Research on the Development Dilemma of China's Platform Economy based on Data Mining Technology

Zhili Dai and Yingge Du

School of Economics and Management, Beijing Jiaotong University, Beixiaguan, Beijing, China

Keywords: Data Mining, Platform Economy, Platform Monopoly, Platform Supervision.

Abstract: The rapid development of the Internet has given birth to a new economic form based on the platform, and the platform economy has gradually become a new engine for development. The platform economy is of great significance in promoting economic development and providing employment opportunities. However, there are still many problems in the development process, which brings many challenges to the original government governance measures. This paper crawls the abstracts and keywords of papers related to platform economic development on CNKI as a data source and uses the TF-IDF algorithm to mine the important features of the obtained texts, and put forward targeted countermeasures to achieve the benign development of the platform economy.

1 INTRODUCTION

Contemporary society has entered the Internet era. The rapid development of the Internet is changing the way people work, live, and think. The development of the Internet has promoted technological innovation, and a large number of platform companies based on the Internet, big data, and mobile technologies have risen rapidly. In recent years, the Internet platform economy has shown a booming development around the world. The platform economy is a digital economy in which the Internet platform organizes production factors to produce products and services. It covers many digital activities in business, politics, and social interaction (Kenney, Zysman, 2016).

The platform economy is regarded as one of the three major events of the "digital revolution". China has a vast Internet market, and the number of netizens ranks first in the world, which provides fertile soil for the development of the platform economy. At the same time, China's rapid development in artificial intelligence, big data, and other technological fields has also played an important role in the rapid development of the platform economy effect.

The platform economy involves all walks of life. While China's platform economy is developing vigorously, it also faces many problems, not only the problems of the platform economy itself but also the problems involving industries and employees. Therefore, based on data mining technology, this paper crawls the abstracts and keywords of CNKI papers related to platform economic development and uses the TF-IDF algorithm to extract important features in the obtained texts. Therefore, it analyses and excavates the main problems of the current development of China's platform economy, and proposes corresponding solutions to the problems, to promote the sustainable development of the platform economy.

2 DATA ANALYSIS

At this stage, Python and its crawler technology are widely used on the Internet. Compared with C++ and Java languages, Python language syntax is simpler and the operation process is relatively convenient. Crawling data with Python-based crawlers is relatively efficient. In addition, Python can provide a more practical framework for program writing, and provide a guarantee for crawlers to quickly obtain data information on web pages. It can quickly complete the writing of various program tasks in a relatively short period, which is beneficial to data mining and data processing. Python technology can be widely used in scientific research, web security,

product development public opinion monitoring, and other fields.

2.1 Data Collection

The data used in this article comes from the abstracts and keywords of papers on the development of platform economy in CNKI. CNKI has strong authority in the domestic academic circle, and the abstracts and keywords can help us quickly understand the main content of the article. The specific steps of data collection in this paper are as follows:

Write a crawler program using the Python programming language. By analysing the HTML grammatical structure of the target website, with the help of the request package in Python, the obtained target text information is located using XPath. Enter "Platform Economic Development" on the CNKI platform for retrieval, carry out data collection for the retrieved content, crawl all the content that meets the retrieval conditions, and save it into a text file. and provide important support for the development of the platform economy.

Import the crawled text data into Excel. Since the conditions for inserting the "$" delimiter between different attribute data are set during the crawling process, after importing, you can successfully store different types of data into different columns through the column-by-column function that comes with Excel to obtain valid data.

Preliminary verification and correction of the data. Check whether the data items and relevant data correspond to each other through the webpage. After the check is completed, the data collection is completed.

2.2 Data Cleaning

After the completion of the data collection, this research collected a total of 302 articles about "platform economic development", and the character attributes include the paper title, abstract, and keywords. Preliminary data preprocessing is performed on the data set according to the collected data, and the processed data are mainly duplicate data, missing data, and wrong data, to achieve a better data utilization effect.

2.3 Data Mining

This paper mainly uses the keywords of the paper to discriminate and separate the keywords with the help of Excel's automatic sorting tool. TF-IDF is a common feature extraction method, which is characterized by being simple and fast, and the results are more in line with the actual situation. This paper uses the TF-IDF algorithm in the Python Chinese word segmentation package 'jieba' to obtain the keywords of the text. Run the program in python to obtain the word frequency, inverse document frequency, and TF-IDF value of the content obtained afterword segmentation. The formulas involved in this program are shown in formulas (1) and (2):

$$\text{Term Frequency (TF)} = \frac{\text{the number of times a word appears in the article}}{\text{the total number of words in the article}} \quad (1)$$

$$\text{Inverse document frequency (IDF)} = \log\left(\frac{\text{total number of documents in corpus}}{\text{number of documents containing the word} + 1}\right) \quad (2)$$

In the program, first count the total word frequency, then calculate the TF and IDF values of each word and finally calculate the TF-IDF value of each word, which is obtained by multiplying the TF value by the IDF value. Among them, we first sorted the word frequency counted by the Python script by sorting, transposing, and descending order in Excel, and obtained the order of each word according to the number of occurrences (due to a large number of words, only part of the display): The word frequency sorting is as follows Table 1 shows:

Table 1: Word frequency order.

Term	Frequency	Term	Frequency
Platform economy	221	…	…
Platform enterprise	40	Antitrust laws	9
Internet platform	24	Network effect	5
Anti-monopoly	20	Monopolistic behavior	4
Government	11	Prudential regulation	2
…	…	Worker protection	2

Finally, arrange and output according to the TF-IDF value of each word from large to small, and the obtained results are shown in Table 2 (due to the excessive amount of actual output, only part is shown here):

Table 2: TF-IDF value of the word.

Term	TF-IDF	Term	TF-IDF
Platform economy	0.014901366	…	…
Platform enterprise	0.014810958	Antitrust laws	0.006786429
Internet platform	0.013226495	Network effect	0.004335937
Anti-monopoly	0.011794409	Monopoli--stic behavior	0.003630275

Government	0.00785033	Prudential regulation	0.003630275
…	…	Worker protection	0.003630275

Because scholars have different definitions of nouns, the words are scattered. After sorting out, it can be found that terms related to platform monopoly, platform supervision, employment protection, and other aspects appear more frequently. Among them, "monopoly" appeared 66 times, "supervision" appeared 22 times, and "labor protection" related terms appeared 10 times. Therefore, these issues are also issues of general concern in the current academic circles. Next, I will elaborate on these three issues.

3 PLATFORM ECONOMY DEVELOPMENT DILEMMA

3.1 Problem of Protecting the Rights of Employees

The platform economy has provided more employment opportunities and absorbed a large number of flexible employment populations. The "China Sharing Economy Development Report (2021)" released by the State Information Center shows that in 2020, the transaction scale of my country's sharing economy market is about 3,377.3 billion yuan, a year-on-year increase of 2.9%, and the number of sharing economy service providers is about 84 million, a year-on-year increase of 7.7%. We can see from Table 3 that the number of platform economy practitioners has gradually increased in recent years.

Table 3: The number of employees of platform enterprises.

Year	Number of employees
2017	5560000
2018	5980000
2019	6230000
2020	6310000

Though the sudden impact of the epidemic, the total pressure on the employment market, and the existence of structural problems have brought huge challenges to the employment situation. However, the platform economy still plays a pivotal role in broadening employment channels, enhancing employment flexibility, and increasing workers' income. Such a large number of practitioners and employment fields have also made the protection of labor rights and interests related to the platform

economy a hot spot and focus of the government and society.

A large part of China's platform economy practitioners have chosen the employment form relying on the Internet. One is e-commerce logistics, takeaway, and online car-hailing in the field of life services; the second is the field of knowledge and creativity, including online live broadcasts, e-sports, etc.; the third is the new employment forms derived from the platform, such as self-media. The business model of the platforms involves the participation of three parties: the user or customer of a certain service; the practitioner (provider or crowd-worker); and the platform which connects the user and the worker (Prassl, Risk, 2015).

The working model of platform economy practitioners is as follows: the enterprise assigns work tasks to the platform workers, and the workers provide goods and services to consumers and users. Workers have no fixed workplace, no fixed working hours, and no stable source of income. Most of their income depends on commissions from orders. The evaluation and complaints of consumers and users are of great significance to the wages of workers. The platform system is based on rating mechanisms, in the light of which customers score and comment on the service received (McDonald 2016). Take food delivery as an example, the uncertainty of working hours enhances the "platform stickiness" of delivery workers, and they dedicate body and all their labor to work (Ping 2019), the takeaway courier does not hesitate to work overtime to take orders to get high pay and violate the traffic rules to deliver the meals on time and get praise from customers, and the takeaway courier has a traffic accident and casualty rates remain high. Practitioner's behavior is an important reference point for customers to evaluate their services, and practitioners who receive multiple negative reviews may affect their wages or lose opportunities to continue working in the platform. That is to say, the platform uses its advantages to completely transfer the risk of being unable to perform its obligations to the employees, thereby keeping the platform from being damaged (Gramano 2020).

At the same time, the social insurance coverage rate of the takeaway couriers is low, and their basic rights and interests cannot be guaranteed. At present, most of the takeaway couriers sign "cooperation agreements" with third-party agencies, rather than labor contracts. The agreement signed by the takeaway couriers and the third party does not belong to the scope of labor relations, and there is no such thing as social security. The characteristics of the

platform economy's de-employment and flexible labor relations also make it impossible for platform economy practitioners to obtain the protection of the current work-related injury insurance system.

3.2 Difficulty of the Platform Supervision

Some scholars believe that although government regulation has defects, it still needs government regulation, and there is information asymmetry in the market, which is likely to lead to market failure or market collapse (Akerlof 1970). Compared with the traditional model, it is more difficult to supervise the platform economy, and it is tough for the original government laws and regulations to effectively restrain it.

At present, there are still many problems that are difficult to supervise in the process of platform economic development, such as malicious negative reviews of platform transactions, platform big data killing, user data privacy leakage, false propaganda, traffic fraud, tax evasion by employees, etc. The loose Internet censorship mechanism and low cost of dissemination also make supervision more difficult. Notable among these issues is the leakage of user data privacy and tax regulation. An important resource for the operation of the platform economy is user data. Now platform companies on the market collect, use and transfer user data without the permission of users, which brings many problems to government supervision. For example: On March 20, 2019, Douyin provided WeChat/QQ platform authorized login services to affiliated companies, and WeChat sued it in court. On July 4, 2021, the State Office released a message saying that the "Didi Chuxing" APP has serious violations of laws and regulations to collect personal information. According to relevant regulations, the Cyberspace Administration of China notified the App Store to remove the "Didi Chuxing" APP.

The operation of the platform economy involves multiple links such as information exchange, demand matching, paperless digital transactions, and delivery of goods between two or more parties using the platform as a medium, which makes it difficult to determine the elements of the tax system. Regulatory measures such as access conditions are still incomplete, which has brought a huge impact on tax supervision. The form of the platform economy has brought about major changes in taxation objects and taxpayers. The platform economy has the characteristics of virtuality, decentralization, cross-regional nature, and privacy of transaction data. The traditional tax supervision model shows lag and inappropriateness.

3.3 Problem of the Platform Monopoly

The rapid development of the platform economy has accelerated the monopoly of the market and industry.

"Platform monopoly", that is, the common phenomenon of "winner takes all" in the platform economy may evolve into a few monopoly platforms that maintain the status of winners for a long time. The platform mainly relies on the Internet. The platform economy has a strong network effect, and its initial advantages are amplified. The network effect of the platform makes the utility obtained by users on one end of the platform subject to the influence of the scale of users on the other end. The larger the scale of users, the stronger the utility with the expansion of the scale, resources will be gathered to a few platforms, and the strong will become stronger (Eisenmann, Parker, Alstyne, 2006). At the same time, the platform's multilateral market is complex and the relevant market is difficult to define. Traditional anti-monopoly methods are not applicable to platform monopoly issues.

Platform monopolies usually take two forms: first, mergers or acquisitions between platforms. After stable development, platform companies usually occupy a leading position in the market through price wars. When price wars are in trouble, platform companies often form new monopoly platforms in the form of mergers or acquisitions. The second is the exclusivity strategy. When platform companies compete with other competitors of the same type, they rely on their dominant positions to allow platform users to "choose one". For example, the two giants in the e-commerce field, Tmall, and JD. pick one of two". The platform occupies a dominant position in the market, which means that the platform has a priority in terms of pricing, transaction methods, and responsibility allocation. The expansion and extension of the power of the platform make it have jurisdiction over multilateral customers and key stakeholders, which means that other participants are in a relatively weak position, lacking the ability to bargain and make independent choices, which also creates a series of problems. The commission is too high, "big data kills cooked", etc. (Liu, Serfes, 2013).

4 COUNTERMEASURES FOR SUSTAINABLE DEVELOPMENT OF PLATFORM ECONOMY

We should face the development dilemma of the platform economy, and propose solutions from the government, the platform, and practitioner. The three parties are supposed to work together to achieve the sustainable development of the platform economy. Their relationship is shown in figure 1.
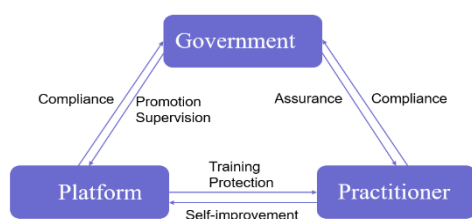

Figure 1: Relationship between the three parties.

4.1 Government

The government should take active measures to promote the development of the platform economy.

Strengthen the construction and maintenance of platform infrastructure, promote the development of technologies such as big data and artificial intelligence, and provide important support for the development of the platform economy.

Formulate and improve relevant policies, laws, and regulations. On the one hand, it needs to solve the problem of protecting the rights and interests of current platform practitioners, and reasonably define the nature of workers in the new form of employment. On July 16, 2021, eight departments including the Ministry of Human Resources and Social Security jointly issued the "Guiding Opinions on Safeguarding the Rights and Interests of Laborers in New Employment Forms" and proposed the "Labor Trichotomy", which is considered to realize the development of the platform economy and new forms of employment. An important measure to balance the protection of labor rights and interests. The government should actively promote the implementation and implementation of relevant policies. On the other hand, through the revision of the Anti-Monopoly Law, the problem of platform monopoly in the process of platform economic development is solved. The government should build a tax legal system that meets the needs of platform economic development, learn from the tax practices

of other countries, and revise and improve my country's existing tax laws according to local conditions.

Optimize regulatory measures. Strengthen the supervision of the behavior of the platform, actively identify potential harmful behaviors and urge and assist the platform to solve them, protect the privacy of users, and maintain a fair market competition environment and a good consumption environment.

4.2 Platform

Establish bottom-line thinking and clarify their corporate rights, responsibilities, and obligations.

For employees, it should protect their legitimate rights and interests by laws and regulations, improve the income system of employees, and at the same time strengthen investment in education and training for employees to improve their skills.

For the platform, it is necessary to strengthen innovation, promote its integration with other industries, and maintain its market position in the market, rather than maintaining its competitive advantage in unreasonable ways such as market exclusivity strategies. And the platform should operate and pay taxes following the law. Comply with various government regulatory measures and strengthen the platform's governance and supervision of the entire business ecosystem. Protect the rights and interests of consumers and create a good atmosphere for consumption.

4.3 Practitioner

Make good use of legal weapons to protect their legitimate rights and interests, platform practitioners, as direct providers of services, should study the law, understand the law, abide by the law, and use it.

Improve personal skills. In the future, more industrial populations will be converted into practitioners of the platform economy, and the platform labor population will become increasingly prominent (Ping, Magasic, 2016). Such a large number of employed people contribute their vitality to the social economy. However, many practitioners are caught in "youth meal anxiety". Although the platform economy provides more employment opportunities and lowers the threshold for employment, knowledge is still essential for the workers. The platform economy will lead to a polarization of the employment structure, with the employment and salary of high-skilled and low-skilled workers increasing, while the employment and salary of medium-skilled workers decrease.

Therefore, practitioners should strengthen their re-education and improve their knowledge and skills.

Possess good professional quality and moral awareness, they should be responsible for their work, adhere to their professional bottom line, and abide by national laws and regulations.

5 CONCLUSIONS

The vigorous development of the platform economy has not only changed people's production and lifestyle, but it is also of great significance in promoting economic development and providing employment opportunities. As a new thing, it still faces many difficulties in the process of development. To promote its long-term development, it is necessary to work together to create a good environment for the development of the platform economy.

The Chinese government should not only promote its development but also strengthen supervision and legal constraints on the platform economy. Platforms should also strengthen their construction and maintain social responsibilities. Platform practitioners should also focus on improving their skills. In a word, as a new economic and employment form, the platform economy has broad for development, but attention should be paid to its construction and management.

REFERENCES

Akerlof G A. (1970). Quality uncertainty and the market mechanism. J. Quarterly journal of Economics. 84 (3), 488-500.

Eisenmann T & Parker G & Van Alstyne M W. (2006). Strategies for two-sided markets. J. Harvard business review. 84 (10), 92.

Gramano E. (2020). Digitalisation and work: challenges from the platform-economy. J. Contemporary Social Science. 15 (4), 476-488.

Kenney M & Zysman J. (2016). The rise of the platform economy. J. Issues in science and technology. 32 (3), 61.

Liu Q & Serfes K. (2013). Price discrimination in two-sided markets. J. Journal of Economics & Management Strategy. 22 (4), 768-786.

McDonald D K. (2016). Reputation will teach the sharing economy to share. J. U. Fla. JL & Pub. Pol'y. 27, 219.

Ping S S & Magasic M. (2016). Knowledge workers, identities, and communication practices: Understanding code farmers in China. J. tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society. 14 (1), 312–332.

Prassl J & Risk M. (2015). Uber, taskrabbit, and co.: Platforms as employers-rethinking the legal analysis of crowdwork. J. Comp. Lab. L. & Pol'y J. . 37, 619.

Sun P. (2019). Your order, their labor: An exploration of algorithms and laboring on food delivery platforms in China. J. Chinese Journal of Communication. 12 (3), 308-323.