Construction and Analysis of a Diachronic Knowledge Base of Traditional Values in Chinese Idioms

Hao Sun
College of Information Science
Beijing Language and Culture University
Beijing, China
202221198129@stu.blcu.edu.cn

Huidong Du
College of Information Science
Beijing Language and Culture University
Beijing, China
yzqtdu@gmail.com

Yue Wang
College of Information Science
Beijing Language and Culture University
Beijing, China
wangy_0062@163.com

Xupeng Liu
Department of History
University of Chinese Academy of Social Sciences
Beijing, China
liuxupeng@ucass.edu.cn

Dong Yu
College of Information Science
Beijing Language and Culture University
Beijing, China
yudong_blcu@126.com

Pengyuan Liu*
College of Information Science
Beijing Language and Culture University
Beijing, China
liupengyuan@pku.edu.cn

Abstract—Chinese idioms are the distillation of ancient Chinese culture, spanning various dynasties and rich in traditional values. Studying the diachronic evolution of these traditional values in idioms is of great significance for understanding Chinese culture. Due to the limited resources on diachronic values, there has been relatively little research on the diachronic evolution of Chinese traditional values, especially a lack of studies involving quantitative and qualitative analysis. In this paper, we design a framework and construct the first diachronic knowledge base of Chinese traditional values in idioms, which includes 2,472 idioms that embody traditional values. Each idiom is annotated with rich related information by GPT-4 Turbo and verified by humans. This method of using a language model for annotation, with human-assisted verification, improves the efficiency and accuracy of the knowledge base annotation.Subsequently, we provide a detailed analysis and discussion of the specific content changes and thematic content (taking ”仁义” as an example) of traditional values in idioms, and conduct standardized processing of the overall quantity of traditional values idioms, analyzing their change trends over time. The analysis reveals that traditional values embodied in Chinese idioms demonstrate strong continuity and are reflective of the era in which they were used. Additionally, Chinese society has consistently upheld the principle of ”altruism.” Furthermore, the cyclical and wave-like patterns of Chinese traditional values closely align with the characteristics of various historical periods in China.

Index Terms—Chinese idioms, Chinese traditional values, diachronic knowledge base, diachronic analysis

I. INTRODUCTION

In recent years, there has been a surge of research in academia on the intersection of values with Chinese traditional culture [8] [9], as well as on the development of value scales, theoretical frameworks, and related constructs [10] [11]. Simultaneously, with the rapid advancement of artificial intelligence, attention has been drawn to the alignment and understanding of values in AI language models [12] [13]. This indicates the significant role of values across various levels of society [1] [2] [3]. Chinese traditional values, primarily rooted in Confucianism and encompassing concepts such as ”仁” (benevolence) and ”礼” (propriety), play a vital role. Amidst the development of a multipolar world and profound changes brought about by globalization, there has been a notable increase in the uncertainty and instability of global governance systems and international order. Humanity faces unprecedented challenges and difficulties. In response, numerous experts in China have initiated discussions, emphasizing the importance of traditional Chinese values in fostering consensus amidst a changing and uncertain world [14]. Idioms, as the essence of Han Chinese language and culture, are derived from ancient Chinese cultural classics and constitute a significant component of China's rich traditional culture [15]. Studying the diachronic evolution of Chinese traditional values through idioms provides a promising approach to understanding traditional Chinese culture and addresses the current lack of research on Chinese traditional values. Therefore, in this paper we utilize a combined approach of GPT-4 Turbo and human expertise to construct the first diachronic knowledge base of traditional values in Chinese idioms. Taking ”仁义” (benevolence and righteousness) as an example, we conduct a diachronic analysis of the traditional values embodied in Chinese idioms. The analysis reveals that the values reflected in Chinese idioms exhibit both continuity and characteristics specific to different historical periods while consistently emphasizing the concept of ”altruism.” Furthermore, a standardized processing of idiom quantities

under traditional values across different dynasties reveals that numerical changes over time reflect the phased and wave-like characteristics of Chinese traditional values, which closely align with the characteristics of different historical periods in China.

II. PREVIOUS RESEARCH

A. Research on Chinese Idioms

Idioms are fixed phrases that have been passed down through generations, rich in meaning, and carry the color of written language, reflecting a wealth of traditional values [4] [16] . Currently, there is limited diachronic research on Chinese idioms in the academic community: Zong [17]conducts a corpus analysis of idioms in "The Book of Songs" and found that the semantics of these idioms have evolved over time; Yang [18]compares and analyzes the distribution and usage characteristics of idioms with the same origin in Chinese and Japanese from a quantitative and diachronic perspective.

B. Research on Values

"Values" are beliefs that involve desirable ultimate states or modes of behavior, transcending specific situations and guiding the selection or evaluation of behaviors, people, and events, as well as prioritizing to form a system of value priorities [1]. In short, values are our judgments about what is right, what is wrong, and what is valuable in terms of things and behaviors [2].

Regarding the theoretical research on values:Wen [19], He [20] conduct studies on the characteristics of value changes in adolescents using survey questionnaires, Schwartz S H [21] proposes a universal theory of values, and Zhang Z [22] uses methods such as factor analysis to define and rank 19 continuous values and create a value scale.

For diachronic research on values: Wen [19] conducts a values survey of 1,080 middle and high school students and college students in different Chinese cities and made diachronic comparisons with survey results from the 1980s and 1990s. Kochetova [23]study the values in early advertisements published in British newspapers from 1788 to 1900 from a diachronic perspective. These studies are limited to specific areas of value research and do not examine the diachronic changes in universal values.

Regarding the construction of value-related knowledge bases: Liu [24] builds and analyzes a knowledge base for Chinese core values, and Wang [25]constructs a Chinese moral semantics knowledge base and conducts a diachronic analysis. There is a lack of research on Chinese traditional values.

Thus, our research fills the gap in the construction and research of a knowledge base for traditional values in Chinese idioms.

III. CONSTRUCTION OF A DIACHRONIC KNOWLEDGE BASE FOR TRADITIONAL VALUES IN CHINESE IDIOMS

We initially select 2,472 Chinese idioms that distinctly embody Chinese traditional values from an authoritative Chinese idiom dictionary [4]. By integrating the provenance and
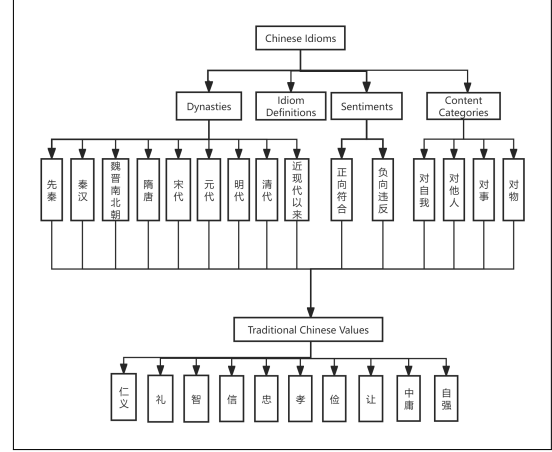

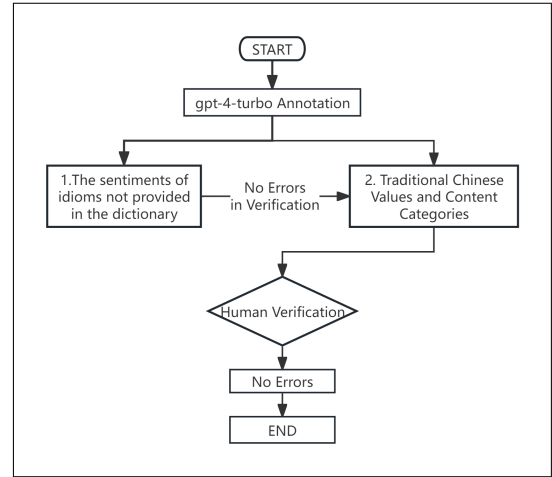Fig. 1. Knowledge Base framework diagram

Fig. 2. Annotation process diagram

usage information provided in the dictionary, we determine the corresponding nine dynasties [6] and two sentiments [4] [16]. Following that, we consult relevant literature to determine ten traditional values [3] [5] and the four content categories [7]. Subsequently, utilizing the GPT-4 Turbo model for annotation and manual review, we annotate the sentiments, traditional value categories, and idiom content categories associated with different traditional values. This process culminates in the establishment of the first diachronic knowledge base of Chinese idiom values.

A. Framework for the Diachronic Knowledge Base

As shown in Figure 1, this diachronic knowledge base includes 2,472 Chinese idioms , ten categories of Chinese traditional values, nine dynasties where the idioms are found, idiom definitions [4], two sentiments, and four content categories.

We utilize GPT-4 Turbo to annotate the ten categories of Chinese traditional values, determine the sentiments of specific idioms, and classify the four content categories within

the knowledge base. Subsequently, three professional trainers (each holding a master's degree) verify and refine these annotations to ensure accuracy. The origins of the nine dynasties and idiom definitions are sourced from authoritative Chinese idiom dictionaries. Figure 2 illustrates the annotation process.

There are several reasons for utilizing GPT-4 Turbo in annotating elements of the diachronic resource library:

Firstly, although the final annotation results still require human inspection and modification, GPT-4 Turbo can significantly reduce the initial annotation workload. This allows human annotators to allocate more time and energy to verification and correction, thereby enhancing overall efficiency.

Secondly, GPT-4 Turbo follows predefined standards and rules for annotation, which mitigates the subjective bias and inconsistency that may arise in manual annotation. While human review remains necessary, the consistency of the preliminary annotations contributes to improved quality in the final results.

Moreover, GPT-4 Turbo can rapidly process large volumes of data, rendering the entire process more efficient, particularly when dealing with extensive idioms and texts. Relying solely on human effort would require considerably more time and resources. GPT-4 Turbo also serves as an auxiliary tool for human annotators by providing initial classification and annotation suggestions, thus offering a reference framework that enhances annotation accuracy and efficiency.

Furthermore, GPT-4 Turbo, trained on an extensive corpus of texts and knowledge bases, can offer diverse perspectives and insights. This assists human annotators in identifying and understanding the background and meaning of certain idioms and values, which is particularly beneficial for unfamiliar areas.

Therefore, the combined use of GPT-4 Turbo and human verification not only improves annotation efficiency and consistency but also provides valuable references for human annotators, enhancing the overall quality and reliability of the process.

Our approach also offers methodological support for the subsequent construction of the resource library.

Detailed information on the construction of the specific knowledge base, annotation, and other aspects can be found in the appendix.

B. Annotated Results of the Diachronic Knowledge Base

Table 1 shows the accuracy rates of annotation and modification, with Model Accuracy (M1), Human Verification Consistency (P1), and Human Modification Consistency (P2).As shown in the table, after GPT-4 Turbo completes the annotation, three annotators review the results and identify discrepancies with their own opinions and calculate Model Accuracy (M1). They then calculate the overlap rate of discrepancies among the three annotators regarding the GPT-4 Turbo's annotation results(P1). Based on this, the three annotators correct the erroneous results from GPT-4 Turbo and calculate the consistency rate of the modifications(P2). If there is a disagreement among the three annotators, we adopt the
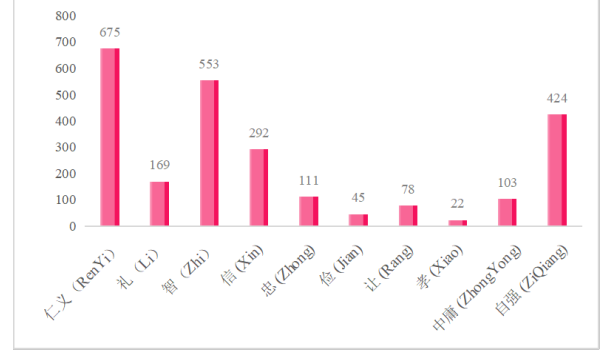

Fig. 3. Statistics of Chinese Idioms Representing Traditional Values in the Knowledge Base: Detailed information on the dynasties, sentiments, and content categories of the idioms can be found in the appendix.

principle of the minority yielding to the majority, preferentially selecting the result agreed upon by two annotators.

Figure 3 presents the overall distribution of traditional values in Chinese idioms by dynasty in the repository. The distribution of other content in the repository can be found in the appendix.

TABLE I
ANNOTATION CONSISTENCY RATE AND ACCURACY RATE

Annotation Model	GPT-4 Turbo		
Sentiments	M1: 85%	P1: 95%	P2: 95%
Traditional Values Conforming Idioms	M1: 83%	P1: 90%	P2: 88%
Traditional Values Violating Idioms	M1: 81%	P1: 89%	P2: 87%
Content Category	M1: 83%	P1: 86%	P2: 95%

IV. DIACHRONIC EVOLUTION ANALYSIS OF TRADITIONAL VALUES IN IDIOMS – TAKING THE "仁义·对他人(BENEVOLENCE AND RIGHTEOUSNESS · FOR OTHERS)" AS AN EXAMPLE

The historical discussion in this section and the next is derived from references [19] — [44].

The interpretation of idioms provides a detailed explanation of the idioms. We use the interpretation of idioms to analyze the diachronic evolution of specific content under traditional values. The distribution of idioms under the "仁义" value has the most instances, and the content category of "对他人(For Others)" under this value is particularly illustrative and representative. Therefore, this section selects the "对他人" content category under the "仁义" value to observe the diachronic evolution of specific content. After screening the "仁义" value and content category, we segment the interpretation of idioms, and use stopwords and manual methods to remove words in the interpretation that are unrelated to expressing the central content of the idiom. For example, in the interpretation of the idiom "爱屋及乌(Love the house and the crows on it)" which reads "比喻爱一个人而连带地关爱与他（她）有关系的人或物。" words like "比喻","一个","而" "连带" etc., which have a smaller relationship with expressing the central

Fig. 4. Diachronic Evolution Diagram of Specific Content under "仁义·对他人(Benevolence and Righteousness·For Others)": The top of the dynasty in this figure represents positive content, while the bottom represents negative content. The content in the same row is the same or similar discussion content. For example, the first row at the top of the dynasty represents "正向-他人(positive- others)" which conforms to the "仁义," while the first row at the bottom represents "负向-他人(negative- others)" which violates one.

content of the idiom, are removed. The central content that remains is "爱人、关爱人、物" The specific content under "仁义·对他人(Benevolence and Righteousness · For Others)" is shown in Figure 6 to analyze the diachronic evolution of specific content.

Next, we use GPT-4 Turbo to further classify the idiom content into topics as shown in Table 2, and then create a scatter plot of the frequency of specific content under each topic category, as shown in Figure 7（Among them, the size of the different points in Figure 7 represents the frequency of detailed category discussions in different dynasties）. This allows for a more direct visual observation of the continuity and modernity of different topic categories across different dynasties.

A. Diachronic Analysis of Specific Content under "仁义·对他人

As shown in Figure 4, during "先秦(the Pre-Qin period)", the positive content that conforms to the "仁义" values mainly includes educating and caring for others, emphasizing the purity of relationships with others; caring for friends, brothers, elders, and children while adhering to principles; being responsible for and caring for the people and the commoners; and valuing talents and virtuous people, choosing virtuous people over beauty when faced with a choice; respecting the weak and not fearing the strong; and officials helping each other within the ranks.The negative content that violates the "benevolence and righteousness" values includes behaviors such as slandering, using, and flattering others to undermine the concept of altruism; treating those who have helped oneself with "destroying the bridge after crossing the river," which may be related to the unpredictable military situation

among the various states in the Pre-Qin period; attempting to eliminate "adversaries" for political interests, considering it necessary to "uproot the weeds"; dealing with the powerless and the elite by colluding and flattering for personal gain; and treating gentlemen with disdain, suppressing and excluding them due to differing interests.

In the'秦汉(the Qin and Han dynasties)', the positive discussion of "others" remains strong; the change is that the content of "friends and brothers" disappears; the continuity of "commoners, gentlemen, and talented and people in distress" indicates that the thoughts of people-oriented governance, valuing gentlemen and talents, and helping others are important; the content of officials helping each other disappears, as '秦汉' are unified periods, and the changes of dynasties are accompanied by the struggles of different forces, leading to thoughts about dealing with prisoners of war and enemies.The negative content of "others" still exists, compared to the Pre-Qin period, the content of those who have helped me and adversaries disappears, while the content of the powerful and the elite and the virtuous people remains, and these contents are still discussed in the Qin and Han dynasties.

During'魏晋(the Wei and Jin dynasties)', the positive content that continued from the Pre-Qin period includes "others" and "friends, brothers, and commoners." However, in this period, the concept of "friends" emphasizes adhering to principles, such as "cutting the mat and sitting separately."The negative content has continued with "others" and the adversaries from the Pre-Qin period, with discussions about adversaries being related to the turmoil of this period; there has been a lack of benevolence towards relatives and brothers, and an imbalance in internal relationships.

In '隋唐(the Sui and Tang dynasties)' ,the positive content continues to discuss "others, commoners, friends, and virtuous people." Helping the poor and exposing the bad continues from the Qin and Han dynasties, and this period's discussion of virtuous people is related to the imperial examination system and the prosperity of the state in the Sui and Tang (mainly Tang) dynasties.The negative content of "others," "authorities and petty officials," "virtuous and talented people," and "gentlemen and junzi," as well as "oneself and internal people," has been carried over, and there has been a new addition of "persecuting the common people." The discussions of this period about virtuous and talented people and the common people are related to the imperial examination system and the prosperity of the state in the Sui and Tang dynasties (mainly the Tang dynasty).

During'宋代(the Song dynasty)' , in addition to the continuity of "others," the content of "close relatives" from the previous period is also continued, but it is wives and children, indicating that the Song dynasty paid attention to the family.The negative content, in addition to continuing with "others," also extends to "enemies" from the Wei and Jin dynasties and "common people" from the Sui and Tang dynasties, which may be related to the instability of the borders and changes in governance style in the Song dynasty.

During '元代(the Yuan dynasty)' , the positive content is

"仁义·对他人" Topics			GPT-4 Turbo Accuracy
帮助自己的人 Someone has helped me	同事或内部 Colleagues or Internal	人民百姓 The People and the Commoners	M1:82%
亲近关系 Close Relationships	罪犯敌人权贵 criminals and enemies of the elite	弱小贫困 the weak and the poor	P1:85%
君子贤能 Gentlemen and Talented People	他人 others	异己 adversaries	P2:95%

more concentrated on continuing "others, brothers, friends, elders, neighbors, and the poor and weak," and new content such as "elders and neighbors" appears. Although it was a minority regime, the basic concepts of "filial piety" and "harmony" are inherited.The negative content extends to "others" and "those who have helped me, the weak, and the strong" from the Pre-Qin period, as well as "the internal relations within institutions, namely the collusion between officials," from the Sui and Tang dynasties.

During '明代(the Ming dynasty)' , the positive content continues to discuss "others, virtuous people, the elderly, children, and the poor and weak." This period had a highly centralized system, and the inheritance and promotion of Confucian culture were also good.The negative content has continued with "others," "evil people," and the relationships between superiors and subordinates, as well as "oneself and one's own people."

During '清代(the Qing dynasty)' , the positive content continues to discuss "others, the elderly, children, private relationships, and the poor and weak," with new content such as not bending to private relationships, which may be related to the fairness and strictness of the Qing dynasty's imperial examination system.The negative content has continued with "others," "evil people," "enemies," "respectable women," "good people," "superiors," and "commoners."

Since the Modern Era (1912—Present), the positive content has continued to discuss "others, the people, kind and weak people, and enemies," which is related to the military strategy of this period.The negative content has continued with "plotting against others," "betraying friends," "plotting against those who have helped me," and "seizing property to aid the poor and destitute."

From this, it can be seen that taking "benevolence and righteousness·for others" as an example, the specific content of Chinese traditional values reflected in Chinese idioms, it demonstrates stability and continuity over different eras, and the related content can reflect the characteristics of each era.

B. Diachronic Analysis of Scatter Distribution of "仁义·对他人"

This subsection performs word frequency statistics by segmenting the idiom interpretations under different topic contents, resulting in a scatter distribution shown in Figure 7. The

size of the points represents the frequency of occurrence, such as the largest point for "弱小贫困(the weak and the poor)" in "元代（the Yuan dynasty）", indicating a higher distribution of idioms on this topic during that period, reflecting the degree of emphasis on this topic in idioms.
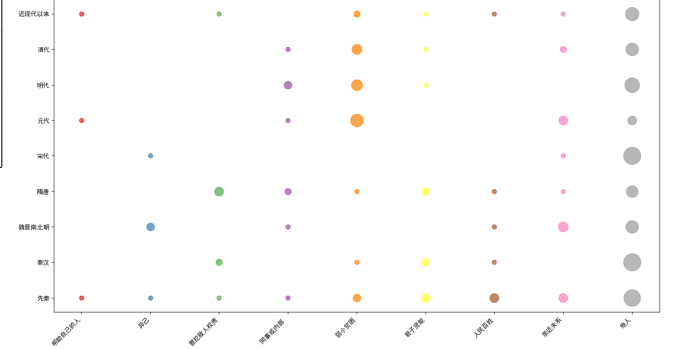

Fig. 5. Scatter Plot of Distribution of Topics under "Benevolence and Righteousness · For Others"

As shown in Figure 5, the topic content of "他人(others)" has the best continuity, followed by "亲近关系 、弱小贫困 、君子贤能及人民百姓类(close relationships, the weak and the poor, gentlemen and talented people, and the people and the commoners)." However, the topics of "异己、帮助自己的人和罪犯敌人权贵类(adversaries, those who help oneself, and criminals and enemies of the elite)" show a sense of discontinuity across different dynasties. The idea of "利他(altruism)" has always been an important part of Chinese traditional values. For example, the topic of "异己(adversaries)" was most prominent in"魏晋 (the Wei and Jin dynasties)", while the topic of "罪犯敌人权贵(criminals and enemies of the elite)" was most prominent in "隋唐(the Sui and Tang dynasties)". This is related to the political instability in "魏晋 (the Wei and Jin dynasties) "and the governance of "隋唐(the Sui and Tang dynasties)". The topics of "异己与罪犯(adversaries and criminals)" are the most scattered, indicating that these topics are less discussed in ancient China. This suggests that Chinese society has always been concerned about "他人(others)" and"他人(others)","弱小贫困(the weak and the poor)" as well as "亲近关系(close relationships)" ."The continuity and discontinuity of the content are related to the characteristics of the times.

V. Z-SCORE STANDARDIZATION ANALYSIS OF TRADITIONAL VALUES IN CHINESE IDIOMS OVER TIME

In this analysis, we employs a standardized analysis method to compare the importance and characteristics of various traditional values across different dynasties in China, aiming to obtain the overall characteristics of stage-by-stage changes. The calculation formula is as follows:

$$\mu_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}; \sigma_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_{ij}-\mu_i)^2} \quad (1)$$

$$z_i = \frac{x_{ij} - \mu_i}{\sigma_i} \qquad (2)$$

Where μ_i represents the mean number of idioms representing the value, n represents the number of idioms in each dynasty, x_{ij} represents the number of idioms representing the value in the current dynasty, σ_i represents the standard deviation of the number of idioms representing the value, and z_i represents the z-score of the number of idioms representing the value.
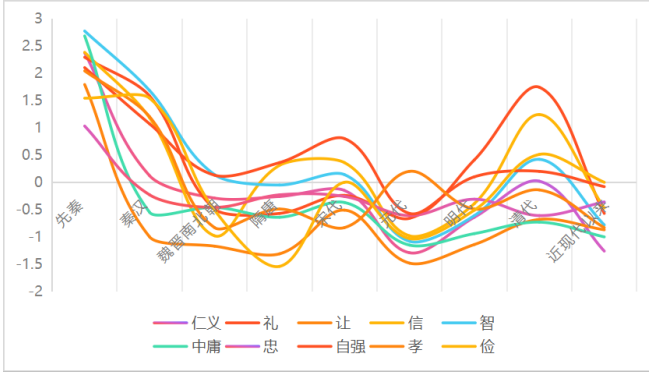

Fig. 6. Diachronic Trend of Z-score Standardized Values of Traditional Values in Idioms:We have created a temporal distribution based on the calculated values, with higher values indicating that the traditional values were more highly regarded in that dynasty.

As shown in Figure 6,The Pre-Qin period was a time of cultural prosperity, with Confucian traditional values abundant in idioms, leading to the highest standardized values in all dynasties. The Spring and Autumn Period and the Warring States Period were both times of ritual and musical warfare, with ritual, music, benevolence, and righteousness being replaced by interests. Therefore, from the Pre-Qin period to the Qin and Han dynasties, the standardized values of various traditional values are in a declining state.

During the Qin and Han dynasties, China enters the era of the unified dynasty. The confluence of Confucianism and Legalism become the mainstream of intellectual development in the two Han dynasties. Therefore, the decline in standardized values during the Han Dynasty is relatively small.

During the Wei and Jin dynasties, the orthodox Confucian classics of the Qin and Han dynasties need to open up new fertile soil for ideas. The speculative nature of the Laozi and Zhuangzi Daoist theories is highly compatible with the pursuit of the Wei and Jin scholars, leading the scholar-officials of the late Han Dynasty to transition from Confucianism to Daoism, ushering in a new era of thought. Confucianism, relative to the Confucianism of the Han Dynasty, underwent a significant transformation. During the the Wei and Jin dynasties, which coincided with the turmoil of the Five Barbarian invasions, social chaos, and cultural setbacks, the standardized values of various traditional values are at their lowest.

From the Wei and Jin dynasties until the Song Dynasty, the standardized values show an increasing trend.

During the Sui and Tang Dynasties, the standardized values of "仁义(benevolence and righteousness)," "信(trust)," "忠(loyalty)," and "自强(self-improvement)" rebound compared to the Wei and Jin dynasties. However, during this period, the standardized values of "礼(Propriety)" and "俭(frugality)" are far below the average level, possibly related to the prosperity of the Sui and Tang (mainly Tang) dynasties, with a prevailing trend of extravagance.

The Song Dynasty sees the highest distribution of standardized values compared to the Wei Jin Dynasties, and the Sui and Tang Dynasties. Among them, "自强(self-improvement)" has the highest standardized values, followed by "信(trust)" and "智(wisdom)." After the turmoil of the Five Barbarian invasions, the Tang Dynasty restore order and, with the efforts of figures like Han Yu in the movement for ancient prose, culture flourished, reaching a cultural peak by the Song Dynasty, with a situation where officials and scholars jointly governed the world.

During the Yuan Dynasty, the overall standardized values of values are relatively low, with only "自强(self-improvement)" slightly above the overall level, while the rest are far below the average level. This period marks the second turning point, and Chinese culture may have experienced a period of setback. The Yuan Dynasty, although adopting a policy of tolerance towards cultural and religious affairs, places the status of Buddhism and Daoism far above Confucianism, and its harsh governance methods can not achieve long-term stability, leading to ineffective governance. Therefore, cultural creativity is not high.However, the standardized values of "孝(filial piety)" in the Yuan Dynasty differs from other traditional values, which can be seen as evidence that the concept of "孝(filial piety)" holds a very high status during the rule of this minority regime.

Subsequently, the Ming and Qing Dynasties see a slight increase in standardized values. In the Ming Dynasty, the standardized values of "礼(Propriety)" and "自强(self-improvement)" are above the average level, with "礼(Propriety)" having a higher value. In the Qing Dynasty, the standardized values of "礼(Propriety)," "trust," and "frugality" are more advantageous, especially "礼(Propriety)" and "trust." The standardized values of "benevolence and righteousness" and "self-improvement" are relatively small. Other values such as "礼(Propriety)," "让(Modesty)," and "忠(loyalty)" are far below the average level. In terms of standardized values change trends, this is consistent with historical development, from the Ming to the Qing Dynasties being another period of cultural development. The Ming and Qing Dynasties are periods of stable recovery for various traditional values.

In "近现代" (1912—Present), the overall standardized values are relatively low, and Chinese society has experienced a turbulent change, both impacting and inheriting traditional values.

CONCLUSION

We create the inaugural diachronic knowledge base of traditional values in Chinese idioms, augmenting it with a

wealth of information using the large language model, GPT-4 Turbo, and manual annotation. Initially, idioms' interpretations are segmented, and GPT-4 Turbo categorizes the content into specific topics. Subsequently, a content analysis traces the evolution of Chinese traditional values, examining the evolution of topic content and the distribution of specific content.Using "benevolence and righteousness" as an example, our findings demonstrate the continuity and distinct characteristics of traditional values within Chinese idioms. The concept of "altruism" consistently permeates the topic content, reflecting its enduring presence in Chinese society. Furthermore, a diachronic analysis of standardized numerical values of traditional value idioms across different dynasties reveals the phased and wave-like nature of Chinese traditional values, aligning closely with the characteristics of various historical periods in China.Our research bridges a critical gap in the study of Chinese idioms and traditional values, enhancing the understanding of the diachronic evolution and attributes of Chinese traditional culture. In the future, we aim to broaden our scope by incorporating newly emerged idioms from online sources and expanding our research to encompass the usage of idioms in corpora, enabling more diversified investigations into idioms and values.

REFERENCES

[1] S.H. Schwartz ,"Are there universal aspects in the structure and contents of human values?" Journal of social issues, vol. 50, no. 4, pp. 19–45, 1994.

[2] X.T. Huang ,Y. Zheng , "Contemporary Chinese Youth Values Research," Beijing: People's Education Press, 2005.

[3] C.Jiang , "Chinese Traditional Values and Their Modern Transformation (Volumes I and II)," Beijing: Social Sciences Academic Press, 2020.

[4] "Dictionary of Idioms (Revised Edition · Bicolor Edition)," Beijing: Commercial Press, 2013.

[5] Z.W. Liu, "translator and annotator," The Analects, Beijing: People's Education Press, 2015.

[6] Dictionary Compilation Office, Institute of Linguistics, Chinese Academy of Social Sciences, "Modern Chinese Dictionary," Beijing: Commercial Press, pp. 1769-1785,2016.

[7] Marx, Engels, "Marx Engels Complete Works," Beijing: People's Publishing House, 1998.

[8] Y. Wu , "An Analysis of the Integration of Chinese Idioms and Socialist Core Values in Education," Modern Education, vol. 4, no. 7, pp. 178–180, 2017.

[9] Yuan Bin, "Idioms in Mencius and the Core Values of Chinese Culture," New Century Library, no. 6, pp. 14–15+44, 2015.

[10] A.S.Cheng , K.R.Fleischmann ," Developing a meta-inventory of human values," In:Proceedings of the American Society for Information Science and Technology, vol. 47, no. 1, pp. 1–10, 2010.

[11] S.H. Schwartz, J. Cieciuch , M.Vecchione , et al .,"Refining the theory of basic individual values," Journal of personality and social psychology, vol. 103, no. 4, pp.663, 2012.

[12] Zhang Z, Bai F, Gao J, et al .," Measuring value understanding in language models through discriminator-critique gap," unpublished .

[13] Yi X, Yao J, Wang X, et al .," Unpacking the ethical value alignment in big models," unpublished .

[14] C. Jiang, N.S. Wang , Michael Slote, J.L. Li , J.L.Li , C.J. Liu , M.Thomas , Wu Fan-ming, Z.Y. Wan , H.M. Qu , F.Q. Zhang ,"The Contemporary Relevance of Chinese Traditional Values - Expert Speeches Selected from the International Academic Conference on Values in a Changing World 2022," Contemporary Chinese Values Studies, vol. 8, no.1, pp. 13-38, 2023.

[15] Y. Wu ,"Exploration of Chinese Idioms Culture and Socialist Core Values in Education," Modern Education, vol. 4, no.7, pp. 178-180, 2017.

[16] B.R. Huang Bo-rong, X.D. Liao,"Modern Chinese (Sixth Edition, Revised and Enlarged)," Beijing: Higher Education Press, 2017.

[17] X.F. Zong , S.X. Wu ," Diachronic Semantic Evolution of Metaphorical Idioms in 'The Book of Odes'," Foreign Languages and Literatures, vol. 27, no.4, pp. 236-240, 2010 .

[18] Y. Yang , H. Liu , "A Diachronic Study of the Temporal Generalization of Chinese and Japanese Idioms with the Same Origin Based on a Corpus," Jinggu Wenchuang, no.27, pp. 84-86, 2020 .

[19] P. Wen , H. Li , K.B. Ma , "A Diachronic Study of the Changes in the Values of Chinese Youth in Different Periods," Youth Research, no.12, pp. 1-8, 2005.

[20] Y.Q. He , B.X. Yao,"A Diachronic Analysis of University Students' Life Values from 1998 to 2008," Chinese Health Service Management, vol. 25, no.12, pp. 846-848,2008.

[21] S.H.Schwartz, J.Cieciuch, M.Vecchione, et al .,"Refining the theory of basic individual values," Journal of personality and social psychology, vol. 103, no.4, pp. 663,2012.

[22] Z. Zhang, F. Bai , J. Gao, et al.,"Measuring value understanding in language models through discriminator-critique gap,"unpublished .

[23] Kochetova, A.Larisa,"Linguistically-Relevant Diachronic Study of Cultural Values in Early British Advertising Discourse," International Journal of Environmental and Science Education, 2016.

[24] P.Y. Liu P, S.L. Zhang , D.Yu , et al .,"CoreValue: Chinese Core Value-Behavior Frame and Knowledge Base for Value Computing", Proceedings of the 21st Chinese National Conference on Computational Linguistics,pp. 417-430,2022.

[25] H.R.Wang ,D. Yu ,"Construction of a Fine-Grained Chinese Moral Semantic Knowledge Base for Machine Moral Judgment Tasks," Chinese Journal of Information Science, vol. 36, no.7, pp. 59-68,2022.

[26] Zhonghua Book Company Editorial Department ,"The Twenty-Four Histories，" Zhonghua Book Company,2000.

[27] Karen Armstrong," The Great Transformation: The Beginning of Our Religious Traditions," Hainan Publishing House,2010.

[28] Y.G.Wu," General History of Chinese Customs: Sui, Tang, and Five Dynasties Volume," Shanghai: Shanghai Literature and Art Publishing House, 2001.

[29] H.Cheng,Y.Cheng,"Collected Works of the Two Chengs,"Beijing: Zhonghua Book Company, 1981.

[30] X.P.Yuan,"History of Chinese Literature (Second Edition)," Beijing: Higher Education Press,2005.

[31] L.Song et al.," History of the Yuan Dynasty, Volume 105, Treatise on Law," Beijing: Zhonghua Book Company,1976.

[32] T.Y.Zhang,"History of the Ming Dynasty,"Beijing: Zhonghua Book Company,1974.

[33] L.Wang,"Ancient Chinese Language (Volume 2),"Beijing: Zhonghua Book Company,1999.

[34] L.Yuan,"Literati in Hangzhou and Hangzhou Literary Circles under the Background of the Three Feudatories Rebellion,"Journal of Suzhou University of Science and Technology (Social Science Edition), vol. 39, no.4, pp. 61-67,2022.

[35] M.Q.Si,"Records of the Grand Historian,"Beijing: Zhonghua Book Company,2019.

[36] Z.Jin,"The Imperial Examination System and Chinese Culture,"Shanghai: Shanghai People's Publishing House,1990.

[37] J.W.Zou,"General History of China: History of the Wei, Jin, and Southern and Northern Dynasties,"Beijing: Jiuzhou Publishing House,1992.

[38] F.Z.Wang,"Critique of the Comprehensive Mirror,"Beijing: Zhonghua Book Company,2022.

[39] Y.S.Huang,"Traditional Chinese Legal System and Thought,"Taipei: Wunan Books Publishing Co., Ltd,1998.

[40] S.L.Shu,"Basic Knowledge of Confucianism,"Shaanxi: Shaanxi Normal University Press,2012.

[41] Y.K.Chen,"On Han Yu,"Historical Research,no.2,1954.

[42] X.D.Jiang,"The Neo-Confucianism of the Five Masters of the Northern Song Dynasty,"Literature, History and Philosophy,no.5, 2007.

[43] Y.M.Wang,"The Record of Teaching and Practice,"Shanxi: Sanqin Publishing House,2018.

[44] Zhao Erxun et al. Draft History of Qing. Beijing: Zhonghua Book Company. 1998.

A. Introduction to the Scale of the Knowledge Base

TABLE III
THE QUANTITY STATISTICS OF IDIOMS BY CHINESE DYNASTIES

Chinese Dynasties	Number
先秦(Pre-Qin)	560
秦汉(QinHan)	376
魏晋南北朝(WeiJin)	222
隋唐(SuiTang)	242
宋代(Song)	276
元代(Yuan)	130
明代(Ming)	213
清代(Qing)	283
近现代以来 (Since the Modern Era (1912—Present))	170

TABLE IV
THE QUANTITY STATISTICS OF CONTENT CATEGORIES OF CHINESE IDIOMS

Content Categories of Chinese Idioms	Number
对自我(For Myself)	560
对他人(For Others)	347
对事(For Affairs)	1170
对物(For Objects)	395

TABLE V
THE QUANTITY STATISTICS OF SENTIMENT OF CHINESE IDIOMS

Sentiments of Chinese Idioms	Number
符合 (Positive Idioms Conforming to Traditional Values)	1320
违反 (Negative Idioms Violating Traditional Values)	1152

B. Annotated Model Information and Prompt

Through careful comparison by three verification personnel of the annotations of traditional values in Chinese idioms using three models, GPT-4 Turbo, ERNIE-Bot 4.0, and GLM-4, the average accuracy rates are as Table 3:

TABLE VI
COMPARISON TABLE OF ACCURACY RATES FOR ANNOTATING TRADITIONAL VALUES IN CHINESE IDIOMS USING THREE MODELS

	ERNIE-Bot 4.0	GLM-4	GPT-4 Turbo
Traditional Values Conforming Idioms	51%	50%	83%
Traditional Values Violating Idioms	50%	60%	81%

Therefore, we have chosen GPT-4 Turbo as the annotation model for the resource library.

The GPT-4 Turbo model annotation prompt is as shown in Table 4:

TABLE VII
THE GPT-4 TURBO PROMPTS FOR ANNOTATING TRADITIONAL VALUES IN CHINESE IDIOMS

GPT-4 Turbo			
prompt	Traditional Values · Conforming Idioms	请问成语{name}符合 "仁义"、"礼"、"智"、 "信"、"让"、"忠"、 "孝"、"俭"、"中庸"、 "自强"中的哪一个? 基于你对成语和传统价值观的理解, 给出最精确一个答案。 只要结果,不要解释。	83%
	Traditional Values · Violating Idioms	请问成语{name}违反 "仁义"、"礼"、"智"、 "信"、"让"、"忠"、 "孝"、"俭"、"中庸"、 "自强"中的哪一个? 基于你对成语和传统价值观的理解, 给出最精确一个答案。 只要结果,不要解释。	81%

C. Definition of the Elements in the Diachronic Knowledge Base of Chinese Idioms

(1) 中国成语（Chinese Idioms)

Due to the fact that traditional values are more about the requirements for human qualities and behaviors, such as "rites, wisdom, and trust," which are all norms for human qualities and actions. Therefore, we choose idioms that can clearly reflect traditional values and are not controversial about being a person and doing things.

(2)成语释义 (Interpretation of Chinese Idioms)

Since this resource library needs manual screening of idioms and manual annotation of traditional values, in order to prevent the screeners and annotators from being unfamiliar with the relevant idioms, we choose to include the interpretation of idioms as a reference for our screening of idioms and subsequent annotation of traditional values, as well as summarizing the corresponding traditional value topics. For example, the idiom "哀毁骨立(Ai Hui Gu Li)" has the interpretation: "指孝子在守孝期间由于悲哀过度而损坏身体,以致消瘦得仅有骨架支撑着,现泛指悲伤过度而过分影响健康。(Refers to a filial son who, due to excessive sorrow during the mourning period, damages his body to the point of emaciation, with only the skeleton supporting him, now generally refers to excessive sorrow that overly affects health.)" Based on the comprehensive interpretation, we can quickly screen and annotate traditional values.

(3) 感情色彩（Sentiment of Chinese Idioms）

"Ethics is a philosophical reflection on beliefs about what is right and wrong, good and evil, and what should and should not be done." "Ethics must critically examine and transform moral norms." Combining the above theories and the definition of values, we believe that as long as it reflects people's "support or opposition" to the content, it has values.

For example, the idiom "偷鸡摸狗(steal a chicken and steal a dog)" has a negative sentiment, from which we can clearly see that people oppose the behavior of "偷鸡摸狗(steal a

chicken and steal a dog)," thus reflecting the corresponding values. Sentiments include positive and negative: "A positive sentiment indicates the speaker's approval and praise for the relevant matter, while a negative sentiment indicates the speaker's aversion and criticism for the relevant matter."

Therefore, by combining the sentiment of the idiom, we can see people's support or opposition to the content of the idiom. Idioms in the dictionary may not have a clearly defined positive or negative sentiment, but they can be given the corresponding sentiment based on their synonyms or the contexts we commonly use. For convenience, we will refer to positive idioms as "正向成语（positive idioms）" and negative idioms as "负向成语（negative idioms）."

(4) 朝代（Chinese Dynasties）

Dynasties can provide us with the diachronic distribution information of idioms, from which we can conduct a diachronic analysis of traditional values in idioms. Most idioms in the dictionary have clear dynasty information. For idioms with missing dynasty information but containing a source, we can find specific dynasty information based on the source, such as the idiom "白面书生(Bai Mian Shu Sheng)" comes from "宋书(Song Shu)," which was written during the Liang Dynasty of the Southern Dynasties.

To avoid controversy, the dynastic division refers to the Historical Chronology published by the Dictionary Compilation Office, Institute of Linguistics, Chinese Academy of Social Sciences, as compiled in the "Modern Chinese Dictionary".

At the same time, due to the uneven distribution of idiom dynasties, some periods only have less than ten idioms, such as those distributed in the Jin Dynasty (1115-1234) period, which only have "海枯石烂、井然有序、昧地谩天、人云亦云、一相情愿 (Hai Ku Shi Kan, Jing Ran You Xi, Mei Di Man Tian, Ren Yun Yi Yun, YI Xiang Qing Yuan)" five idioms. Therefore, in order to facilitate statistical explanation, this article does not make overly detailed divisions of historical periods. Instead, it is unified to refer to the period from the pre-Qin period (about 2070 BC - about 1600 BC) to the Eastern Zhou period (475 BC - 221 BC) as 先秦（the pre-Qin period）, and from the Western Han Dynasty (221 BC - 207 BC) to the Eastern Han Dynasty (25 - 220) as 秦汉（the Qin and Han Dynasty）. The period from the Three Kingdoms (220-280), Western Jin (266-316), Eastern Jin (317-420), Southern Dynasties (420-589) to Northern Dynasties (439-581) is referred to as 魏晋南北朝（the Wei, Jin, Southern and Northern Dynasties）. The period from the Sui Dynasty (581-618) to the Tang Dynasty (618-907) is referred to as 隋唐（the Sui and Tang Dynasties）. The period from the Northern Song Dynasty (960-1127) to the Southern Song Dynasty (1127-1276) (during which there was the Western Xia Dynasty (1038-1227) and the Jin Dynasty (1115-1234)) is referred to as 宋代（the Song Dynasty）. The period after 清朝（the Qing Dynasty）（1911) (excluding the Qing Dynasty period) is referred to as 近现代以来(Since the Modern Era (1912—Present). Finally, nine historical periods are determined: "先秦、秦汉、魏晋南北朝、隋唐、宋、元、明、清、近现代以来（Pre-Qin, Qin and Han, Wei, Jin,

Southern and Northern Dynasties, Sui and Tang, Song, Yuan, Ming, Qing, and Since the Modern Era (1912—Present)）."

(5) 中国传统价值观（Chinese Traditional Values）

We refers to relevant monographs (the references are provided in the original text) and, based on the research characteristics, selects traditional values that are more important and universal in reflecting the requirements for human qualities and behaviors. These traditional values all meet the characteristics of the idioms selected by us. At the same time, to ensure the distinction between value categories, we finally determines ten traditional values: "仁义、礼、智、信、忠、俭、让、孝、中庸、自强" The definitions of each traditional value are as follows:

① 仁义(Benevolence and Righteousness): "仁（Benevolence）" refers to "仁爱、爱人（loving and caring for others）." The content of benevolence is very broad and includes" 德、利他、包容、公正（virtue, altruism, tolerance, justice）," and anything that meets these qualities can be considered "benevolence." "义（Righteousness）" means "道义（morality）," as Mencius said, "不义而富且贵，于我如浮云。（Unrighteousness, yet rich and honored, is to me as floating clouds.）" In other words, "义（righteousness）" means considering "others" when making choices, adhering to the principle of "virtue" in actions, and doing what is just and principled. Since the concepts of "benevolence" and "righteousness" are relatively close, they are combined for the convenience of categorizing and explaining idioms.

② 礼(Propriety): The basic norms of society. It is a complete system of etiquette and a mechanism of constraint, but also a spiritual connotation. In fact, the discussion of "rites" is very complex, and this article can only take the basic meaning of "rites," that is, "behavioral social norms."

③ 智(Wisdom): Confucianism discusses "wisdom" less, mostly focusing on "道德之知(moral knowledge)." In traditional culture, " 明白是非、曲直、邪正、真妄 (understanding right and wrong, fairness, evil, and truth)," "知万物阴阳之变化，对事物的过去、现在和未来的变化胸有成竹(knowing the changes of all things yin and yang)," and "having a clear plan for the past, present, and future changes of things." Therefore, we defines "wisdom" as the pursuit of knowledge, distinguishing right from wrong, and adapting to changes flexibly.

④ 信(Trust): "主忠信，无友不如己者(Hold faithfulness and sincerity as primary principles)" (from "The Analects of Confucius, Chapter "). "Trust" means integrity, that is, being consistent in words and actions, and trusting and uniting with each other.

⑤ 忠(Loyalty): "Confucius spoke of faithfulness and trust, which is '忠实诚信(faithfulness and integrity.)' '尽己之心以待人谓之忠，推己之心以及人谓之恕。(To do one's best for others is called faithfulness, to extend one's best intentions to others is called tolerance.)' Therefore, we define "loyalty" as "faithfulness," which is a deeper level of "commitment" than "trust," with a longer time span and a deeper traditional value,

that is, loyalty to justice and faith, loyalty to the country, and loyalty to the family.

⑥ 俭(Frugality): "Frugality" means thrift, living a frugal life and maintaining inner calmness in thought.

⑦ 让(Modesty): Courtesy, respect, and tolerance. "让，礼之主也。(Courtesy is the main principle of etiquette.)"

⑧ 孝(Filial Piety): Filial piety and obedience.

⑨ 中庸(Moderation or the Middle Way): "The Confucian concept of 'moderation' generally means a person's '致中、用中的品质(quality of being in the middle and using the middle.)'" "中庸 is harmony." Therefore, we define moderation as "seeking the middle and being moderate," emphasizing that being a person and doing things should not be excessive.

⑩ 自强(self-improvement and perseverance): "天行健，君子以自强不息（The sky is vigorous, and the gentleman is never weary of self-improvement）"（from "The Book of Changes, The Commentary of the Great Image, The Heavenly Stem"），"士不可以不弘毅（A gentleman must be robust and resolute）"（from "The Analects of Confucius, Chapter 7"）. These are manifestations of "自强不息（self-improvement without ceasing）." Therefore, we abbreviate "自强不息（self-improvement without ceasing）" as "self-improvement," mainly reflecting the Chinese traditional value of enterprising and pioneering.

(6) 内容分类（Content Categories of Chinese Idioms）

As mentioned earlier, traditional values often require people to maintain certain qualities and behaviors. In the "1857-1858 Economic Manuscripts," the overall relationship between "nature, human beings, and society" is depicted, describing a diachronic blueprint from "the dependence of human beings on each other" to "the independence of human beings based on the dependence on objects" to "the free individuality of human beings." Based on this theoretical foundation and the characteristics of idioms under traditional values, the content categories of idioms are determined as follows:

①对自我（For Myself）: This refers to the requirements for one's own qualities. In idioms, it is manifested as descriptions of human qualities, character, and pursuits, such as "士可杀不可辱（A scholar may be killed but must not be humiliated）" and "安分守己（Be content with one's lot and keep to one's own place.）"

② 对他人（For Others）: This refers to one's attitude and behavior towards others. In idioms, it is manifested as ideas, actions, and interpersonal relationships, such as "斑衣戏彩（Play in multicolored clothes to make parents happy.）" and "阿谀奉承（Flatter and fawn upon）."

③ 对事（For Affairs）: This refers to how to conduct oneself in doing things. In idioms, it is manifested as the values that should be adhered to when dealing with problems, such as "安不忘危(Be not forgetful of danger even in times of safety)" and "安常习故(Be accustomed to peace and be accustomed to the old)."

④ 对物（For Objects）: This refers to the attitude towards fame, wealth, righteousness, and occupation. In idioms, it is manifested as how to face and handle these matters, such as

"爱日惜力(save time and cherish one's strength)" and "爱惜羽毛(Love one's fame and protect one's reputation)."

D. Selection of Chinese Idioms in the Diachronic Knowledge Base of Chinese Idioms

（1）Source of Idioms and Their Definitions

The idioms are selected from the latest revised edition of the bicolor "Dictionary of Idioms" published by the Commercial Press of China. We chose this dictionary as the basis for constructing the resource library of traditional values in idioms for the following reasons:

① The dictionary is authoritative, containing over 18,000 idioms, which are relatively comprehensive.

② Each entry in this dictionary clearly states the source, which allows us to determine the dynasty in which the idiom appeared based on the source.

③ Most entries in this dictionary provide the corresponding sentiment, which is convenient for us to label in the subsequent steps.

（2）Selection Criteria for Idioms

Idioms are the carriers of Chinese traditional culture, and the concepts and behaviors embodied in idioms reflect rich traditional values. For example, the concept and behavior embodied in the idiom "百闻不如一见(Better see once than hear a hundred times)" contain the value of "智(wisdom)." Values are based on human choices, cognition, and other content, and traditional values are more about the requirements for human qualities and behaviors. Therefore, we will choose idioms that are primarily about individuals and are related to personal qualities and actions to build the resource library of traditional values in idioms. For example, "阿谀取容(flatter for favor)" mainly refers to ingratiating oneself to others for their favor; "背信弃义(betray trust and righteousness)" mainly refers to breaking promises and not adhering to moral principles, selling out friendship between friends. Such idioms are related to both actions and personal qualities, making it easier to see traditional values. Therefore, we have established the following criteria for selecting idioms:

① Idioms unrelated to personal behavior and actions are not selected, such as "斑驳陆离、半斤八两" etc.

② Idioms whose subject is not an individual are not selected, such as "不得人心(Not pleasing to the people)," "闭关锁国(Close the country and lock the state)," etc.

③ Idioms that are clearly marked with sentiments in the dictionary or whose sentiments can be determined based on the sentiments of synonyms, such as "察察为明(Chat Chat Wei Ming)," which is a synonym for "斤斤计较(Jin Jin Ji Jiao)," a negative idiom.

④ Idioms that only reflect values in terms of usage are not selected, such as "班门弄斧、不揣冒昧" which is used as a modest usage in the dictionary.

⑤ Idioms with ambiguous meanings or whose original meaning is unrelated to personal behavior and actions are not selected, such as "白龙鱼服(white dragon in fish attire)," "避实就虚(avoids the solid and attacks the void)," etc.

⑥ Idioms that do not explicitly state specific principles for guiding personal behavior and actions are not selected, such as "败国丧家(ruin the state and lose the family)," which refers to "ruining the country."

⑦ Idioms unrelated to describing personal qualities are not selected, such as "白发红颜(white hair and red complexion)," which describes a person's state, "弊绝风清(extremely pure and honest)," which describes a society's honest atmosphere, "冰肌玉骨(delicate and fair complexion)," which describes appearance, "六根清净(clean and clear six roots)," etc., which are Buddhist content idioms.

⑧ Idioms with ambiguous meanings are not selected, such as "不明不白(not clear and not clear)," which describes confusion, and can also describe ambiguous, unclear, and immoral relationships, or actions. Although it is a negative idiom, this idiom can describe a person's character, interpersonal relationships, and actions, with a wide range of uses.

E. Annotation Specifications for the Content of the Diachronic Knowledge Base of Traditional Values in Chinese Idioms

We utilize the language model GPT-4 Turbo to annotate traditional values and then conduct manual verification and modification. When making modifications, we follow the following specifications:

(1) Sentiment Verification and Annotation: First, judge whether the model's annotation result is acceptable. If there is disagreement, it is necessary to make a judgment based on the information of synonyms given in the dictionary.

(2) Modification and Annotation of Unfamiliar Idioms: Make a judgment based on the interpretation. For example, for the idiom "斑衣戏彩," the annotator needs to refer to the corresponding interpretation "pursue filial piety and make parents happy," and thereby determine the traditional value of "孝."

(3) Confusion of Traditional Value Categories: It is necessary to strictly follow the definitions of traditional values and the content of idioms for verification and annotation, and not to rely solely on the literal meaning. For the idiom "搬弄是非," the concept and behavior it embodies are relatively ambiguous between violating "仁义" and violating "礼." In this case, we need to refer to the definitions of "benevolence and righteousness" and "礼" to find that "仁义" is not a selfless act, but it does not violate social norms. Therefore, it is determined as "仁义."

(4) Confusion of Content Categories: This is the same as (3), and both should be strictly distinguished according to the definitions.

F. Annotation Process for the Content of the Diachronic Knowledge Base of Traditional Values in Chinese Idioms

(1) Sentiment Annotation by Language Model: We use the language model GPT-4 Turbo to annotate the sentiment of idioms. There are some idioms in the dictionary that inherently have a specific sentiment, but this information is not provided in the dictionary. We can use the information of synonyms listed in the dictionary to determine the sentiment of these idioms. For example, the synonym for "察察为明" is "斤斤计较," and the dictionary does not provide the sentiment for "察察为明," but "斤斤计较" is clearly marked as negative. There are a total of 816 idioms like this in our resource library.

(2) Manual Verification and Modification of Sentiment: We recruit three students to verify and modify the sentiment annotations made by the language model, and select the consensus result. For results that are controversial, we will prioritize the result that is consistent between two people.

(3) Traditional Values Annotation by Language Model: We use the language model GPT-4 Turbo to annotate the traditional values of idioms. Since values are more personal and related to individual choices and motivations, values include both positive and negative values. However, traditional values are those that we advocate and praise. Therefore, during the annotation process, based on the sentiment annotations from the previous step, we categorize the idioms into two types:正向成语 (positive idioms) and 负向成语(negative idioms). When annotating positive idioms, we mainly judge which category of traditional values the idiom conforms to. When annotating negative idioms, we mainly judge which category of traditional values the idiom violates. For example, "斑衣戏彩" is a positive idiom that conforms to "孝," and "搬弄是非" is a negative idiom that violates "仁义."

(4) Manual Verification and Modification of Traditional Values: The process is the same as that for verifying and modifying sentiments.

(5) Content Category Annotation by Language Model: We use GPT-4 Turbo to annotate the four content categories of idioms. "对自我(For Self)" mainly focuses on idioms that describe human qualities, such as "大仁大义(Great benevolence and righteousness)"; "对他人(For Others)" mainly focuses on attitudes, opinions, and behaviors towards others, such as "斑衣戏彩(Play in multicolored clothes)"; "For Affairs" mainly focuses on methods of doing things, such as "表里相济(External harmony and internal integrity)"; and "For Objects" mainly focuses on views on wealth, justice, the state, and laws, such as "财迷心窍(Money-mindedness)."

(6) Manual Verification and Modification of Idiom Content Categories: The process is the same as that for verifying and modifying sentiments.

(7) Final Verification of All Results for Correctness and Reasonableness.

G. Diachronic Analysis of the Overall Proportion of Traditional Values in Chinese Idioms

Overall, as depicted in Figure 5, we can observe that the proportion of each traditional value within a dynasty is relatively balanced. Despite the changes through different eras, the excellent traditional values of China still exhibit strong stability and continuity.

"仁义" (benevolence and righteousness) is predominant in various dynasties, followed by "自强" (self-improvement) and "智" (wisdom). This indicates that "仁义," as a value that has been enduring, has had a profound impact on us.
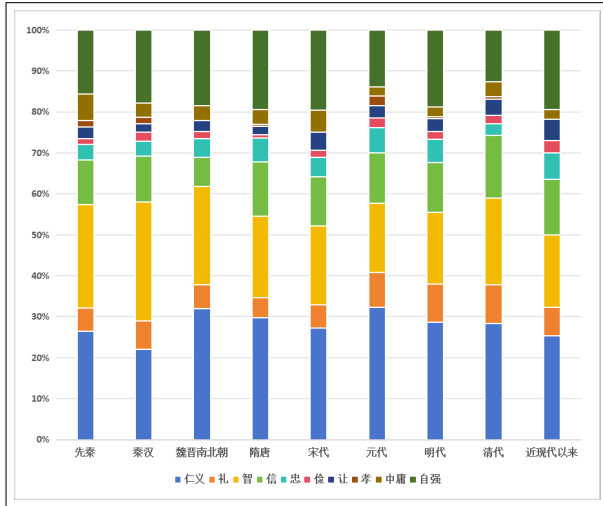
Fig. 7. Overall Proportion of Traditional Values in Chinese Idioms over Time

The Pre-Qin period, as a significant period in the development of Chinese civilization, stands out compared to other periods. For example, idioms from the Song Dynasty lack content related to ”孝” (filial piety). Each traditional value has a proportion, indicating that the various thoughts of the Hundred Schools of Thought in the Pre-Qin period, especially Confucian culture led by Confucius and Mencius, have had a profound impact on China. Idioms like ”哀而不伤” (”Sorrow without pain”) and ”不在其位不谋其政” (”Not thinking about state affairs if not in office”) can be seen as discussions of ”中庸” (moderation) and ”礼” (rites), from which modern Chinese concept of ”moderation” and ”acting with principles” originates.

During the Wei, Jin, Southern and Northern Dynasties to the Sui and Tang Dynasties, the proportions of the traditional values of ”俭,” ”礼,” ”让,” and ”智” decreased compared to previous periods, while the proportions of ”仁义,” ”自强,” ”信,” and ”忠” showed a more obvious increase trend during the Sui and Tang Dynasties, which was related to the cultural prosperity and the strict establishment of the city and market system during that period. People paid attention to ”忠” and ”信” in dealing with others, and ”自强” was more closely related to the flourishing national strength of the Sui and Tang Dynasties. In the Song Dynasty, we can observe a clear trend of change in traditional values, and the proportions of various traditional values became more balanced in the Song Dynasty, such as ”自强,” ”让,” ”俭,” ”忠,” and ”让” all had proportions, which was greatly related to the emergence of ”Neo-Confucianism” during the Song Dynasty.

From the Song Dynasty to the Yuan Dynasty, the number of traditional values compared to other periods showed a significant decrease. During the Yuan Dynasty, most literary works were dramas, and Yuanqu (Yuan dynasty opera) was mainly characterized by colloquial language. Additionally, the Mongol rulers promoted the policy of Mongol supremacy to prevent their assimilation into Han culture. Therefore, they did not emphasize the inheritance of Confucian culture. In the Yuan Dynasty, if a Mongol killed a Han or Southern Chinese person, they would have to pay with their lives, whereas if a Mongol was killed by a Han or Southern Chinese person, they would only have to serve a military campaign and pay a fine, along with the full cost of the funeral.

In the Ming Dynasty, the proportions of the traditional values of ”自强” and ”忠” increased significantly compared to the Yuan and Qing Dynasties. This was related to the unprecedented strengthening of the monarchy in the Ming Dynasty and the further unity and integration of the multi-ethnic nation.

The Qing Dynasty, following the Song Dynasty, was another period where the proportions of traditional values increased. This was highly related to the prosperity of traditional ”小学” (elementary education) during the Qing Dynasty. Interestingly, the proportion of idioms reflecting the value of ”忠” showed a decreasing trend in the Qing Dynasty, which was somewhat related to the ”Three Feudatories' Rebellion.”

Since the Modern Era (1912—Present), in addition to the proportion of idioms reflecting ”自强” and ”Loyalty” being different, the proportions of idioms reflecting other categories of traditional values have also changed. Regarding the proportion of the traditional value of ”孝,” it has been very rare since the Wei, Jin, Southern and Northern Dynasties. The proportion of idioms reflecting the traditional value of ”俭” has generally remained stable, which can be understood as the concepts of ”孝” and ”俭” being deeply influenced by the Pre-Qin period and remaining unchanged.

However, a comprehensive observation reveals that regardless of the circumstances of the various dynasties in China, the traditional values represented by Chinese civilization have a certain continuity and stability. The number of idioms reflecting Chinese traditional values, represented by Confucian culture, has been relatively balanced across different time periods, reflecting the strong stability and vitality of these excellent traditional values.