

# ***Analysis of Public Opinion and Sentiments Towards Gap Year in China Based on Weibo Data Mining***

**Pengyi Jiang<sup>1,a,\*</sup>, Lingyun Lu<sup>2,b</sup>**

<sup>1</sup>*School of Journalism and Communication, Liaoning University, Liaoning, 110000, China*

<sup>2</sup>*School of Literature and Journalism, Shandong University of Finance and Economics, Shandong, 250000, China*

*a. 165218929@qq.com, b. 503352137@qq.com*

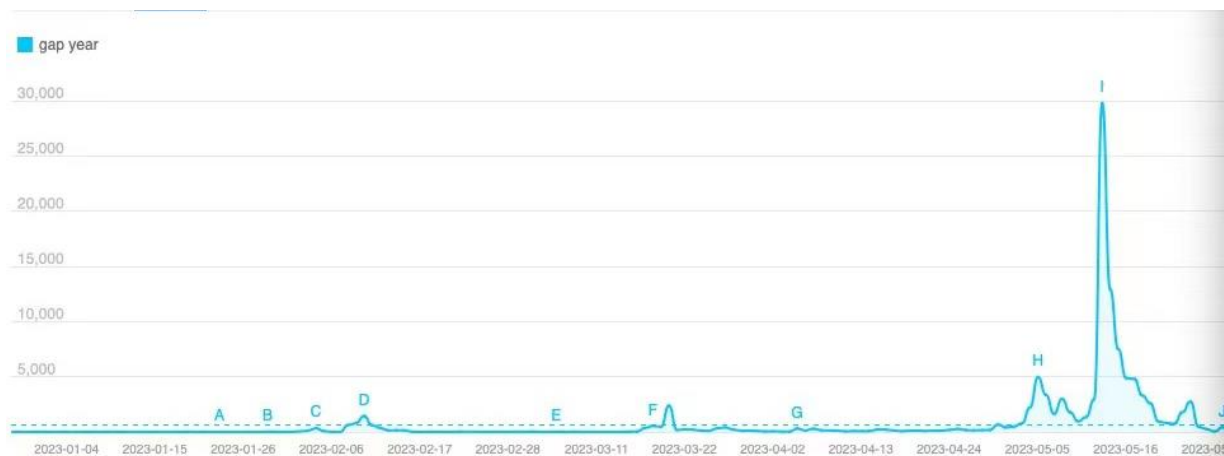
*\*corresponding author*

**Abstract:** In May 2023, there was a heated discussion about gap year popped up on Chinese social media. To better understand this phenomenon, we investigated public opinion on gap year in China based on Weibo in this study. A relatively novel study approach in this field, social media data mining, was proposed in this essay. Then, we conducted some analyses based on the collected data, especially the sentiment of each individual who participating in the heated discussion. Last, we found that the discussion was heavily focused on the conflict between gap years and life schedule recognized by public, and negative emotions almost dominated public opinion. Most previous study about Gap Year in China focused on the small samples, the factors that individuals consider as barriers to participating in a gap year, and the motivation to participate in the gap year. Different from them, the result from this study can provide a new and relatively objective insights with a larger sample size into the real situation and public opinion of gap year in China.

**Keywords:** Gap Year, public opinion, Sina Weibo, sentiment analysis, data mining

## **1. INTRODUCTION**

The concept of Gap Year was first proposed in the UK in the 1960s, and had grown popularity in western countries [1,2]. However, due to cultural diversity, taking a Gap Year as well as interrupting the education career was not well accepted in China during the past decades [3,4]. But surprisingly, in May 2023, discussions about Gap Year erupted on Chinese social media. Baidu Index, the Baidu's data analysis platform based on Baidu's massive netizen behaviour data, shows that Baidu search index of Gap Year reached a peak of 29787 on May 4th, which was 2.2 times higher than that at the same period last year. The information reporting index of Gap Year from Baidu data also increased by 1396 times year on year.



**Figure 1.** Baidu index of keyword “gap year” on June 21.

Furthermore, netizen have been engaged in heated discussions on Sina Weibo, one of the largest social media in China. Most of the young Chinese involved in the discussion expressed their attitude towards Gap Year. They even proposed a series of buzzwords to describe "the version of Chinese Gap Year ", such as “Gap Day”, “Gap Hour”, “Gap Minute”, and even “Gap Second” to mock themselves. China Youth Daily, one of the official mainstream media, reported that 73.1% of college students said that some classmates around them proactive chose to delay graduation, which can be seen as a substitute for a real Gap Year [5].

Although the phenomena mentioned above indicated that the concept of Gap Year had attracted intensive attention, the performance of this activity in China is significantly different from that of Western countries [6]. For college graduates who interrupted their educational career in China, preparing for civil servants’ examinations is their main purpose. In this context, they actually took a Gap Year that extended their studies. It is important to note that most of the participants in the discussion on Weibo only expressed their views and few implemented a “real Gap Year”. The contrast between the high heat discussions and low practice rates reflects the real attitude and concerns of young people in China towards the Gap Year which deserves attention.

In previous reports, public opinion towards Gap Year was studied mostly in western countries, with the traditional research methods including telephone interviews, questionnaire surveys, etc. With the rise of the internet, young people are increasingly inclined to express their opinions on social media, which provides extensive data and efficient methods for the study of public perspectives. In order to fill out the research gap, this study aims to objectively investigate the public opinion and sentiments towards “Gap Year” in China with text mining, based on the data collected from Sina Weibo.

## 2. LITERATURE REVIEW

### 2.1. Gap Year in Western Countries

The concept of Gap Year is believed to first be emerged in the UK in the 1960s [2]. The birth of the concept in the UK has its historical origin, which can be traced back to the aristocratic tours and the Oxford Cambridge Examination System in the 18th and 19th centuries [7,8]. Subsequently, this behaviour gradually spread and developed in other English-speaking countries such as Australia, North America, and other regions such as China [9,10]. Today, taking a Gap Year has been promoted in newspapers, books, and the internet as a means of developing themselves to be more mature, independent, confident, and potentially more employable. There is an increasing trend for young people to delay their higher education and choose to take a Gap Year as a substitute [11].

## 2.2. The Rise and Development of Gap Year in China

The rise and development of Gap Year in China was relatively late in China. It was not until 2009 that Chinese people became aware of the concept from a book entitled 'My Late Arrival Gap Year'. Thereafter, taking a Gap Year has gradually become popular in the last decade in China, with more and more participants sharing their Gap Year experiences on the Internet. In 2014, the China Youth Foundation China Gap Year Public Welfare Fund was established, which is the first special fund in China to support the implementation of the Gap Year for young people. The fund aims to spread the concept of Gap Year among young people in China and help them to explore the world and understand themselves through the implementation of the China Gap Year Plan. Aiming at students aged 18 to 28 in universities across the country every year, the fund provides rewards and funding for outstanding programs submitted and selected, helping them complete their three-year to one-year gap plan in China or overseas [12].

With China's economic and social development and the promotion of the government, universities and other organizations, the rise of the gap year in China is an inevitable trend and a positive phenomenon in line with the trend of youth development. However, due to the lack of supporting systems in China, the implementation of Gap Year faces some difficulties and risks [3,12]. For example, if fresh graduates did not seek employment and further education in a timely manner, they may encounter some problems or lose some opportunities. It is also not easy to take a break from school during the university stage. In addition, the parents' recognition of Gap Year is still relatively low, which greatly hinders the further development of gap year in China.

## 2.3. Mining Public Opinion and Sentiments Towards Gap Year in Social Media

Previous studies focus on Gap Year in Western countries while there are few empirical studies on the current situation of Gap Year in China. Two papers were found to explore the rise, motivations, identities of participants and the constraints of the Gap Year in China. Wu use netnography and in-depth telephone interview approaches to analyze the identities and motivates of Chinese Gap Year takers, and analyze the barriers to Gap Year briefly [4]. It is revealed that the barriers include the Chinese higher education system, the plentiful labour force and the tradition of natural transition by family and society. Wu use the questionnaire focus on analyzing the constrains of Gap Year in China and found that there are six factors have prevented the Chinese from implementing Gap Years: "financial and time concerns", "social pressure", "social responsibility", "personal skills and safety concerns", "competing interests and effort" and "travel companions" [13].

According to the research of the previous literatures, it is noted that existing research also lacks research that focuses on the sentiments of public opinion about the Gap Year, especially the sentiment analysis of public opinion on Weibo, the biggest social media in China [14]. According to the 2022 annual report released by Sina, the number of monthly active users of Weibo reached 586 million, which can provide a sufficient sample capacity for the study. Furthermore, most Weibo tweets are the direct attitude expression from individuals, which enables us to explore the real situation of Gap Year in China and gain insight into the underlying reasons. We captured 9,471 discussions about Gap Year from Weibo from April 21th, 2023, to June 21th, 2023. The data shows that more and more people desire to take a Gap Year, indicating the increasing recognition of Chinese youth towards Gap Year. Unfortunately, it can be observed that the proportion of negative posts about Gap Year is also growing. Therefore, to further analyze the ratio of positive to negative emotions and the reasons behind this phenomenon is of great significance for understanding Chinese youth's current mental state.

More specifically, this research aimed to: (1) collecting the public opinion about Gap Year on Chinese social media, Sina Weibo. (2) Analyzing the sentiment of Chinese people conveyed on

Weibo on the issue of the Gap Year. (3) Identifying the motivations and thoughts behind this emotional expression. Firstly, this study can fill the gaps in previous research on sentiment analysis of Gap Year on Chinese social media. Secondly, by providing a reference for government departments and thus setting up a gap so that young people can release their social and labour pressure as well as grow up with the pace of the times. Finally, this study is expected to arouse the attention of some parents, educational experts, and entrepreneurs who can help Gap Year to advance practice in a way that adapts to China's national conditions in the future.

Therefore, focusing on what Weibo users were discussing about the gap year in China and the emotions behind it, two research questions are posted:

RQ1: What is the most frequently mentioned content when people talk about China's gap year on Weibo?

RQ2: What is the most prominent emotion in discussion when people talk about China's gap year on Weibo?

### 3. METHODS

#### 3.1. Data Collection

To better understand the public opinions and discussions about Gap Year in China, researchers searched with the keyword of "Gap Year" on Sina Weibo and downloaded all texts of users' posts from April 22 to June 22. A Chinese crawler software "Octopus Collector", which had been proved to be efficient and convenient in Weibo data collection, was used to collect information of users' posts containing information of User ID, posting time and content. After cleaning and deleting irrelevant data (advertisements, meaningless messy texts, etc.), a total of 9471 posts are left for the following study.

We set the timeframe between April 22 and June 22 because at the beginning of the study, it happened to be the National College Entrance Examination season and graduation season in China, when people's anxiety and stress concerning entering university or finding jobs led to a heated discussion about Gap Year. According to Baidu Index, Chinese netizens' behaviour on the topic of Gap Year has surged since May 2023, as shown in Figure 1. Mining and analyzing the user-generated contents on Sina Weibo will help us fully understand the public opinion and sentiment on the topic of Gap Year.

#### 3.2. Data Analysis

The study focused on the content of Weibo posts, trying to address the RQ1——what people talked about under the topic of gap year. As textual data on social media platforms are not coded, they are unstructured, and users can combine words and phrases freely to convey their ideas. Thus, to further analyze the unstructured raw data and make out their meanings, R and its packages are used to process the data. JiebaR package was applied to accomplish the Chinese text segmentation and keyword extraction. To raise the accuracy of jiebaR's segmentation, researchers improved the stopwords dictionaries to better fit in the study. Then the frequency of keyword occurrences was counted. Words with a frequency greater than 100 were considered high frequency and were subsequently translated into English to build a database. After that, wordcloud2 package was used to make a word cloud map of high frequency words to figure out the most frequently mentioned words when people talk about Gap Year on Weibo.

Furthermore, the study dug deeper into the emotions behind the content to answer RQ2, attempting to find people's sentiment on Weibo discussion. Researchers manually classified the posts into three emotional tendencies: (1) positive (happy, excited, hopeful, etc.), (2) negative (sad, anger, upset, etc.) and (3) neutral (posts without a clear emotional tendency). To improve the accuracy of the





Table 1: (continued).

11	需要	need	760
12	感觉	feel	751
13	实习	internship	741
14	day	day	705
15	知道	know	698
16	宝宝	baby	660
17	喜欢	like	658
18	事情	thing	647
19	适合	fit	636
20	想要	would like	633

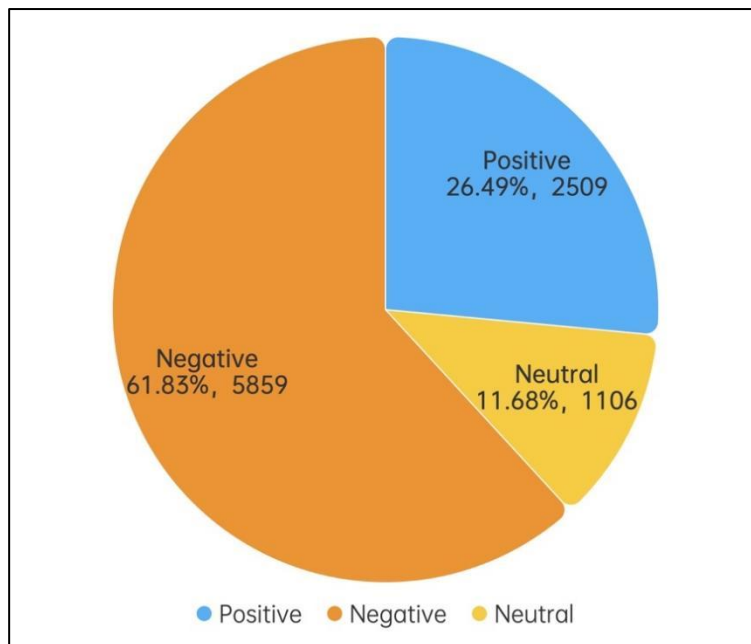
To address RQ1, we counted and produced word frequency data and word cloud map to figure out what people were talking about on Weibo. In the word cloud map shown in Figure 2, the font size of a word is positively correlated with its frequency. The top 20 high-frequency words are shown in table 1 as examples. Combining word cloud map and word frequency statistics, it was found that when people discussed gap year in China on Weibo, the most mentioned words included “want”, “none”, “work”, “China”, “a year”, “living”, “now”, “life”, “time”, “graduation”, “need”, “feel”, “internship”, “day”, “know”, “baby”, “like”, “thing”, “fit”, “would like”, etc. Referring to Ji’s classification of discussions on Weibo[15], the high-frequency words were mainly classified into three categories labelled as “factual attribution”, “causal attribution” and “emotional release”.

The words of factual attribution are few but concentrated, involving words such as “none” (n=2258), “China” (n=1839), “a year” (n = 1345), etc. Analyzing the data further, it suggested that most posts involving these words tried to state the fact that taking a gap year is difficult in China, such as “今天学到一个新概念, 中国没有 gap year (Learnt a new concept today, there is no gap year in China)”. But a small number of posts refuted such view, arguing that Chinese can also choose to take a gap year.

Words of causal attribution mainly focused on explaining or complaining about Chinese people’s inability to take a gap year, which extended Wu’s conclusion on six constraints factors of gap year in China in the context of Weibo [13]. Among them, “financial and time concerns” and “travel companions” are also found on Weibo discussion, reflected by the words “money” (n = 352), time (n = 927) and friend (n = 524). In addition, the factor “social pressure” noted by Ji was interpreted more clearly as the limitation of the “social clock”. The Weibo data revealed that people often used words related to the life schedule when talked about gap year in China, containing “work” (n = 2050), “graduation” (n = 887), “internship” (n = 741), etc. This may be explained by the “social clock”, just as the Chinese proverb goes “Do the right thing at the right age”. Activities that deviate people from social schedule, like gap year, is considered useless and unacceptable to the wider community, such as parents and employers. Defying social clock to take a gap year means taking on the enormous strain of family relationships, job opportunities, economics, and criticism from traditional values. Ji’s other three constraints factors of “social responsibility”, “personal skills and safety concerns” and “competing interests and effort”, however, did not significantly affect the results.. This finding may indicate people’s updated focus on limiting factors in the context of contemporary social media.

It was also notable that words of emotional release were significant in the word cloud map. Gap year in China was associated with sentiment of “happy” (n = 616), “anxious” (n = 499), “hope” (n = 488), etc. Besides, many expressions that implied emotional tendency were also classified as emotional release including the most frequent word “want” (n = 2661). Emotional release took up a large part of the discussion on Weibo, which was ignored in previous studies due to data sources and research methods. The findings pointed out that Chinese netizens have mixed feelings, involving imagination and longing for the good image of gap year, hesitation to participate in gap year and dissatisfaction with the current situation.

#### 4.2. Public Sentiment: Negative Emotion Dominance



**Figure 3.** Sentiment classification chart for Weibo posts.

**Table 2.** Examples of sentiment classification sheet for Weibo posts.

Classification	Number	Posts (Chinese)	Posts (English Translation)
Positive	1	今年,必是我的gap year !冲!	This must be my gap year! Here we go!
	2	我宣布 从此刻开始 我将进入人生中的第一个gap year 我要大玩特玩 大睡特睡 我不去上课啦!我要睡觉!	I announce that from now on, I will enter the first gap year of my life. I'm going to have a lot of fun and sleep. I won't go to class! I want to sleep!
	3	想有个gap year去体验我想过的生活,去我想去的地方,去跟不同行业、不同年龄的人交流,去汲取千奇百怪的思想丰富我的大脑这真的太吸引人了!	Want to have a gap year to experience my dreaming life, to go to my dreaming places, to communicate with people of different industries and ages, to absorb strange ideas to enrich my brain. This is so attractive!
	4	我想要gap year~实在不行每天的gap time也行大学生活许多开心的时光也是因为拥有gap time和朋友“载歌载舞”的疯闹	I'd like a gap year, or a gap time every day if I have to. Many of the happiest moments of college life are due to having gap time and “singing and dancing” crazy with friends.
	5	那就享受一下gap year 吧^^	Then enjoy the gap year^^

Table 2: (continued).

Negative	1	买了一些破烂,做梦有gap year,然后回到现实	Bought some junk, dreamt of a gap year, and then came back to reality.
	2	确实挺讽刺的,大谈特谈gap year。还说什么不用几年啦,三到六个月也行,什么年龄段都可以gap,我来考考你怎么准备gap,我又从gap中得到了什么.....知道坐在这儿的都经历了什么吗	It's kind of ironic to about a gap year, saying that it won't be a few years, three to six months are also fine, and one can gap whenever. Let me ask you how to prepare the gap, and what I get from the gap... Do you have any idea what people sitting here went through?
	3	gap尼玛的year	Gap your fucking year.
	4	还谈gap year,极其搞笑。	Yet talking about gap year, how ridiculous.
	5	中国人是不是真的没有gap year去年没考上研 家里人都想让我再考编去年离过线差了40分我觉得我还要再试一次但是爸爸发了好大一通火	Is it true that Chinese do not have the gap year? I failed to enter the graduate school last year, and my family all wanted me to take the civil service examination. I missed the mark by 40 points last year, and I wanted to try again, but Dad got so angry.
Neutral	1	gap year,我也算是经历了	Gap year. I kind of went through it.
	2	为什么中国大学生很少选择间隔年(gap year)?	Why do Chinese college students rarely choose to take a gap year?
	3	好像找不到gap year准确的中文翻译	I can't seem to find an accurate Chinese translation of gap year.

To answer RQ2, all Weibo posts were used for sentiment classification. The pie chart in Figure 3 shows the proportion of positive, negative and neutral emotions in the public discussion about gap year in China, and table 2 provides some typical cases of the emotion classification. The inquiry illustrated that negative emotions dominated Weibo's discussion, accounting for 61.83% of all Weibo posts. Only 26.49% of posts delivered positive emotional tendency. What is more, 11.68% of posts were found with no obvious sentimental expression, which can be mainly attributed to factual attribution. There is strong evidence that negative emotions such as disappointment, anger, sadness shaped both the form and content of discussion, and the topic of gap year on Weibo became a public place for people to vent their grievances collectively. Such sentiment is strengthened through the gathering of viewpoints and people's interaction, even interspersed with criticism of the government and society. It's also worth noting that the TEM-8 (Test for English Majors-Band 8, a national English test held by ministry of education) was held during the timeframe of the research, whose listening section talked about gap year, leading to an emotional flashpoint on Weibo. Many users expressed negative or critical words under the hashtag "TEM-8 on Weibo, such as "gap year 听得我心累,哪来的gap啊?哭了 (The conversation made me frustrated, where is my gap year? Crying).

## 5. CONCLUSION

Different from previous studies concerning gap year in China, this study use a relatively novel approach of social media data mining to explore the research gap about Gap Year in China. A large number of samples collected from Weibo, enable researchers to get more extensive and direct perspectives than previous studies which use traditional survey methods. A range of research methods on the online posts give diverse insights into the public opinion about gap year in China.

Our research mainly draws three conclusions:



(1) The contents of Weibo posts were found to be classified into three categories, “factual attribution”, “causal attribution” and “emotional release”. Though with some debates, most posts tended to believe that it is harder for Chinese to take a gap year than Westerners.

(2) “Financial and time concerns”, “travel companions” and “social pressure” were found to be the most significant factors preventing Chinese people from taking a gap year. Overly tight life schedules and societal expectations made it difficult for people to take a gap year, but also made people more eager to take one to relieve their stress.

(3) Negative sentiment involving upset, pessimism, anger, dominated discussions about gap year on Weibo. Many Weibo users complained and even criticized the high-pressure environment of Chinese society that kept them from taking a gap year, and such emotion-releasing posts occupy a large part of the public opinion on Weibo.

Our study better summarizes the public opinions and emotions on gap year in China reflected in Weibo data, filling the gap of previous research on public emotions, especially the negative emotion research, and promotes the application of data mining and sentiment analysis based on microblog data under such topics, which can provide inspiration for analyzing public opinions and emotions in the context of social media in the future. The research, however, has limited at the practical level. we only analyze current public perspective and dissatisfaction on the gap year in China, which limited to just one platform. The situation and solutions of gap year in China demands much further study. This topic may be worthwhile to extend our study, collect more data from various platforms, lookout for the data from various industries, and dig deep the actual social situation in China, to explore the real dilemma of gap year in China, and a gap year that is more coincide with China condition.

## References

- [1] Hillman, K. 2005. *The first year experience: The transition from secondary school to university and TAFE in Australia (Research Report No.40)*. Melbourne, Australia: Australian Council for Educational Research.
- [2] Jones, A. 2004. *Review of Gap Year provision (Research Brief No. RB555)*. London, England: Department for Education and Skills.
- [3] Li Y. 2020. *Gap Year in the perspective of evolution psychology [J]. The Guide of Science & Education*, 404(08):168-169+184.
- [4] Wu M-Y, Pearce P., Huang K, et al. 2015. ‘Gap Year’ in China: views from the participants and implications for the future [J]. *Current Issues in Tourism*, 18(2): 158-174.
- [5] Wang P., Sun L. 73.1% of the surveyed students said that some of their classmates chose to postpone graduation. *China Youth*, 2023-05-09
- [6] Sun Y. 2022. *Gains and losses of Gap Year [N]. People’s Daily Overseas Edition*, 2022-05-12 (008).
- [7] Heath, S. 2004. *Widening the gap: pre-university gap years and the ‘economy of experience’*. *British journal of sociology of education*, 28, 89-103.
- [8] Simpson, K., 2004. *Broad horizons geographies and pedagogies of the gap year*. Unpublished PhD thesis. University of Newcastle.
- [9] Curtis, D. D., Mlotkowski, P., Lumsden, M., & NCVER. 2012. *Bridging the gap: who takes a gap year and why longitudinal surveys of Australian Youth*. Adelaide: National Centre for Vocational Education Research.
- [10] Qian, R. 2013. *Benefits and constraints of gap year*. (M.S.). Kent State University, Kent.
- [11] O’Shea, Joseph. 2011. *Delaying the academy: A Gap Year education*. *Teaching in Higher Education*, 16: 565-577.
- [12] Gong H. 2016. “Gap Year” for Chinese Young People. *China Daily*, 2016-08-17.
- [13] Wu M-Y., Pearce P. 2018. *Gap time and Chinese tourists: Exploring constraints*, *Current Issues in Tourism*, 21:10, 1171-1186.
- [14] Wang Y, Li H., Wu Z. 2019. *Attitude of the Chinese public toward off-site construction: a text mining study*. *Journal of Cleaner Production*, 238:117926.
- [15] Ji, P. (2016). *Emotional criticism as public engagement: How weibo users discuss “Peking University statues wear face-masks”*. *Telematics and Informatics*, 33(2), 514-524.