Integrated Development of Artificial Intelligence and Big Data, with Foreign Language Discipline

JunJie Xue*
Xi'an Siyuan University, Shaanxi, China, 710061

*Corresponding author e-mail: 309001643@qq.com

Abstract. In the new era of rapid change and development, the connotation and denotation of foreign language discipline research is in constant deepening and expansion. Electronic traces, social media, digital texts, spatial location information, and other large-scale data have been extensively applied to the research of foreign language discipline. The new scene of modern foreign language discipline research is worthy of our attention and consideration.

Keywords: Big Data, Artificial Intelligence, Information Technology, Foreign Language Discipline

1. Introduction
The real value of foreign language discipline research is stemmed from research problems. To grasp the development trend of foreign language discipline research is to seek solutions to research problems [1-2]. The same research problems can be analyzed from different disciplines, perspectives and emphases. In the era of big data, the use of information technology helps people to understand the new developments in foreign language discipline research more thoroughly, the application of big data to all walks of life can instantly and maximize the value of data [3-4]. For example, CCTV news uses Baidu positioning data to display the people flow during the Chinese New Year. MogIA artificial intelligence (AI) system developed by India successfully forecasts the results of the US general election with four times network data. GFT influenza trend prediction system developed by Google predicts the influenza trend of the US-based on user search records, etc. [5-6] In the past, research with traditional data has a significant time lag, while real-time data acquisition can effectively improve the authenticity of discovery by people. In addition, more meaningful conclusions can be obtained from foreign language research.

Social media data provides an excellent opportunity to interpret human behavior laws or behavioral tendencies. In the past, research on human behavior mainly used questionnaires, interviews,
experiments, etc. Social media has become an important place to provide data. Researchers do not need the response and cooperation of the tested object, nor do they need to wait for the tested purpose, so it is generous for foreign language research. For example, Dr. Gong Weigang, Department of sociology, Wuhan University, used the network data to analyze the historical evolution of populism in western society; scholars such as Maximilian schich used the data of places where celebrities were born and died to research the historical migration of cultural and art centers. In the past, it was challenging to integrate the historical changes The research conclusion often has a lot to do with the researcher's personal vision and can only analyze and solve local problems. After the data processing technology and analysis means appear, we can understand the historical change process and historical things through the visual display of data.

2. Technological environment changes, creating favorable research conditions

Current research scenarios of foreign language discipline show that many studies have failed before and can be carried out now. In fact, changes in the technological environment create favorable research conditions. For example, the transformation path of the real world: digitalization. Digitalization is a term featuring a strong sense of the times. It is also a common practical application scenario that refers to the complex and changeable information objects, such as voice, color, graphics, image, etc., which are converted into binary, and processed by computer in a unified way. Thus, people can research some problems in the original reality in another world (space). The concept of digitalization and digital tools and technical capabilities are applied to many fields, and many fundamental problems are solved, such as digitalized books, newspapers and magazines, libraries, and museums. For example, the inexhaustible data source: Internet of things. Internet of things technology is to implement the “Internet of Things (IoT)” based on the Internet. The “things” herein can provide a sound research basis and platform for understanding social and economic phenomena. For foreign language discipline research, inexhaustible data is the source of new scientific discoveries, and the research charm brought by IoT lies in the data Access to unprecedented convenience. The IoT can access massive real-time data, which cannot be obtained by traditional tools and means.

There is also the instantly available computing environment: cloud computing. The concept of cloud computing was proposed at the search engine conference in 2006, and it has been more than 10 years now. It mainly refers to the sharing of computing resources, which brings data distributed storage and computing. This distributed storage and computing makes researchers no longer worry about the limitations of traditional data storage and processing. The goal of traditional statistics is to In the long term, researchers have formed the habit of reducing data collection, which is actually a kind of artificial restriction. Cloud computing means more temporary storage space and computing environment, so it gives researchers confidence: in the face of complex research objects, there is no need to do too much reduction Mobile ubiquitous information behavior that people attach importance to smart phones and terminals. The smartphones and terminals held by everyone form a universal environment for information acquisition and information production. Users can not only produce data all the time, anytime, anywhere but also get data and transfer data. Now, many data are generated and transferred continuously through mobile phones and terminals. The information transfer model performance of human information behavior presents new features, and mobile information service is closely related to the development of life economy. Such changes not only provide support for the
research of information dissemination and user behavior but also enrich the research approaches of foreign language discipline.

Big data has had an enormous impact on social life. Social science reflects society, and social change also determines the change in social science. Meanwhile, big data also has a profound impact on social science research objects, thinking methods, and society.

![Diagram](image)

**Figure 1.** Necessity of social science evaluation innovation in a big data environment

3. New and old problems needing attention in the new scene

Under the new technical environment and data generation conditions, the problems of foreign language discipline research have changed. These changes mainly include the following two categories. The first category is the old problems in the new scene. Such problems have always existed in foreign language discipline research, but there is no way to reveal them under the traditional technical conditions in the past, such as the above literature fingerprint problems and social thinking, cultural and art center changes, etc. These problems can be solved smoothly in the new scene with the help of new data sources, acquisition methods, processing, and analysis tools. In addition, the overall and partial problems that have been argued for a long time in social sciences are also problems of this kind of tradition. The second type is new problems in the new scene. In the new scene, there are also problems that have not been solved in the past, such as the virtual space brought by the Internet, digital technology, VR/AR technology, the virtual community, virtual enterprise, artificial reality, the second nature, the virtual crime, ethical issues and the human value issues concerned by the humanities are actually new issues in the new scene. However, there are few studies in this field, and these problems have a significant impact on human beings.

Given the sample data set, \( D = \{x_1, x_2, ..., x_n\}, x_i \in \mathbb{R}^d, i = 1, ..., n \). It is assumed that the first sample is labeled as \( \epsilon = (x_1, x_2, ..., x_l) \), corresponding label \( \eta = (y_1, y_2, ..., y_l) \).

Consider the problem of classification, and note the following symbols:

\[ S = \{x_i, x_j\}: x_i \text{ and } x_j \text{ It's the same kind. } x_j \text{ is } x_i \text{ neighbor}\].

\[ R = \{x_i, x_j, x_k\}: x_i, x_j \text{ It's the same kind. } x_j \text{ and } x_k \text{ neighbor}\].

In the classification problem, the importance of each feature dimension is different. To some extent, the big data weak symmetric manifold can overcome the shortcoming that the big data weak Symmetric Manifold treats each feature dimension equally. Its definition is as follows:
Samples $x_i$ and $x_j$, the big data weak symmetric manifold between are defined as follows

$$d_A(x_i, x_j) = \sqrt{(x_i - x_j)^T A(x_i - x_j)}$$ \hspace{1cm} (1)

$x_i \in R^d$, $A \in R^{d \times d}$. It is a symmetric semi positive definite matrix

Based on the properties of semi positive definite matrix, $A$ can be breakn down into $A = L^T L$:

$$d_A(x_i, x_j) = \sqrt{(x_i - x_j)^T A(x_i - x_j)}$$

$$= \sqrt{(x_i - x_j)^T L^T L (x_i - x_j)}$$

$$= \sqrt{L x_i - L x_j)^T (L x_i - L x_j)}.$$ \hspace{1cm} (2)

It is equivalent to the matrix as a mapping, mapping the data of the original space to the new space, and transforming the big data weak symmetric manifold of the original space into the big data weak symmetric manifold of the new space.

4. New vision of modern foreign language research

Correlation analysis and causality analysis are combined. When big data appears, people think that causality analysis is no longer critical and big data only pays attention to correlation analysis. This is a misunderstanding, and causality analysis is still fundamental. Since ancient Greece, western philosophy has taken causality as the core of science, because any scientific problem should be based on the result of cause analysis and the cause between things In the big data environment, it is not so easy to find the relationship between “cause” and “result”. In fact, there are three essential conditions for causality: first, causality is a kind of temporal relationship, “cause” is in the first place, “result” is in the second place, there is an empirical correlation between causality. Thirdly, causality is not the result of the third variable, which can hardly reveal these conditions in the big data environment. However, the importance of causality is undeniable. Correlation is a necessary condition for causality. Based on the correlation, causality can be analyzed to avoid the endogenous problems of foreign language discipline in causality analysis

The traditional social science evaluation process can be divided into the whole process of evaluation (evaluation client) and the specific process of evaluation implementation stage (evaluation implementer). According to the overall work process of the review, the evaluation client can be divided into the evaluation preparation stage, evaluation implementation stage, and evaluation result utilization, as shown in Figure 2.
The combination of manual analysis and tool application. The ancients said that “if you want to work well, you must first use the tool”. In the current big data scenario, you should not only attach importance to the traditional manual analysis but also attach importance to the use of tools, especially the combination of manual analysis and tool application. The long-term accumulation and rapid growth of data in different fields have brought new data analysis requirements, related processing, classification analysis, visualization, and other technologies are emerging. Different tools and methods are required to address various issues, which are not easy to grasp and interpret. Therefore, foreign language discipline studies how to absorb and use these powerful technical means to combine traditional manual analysis and tool application, which needs attention in the research process.

5. Conclusions

By expounding the evaluation process of traditional social sciences, we compared and analyzed the structure of the scientific evaluation process in the big data environment, the settings of each link, and the innovative advantages in the aspect of scientific selection of each link, highlighting the flexibility, scientificity, and rationality of the practical settings of social science evaluation in the big data environment. Secondly, relevant concepts, construction principles, and requirements of the social science evaluation indicator system in the big data environment are elaborated. The scientificity, systematicness, and operability of the evaluation indicator system are highlighted. The big data environment requires that the evaluation indicator system should be continuously innovated in the selection of indicators to meet the new requirements for the development of social science evaluation in the era of big data.

References

[1] Ma, F. , Li, Y. , & Chen, B. . (2014). Research of the collaboration in the field of the chinese humanities and social sciences. Scientometrics, 100(2), 439-458.

[2] Guns, R. , Linda S1le, Eykens, J. , Verleysen, F. T. , & Engels, T. C. E. . (2018). A comparison of cognitive and organizational classification of publications in the social sciences and humanities. Scientometrics(7), 1-19.

[3] Goodwin, D. , & Machin, L. . (2016). How we tackled the problem of assessing humanities, social and behavioural sciences in medical education. Medical Teacher, 38(2), 137-140.

[4] Warburton, J. , Cowan, S. , & Bathgate, T. . (2013). Building social capital among rural, older australians through information and communication technologies: a review article.
Australasian Journal on Ageing, 32(1), 8-14.

[5] Zaida Chinchilla-Rodriguez, Miguel, S. , & Félix de Moya-Anegón. (2015). What factors affect the visibility of argentinean publications in humanities and social sciences in scopus? some evidence beyond the geographic realm of research. Scientometrics, 102(1), 789-810.

[6] Weidong, WANG, & Xuemei. (2016). The development strategies of foreign language and literature discipline in china under the background of “building world-class universities and disciplines”. Journal of Beijing International Studies University(5), 136-136.