

## Retraction

# Retracted: Implications of Artificial Intelligence Technology for the External Communication of Chinese Ethnic Cultures

### Mobile Information Systems

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

Copyright © 2023 Mobile Information Systems. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] Y. Qiao and L. Zhou, "Implications of Artificial Intelligence Technology for the External Communication of Chinese Ethnic Cultures," *Mobile Information Systems*, vol. 2022, Article ID 3539332, 13 pages, 2022.

## Research Article

# Implications of Artificial Intelligence Technology for the External Communication of Chinese Ethnic Cultures

Yanna Qiao <sup>1</sup> and Lili Zhou<sup>2</sup>

<sup>1</sup>School of Humanities and Education, Xijing University, Xi'an 710123, Shaanxi, China

<sup>2</sup>Academic Office, Xijing University, Xi'an 710123, Shaanxi, China

Correspondence should be addressed to Yanna Qiao; 20120005@xijing.edu.cn

Received 6 June 2022; Revised 28 July 2022; Accepted 18 August 2022; Published 13 September 2022

Academic Editor: Imran Shafique Ansari

Copyright © 2022 Yanna Qiao and Lili Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

At present, artificial intelligence technology is experiencing rapid development, and artificial intelligence has a significant impact on human production and lifestyle, as well as the political and cultural aspects of society. In recent years, the rapid development of artificial intelligence, in the context of this era of development, for the national cultural communication, how to faster and better adapt to the general trend of the times, and constantly interchange and improve become an important issue. As an indispensable part of the world's civilization, Chinese ethnic cultures cannot only focus on their domestic inheritance but also should establish an external spreading pattern, thus promoting their own development as well as the diversity of the world's cultures. This article discusses the implication of AI technology for the promotion of Chinese ethnic cultures' external communication. AI technology can be used in text translation, data mining, communication channel analysis, etc. It is more efficient and accurate than experienced journalists and editors in collecting big data to develop selected topics and research, judge the direction of ethnic culture communication, and discover the content that users are interested in. At the same time, the instrumental and auxiliary roles of AI translation can be used to improve the construction of ethnic culture corpus. Through the analysis of the data from the survey and experiment, it is found that the media of ethnic culture communication has changed dramatically since 2016. In 2016, the highest proportion of television channels was 70%, but in 2021, the proportion of television channels dropped to 23%, and Internet and book communication have become the "main force," of which the proportion of Internet is 59% and the proportion of book channels is 48%, and the popularity and growth speed of the mass media ranked first. The modern intelligent technology can be used to spread Chinese ethnic cultures more widely and effectively. It has a far-reaching influence in that, it can promote the inheritance and innovation of Chinese ethnic cultures, and to improve ethnic groups' confidence in their own ethnic cultures.

## 1. Introduction

China is a country with a large number of ethnic minorities. The rich and colorful ethnic cultures provide a natural soil for the study of ethnic minority cultures. Every country in the world has its own unique ethnic culture, no matter whether it is strong or weak. Culture is the precious spiritual wealth left by ancestors, and has always helped people to create material wealth. At the same time, culture also reflects the wisdom of human beings. Therefore, the study of culture is also a stimulus to today's productivity. In the conditions of the market economy, the production of cultural products and the provision of cultural services is one of the important

forms of contemporary economic life. Globalization is profoundly affecting human life, and with the increasing exchanges between different peoples, cultures are constantly changing. Many traditional ethnic cultures are facing unprecedented difficulties and challenges.

China has a vast territory and 56 ethnic groups at the same time. Different ethnic groups have created different cultures at different times and in different geographical locations. The uniqueness of these ethnic cultures should be widely appreciated and protected. Therefore, the study of ethnic minority cultures is not only related to the rise and fall of ethnic minorities but also to the development of the national strategic pattern and the long-term sustainable

development of the country. In a rapidly changing society, research is extremely important for the integration and fusion of traditional civilization and modern civilization, because the progress and development of traditional ethnic cultures should be spread to the outside world on the basis of cultural protection. The dissemination of national culture is a cross-ethnic, cross-national, and cross-regional communication activity. In the context of artificial intelligence, the identities of cultural disseminators and cultural receivers are interchangeable, and cultural receivers are no longer single cultural consumers.

Ethnic traditional culture serves as the identity card of a nation, and it is the diversified wisdom accumulated by people in long-term production and life. This article firstly analyzes the present situation and the means of Chinese ethnic cultural external dissemination, as well as cultural changes and the influence of cultural communication media on external dissemination. The innovation of this article lies in that it studies the external communication of ethnic cultures from the perspective of artificial intelligence technology, which facilitates the text translation and the transmission of messages on a large scale. The external dissemination of ethnic culture is different from the dissemination of major culture. It has unique methods and ideological systems, which should be understood from a philosophical level.

## 2. Related Work

The protection and dissemination of intangible cultural heritage in China have evolved from a single industry act to a social work involving a wide range of participants from multiple fields. After analyzing the current situation of digital games, Cui et al. proposed that digital games should be used as a new media for the protection and dissemination of intangible cultural heritage in China. Taking Hangzhou traditional food culture as an example, a light game design based on the WeChat platform was introduced to demonstrate the skills of effectively integrating elements of intangible cultural heritage into game design, but it has not been applied on a large scale, but the research scope is not extensive [1]. Watt P seeks to examine, from a cultural-historical perspective, how the “graduates” has developed into a central figure in a significant part of the contemporary labor market. It is first shown how the rise of the “new” or “knowledge economy” (throughout the 1990s and 2000s) became a new source of pressure for generations entering the world of work. Gradually, the vocabulary and characteristics of the “graduate” became more visible through complex and sophisticated patterns of cultural transmission. The themes that articulate this identity today have cultural roots that are not entirely new [2]. Brochet et al. examine how managers’ ethnic cultural backgrounds affect their communications with investors. Through earnings calls with executives from 42 countries, he found that managers from ethnic groups with individualistic cultures used a more optimistic tone and displayed greater self-recrimination. A manager’s ethnic culture has lasting effects that persist for executives whose work experience is later exposed to a different ethnic culture.

In conversations that reflect real-time interaction (i.e., question-and-answer), the influence of ethnic traditions is observed, whereas in management discussions, the influence of ethnic traditions is less pronounced. The research results show that the ethnic background of managers has a great influence on the way they communicate with the capital market and the market’s response to information disclosure, and to a certain extent reflects the influence of ethnic culture on individuals [3]. Big data thinking has important practical significance for promoting people’s cognition of ethnic cultural characteristics and realizing the sharing of ethnic cultural resources. Zhang and Yang analyzed the development of northern ethnic culture based on big data mining. Compared with the cultural characteristics of the southern plains, the northern peoples originated in the mountains or deserts, so they have a tenacious will and a strong conscious spirit. With the application of data technology, the culture of northern nationalities can be effectively protected and popularized, and the way of cultural inheritance can be changed, but the research method has not been proven [4]. The psychology of cultural dynamics is a psychological investigation of the formation, maintenance, and change of culture over time. Kashima et al. chart this terrain review the existing literature and point to potential future directions for this research. It is divided into three parts. The first part focuses on microcultural dynamics, referring to the social and psychological processes that contribute to the transmission and retention of cultural information. The second part, on micro-macro dynamics, examines how microlevel processes give rise to macro-cultural dynamics. The third part, concerned with macro-cultural dynamics, refers to the distribution and long-term trends of cultural information among populations, which in turn enable and constrain microlevel processes. Finally, concluding this review with a consideration of future directions, the study of behavior change is suggested as a study of cultural dynamics from a translation perspective. But the research questions do not go far enough [5]. The purpose of Popov et al.’s study was to clarify the cultural attributes that influence interfirm cooperation, advancing the traditional economic-based theoretical framework. The assessment of concepts related to parametric cultural assessment as a theoretical basis for structural aspects of interfirm cooperation was documented in an empirical study investigating cross-cultural analysis. Cultural attributes such as freedom of choice; advancement of norms of equal communication power; low perceived uncertainty; strategic orientation to the future; and humanistic-oriented communication norms were found to further develop interfirm cooperation. The current research provides a systematic and practical definition of attributes in the culture of interfirm cooperation. The advantages of this advanced framework are a more sustainable model of cooperation and a reduction in the costs of interfirm cooperation. But the research is not very practical [6]. These studies have provided detailed analyses of ethnic culture and mental health. It is undeniable that these studies have greatly contributed to the development of the corresponding fields. We can learn a lot from the methodology and data analysis. However, there are relatively few and not thorough enough

studies on the foreign communication of ethnic culture in combination with artificial intelligence techniques, and there is a need to fully apply these techniques to the research in this field.

### 3. Artificial Intelligence Technology for External Communication of Chinese Ethnic Cultures

Ethnic culture is the sum of material and spiritual wealth produced and created by a certain ethnic group during its long-term common production and living practices, which can reflect the characteristics of the ethnic group. National culture reflects the level of the historical development of the nation. The emergence of many ethnic cultures is to solve the difficulties in production and life. Therefore, for the folk, ethnic culture serves the basic functions of comforting, edifying, and gathering [7]. The main purpose of ethnic culture is to drive economic development and improve the quality of life of local people. When disseminating to the outside world, it is necessary to disseminate the meaningful parts according to the characteristics of each culture [8]. There are some problems existing in the spread of ethnic cultures to the outside world. For example, firstly, the content of external communication of culture is relatively one-sided. The external dissemination of Chinese culture includes both traditional culture and new socialist culture [9]. However, the cognition of Chinese culture around the world is mainly limited to traditional culture, and a comprehensive and rational cognition of the new culture formed since the founding of New China has not been formed [10]. Most of the cultural products exported from China are from traditional culture. On many occasions, it is also the traditional elements of Chinese culture that are accustomed to being presented to the world [11]. Secondly, the leading force of external communication has been the government, and the communication activities form tends to be single. Although the current nongovernmental cultural exchanges are becoming more and more frequent, the government is still the main body of cultural external communication and the main promoter of external cultural communication. As the main body of communication, the cultural exchange activities organized by the government are inevitably considered to have political overtones and are rejected and resisted [12]. Even the international cultural exchange activities carried out by cultural and art institutions or nongovernmental organizations are also subject to the suspicion of governmental interference. Thirdly, the communication method is traditional and lacks innovation, which can always be triggered by the outcome of new technology. Finally, the audience analysis is ignored, and the goal is not clear. Foreign cultural communication is not a single cultural indoctrination. The communication process requires the interaction of both parties, and the communication effect is reflected in how the audience responds and how much influence cultural communication has on the audience. Different audiences have different preferences and requirements for communication content, methods, media,

and many other factors. Artificial intelligence technology could serve as a solution to some of these problems.

Over the past decades, the method of communicating Chinese culture to the rest of the world has been changing with social and technological signs of progress. Cultural exchange activities, films, television programs, publications of books and periodicals, sports competitions, cultural exhibitions, online forums, and other communication methods have jointly promoted the development of Chinese cultural communication. However, with today's fast-changing science and technology, the communication method should also keep up with the development of the times and constantly keep being updated [13]. The traditional communication method not only limits the content of the communication but also inevitably leads to the audience's aesthetic fatigue after a long time. The continuous development of Internet media has changed the means and environment of communication. The emergence and development of new media provide a new platform and channel for cultural communication. At the same time, the communication mode of culture also needs to keep pace with the times for continuous innovations, and rely on the Internet to provide new ideas for the innovation of communication mode [14].

*3.1. Effectiveness of the Network in Ethnic Cultural Communication.* Artificial intelligence is mainly a new technical science that conducts research, development, simulation, extension, and expansion of functional theories, approaches, technologies, and application systems. The foreign communication of national culture in the context of this new era must be constantly innovative and must incorporate certain characteristics of the era of artificial intelligence so as to enhance the advanced nature of national culture. The pressure that traditional culture encounters today is not the culture itself, but the dilemma in the form of communication. Inheriting and promoting the excellent Chinese traditional culture needs to stimulate the modern public's sense of participation, which cannot be separated from the participation and rational use of new technologies. The rapid development of artificial intelligence in recent years is exciting and the related research results are rich, so the analysis of artificial intelligence concepts is also diverse, and in its essence, it is a simulation of human thinking. Thinking is a subjective reflection of objective reality, so thinking is both subjective and objective. And according to the diversity of thinking and the different perceptions of thinking, the human-centered approach must be part of the empirical sciences, involving observations and assumptions about human behavior. The rationalist approach involves a combination of mathematics and engineering. For the research on the application status of the network in the dissemination of ethnic culture, the usefulness, ease of use, and influence of the website can be analyzed. Figure 1 shows the effectiveness of the network in ethnic cultural communication.

*3.2. Artificial Intelligence Technology for Analyzing the Internet Users' Interest.* Artificial intelligence technology is the use of computers to achieve computing similar to that of the

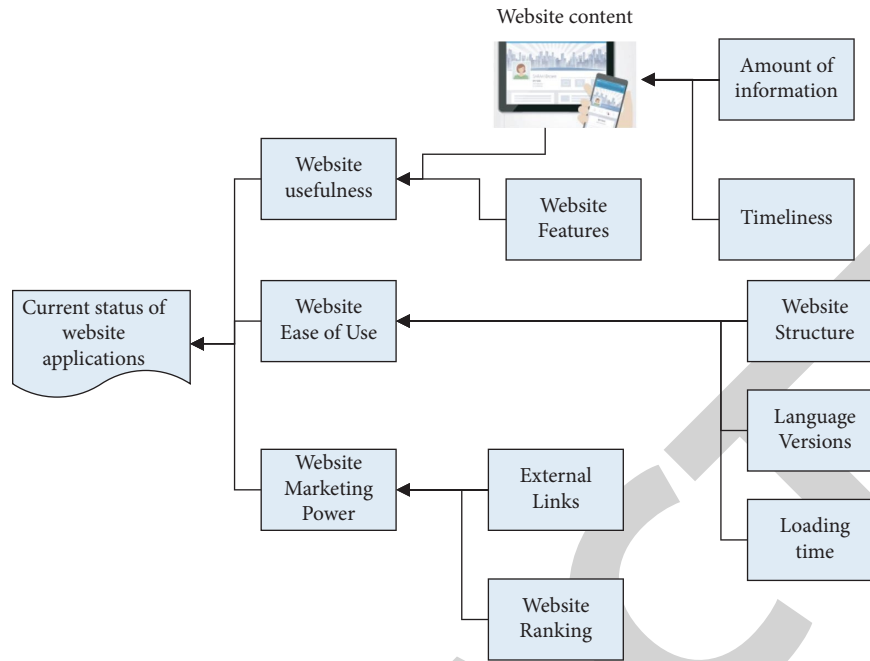


FIGURE 1: The effectiveness of the internet in the communication of ethnic culture.

human brain, thus realizing the use of computers at a higher level. Artificial intelligence is the product of human brain work, and human intelligence is applied to production through practice, thus materializing the essential power of human knowledge and wisdom in products. Artificial intelligence technology is more efficient and accurate than experienced journalists and editors in using big data to develop topics, judge the direction of ethnic culture communication, and discover the content of interest to users. In the process of cultural dissemination on the Internet, users play an important role. Different from traditional social network communication based on information content, information communication in online society mainly depends on the relationship between users. Therefore, in the study of information dissemination, user characteristics should be discussed. Figure 2 shows the change in user relationship status in the process of external communication of ethnic culture by means of online forums. For example, when user A creates a post, some users, including user A's followers, will soon receive the post and comment or retweet it. If the user is a general user with few followers, the information may spread very slowly, or even disappear quickly [15]. On the contrary, if the user is a star user and has a large number of fans, the news may spread rapidly, reach a peak in a short time, and may become a hot topic in the near future. The users in the user's friend list are likely to be friends in real society. They follow each other and become friends with each other. They have a higher probability of forwarding or commenting on information to complete the dissemination of information. The number of user comments indicates the time users spend on Weibo, and users with high comments are more active on the Internet and are more likely to spread information. In the process of information dissemination, people with different process characteristics have different

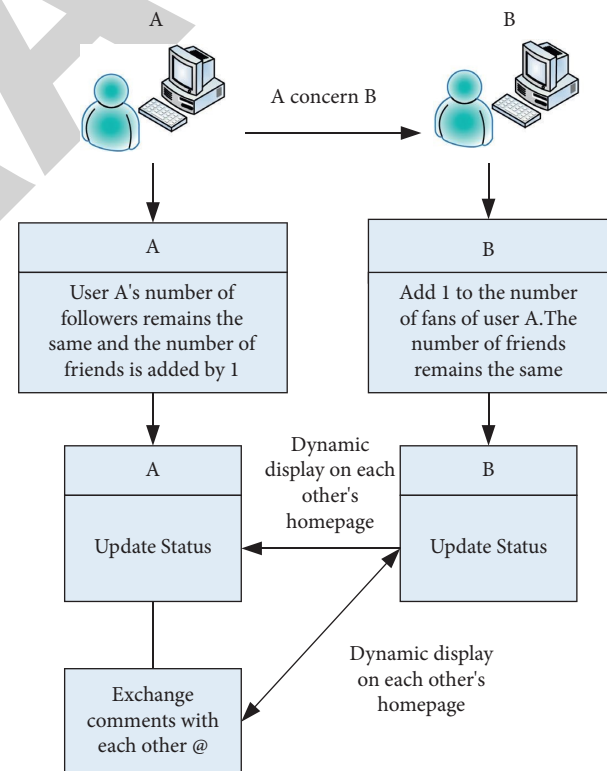


FIGURE 2: Changes in the state of user relationships in the process of cultural communication.

influences, and these characteristics should be taken into account when researching and establishing predictive information dissemination models [16].

For user  $Q$ , consider his/her characteristics  $T$ : number of followers  $T_1$ , number of friends  $T_2$  and number of

comments  $T_3$ . The number of fans, friends, and comments represents the user's popularity, interest, and activity, respectively. High popularity allows information to spread quickly. The degree of interest represents the user's hobbies [17]. Research shows that the interests of users are very similar to those in real life. Users among friends may be friends in real life. Once they receive information from them, the probability of spreading is very high. Activeness represents the time spent on the network, generally, the more time spent on the network, the higher the activeness. Different features have different effects on dissemination. In the process of information dissemination, user features are defined as follows:

$$G(m) = \sum_{n=1}^k \alpha_n^m x_n^m, k, m, n = 1, 2, \dots \quad (1)$$

Among them,  $G(m)$  represents the development trend of information at time  $m$ , that is, the number of posts at  $m$  time;  $x_n^m$  represents the characteristics of users participating in information dissemination at  $m$  time;  $\alpha_n^m$  represents the influence probability of  $x_n^m$  at  $m$  time.

To this end, an information propagation model based on user characteristics is proposed. The details of the model are as follows: Time is discrete, with a minimum step size of 1. In the network, the speed of information dissemination is much higher than that of traditional media. Information may reach a dissemination peak in a short period of time, and then quickly decline after reaching the peak. Therefore, minutes are used as the minimum time unit in the experiment [18].

Because the information spreads very fast with the network, after commenting or forwarding the information at the time  $t_m$ , most users can receive the information and process it in time, but some users cannot receive the information in time due to other reasons. For example, information is posted at midnight, during that time period, when most people have taken a break, or during working hours, the information cannot be viewed in time. The contribution of this part of users to information dissemination cannot be reflected in the moment  $t_m$ , and he may appear in subsequent  $t_{m+q}$ ,  $q = 0, 1, 2, 3, \dots$ . This part of the impact is defined as follows:

$$H(m) = \sum_{m=1}^q \beta_{m-1} Y_{m-1} q, m = 1, 2, 3, \dots \quad (2)$$

Among them,  $H(m)$  represents the delay of the user's influence on information dissemination at time  $m$ ,  $\beta_{m-1}$  represents the weight of  $Y_{m-1}$ 's influence on the present at time  $m-1$ , and  $Y_{m-1}$  represents the number of information comments or reposts at time  $m-1$ . The model is represented as follows:

$$Y(m) = G(m) + H(m). \quad (3)$$

Substituting  $G(m)$  and  $H(m)$  into  $Y(m)$  to get

$$Y(m) = \sum_{n=1}^k \alpha_n^m x_n^m + \sum_{m=1}^q \beta_{m-1} Y_{m-1} k, q = 1, 2, 3, \dots \quad (4)$$

Assuming the number of posts  $Y(m)$  per hour on a topic hotspot of ethnic culture as the output variable, and the attributes of user  $T$ , namely the number of fans  $T_1$ , the number of friends  $T_2$ , and the number of comments  $T_3$  as the input variable.

$$Y(m) = \sum \alpha^m x^m + \sum_{m=1}^q \beta_{m-1} Y_{m-1} q = 1, 2, 3, \dots \quad (5)$$

Among:

$$T_m = (T_1^m, T_2^m, T_3^m). \quad (6)$$

To test the proposed model, the model is validated using the root mean square error (RMSE), which is defined as follows:

$$K_{\text{RMSE}} = \sqrt{\frac{\sum_{m=1}^Q (\hat{y}_m - y_m)^2}{Q}}, \quad (7)$$

where  $N$  represents the number of samples, and  $\hat{y}_m$  and  $y_m$  represent the estimated and true values of the output, respectively.

In the network, when a user posts information, his/her fans or others may not be online due to work reasons, so they cannot see the information immediately. This information can only be seen and commented on when they are free online, so there will be a lag in the dissemination of information. However, in an online social network, information is updated very fast, and people's attention is changed very fast. After the information is released for a period of time, users will lose interest in the topic, so the probability of user comments drops significantly [19]. In experiments, it has been shown that if people have not commented for more than 24 hours after receiving a message, the probability of commenting after that is very low, so here,

$$H(m) = \sum_{m=1}^q \beta_{m-1} Y_{m-1} q = 1, 2, 3, \dots \quad (8)$$

A text classification system must be used to classify cultural texts when ethnic cultures are spread on the Internet. The function of the text classification system is to automatically determine the number of text links according to the text content under a specific classification system [20]. Mathematically, text classification is a mapping process that maps undivided text into existing categories. Mathematically expressed as follows:

$$f: A \longrightarrow B. \quad (9)$$

Among them,  $A$  represents the text set to be classified;  $B$  represents the division of the classification system.

Text classification mapping rules refer to the types and rules established by the system based on multiple copies of each type of data to summarize the consistency of classification. When testing a new text, the text category is determined by type and rules [21]. At this stage of the text processing process, words, and phrases can be selected as elements of the text, but it is generally considered preferable

to select words. Therefore, when the classification system is based on the vector space model, to represent text in vector form in vector space, the text must first be subdivided into words, and these words must be used as vector elements to represent the features of the text. However, in each entry record, there are a large number of different words separated by Chinese entries. After searching, it is found that these words not only convey less textual information but also have a certain exclusion effect. In order to improve the efficiency and accuracy of classification system processing, all independent words should be removed in text processing [22]. Furthermore, Chinese characters containing mathematical symbols or English characters did not contribute much to the correct classification, nor did they convey a negligible level of information. Finally, the Chinese text classification system needs to filter English words, such as English abbreviations in Chinese text, to ensure that the features represented by the Chinese text are Chinese words. The text processing process described in this paper can be summarized into three parts, as shown in Figure 3.

**3.3. Artificial Intelligence Technology for Cross-Cultural Text Translation.** The dissemination of culture is inseparable from language translation. In the process of translation, translators often encounter various translation problems caused by cultural differences. This article introduces a cross-language document classification method based on machine translation and word frequency features from both an intuitive and formal perspective. Intuitively, with this approach, a source-to-target machine translation model is used to translate the test set into the same language as the training set. In this way, the cross-lingual document classification problem is transformed into a standard monolingual document classification problem [23]. After translation, documents on both training and test sets are represented using discrete word frequency features and a bag-of-words model. An average perceptron is used for document classification. This process is shown in Figure 4.

This paper records a sentence in the source language as follows:

$$x = \langle x_1, x_2, \dots, x_{|x|} \rangle. \quad (10)$$

And record a sentence in the target language as follows:

$$y = \langle y_1, y_2, \dots, y_{|y|} \rangle. \quad (11)$$

The cross-language document classification task view predicts the label  $c$  of the sentence  $y$  in the target language given the training set  $A = \{\langle x_m, c_m \rangle | m \in [1, N]\}$  in the source language. In this formal description, documents are treated as long sentences composed of multiple sentences.

In the cross-lingual document classification task, an arbitrary model can be regarded as using formula (12) to predict the class  $c$  of a document.

$$P(c | y, A) = \sum_{\alpha} P(c | y, \alpha) P(\alpha | A). \quad (12)$$

Here  $\alpha$  is the model parameter of the conditional probability  $P(c | y, \alpha)$ .

The baseline approach to machine translation used here trains the classifier directly on the training set in the source language while making predictions on the translated test set. Therefore  $\alpha = \alpha_w$  is estimated in the source language. By Bayesian formula, formula (12) can be decomposed into formula (13).

$$P(c | y, A) = \sum_{\alpha} \sum_x P(c | x, \alpha_w) P(x | y) P(\alpha_w | A), \quad (13)$$

$$P(c | y, A) \approx P(c | \hat{x}, \hat{\alpha}_w) P(\hat{x} | y) P(\hat{\alpha}_w | A), \quad (14)$$

$$P(c | y, A) \propto P(c | \hat{x}, \hat{\alpha}_w), \quad (15)$$

where the parameter  $\hat{x} = \operatorname{argmax}_x P(x | y)$ ,  $\hat{\alpha}_w$  is obtained by maximum likelihood estimation (MLE) on  $A$ ,  $\hat{\alpha}_w = \operatorname{argmax}_{\alpha_w} \log P(\alpha_w | A)$ , and “ $\approx$ ” by 1-best estimation.

Specifically, when estimating the classifier parameter  $\alpha_w$  with MLE, formula (16) is used as the objective function:

$$\log P(\alpha_w | A) = \sum_m \log P(c_m | x_m, \alpha_w), \quad (16)$$

$$\log P(\alpha_w | A) = \sum_m \log \frac{\exp(\sum_x \alpha_w(x, c_m) \phi(x, x_m, c_m))}{\sum_c \exp(\sum_c \alpha_w(x, c) \phi(x, x_i, c))}, \quad (17)$$

where  $\phi(x, x_i, c)$  is the feature function corresponding to the word  $x$  and category  $c$  in the document, and  $\alpha_w(x, c)$  is the perceptron weight corresponding to this feature function. In the baseline method using machine translation,  $\phi(x, x_i, c)$  is specified as the word frequency of  $x$  in document  $x_i$ .

A cross-language document baseline classification method for machine translation uses word frequency features and a bag-of-words model to represent the documents to be classified. Using monolingual word vectors instead of word frequency features in the baseline method requires choosing a word vector-based document representation. In the proposed method for bilingual word vectors, the TF-IDF weighted sum of word vectors of words in the document is used as the distributed representation of the document [24]. Simply put, the TF value describes the importance of a word to a document. The IDF value describes the ability of a word pair to distinguish documents on a particular set of documents. In the subsequent research on cross-language document classification, in order to compare the cross-language performance of the model and exclude the interference of document representation performance, this document representation is used to obtain results comparable to previous work. The TF-IDF value and the above document representation are formally described below.

Here first define the word frequency statistical function  $f(x, x_i)$ :

$$f(x, x_i) = \begin{cases} 0, & x \neq x_i, \\ 1, & x = x_i. \end{cases} \quad (18)$$

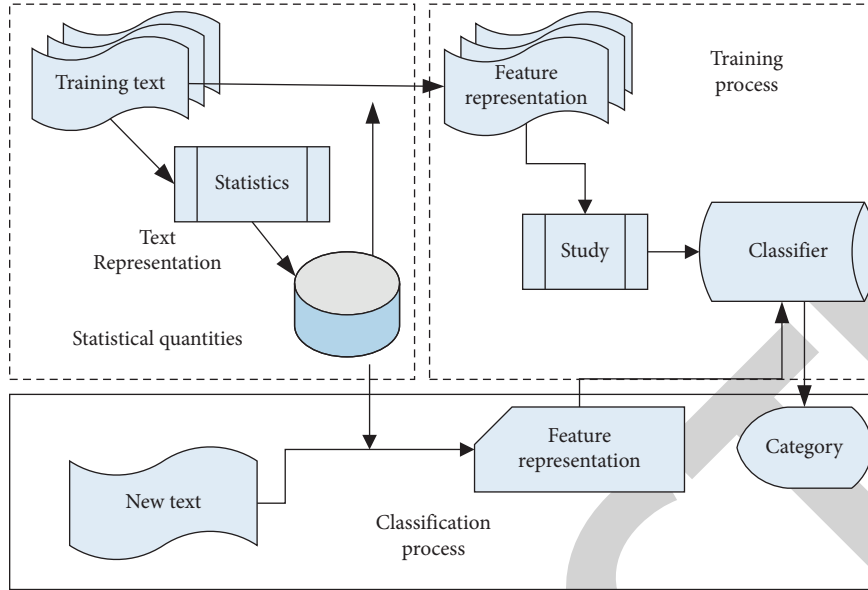


FIGURE 3: Structure block diagram of the automatic text classification system.

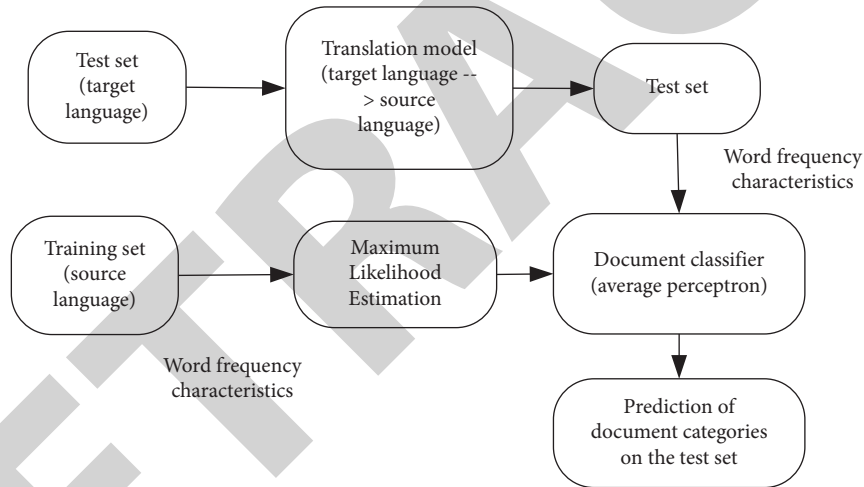


FIGURE 4: Schematic diagram of the baseline method.

Given a document  $x = \langle x_1, x_2, \dots, x_{|x|} \rangle$  and word  $x$ , term frequency (TF) is defined as follows:

$$TF(x|x_i) = \frac{\sum_{x_i} f(x_i, x)}{|x|}. \quad (19)$$

Given a set of documents  $B = \{x_i | i \in [1, N]\}$  and a word  $x$ , the inverse document frequency (IDF) is defined as follows:

$$IDF(x, B) = \log \frac{|B|}{1 + |\{x | x \in X, X \in B\}|}. \quad (20)$$

Using TF-IDF values, document  $x$  is represented as follows:

$$\text{rep}(x) = \sum_x^{\{x|x \in X\}} TF(x, X) \cdot IDF(x, B) \cdot \text{emb}(x), \quad (21)$$

where  $\text{emb}(x)$  is the word vector pretrained on the monolingual corpus.

Since the performance of machine translation models is not perfect, each translation process introduces additional translation noise. In order to minimize the introduction of machine translation noise, this paper only considers the case where the common language is selected as the source or target language in engineering. In this way, only one translation is required, and the training set and test set are mapped to the same language space. Figure 5 is a unified framework for applying machine translation to cross-lingual document classification.

In terms of translation efficiency, artificial intelligence translation is slightly better. However, when complex language logic or cultural elements are involved, it has to be supplemented by human translation. Under the upsurge of communication of ethnic cultures, it is important to give full



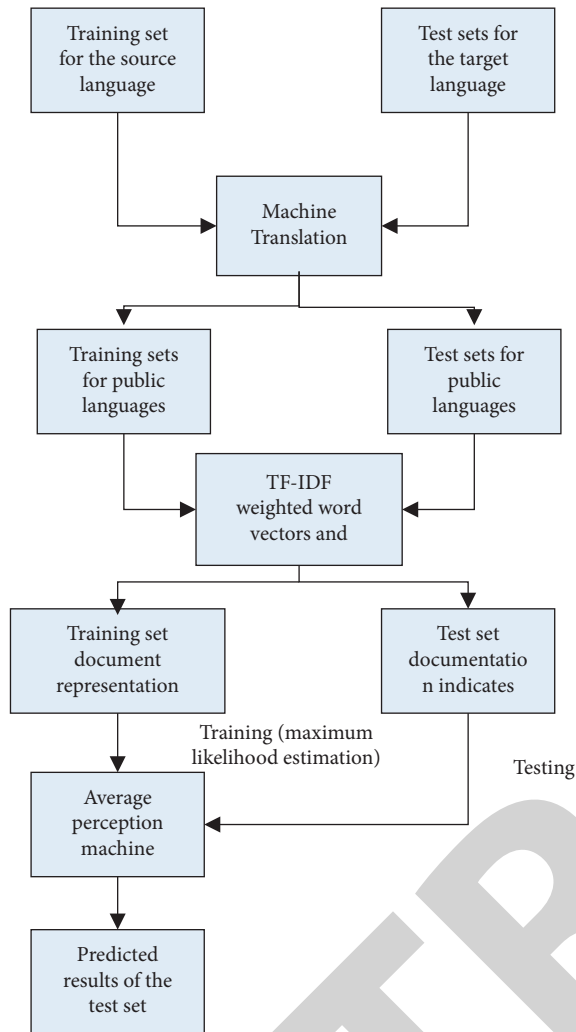


FIGURE 5: A unified framework for applying machine translation to cross-lingual document classification.

play to the instrumental and auxiliary role of artificial intelligence translation to help Chinese ethnic cultures “go global.” On the other hand, artificial intelligence can be used to improve the construction of ethnic culture corpus. Although China has made some great achievements in translation corpus construction, it mainly focuses on the text of science and technology, and the construction of ethnic culture corpus is still at its primary stage. The construction of ethnic culture corpus under the background of artificial intelligence translation is a systematic project, which aims to design and construct a relatively Chinese-English parallel corpus of ethnic culture on the basis of revealing the characteristics and development law of Chinese ethnic culture by means of data analysis and mining of artificial intelligence technology, so as to lay a good foundation for artificial intelligence translation.

*3.4. Artificial Intelligence Technology for Communication Data Acquisition and Analysis.* With the rapid development of mass media, especially the rapid advance of new media

represented by the Internet in the world, today’s audiences are no longer just readers, listeners, or viewers wandering between different media, nor are they mere recipients of the information. The data used in the article come from a website. Programming language is used to achieve deep extraction based on AJAX, and time is used as the incremental factor to achieve incremental extraction. The data acquisition methods include web page parsing and API interface acquisition. The combination of Python Scrapy architecture and API interface is adopted. Since the website limits the number of visits per hour, this brings great inconvenience to the data extraction work. Therefore, multithreaded concurrent operation is adopted. Since the information transmission in the network requires a period of time, the topic may last on the network for a certain period of time, so the same topic needs to be extracted regularly, which requires incremental extraction of the topic. In order to achieve incremental extraction, it is first to parse the crawled source files and use regular expressions to grab relevant fields, then to format the grabbed fields and tread the post list to determine whether the information has been crawled. If it has not been crawled, it needs to store the post in the database and update the relevant post-table and user post list. If it has been fetched, then it needs to judge whether the next information has been fetched. Next, it is to judge whether the crawling is completed. If not, the page value will be +1, and the URL will be updated to continue crawling. If the crawling is complete, the program will be exited, realizing incremental crawling. Regular analysis and DOM analysis are used to crawl the obtained source files. The data are sorted out after the analysis and saved to the local database. The whole process is written in Python language. The collection process is shown in Figure 6.

Three ethnic cultural hot topics were randomly selected as the experimental objects, and the user IDs participating in the hot topic discussion, the time of commenting, the number of fans and friends of the user, and the number of microblogs published by the user were selected. The specific data format of the experiment to be carried out is shown in Table 1.

In order to study the number of features that cannot be used in the test set translated by the machine translation model due to the sparse word frequency feature problem, Table 2 counts the unregistered words in the English-Chinese and Chinese-English directions. Here what is counted is how many words in the translated English and Chinese test sets, respectively, have never appeared in the Chinese and English training sets.

As can be seen from Table 2, on the cross-language document classification dataset, there are a large number of unregistered words when translating the test set with the machine translation model and directly using the word frequency feature. In the English→Chinese transfer direction as an example, 68.5% of the words in the test set vocabulary did not appear in the training set. Although these words are not all “key features” in the above example, such a large proportion of unregistered words still implies that there are a large number of effective features on the test set that are underutilized due to the sparse nature of word frequency features.

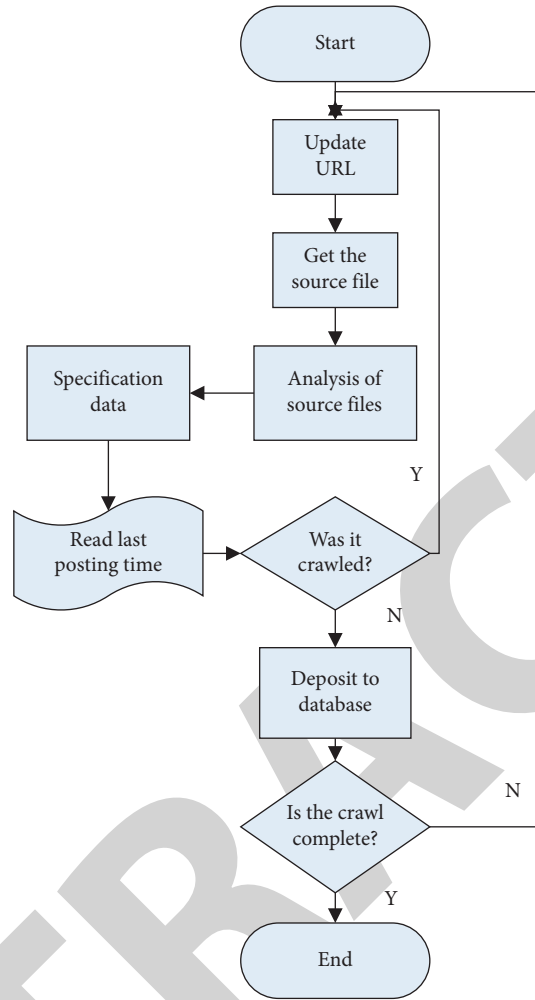


FIGURE 6: Data extraction flowchart.

TABLE 1: User data format.

User ID	ID 1	ID 302	ID 5241
User comment time	2021.03.02 10:23:10 AM	2021.04.12 03:56:00 PM	2021.11.05 11:33:20 AM
Number of user followers	13452	23	352
Number of user friends	1204	35	258
Number of user comments	4720	87	1295

TABLE 2: Statistics of unregistered words in translating the test set.

Direction	English→Chinese	Chinese→English
Training set word list size	25012	23140
Translated test set word list size	30214	64215
Number of unlisted words	21045	53246
Percentage of unlisted words	69.7%	82.9%

3.5. Artificial Intelligence Technology for Communication Channel Analysis. In recent years, with the vigorous development of new media and the rise of social networks, communication channel has been experiencing much

innovation. With the development of new media, the social network is taking its dominance in the external dissemination of ethnic culture. The main media for people to learn about ethnic culture are television programs, newspapers, word of mouth, the Internet, the radio, general magazines, and books. Figure 7 shows the specific data for obtaining ethnic cultural information in 2016. Figure 8 shows the specific data for obtaining ethnic cultural information in 2021.

From Figures 7 and 8, it can be found that the media of ethnic culture has undergone great changes. In 2016, TV and newspapers accounted for the highest proportions, with state media leading the way and having a long memory cycle. But it can be found that in the current era, the proportion of TV channels has dropped from 70% to 23%, and Internet

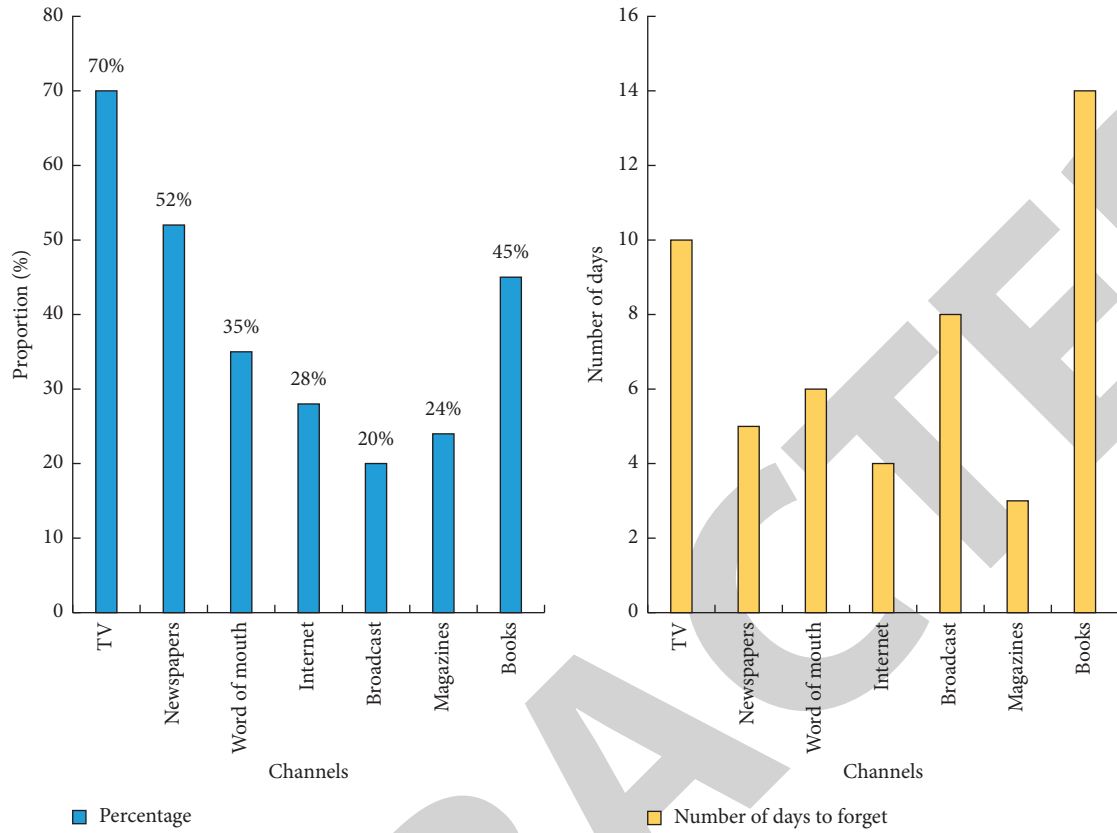


FIGURE 7: Channels to obtain ethnic cultural information in 2016.

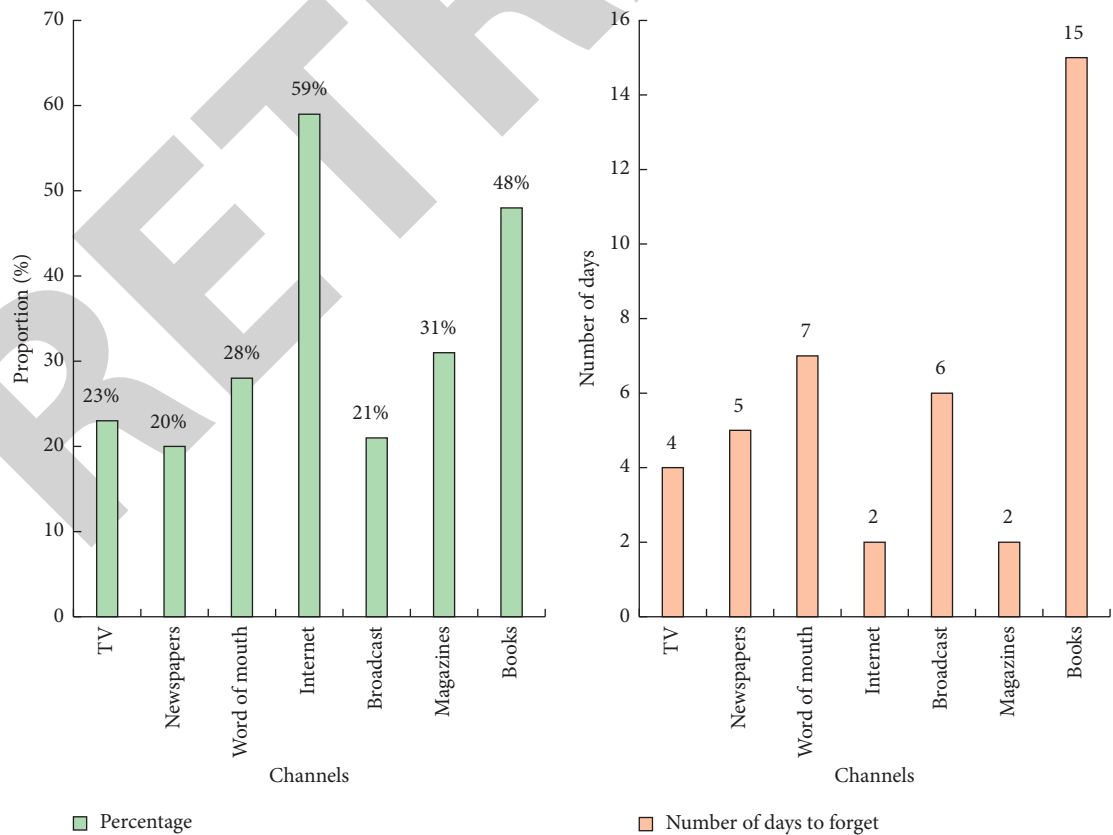


FIGURE 8: Channels to obtain ethnic cultural information in 2021.

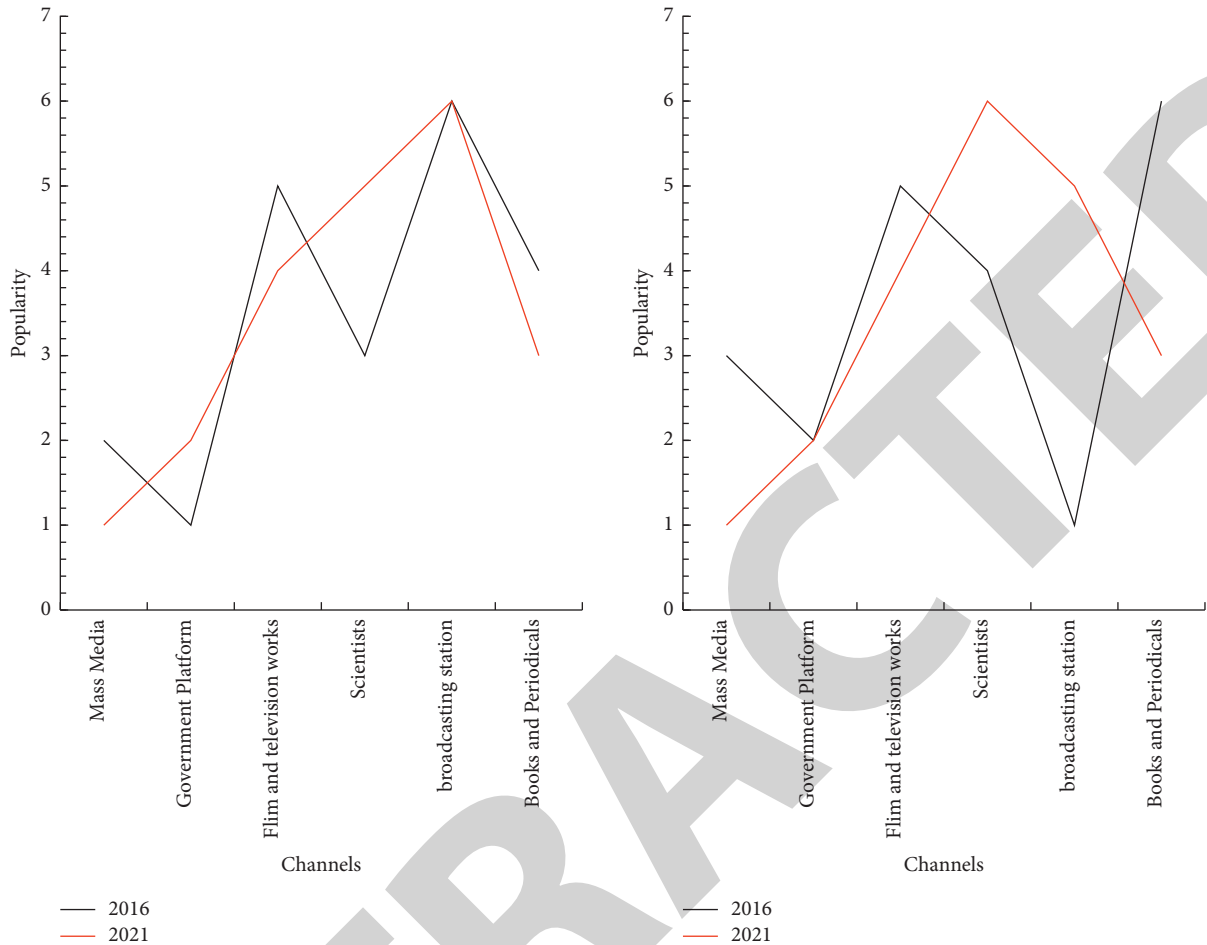


FIGURE 9: The effect of various channels in the dissemination of ethnic culture through artificial intelligence technology in 2021 compared with 2016.

communication and book communication have become the “main forces.” Among them, the proportion of the Internet is 59%, and the proportion of book channels is 48%, but the memory period of the Internet is very short, only 2 days, which is related to the rapid replacement of Internet hot spots and the large amount of information. Even so, it reflects that the channels to gain ethnic culture knowledge are not only the official media, but ordinary citizens will also take the initiative to discuss related culture, which is a good trend.

Artificial intelligence technology can be extensively used to improve the effectiveness of ethnic culture dissemination. For example, after an ethnic cultural text is published abroad, the AI big data collector can be used to count the data of target reader groups and analyze the dissemination effect, etc. Based on the evaluation of public communication of artificial intelligence technology, the popularity of communication subjects reflects the communication effect to a certain extent. Figure 9 shows the popularity and dissemination speed of mass media, government platforms, scientists, film and television works, and radio stations in the dissemination of ethnic culture through artificial intelligence technology in 2021 compared with 2016.

It can be seen from Figure 9 that in 2021, the popularity and transmission speed of mass media ranks first. It does not only show that the mass media is efficient and fast, but more importantly, the mass media itself is highly interactive, which is convenient for netizens to comment and express their opinions and is the most popular with the public. Because it is often transmitted through official media, the dissemination on government platforms produces a better dissemination effect. Today, with the popularity of the mass media, scientists have moved from the forefront of scientific communication to the rear of the mass media and official media. Film and television works have dropped from third place in 2016 to fifth place due to the poor availability of communication channels, and radio stations are the least popular.

#### 4. Conclusion

This article conducts basic research and discussion on the application of artificial intelligence technology to Chinese ethnic culture communication with foreign countries. With the advanced development of mass media, communication

channel has undergone a dramatic evolution. The community has migrated from offline to online, which is largely dependent on the development of information technology. As for Chinese ethnic culture and external communication in this half-real-half-virtual world, artificial intelligence technology can offer great implications. The use of digital and intelligent technology of artificial intelligence to disseminate national culture not only meets the cultural needs of the young generation audience around the world but also meets the trend of social development. This technology can be used to analyze Internet users' interests, which is helpful for selecting appropriate materials for certain target receivers. AI technology of machine translation can boost the construction of ethnic culture corpus, which is useful for the cross-cultural communication of Chinese culture. With a series of surveys and experiments, it is found that the media is playing an increasing role in ethnic culture communication, both in terms of media technology and the influence of the media itself. It is changing the process of ethnic culture communication, and Internet communication and book communication have become the mainstream channels of ethnic culture communication, so ethnic culture communication cannot stick to the old modes and should initiate its own changes with the emergence of new media.

### Data Availability

No data were used to support this study.

### Conflicts of Interest

The authors declare no potential conflicts of interest in this study.

### Acknowledgments

This work was supported by the project Research on the External Communication Path and Media of Northern Shaanxi Culture under the Background of "One Belt One Road" funded by the Research Fund of Xijing University (project number: XJ180207, 2018).

### References

- [1] C. Cui, Y. Zhao, and L. Wang, "Protection and dissemination of Chinese intangible cultural heritage based on digital games," *International Communication of Chinese Culture*, vol. 8, no. 4, pp. 483–491, 2021.
- [2] P. Watt and B. Costea, "Reappraising Charles Webb's the Graduate (1963): e(1963): Exploring cultural and historical elements of a character in the contemporary economy," *Organization*, vol. 27, no. 1, pp. 140–161, 2020.
- [3] F. Brochet, G. S. Miller, P. Naranjo, and G. Yu, "Managers' cultural background and disclosure attributes," *The Accounting Review*, vol. 94, no. 3, pp. 57–86, 2019.
- [4] G. Zhang and X. Yang, "Research on the development of northern ethnic culture based on big data mining," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 15, pp. 734–739, 2017.
- [5] Y. Kashima, P. G. Bain, and A. Perfors, "The psychology of cultural dynamics: what is it, what do we know, and what is yet to be known?" *Annual Review of Psychology*, vol. 70, no. 1, pp. 499–529, 2019.
- [6] E. Popov, J. Stoffers, and V. Simonova, "Cultural factors of network inter-firm cooperation," *Review of International Business and Strategy*, vol. 29, no. 2, pp. 103–116, 2019.
- [7] X. Qin and N. A. Fedorovskaya, "Specifics of Russian-Chinese cross-cultural communication in the field of fine art of the second half of the 20th — early 21st century," *Observatory of Culture*, vol. 17, no. 6, pp. 582–593, 2021.
- [8] K. G. Han, "Implications of artificial intelligence technology for educating students with severe and multiple disabilities," *Korean Journal of Physical Multiple & Health Disabilities*, vol. 60, no. 3, pp. 47–65, 2017.
- [9] M. M. Rahman, F. Khatun, A. Uzzaman, S. I. Sami, A. A. Bhuiyan, and T. S. Kiong, "A comprehensive study of artificial intelligence and machine learning approaches in confronting the coronavirus (COVID-19) pandemic," *International Journal of Health Services*, vol. 51, no. 4, pp. 446–461, 2021.
- [10] Gretchen and J. Jackson, "Artificial intelligence in health in 2018: new opportunities, challenges, and practical implications," *Yearbook of medical informatics*, vol. 28, no. 1, pp. 52–54, 2019.
- [11] H. Devaraj, S. Makhija, and S. Basak, "On the implications of artificial intelligence and its responsible growth," *Journal of Scientometric Research*, vol. 8, no. 2s, pp. s2–s6, 2019.
- [12] E. Farrow, "Organisational artificial intelligence future scenarios: futurists insights and implications for the organisational adaptation approach, leader and team," *Journal of Futures Studies*, vol. 24, no. 3, pp. 1–15, 2020.
- [13] "ZelenkováAnnazelenkova.anna@centrum.cz|Institute of slavonic studies of the academy of sciences, prague, Czech republic. Comparative literature as educational means of understanding and communication," *Journal of Language and Cultural Education*, vol. 6, no. 1, pp. 108–116, 2018.
- [14] Z. Sydykova, N. Medvedev, S. Suleymanova, E. Nazarova, and V. Volokh, "Governance of cross-border migration in asia," *Space and Culture India*, vol. 7, no. 4, pp. 264–273, 2020.
- [15] S. Huang and B. Yang, "A comparative study on Chinese teachers' implicit theories of creativity in the context of cross-cultural communication," *International Journal of Learning and Teaching*, vol. 7, no. 1, pp. 20–24, 2021.
- [16] Z. Jingjing, X. Yutong, and W. Qinhan, "A review on the study of confucius institutes and their significance in foreign cultural communication — research based on national social science fund project," *Linguistics*, vol. 3, no. 1, pp. 137–144, 2021.
- [17] T. A. Sharypina and M. K. Men'Shchikova, "Literary canon in the context of cross-cultural communication," *Proceedings of Petrozavodsk State University*, vol. 43, no. 1, pp. 119–121, 2021.
- [18] W. Chen and B. Jin, "Modeling and simulation of cultural communication based on evolutionary game theory," *Complexity*, vol. 2021, no. 2, 12 pages, Article ID 6115606, 2021.
- [19] V. Jakuionyt, "Cross-cultural communication: creativity and politeness strategies across cultures. a comparison of Lithuanian and American cultures," *Creativity Studies*, vol. 13, no. 1, pp. 164–178, 2020.
- [20] A. V. Lentovskaya, "Information and communication technologies and the formation of cross-cultural communication skills in a foreign language environment," *Moscow State University Bulletin Series 18 Sociology and Political Science*, vol. 26, no. 4, pp. 76–88, 2020.

- [21] V. Ahieieva, "Types of interference as the manifestation of cross-cultural communication in translation," *Humanities Science Current Issues*, vol. 1, no. 29, pp. 4–9, 2020.
- [22] Z. Wang, J. L. Peterson, C. Rea, and D. Humphreys, "Special issue on machine learning, data science, and artificial intelligence in plasma research," *IEEE Transactions on Plasma Science*, vol. 48, no. 1, pp. 1-2, 2020.
- [23] I. V. Kononenko, "Cross-cultural communication - lost in translation: a corpus study (based on the material from the Russian-Polish corpus)," *Russian Journal of Linguistics*, vol. 24, no. 4, pp. 926–944, 2020.
- [24] I. Stanislavova and G. Solovyova, "Modernism and the problem of cross-cultural communication," *Herald of Omsk University*, vol. 25, no. 2, pp. 75–80, 2020.

RETRACTED