2) to clearly and intuitively present the research hotspots in the field of hydrogen energy.

From the generated word cloud map, hydrogen energy is a clean, efficient, and safe source with broad application prospects. Hydrogen Fuel Cell Vehicle is a new type of transportation that uses hydrogen as a fuel and converts the energy generated by the chemical reaction between hydrogen and oxygen to drive the vehicle through a fuel cell. Among them, hydrogen is the core fuel of hydrogen energy vehicles, and its cost and supply are the key factors restricting the development of the hydrogen fuel cell vehicle industry. At present, the production cost of hydrogen is high, and it is mainly prepared through fossil fuel cracking and water electrolysis. The hydrogen supply network is also incomplete, and the construction cost of hydrogen refueling stations is high.

Meanwhile, Japan is a leading country in the world in hydrogen energy technology, with a high R&D and industrialization capability. The Japanese government attaches great importance to developing the hydrogen energy industry. It has introduced a series of policy measures to promote the popularization and application of hydrogen energy vehicles. Japan has substantial advantages in hydrogen vehicles and is expected to occupy a vital position globally. Regardless of the country or region, the core technology of hydrogen vehicles is the fuel cell system. The efficiency, life, and reliability of the fuel cell system are essential factors affecting the performance of vehicles. As a core component of electric cars, its performance directly affects the range and cost of use of electric vehicles.

However, there are still some problems with hydrogen fuel cell vehicles, including high costs. The cost of hydrogen fuel cell vehicles is still high, which is the main factor restricting their popularization and application; lagging behind in the construction of hydrogen refuelling infrastructure, the high cost of construction of hydrogen refuelling stations, and the network of hydrogen refuelling stations is still not perfect, which is another important factor restricting the popularization and application of hydrogen fuel cell vehicles; safety, hydrogen is a flammable and explosive gas, and the hydrogen safety, hydrogen is a volatile and combustible gas. Hence, the safety of hydrogen fuel cell vehicles is a critical issue to consider.

***5.2 LDA Topic Modeling and Sentiment Analysis***

Table 1 provides the findings of LDA themes, and Figure 3 provides the sentiment analysis based on LDA themes.

Topic 0: Technology. On the theme of technology, the views held by the public are neutral. Hydrogen, wind, and other green power hydrogen production differ from traditional fossil energy. The technology faces the problem of high volatility of power supply, and the cost faces the problem of comparison of the conventional hydrogen production methods. Second, hydrogen storage and transportation, the rise of fuel cell vehicles mobile users and the wide spatial and temporal distribution of wind and light resources, the storage and transportation of hydrogen technology puts forward a new requirement: large-scale, low-cost, high-efficiency and safety of the Large-scale, low-cost, high-efficiency and safe storage and transportation technologies are particularly urgent; Third, hydrogen utilization, in which technologies such as ammonia, synthetic methanol and petroleum refining are facing the reshaping of the process production flow, and need to re-attain the balance of energy and materials in the system. At the power and transportation level, hydrogen storage and hydrogen power generation are facing the exploration of technical routes and business models, fuel cell technology iteration, and cost reduction issues. China's hydrogen energy development is mainly focused on fuel cell vehicle-related technologies. Power stacks have been laid out, but in hydrogen energy, "production, storage, transportation and addition" and other technology research and application areas have certain limitations.

Topic 1: Energy. The public maintains a relatively positive attitude towards the energy. In terms of energy supply methods, hydrogen comes from a wide range of sources, from coal, natural gas, and other fossil energy sources. In contrast, other forms have potential for large-scale hydrogen supply from biomass, solar thermal, and photovoltaic energy.

Hydrogen-powered vehicles are a new type of transportation that uses hydrogen as fuel and converts the energy generated by the chemical reaction between hydrogen and oxygen into electricity to drive the vehicle through a fuel cell. The Chinese government attaches great importance to developing the hydrogen energy industry. It has introduced a series of policies and measures to promote the popularization and application of hydrogen energy vehicles. China's hydrogen energy vehicle industry has made significant progress in recent years, and the cost, performance, and safety have significantly improved.

With the continuous progress of technology and policy support, people will consider that the cost of hydrogen energy vehicles will be gradually reduced, and safety issues may be resolved. In addition, hydrogen energy vehicles have some advantages: First, cleanliness, hydrogen energy vehicles do not produce any pollution in the process of use, which is a crucial way to achieve decarbonization of the transportation industry; second, they have high efficiency, the energy conversion efficiency of hydrogen energy vehicles is increased up to more than 50%, which is more than twice as much as that of electric cars; and third, long-range, hydrogen energy vehicles can have a range of up to 500-600 kilometres, which is comparable to that of traditional fuel vehicles.

Topic 2: Development (Automotive): Global hydrogen energy development will continue accelerating in 2023, and clean hydrogen energy will be more clearly positioned as an essential tool for achieving "peak carbon and carbon neutrality." We will see more and more hydrogen fuel cell buses and trucks travelling on the road, renewable energy hydrogen production projects are being carried out in many places, and hydrogen is beginning to enter the lives of the public. Although hydrogen energy brings us green and environmental protection, several hydrogen energy accidents result in some casualties and property damage. The root cause is that hydrogen easily leaks, diffuses, burns, and explodes quickly, and other characteristics. For the current development of hydrogen energy vehicles, although the cost of passenger cars will drop to a level approximating that of lithium battery cars, the essential factors constraining the promotion are safety and the lack of other "low-level" goods to pave the way for the consumption of hydrogen fuel cell cars, such as hydrogen refuelling stations and additional infrastructural support, and so on. There is no apparent resistance or rejection of the development of hydrogen vehicles at the social and market levels, and it can even be said that there is an overall friendly attitude (Ng & Zheng, 2018).

Topic 3: Japan. As the global market's concern for the ecological environment grows, environmental technology has become a key technology for automotive product development. Various countries' governments have successfully formulated development strategies for fuel-cell vehicles and introduced incentives for green cars. The world's major automotive giants have actively invested heavily in the research of fuel cell vehicles, which has opened up a global automotive revolution. Among them, Japan is the most popular. Japan has been actively promoting the development of hydrogen energy. As early as 2014, it formulated the "hydrogen society realization strategy," and in 2020, released the "green growth strategy," put forward for the 2050 carbon neutral target of low carbon / "zero" carbon power. In the Green Growth Strategy released in 2020, the strategy proposes development goals and supportive policy measures for emissions reduction in critical areas such as new energy vehicles, hydrogen energy, aviation, and residential buildings. Japan has become one of the largest hydrogen energy markets in the world, and the status of the development of the hydrogen energy industry has attracted a great deal of attention (Kejun et al., 2021).

Meanwhile, Toyota is the most active automaker in promoting hydrogen vehicles, with many hydrogen-related patents and extensive experience in developing hydrogen-powered internal combustion engines. However, there are still public concerns, mainly since Japan is a relatively small island nation geographically with limited land resources. As a result, it faces several challenges in building hydrogen energy infrastructure. For example, a lack of land may constrain the construction of large-scale hydrogen production and storage facilities. Moreover, Japan is in a region prone to earthquakes and typhoons, which increases the risk of constructing and operating hydrogen infrastructure, and the ability to ensure the resilience and safety of hydrogen systems is an important issue. Although Japan is at the forefront of hydrogen energy development, consumer attitudes toward hydrogen vehicles in Japan remain neutral due to resource and location issues.

Topic 4: Markets. Hydrogen energy is an essential alternative to fossil energy to achieve carbon neutrality. By the end of 2020, China's hydrogen energy production and consumption had exceeded 25 million tons, with more than 7,355 hydrogen fuel cell vehicles and 128 hydrogen refueling stations completed. Despite the massive resources invested by academia and the industry, the development of the consumer market for hydrogen fuel cell vehicles in China is not mature because of its technical limitations, insufficient target consumers, and market pricing that deviates from consumers' psychological expectations.

Most of China's total hydrogen production is used as industrial raw materials, and only a tiny portion flows to the energy sector. The supply is not enough to support a large-scale energy consumption market. At the same time, on the demand side, in the industrial sector, China's power network is now basically perfect, forming a pattern of fossil-driven thermal power generation supplemented by wind power, hydropower, and solar energy. Promoting hydrogen power generation requires the additional paving and improvement of many infrastructures, with substantial economic costs. On the other hand, in the residential and commercial sectors, China's consumption is shallow due to the low ownership of hydrogen vehicles, the lack of promotion of small and medium-sized domestic fuel cells, and the lack of a plan to mix hydrogen into the residential natural gas network at a safe rate of 10 per cent for the time being. In addition, compared to pure electric vehicles, the continuous iteration of lithium-ion battery technology makes pure electric cars begin to gradually take the lead, and the manufacturing and use costs of hydrogen vehicles are higher than those of pure electric vehicles, which also makes the market's recognition of hydrogen fuel cell vehicles to be certified over time.

Topic 5: Transportation. There is a negative public attitude toward this theme. Storage and transportation of hydrogen is a problematic part of the hydrogen supply chain and a significant reason for the high-end price of hydrogen. Hydrogen is the first element in the periodic table. It has a very low density in its gaseous state, is also hot-tempered, and can explode if not properly stored. Therefore, hydrogen must be compressed and held in large-scale transportation to present a high-density gaseous or solid, organic form. According to the different hydrogen conditions, hydrogen transportation can be divided into gaseous, liquid, and solid hydrogen. Gaseous hydrogen is usually transported in long tube trailers and pipelines; liquid hydrogen is traditionally transported in tanker trucks; and dependable hydrogen transportation can be directly transported in hydrogen storage metals. However, the high cost of storing and transporting hydrogen is easily solved by using pipeline transportation because existing natural gas pipelines cannot be used to transport hydrogen directly. Generally, natural gas is transported using steel pipes, and hydrogen molecules can dissolve in steel, creating hydrogen embrittlement, which exceeds the steel's strength limit. It is vital to use pipelines to transport hydrogen; materials with shallow carbon content are needed, which typically cost twice as much as natural gas pipeline materials. It is possible to use a mixture of natural gas and hydrogen for transportation. Still, there are strict requirements for the percentage of hydrogen content to be no more than 20%. Also, mixing expensive hydrogen with cheap natural gas results in a loss of value.

Along with the increasing demand for long-distance hydrogen transportation, pipeline transportation will become the optimal choice for future long-distance transit. Still, the initial investment is vast, and the utilization rate of the equipment is also low. The corrosion and strength change of the pipeline will be subject to further research when the liquefied gas pipeline is used to transport hydrogen.

### *5.3 K-means Clustering and Sentiment Analysis*

This paper improves the LDA topic model. A clustering topic model based on the K-means algorithm is proposed to find out K random clustering centroids, and then according to the distance between each data point (Pattern) and some clustering centroids to decide which data points are suitable to be combined into the same cluster, to get the new clustering centroids, and correct the clustering results. The new clustering centres are used to correct the clustering results, and these steps are repeated until the set termination conditions are met. Finally, the clustering results can be obtained successfully (Zhao, 2022).

Assuming the total number of samples in the dataset is denoted as 'n,' and the K-means algorithm performs 't' iterations, it can be deduced that the algorithm's complexity is O(nKt). For the selection of the k value, the elbow method used in this paper is calculated, see Eq. (11), as follows:

$$SSE = \Sigma_{q=1}^{k} \Sigma_{l=1}^{n_q} \left\| P_{ql} - \bar{P}_q \right\| \quad (11)$$

to derive the K-value. Where nq is the number of samples in each cluster, $P_{ql}$ is the lth sample in the q$^{th}$ cluster, $\bar{P}_q$ is the clustering centre in the qth cluster.

Figure 4 illustrates the optimal k values confirmed by the elbow method. Also, Figure 5 provides the K-means thematic clustering. Combined with the above sentiment analysis of LDA, the sentiment analysis of k-means will be for a more detailed and additional explanation of it.

Topic 0-Engine. The public sentiment for energy and engines is negative. In contrast, the opinion for energy in the LDA is positive. The k-means method indicates that the public's concern about hydrogen vehicles lies more in engine technology. They view hydrogen energy positively but are sceptical about hydrogen-related equipment.

Topic 1-International. Consistent with LDA's market sentiment, the public generally believes that the market for hydrogen vehicles lies mainly in China, the United States, and Japan precisely because these three countries have better technology and industries and are prone to produce market-recognized products. Meanwhile, Toyota and Tesla are receiving critical attention from the market. However, the market maintains a neutral and wait-and-see attitude as hydrogen cars are still developing, and no country has yet to develop particularly prominently.

Topic 2- Human Development. In line with the sentiment of LDA, a neutral attitude towards the future is maintained. The public believes that the development of human civilization is a history of the result of energy use; energy has changed the course of human society, and human civilization is constantly exploring new energy sources. For the development of hydrogen energy, people are worried about whether it is safe and controllable in the future, as well as the popularization of hydrogen energy.

Topic 3-Electric Vehicles. LDA mentioned earlier that the share of hydrogen vehicles is tiny compared to electric cars, and the market understanding is also shallow. With the development of electric vehicles and fast charging technology, the public believes hydrogen-fueled vehicles are becoming redundant for road transportation. For the passenger car market, the market size of hydrogen vehicles from the beginning of 2021, there are about 25,000 hydrogen-fueled vehicles in stock globally, of which more than 90% are distributed only in

South Korea, the United States, China, and Japan. And for pure electric and plug-in hybrid vehicles, there could be about 15 million on the road globally by early 2022. Regarding available models, manufacturers worldwide offer only two passenger fuel cell models. In contrast, almost all manufacturers now sell electric vehicles, with more than 350 models available globally.

Topic 4-Storage and Topic 5-Hydrogen Production are both on the topic of hydrogen. The preparation of $H_2$ is the basis for the commercialization of hydrogen fuel cells. The leading industrially mature technologies for $H_2$ production are water electrolysis, natural gas-water vapour reset, vaporization of heavy oils, and partial oxidation of hydrocarbons. 85 per cent of $H_2$ production comes from fossil fuels, with natural gas-water vapour reset being the primary technology for $H_2$ production. -The water vapour reset method accounts for over half of the $H_2$ production market. Although both use truly environmentally friendly energy sources, the cost of high-purity hydrogen for hydrogen fuel cells is still high, and storage is cumbersome, so the market does not have a positive sentiment toward energy.

### *5.4 Discussion and Implications*

#### *5.4.1 Technological Progress and R&D*

Pay more attention to R&D related to hydrogen energy technology, which promotes the entry of HFCVs into the market. As society moves towards the future energy transition, the utilization of hydrogen energy is becoming a key factor. The public is paying more attention to hydrogen technologies in the energy transition, especially those for hydrogen production, storage, and transportation. For the technology research and development stage, it is necessary to further advance the research and development of fuel cell systems and core components and continue to carry out R&D and testing activities to improve the critical performance and economy of products, to promote the performance optimization and upgrading of the enterprise's products. At the same time, it is necessary to make efficient use of the rich experience accumulated in the fuel cell system industry and the in-depth understanding of the downstream market to further commercialize the enterprise's fuel cell-related research, explore the overseas R&D layout, develop the core raw materials of the fuel cell stack, and form a strong competitive advantage through the cooperation of the industry, education, and scientific research institutes to reduce the manufacturing cost of the hydrogen fuel cell system.

Although there are some issues with public perception of hydrogen vehicles compared to EVs, the fact that HFCVs can be driven over longer distances gives HFCVs an optimistic outlook. In fact, due to the more extended drivable range, HFCVs have considerable potential in the commercial vehicle or public transportation sector. Governments or companies could use HFCVs as commercial vehicles in the early market stage to increase consumer familiarity with these vehicles. With a longer maximum drivable distance than other new energy vehicles, it is more appropriate for HFCVs to be used for commercial purposes (e.g., cargo or public transportation) until their mainstream use. In addition, to overcome the HFCV divide, it was felt necessary to reduce consumer uncertainty about hydrogen as a fuel, as expanding the infrastructure would take considerable time and cost. However, it is an essential factor in the proliferation of HFCVs. Therefore, HFCVs are mainly focused on the commercial vehicle or public transportation sector at the beginning of the market before penetrating the private vehicle sector of the general public (Moon et al., 2022).

#### *5.4.2 Market Operations*

China needs to follow or surpass the development of the United States to develop the hydrogen vehicle market. Although Japan was the first to introduce hydrogen energy vehicles, and the public has a basic knowledge of Japan's story, one of the more interesting points is that although the public agrees that Japan and its Toyota brand are doing an excellent job on this, the things that are more concerned about are the Chinese and American markets, and more worried about the technology in these two markets. The public is very excited about this as both are rapidly developing countries with high technology. If hydrogen cars are to be developed, the sector, government, business, social organizations, etc., must be compared with the U.S. In particular, brands such as Toyota, Tesla, BYD, etc. are getting attention.

From a cost-benefit perspective, institutional support measures such as education and awareness are more important than technical and economic incentives because they can change consumer perceptions faster and promote hydrogen energy. Specifically, education and awareness can help people understand the advantages of hydrogen energy and dispel misconceptions and concerns about it. This can be accomplished in various ways, such as conducting hydrogen energy education activities in schools and communities to improve the public's scientific literacy and awareness of the development of hydrogen energy. Hydrogen energy publicity through channels such as the media and social media to convey the advantages of hydrogen energy, such as environmental protection, safety, and efficiency. Technical and economic incentives, such as infrastructure expansion, R&D support, and subsidized support for the proliferation of HFCs while promoting the development of hydrogen energy, will take longer for the effects to be felt. For example, infrastructure development requires significant capital and time, R&D support requires a certain amount of technological accumulation, and subsidy support requires a certain amount of financial aid.

***5.4.3 Strategic Development***

A hydrogen fuel cell vehicle is a zero-emission transportation vehicle with great potential. However, it is still in the early stage of development. China must formulate corresponding supporting measures to promote the industrialization and commercialization of hydrogen fuel cell vehicles: first, develop mandatory laws and regulations, form relevant manufacturing chains, and expand production scale. Second, set clear promotion targets and timetables to guide market demand. Third, increase R&D investment to improve hydrogen fuel cell vehicles' performance and cost competitiveness. Fourth, accelerate infrastructure construction and enhance the hydrogen refuelling network to facilitate consumer use. These supporting measures will firmly guarantee the popularization of hydrogen fuel cell vehicles and help China achieve its energy transition and carbon neutrality goals.

The biggest problem in marketing hydrogen fuel cell vehicles is their poor fuel cell technology and battery durability. Therefore, the key to the marketization of fuel cell vehicles is the breakthrough of crucial fuel cell technology and the construction of an industrial chain. The China Energy Saving and New Energy Vehicle Industry Development Plan (2012-2020) emphasizes that it focuses on the research of critical technologies for new energy vehicles, such as fuel cell electric stacks, fuel cell engines, and related material technologies, and the establishment of cutting-edge technology innovation alliances for related industries and organizations. In this way, a new high-tech ally can be established to combine the industrial chain of teaching, research, and production to promote the development of fuel cell vehicles.

In LDA thematic clustering, the public is concerned about the development of the automobile market. Still, in k-means, the public hydrogen pairs are more worried about the impact of hydrogen energy on human society, such as whether hydrogen energy manufacturing and landing are controllable and future development. To address this concern, the design and production of hydrogen vehicles must follow strict safety standards. Hydrogen tanks must be robust and have devices to prevent leakage. Hydrogen lines must be well-sealed and comply with relevant safety codes. In addition, hydrogen-powered vehicles must be equipped with comprehensive safety systems to prevent fires or explosions in the event of an accident.

**6. Conclusion**

In this research paper, the public perception of hydrogen energy vehicle consumption from Bilibili was analyzed. Through Python and MySQL databases, online comments and real-time comments in the video were captured, and 42,035 words were obtained. In the exploratory analysis stage, the research hotspots in the field of hydrogen energy were demonstrated by a word-cloud map, indicating that hydrogen energy is a clean, efficient, and safe energy source with broad application. As a new type of transportation, hydrogen energy vehicles have received public attention. Then, general views and attitudes under different topics were identified through LDA topic modelling and sentiment analysis. The main themes are technology, energy, development, Japan, market, and transportation. The public has a positive attitude towards hydrogen energy's technology, energy supply, and development prospects. Still, there are some concerns about marketing hydrogen storage and transportation. The themes are further analyzed in more detail by combining sentiment analysis with K-means clustering. The results show that the public holds negative sentiments towards engine technology and battery durability of hydrogen vehicles, relatively positive views towards hydrogen energy, but sceptical towards hydrogen preparation and storage equipment. As a result, this paper proposes recommendations for technology R&D, market operations, and strategic development. In terms of technology R&D, more investment is needed to further improve fuel cell systems and core components' performance and cost competitiveness.

Market operation, education, and publicity must be strengthened to eliminate the public's misunderstanding and concern. Regarding strategic development, corresponding policy measures must be formulated to promote the industrialization and commercialization of hydrogen fuel cell vehicles. The results of these analyses have essential reference value for developing and promoting hydrogen vehicles. However, due to the limitations of data sources and the lack of independent samples, the results of this study need further validation. Future studies can combine other data and methods to provide a more comprehensive analysis of the perceptions.

**References**

Aminudin, M. A., Kamarudin, S. K., Lim, B. H., Majilan, E. H., Masdar, M. S., & Shaari, N. (2023). An overview: Current progress on hydrogen fuel cell vehicles. *International Journal of Hydrogen Energy*, 48(11), 4371-4388.

Barbir, F. (2009). Transition to renewable energy systems with hydrogen as an energy carrier. *Energy*, 34(3), 308-312.

Buttner, W., Rivkin, C., Burgess, R., Hartmann, K., Bloomfield, I., Bubar, M., ... & Moretto, P. (2017). Hydrogen monitoring requirements in the global technical regulation on hydrogen and fuel cell vehicles. *International Journal of Hydrogen Energy*, 42(11), 7664-7671.

Byun, H., Shin, J., & Lee, C. Y. (2018). Using a discrete choice experiment to predict the penetration possibility of environmentally friendly vehicles. *Energy*, 144, 312-321.

Chen, C., Chen, J., & Shi, C. (2018). Research on credit evaluation model of online store based on SnowNLP. *E3S Web of Conferences*, 53, 03039.

Choi, H., Shin, J., & Woo, J. (2018). Effect of electricity generation mix on battery electric vehicle adoption and its environmental impact. *Energy Policy*, 121, 13-24.

Elam, C. C., Padró, C. E. G., Sandrock, G., Luzzi, A., Lindblad, P., & Hagen, E. F. (2003). Realizing the hydrogen future: the International Energy Agency's efforts to advance hydrogen energy technologies. *International Journal of Hydrogen Energy*, 28(6), 601-607.

Hardman, S., & Tal, G. (2018). Who are the early adopters of fuel cell vehicles? *International*

*Journal of Hydrogen Energy*, 43(37), 17857-17866.

Hetland, J., & Mulder, G. (2007). In search of a sustainable hydrogen economy: how a large-scale transition to hydrogen may affect the primary energy demand and greenhouse gas emissions. *International Journal of Hydrogen Energy*, 32(6), 736-747.

Hienuki, S., Hirayama, Y., Shibutani, T., Sakamoto, J., Nakayama, J., & Miyake, A. (2019). How knowledge about or experience with hydrogen fueling stations improves their public acceptance. *Sustainability,* 11(22), 6339.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78, 15169-15211.

Kejun, J., Chenmin, H., Songli, Z., Pianpian, X., & Sha, C. (2021). Transport scenarios for China and the role of electric vehicles under global 2 C/1.5 C targets. *Energy Economics*, 103, 105172.

Kovač, A., Paranos, M., & Marciuš, D. (2021). Hydrogen in energy transition: A review. *International Journal of Hydrogen Energy*, 46(16), 10016-10035.

Lee, Y., Kim, Y. J., & Lee, M. C. (2021). Improving public acceptance of H2 stations: SWOT-AHP analysis of South Korea. *International Journal of Hydrogen Energy*, 46(34), 17597-17607.

Lopez Jaramillo, O., Stotts, R., Kelley, S., & Kuby, M. (2019). Content analysis of interviews with hydrogen fuel cell vehicle drivers in Los Angeles. *Transportation Research Record*, 2673(9), 377-388.

Marchenko, O. V., & Solomin, S. V. (2015). The future energy: Hydrogen versus electricity. *International Journal of Hydrogen Energy*, 40(10), 3801-3805.

McPherson, M., Johnson, N., & Strubegger, M. (2018). The role of electricity storage and hydrogen technologies in enabling global low-carbon energy transitions. *Applied Energy*, 216, 649-661.

Moon, S., Kim, K., Seung, H., & Kim, J. (2022). Strategic analysis on effects of technologies, government policies, and consumer perceptions on diffusion of hydrogen fuel cell vehicles. *Energy Economics*, 115, 106382.

Mourato, S., Saynor, B., & Hart, D. (2004). Greening London's black cabs: a study of driver's preferences for fuel cell taxis. *Energy Policy*, 32(5), 685-695.

Ng, A., & Zheng, D. (2018). Let's agree to disagree! On payoffs and green tastes in green energy investments. *Energy Economics*, 69, 155-169.

Nie, X., Wei, X., Xia, Q., & Zhou, M. (2021). Customers' purchase intention for hydrogen vehicle and industrial agglomeration: Evidence from Jiangsu Province, China. *International Journal of Hydrogen Energy*, 46(34), 18011-18019.

Noussan, M., Raimondi, P. P., Scita, R., & Hafner, M. (2020). The role of green and blue hydrogen in the energy transition—A technological and geopolitical perspective. *Sustainability*, 13(1), 298.

O'Garra, T., Mourato, S., & Pearson, P. (2005). Analyzing awareness and acceptability of hydrogen vehicles: A London case study. *International Journal of Hydrogen Energy*, 30(6), 649-659.

REN21 (2023). *Renewables 2023 Global Status Report*. Paris: REN21 Secretariat.

Wu, Z., He, Q., Li, J., Bi, G., & Antwi-Afari, M. F. (2023). Public attitudes and sentiments towards new energy vehicles in China: A text mining approach. *Renewable and Sustainable Energy Reviews*, 178, 113242.

Zhang, D., Wang, J., Lin, Y., Si, Y., Huang, C., Yang, J., ... & Li, W. (2017). Present situation and future prospect of renewable energy in China. *Renewable and Sustainable Energy Reviews*, 76, 865-871.

Zhang, L., & Gao, J. (2016). Exploring the effects of international tourism on China's economic growth, energy consumption and environmental pollution: Evidence from a regional panel analysis. *Renewable and Sustainable Energy Reviews*, 53, 225-234.

Zhao, H. (2022). Design and Implementation of an Improved K-Means Clustering Algorithm. *Mobile Information Systems*, 2022, 6041484.

Zhao, T., Liu, Z., & Jamasb, T. (2022). Developing hydrogen refueling stations: An evolutionary game approach and the case of China. *Energy Economics*, 115, 106390.

**Figure 1. Framework Diagram**

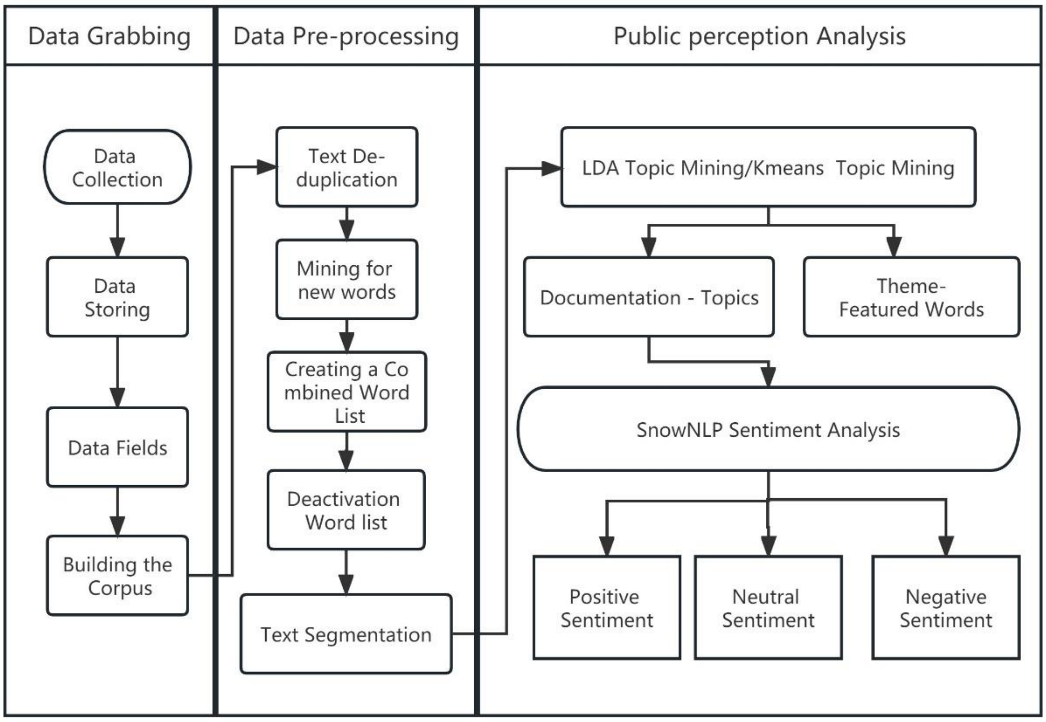
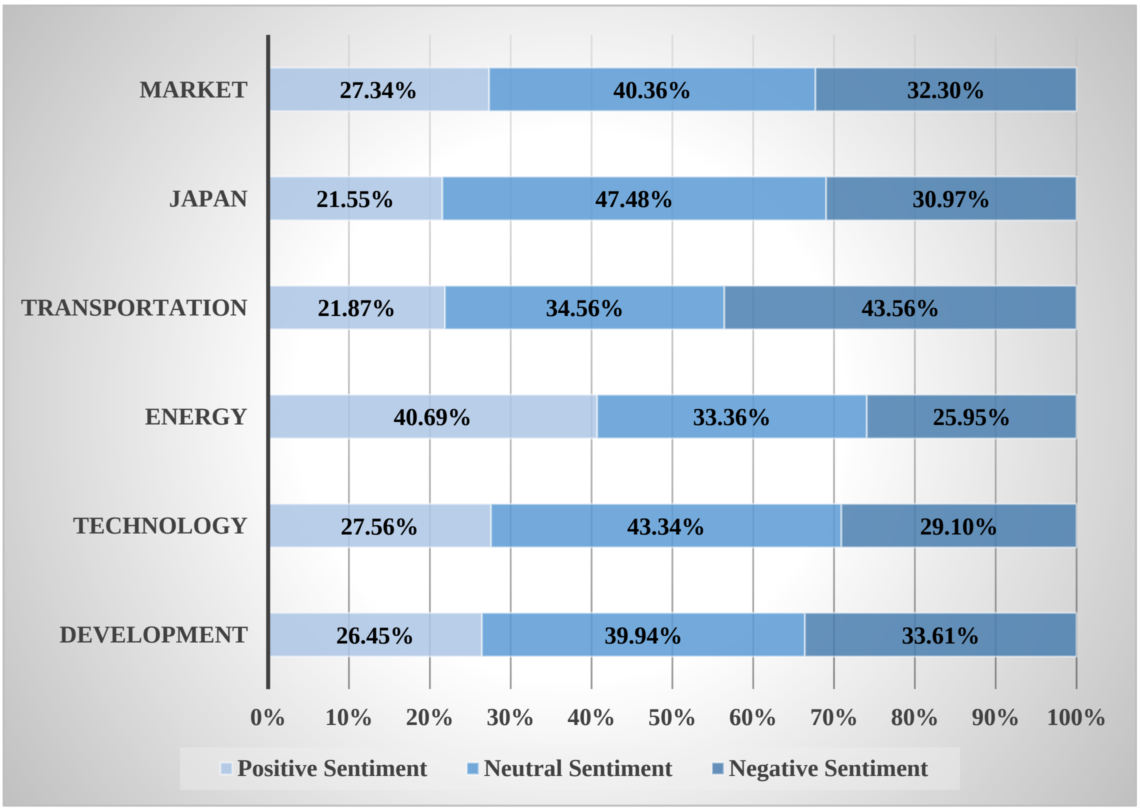



**Figure 2. Word Cloud Diagram**


Note: 能源（Energy, 问题（Issues), 日本(Japan).

**Figure 3. Sentiment Analysis Based on LDA Themes**



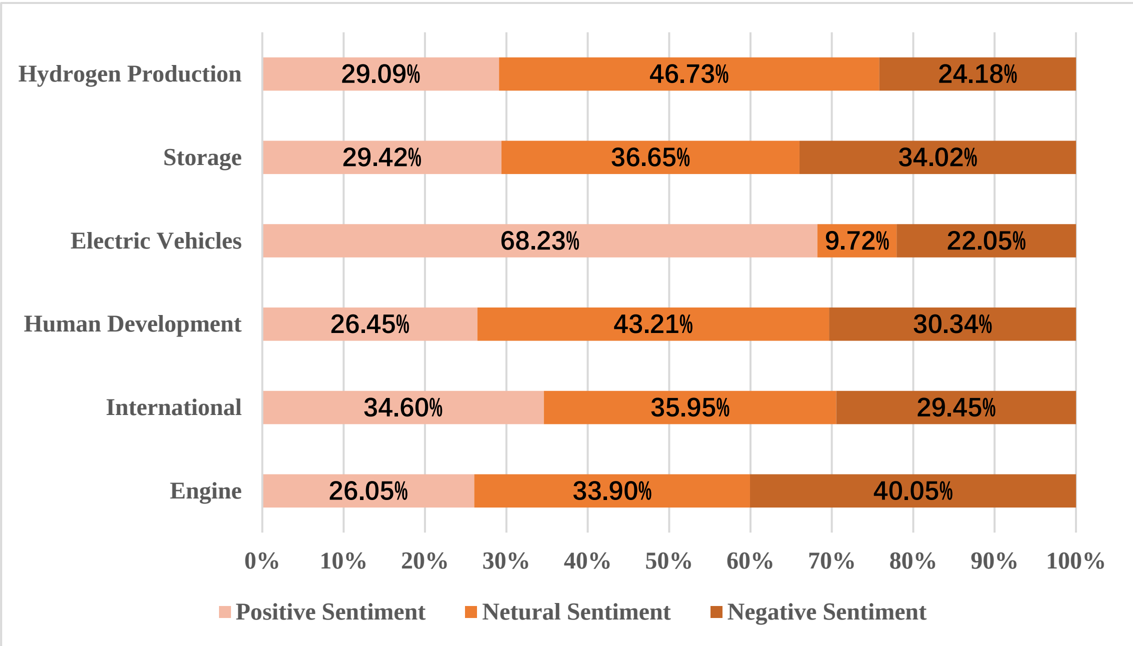




**Figure 4. Optimal K Values Confirmed by the Elbow Method**

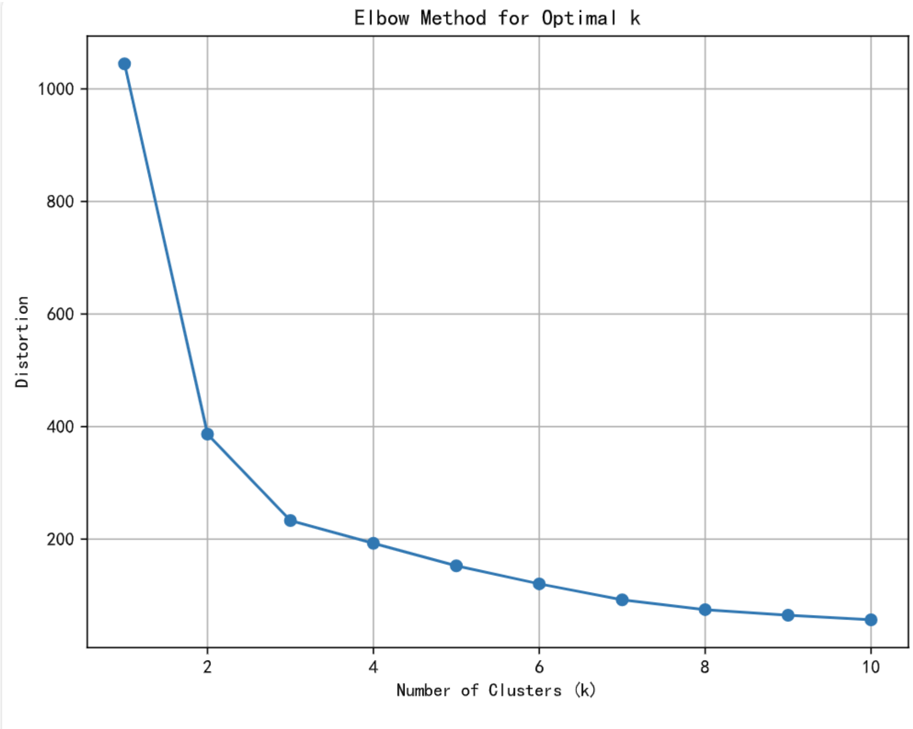



**Figure 5. K-means Thematic Clustering**

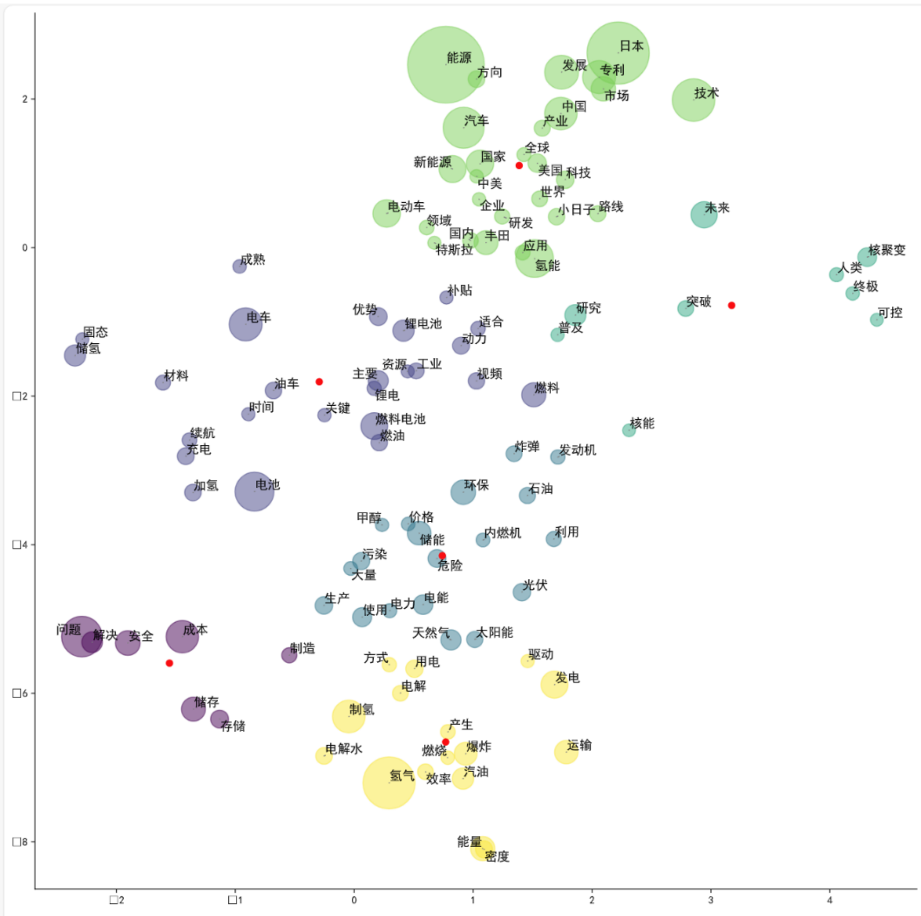



**Table 1. LDA Themes**

| Number of Themes | Characteristic Word | Theme Name |
| --- | --- | --- |
| 0 | Hydrogen, Japan, batteries, hydrogen energy, automobiles, cost, China, hydrogen production, fuel cells, power generation | Technology |
| 1 | Energy, Issues, Batteries, Hydrogen, Cars, Trams, Costs, China, Japan | Energy |
| 2 | Issues, Batteries, Hydrogen, Cars, Trams, Costs, Development, Electric Vehicles, Fuel Cells, Environmental Protection | Development |
| 3 | Japan, Hydrogen, Cars, Trams, China, Electric Vehicles, Toyota, Research, Stuff, Feelings | Japan |
| 4 | Hydrogen, Automotive, Patent, China, Development, Electric Vehicles, National, U.S., Toyota, Markets | Market |
| 5 | batteries, hydrogen energy, transportation, cost, hydrogen production, fuel cells, fuel, safety, storage, energy | Transportation |