

Big Data Analysis and Decision Making of Chinese Paper-Cutting

Cheng Yang^{1,*}, Lingang Wu², Dan Li³ and Zhichao Zhang²

¹Zhejiang university city college, China

²Computer science college, Zhejiang university, China

³Art college, Changzhou university, China

Keywords: *paper-cutting, big data analysis, intangible cultural heritage, online comment*

Abstract. Chinese paper-cutting is a traditional folk craftsmanship that cuts and carves patterns on paper. In urban civilization, how to protect and inherit traditional paper-cutting techniques is an urgent issue to consider. The continuous integration of social science and computing science provides a technical means for big data to analyze social issues. Use the crawler program to gain big amount of online comments about paper-cutting. Apply the Latent Dirichlet Allocation (LDA) to extract subjects of online comments. By using the comment extraction function in the AipNlp module, we can discover dissatisfaction of users, find their needs, and collect as well as filter improved strategies.

Introduction

Chinese paper-cutting is a traditional folk craftsmanship that cuts and carves patterns on paper.^[1] Having a extensive mass base in traditional Chinese society and rich cultural connotations and historical information, Paper-cutting have blended in the social life of the people and become an important part of folk customs.

As a folk art, paper-cutting has a broad public base and lasting vitality, which is mostly taught orally by masters. Due to the changes in the modern social environment, most folk craftsmanship has gradually disappeared in the past 100 years. While paper-cutting still has many participants nowadays mostly because of its unique artistic language and cultural connotation. However, with the social development, paper-cutting is gradually losing original cultural ecology and social environment. It's an urgent issue to consider that how can traditional paper-cutting techniques adapt to the needs of modern people in urban civilization.

In 2009, China paper-cutting was selected for the UNESCO Intangible Cultural Heritage List. The "Intangible" in the concept of intangible cultural heritage is relative to "materiality", but it is not completely insulated from the latter, while referring to its creative activities existing as non-material form that focus on the spiritual domain as well as the result of these activities^[2].

The continuous integration of social sciences and computing science provides a technical means for big data to analyze complex social problems. Big data collection and analysis makes it possible to combine quantitative and qualitative research in intangible cultural heritage work. And the widespread use of computers and the Internet has brought convenience for large-scale social data analysis and statistics. There are billions of user messages and comments on various websites including WeChat, Weibo, Facebook, Twitter and Post Bar, which can be identified and analyzed by existing text analysis software. For example, LIWC text analysis system can identify more than 60 human emotions such as joy, disgust and sadness through text.^[3]

The Popularity of Chinese Paper-cutting

Whether in the streets worldwide or in Hollywood movies, you can see all kinds of "Chinese elements", the most significant elements of which are Chinese intangible cultural heritage. The rising international popularity of Chinese intangible cultural heritage also strongly reflects the awakening process of Chinese culture consciousness.

Have studied the value of intangible cultural heritage, we can easily find its historicity, rarity and nationality. And the “international awareness of intangible cultural heritage” is more a symbol of it’s significant cultural inheritance and transmission .^[4] There are 12 kinds of chinese intangible cultural heritagethat have never been mentioned in the Google Book Corpus (AD 1500-2008, 8.11 million English books). It shows that the contemporary popularity and publicity of them are extremely insufficient, proving that they are in a dangerous state and need urgent attention and protection.

Google is one of the world's largest search engines, and Google Trends is an analysis product based on Internet search records launched by Google. Baidu, a well-known search engine in China, has also launched a similar product, which is called Baidu Index. By using "paper-cut" as a search term through the search tools such as Google Trends and Baidu Index in some large-scale databases, we can analyze the Internet attention of the theme.

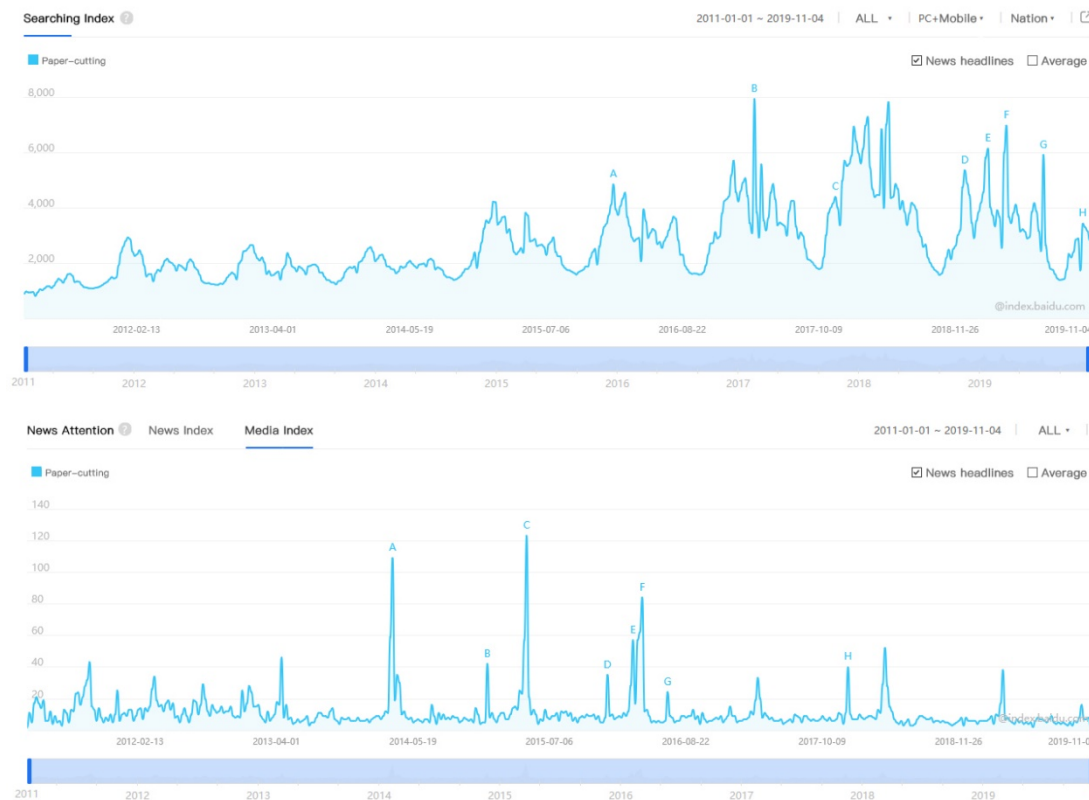


Figure 1. Baidu search index and media index of keyword "paper-cut" from 2011 to 2019

From Figure 1 we can see the number of concerns about “paper-cutting” from January 2011 to November 2019. User search volume fluctuated year by year, from the initial search volume of more than 900 to the volume of 3000 in November 2019. While the media attention rate is generally declining: The average number of reports from 2011 to 2013 was generally high, and the highest peak was seen from 2014 to 2016; from 2017 to the present, the average number of reports fell, and the peak value also dropped significantly. Have analyze the attention rate that the media and netizens paid to "paper-cutting" throughout the past five years, it can be found that the mass audience has begun to actively contact with "paper-cutting culture” under the influence of the mass media.

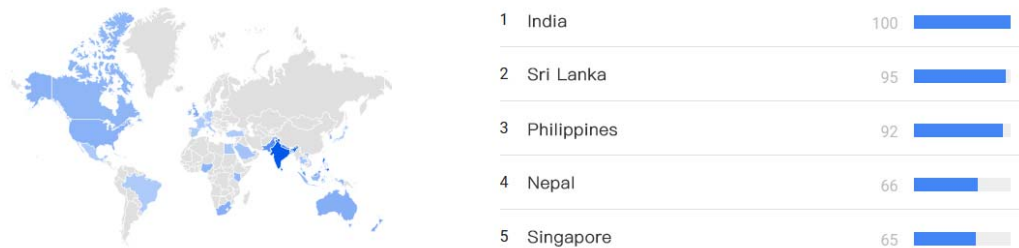


Figure 2. Geographical distribution of users searching "paper cutting" on Google Trend

The highest value of the user search volume appears around the Spring Festival and the Lantern Festival. The fourth quarter of each year is also the peak period of searching, and the period of July to August annually is the least. The media attention reached its peak during the Spring Festival, which was many times of that in other periods of the year. The number of reports in one day of February can reach seventy or eighty, while in other periods it is generally less than ten. Similar user and media attention patterns also appear in other years.

Figures 2 reflect the geographical distribution of users searching "paper cutting" in the world. It can be seen that China's surrounding areas, including South Asia and Southeast Asia, and North America areas have a high searching popularity, which shows the paper-cutting culture spread not only to neighboring countries, but also to Western developed countries. In China, the highest search popularity areas are concentrated in North China and South China.

The number of articles containing the theme of "paper-cutting" or titled "paper-cutting" basically increased year by year before 2016, but there was a slight fluctuation and fall in the past three years. In Google Scholar, the overall trend is the same as that of CNKI except for an abnormal decline due to the fact that the data of the past two years has not been updated. The number of articles on "paper-cutting" has generally increased in the past decade, indicating that the academic attention of "paper-cutting" is also rising overall. However, attention has become saturated during the past three years, which requires new methods and ideas to introduce relevant researches.

Needs Analysis Based on Big Data

When the global market entered excessive era, the relative weak position of user resources in the entire industry chain was completely changed, and user resources gradually became the dominant force in the entire transformation of the industry chain. Therefore, it is necessary to analyze current market need based on data of paper-cutting users. Having large data volume and wide coverage, big data can increase user participation in the development process of intangible cultural heritage products, reduce the time and labor cost of user research, and help researchers understand complex data more quickly and more accurately to find user need faster.

The data used in classical empirical analysis is generally structured data, while the comments big data analysis is based on unstructured data. It's a complex issue to converting massive amounts of unstructured data into structured data, and the quality of the transformation affects subsequent analysis results. In addition, the variable indicators in the classical empirical analysis tend to be relatively clear, and such clear boundaries are often predetermined when collecting data. However, text big data has complicated sources and diverse kinds, which does not contain clear variables. As a result, how to extract effective information and to prove that the author extracts the right target information is the main point of text big data analysis.^[5]

User-concerned Topics

By using the crawler program, we can gain big amount of online comments. To ensure the accuracy of the comment data, we pre-process the crawled comment data, delete duplicate comments, comments with advertisements and comments that are punctuation. Have applied the Latent Dirichlet Allocation (LDA), we extract subjects of online comments to gain user-concerned topics, as listed in Table 1.

Table 1. User-concerned topics about paper-cutting

Topics	Keywords	Topics	Keywords
Styles	18	Modern Production	15
Practicality	12	Souvenirs	24
Craftsmanship	13	Technical Heritage	38
Marketing	11	Cultural Transmission	21
Product Innovation	27	Digital Research	11

Need Extraction

Through the user's negative comments on each topic, we extract the viewpoints to analyze the user's dissatisfaction and needs. The specific steps of the needs analysis are as follows:

- (1) Extract relevant comments on each topic: After obtaining the topic that users pay attention to, screen out online comments describing concerns based on the users' attention.
- (2) Negative comment screening: Obtain negative comments according to the negative sentiment words.
- (3) Negative opinion mining: After obtaining the negative online comment, the commentary extraction function of the AipNlp module is used to extract the views, which are sifted and processed later.

Table 2. Needs mining results

Topics to be improved	Viewpoints
product innovation	Simplex product form; universally for aged people; the products overvalue artistry while overlook practicality; lack of highly integration between traditional culture and modern design and that between traditional crafts and modern technology;
production and processing	Production immigration; rough production; too countrified; less intelligent;
cultural transmission	Low degree of internationalization and informatization; Lack of connotation in cultural stories;
Marketing	Weak brand popularity;
digital research	Lack of connotative digital paper-cutting websites; few dynamic paper-cutting technique videos;
technical heritage	Lack of inheritors; young people lack interest in paper-cutting; the existing inheritors are too old.

Taking Chinese paper-cutting as an example, through comment data analysis, topics that need to be improved are "product innovation", "production and processing", "cultural transmission", "marketing", "digital research" and "technical heritage". The analysis module in AipNlp is used to extract viewpoints from negative comments. And the data processing result are shown in Table 2.

Protection Decision Based on Big Data

Based on big data comments, we can find users' pain points and analyze their needs to determine improvement strategies. However, pain points raised by ordinary users are mostly individual needs and shallow thinking. While academic experts, government administrators, paper-cutting inheritors and paper-cutting companies have deeper and more systematic thinking on protection strategies. Therefore, it is necessary to collect and filter improvement strategies based on academic data from Baidu Academic and CNKI. Details are as follows:

Establish MOOC courses about Chinese paper-cutting, including paper-cutting culture, history, art appreciation, master forums and craft teaching lessons.^[6] The teaching video center software system and remote live broadcast interactive classrooms that can recording and broadcasting automatically are used to spread lessons freely to realize the democratization of educational resources. Integrating resources, revitalizing data and building a paper-cutting MOOC lesson are conducive to transmit paper-cutting culture, popularizing paper-cutting knowledge, cultivating paper-cutting talents and raising public popularity of paper-cutting techniques.

Focus on the concept transformation of paper-cutting inheritors.^[7] Paper-cutting workshops can be opened in primary and secondary schools for students to learn the traditional techniques of paper-cutting and apply new techniques. At the same time, establish a big data network of existing paper-cutting inheritors as well as potential ones.

Apply big data technology to collect and organize traditional paper-cutting patterns.^[8] With the aging of the old generation of paper-cutting inheritors, it is necessary to carry out the collection and organization of traditional paper-cutting patterns. The work can be more comprehensively done by establishing a data network of traditional paper-cutting patterns.

Promote the development of paper-cutting related industries.^[9] Paper-cutting not only needs to be rescued and be passed on, but also needs to be transformed into creative industries through technical innovation and cultural creativity to create industrial value.

Promote the paper-cutting culture and shape the paper-cutting brand festival.^[10] The daily nature of paper-cutting is rooted in the temporal and spatial unity of culture. The reason why paper-cutting and daily life can be closely combined is that daily life is the realm of the survival and development of paper-cutting. There are many national characteristics in paper-cutting itself, which should be carried forward to form some special festivals to make them shine.

Strengthen the connotation protection of the paper-cutting art and pay attention to its artistic application.^[11] The focus of the application protection of paper-cutting is highlighting its unique artistic characteristic. Rather than being industrialized, It must be purposefully preserved of the unique cultural atmosphere and the craftsmanship. In the specific production application, the artistic quality of the product can be improved by hand-finishing production.

Summary

The traditional data is mostly questionnaire survey data, which is highly subjective. While big data is more of the "natural data" formed in the real life, which is highly credible. The famous historian Yuval Herali believes that our generation is in the transitional period from humanism to Dadaism. For a long time in the past, people believed that the authority came from heaven or god, while in the time of humanism, people believed the authority is from the feelings and choices of human beings. Now we are slowly transferring the authority to the things outside, that is, the data terminals.

Acknowledgment

The research program is supported by the fund project of the Humanities and Social Sciences of the Ministry of Education of China (17YJAZH103).

References

- [1] Jianjun Sun. Rediscovering the cultural significance of paper-cutting. [J]. *Hundred Schools in Arts*.2013.
- [2] Bianan Wu. Theory and method of intangible cultural heritage protection [M]. Beijing: Culture and Art Publishing House, 2010.
- [3] Wei Luo. New Computing sociology: sociological research in the age of big data [J]. *Sociological Study*, 2015 (3), 222-241.
- [4] Yan Sun, Ronggui Huang, Yanwei Hong. Chinese intangible cultural heritage in the age of big data: an analysis of international reputation of 300 years [J]. *Academic Forum* (6 issues): 106-112.
- [5] Xiaoguang Qiao, Huan Su. Cultural inheritance in the internet time - observation and thinking based on social practices and phenomena of intangible cultural heritage [J]. *Cultural Heritage*, 2017(1).
- [6] Wei Sun, Fang Liao. Protection and product innovation design of the intangible cultural

heritage of Xiang embroidery in the background of big data [J]. *Creation and Commentary*, 2016(12): 116-119.

[7] Yiru Xu, Jianwu He. The application of big data technology in the cultivation of intangible cultural inheritors in minority women—taking Zhuang brocade as an example[J]. *Comparative Study of Cultural Innovation*, 2019(25):61-62.

[8] Yan Chen. Digital protection and inheritance of intangible cultural heritage under the background of the belt and road initiative [J]. *China National Exhibition*, 2019(03):77-78.

[9] Yonglin Huang. Protection and utilization of intangible cultural heritage in the digital background [J]. *Cultural Heritage*, 2015(1):1-10.

[10] Youzhen Wu, Bo Wu, Lei Ma. Research on internet big data mining and intangible cultural heritage activation [J]. *Journalism Quarterly*, 2013(3):66-71.

[11] Yang Qin, Guanyu Li. The development strategy of the application and inheritance of the regional artistic intangible cultural heritage in the new time [J]. *Hundred Schools in Arts*, 2018(1): 233-239.