Research on the Application of Classical Literature Based on Computer Retrieval

Huiting Dai¹,*

¹School of Literature and Journalism, City College of Dongguan University of Technology, Dongguan, Guangdong, China, 523000

*Corresponding author e-mail: daiht_2020@ccdgut.edu.cn

Abstract. In recent years, computer information technology has brought great impact on the traditional life model and has also brought great influence to the traditional academic research model. At present, in the field of "text", the most important thing is the data mining based on big data, that is, the direction of computer retrieval. In the face of data mining, there are two ideas: either to manually make enough rules to break unstructured text into semi-structured or structured text that machines can understand, or to make machines more and more intelligent through deep learning. This is of course the trend of scientific and technological development. In this general trend, as researchers, we should also study different styles with a positive attitude. For example, the analysis platform of Yuefu poetry can be developed according to the key elements of Yuefu poem, such as title, usage, skill, tune and style. In view of the unique writing style of epitaphs and letters, data mining can be carried out on such styles, and the generated knowledge pedigree is also of great benefit to the related research of other styles.

Keywords: Computer Retrieval, Classical Literature, Applications

1. The influence of computer technology on the study of classical literature

In the preface to Mr. Wang Jing'an's Legacy, Mr. Chen Yinke concludes that Wang Guowei's methods of study are as follows:" one day takes the underground object and the paper's remains to explain each other's evidence ";" two days take the alien's old book and our country's old nationality to complement each other's evidence ";" three days take the foreign idea and the inherent material to consult each other ".These methods represent the most advanced academic ideas at the beginning of the 20th century, and even today they still have important significance. However, the times are developing, in the 30 years of reform and development, the first decade of the 21st century has not much time left today, we reflect on and summarize the achievements and shortcomings of contemporary academic circles in classical literature research. Since the 1980s, despite our constant attempts to use various methods, the most influential information technology in contemporary classical literature research is computer and network[1].

Because the influence of IT, which seems to have nothing to do with the study of classical literature, on the study of classical literature has penetrated into all aspects of the subject. This
influence is not only reflected in the convenience and accuracy of literature retrieval and data collation, but also in the continuous development of technology.

2. Research on three levels of classical literature based on computer retrieval

The study of classical literature, including historical literature related to literature, involves a wide range of levels and research scope. If we make a superficial distinction, it can be roughly divided into three levels: one is literature collation, including the compilation of large literature collections, the collation of other collections, and so on, involving version collation, document punctuation, works compilation and other work. The second is to sort out the research materials of the writers and works of the past dynasties, including collecting the materials of a particular topic, compiling the chronology, writing the history of chronicle literature and evaluating the writers. Third, the third level of classical literature research is mainly the aesthetic appreciation of literary works and the analysis of the development of literary history. At this level, with the current computer can achieve[2].

Table 1. Research on Classical Literature.

|                      |                                                                                                                                  |
|----------------------|----------------------------------------------------------------------------------------------------------------------------------|
| **Documentation**    | The compilation of a large-scale collection of documents, the collation of other collections, and so on, involving the collation of versions, punctuation of documents, compilation of works and so on |
| **Writer's Works**   | Collection of materials on a particular topic, compilation of a chronology, writing of the history of chronicle literature, and assessment of writers, etc |
| **Appreciation of Literary Works and Analysis of the Development Law of Literary History** | It includes the appreciation and analysis of various literary works and the research on the development trend of literary history for a long time |

2.1. Documentation

With the increasing quantity and quality of digitizing ancient books, the advantages of computer in literature collation are being displayed incisively and vividly. For example, the retrieval function of computer can effectively avoid the problems of re-collection and false collection which affect the compilation quality of large-scale literature collection, which has been well verified in the Electronic Analysis system of Song Poetry developed by the Chinese Department of Peking University. For example, the research on the comparison of classical novels by the Digital Center of Chinese traditional Culture of Capital normal University is also very enlightening. With more and more ancient books being digitized, computers are sorting out the literature. The advantages of the aspect will be exhausted. The computer retrieval system based on accuracy, speed, efficiency and intelligence can help us to improve the working efficiency and work level greatly in the collation of ancient books and documents[3].

Let's take the example of sorting out a particular set of writers. With his extensive knowledge and rigorous and realistic academic spirit, the Qing Dynasty has set a good example for future generations in the collection of works of many writers such as Li Bai, Du Fu and Su Shi. However, the erudition and rigor of Qing people no longer have any advantages in today's computer age, because the computer can accurately list more documents that help to explain the source of a word or allusion. This is now a well-known fact, needless to say. What the author needs to explain is that if there is enough digital literature support, many functions of computer based on accurate retrieval can be better played in the collation of ancient books. It is not limited to simply helping scholars find the original source of a word. We know that sorting out a collection of writers requires only a few steps: first, identifying a good version as the base. Second, determine several reference books to compare with the base. Third, from all kinds of collections, books, local chronicles, other people's works and other documents to compile the writer's works. In the meantime, the author's works are wrongly collected
by others or the same works are included in many collections, and the signature can not be judged.

Fourth, for the collection of works received notes. In these kinds of work, the work of determining the base must be carried out by people, but when it is impossible to determine which version is better, all versions are extracted by the computer, and some of the words in this book are compared to see which is more appropriate. It is also an effective auxiliary method. Because the computer can arrange the data arbitrarily according to the relevant instructions, it can be compared according to the whole, the whole sentence or the word for word, so whether it is a separate collection or a chronological collection, it can be compared by the computer. The collection of lost works can be retrieved by computers in a wide range of ancient books databases to extract the maximum number of alternative works, including cases where the same work has different signatures. The computer will submit a list of comparative processing and only final judgment for us. This not only enables us to search for the works that our predecessors may ignore because of the limits of our knowledge, but also avoids all kinds of regrets caused by the lack of knowledge, and the complexity and workload of the work are greatly reduced. As for the work of note for the work, the computer can analyze the processed objects according to the pre-established allusions, words, names, place names and other proprietary databases, and list the results of the analysis under the annotated works according to the set format. In fact, such a way of working in a sense or according to the traditional way of reading. Today, the Internet has entered the Web2.0 era, because of the use of Wiki and other technologies, computers can extract relevant data from relevant allusions, words and other databases according to the needs of reading at any time to interpret the text instantaneously\([4]\).

2.2. Collation of author's works
This includes the compilation of materials on a particular topic, the compilation of chronologies, the writing of the history of chronology and the assessment of writers. These work itself not only has considerable scientific research academic value, but also lays the foundation for subsequent research work. At this level, computers can also be useful. For example, the use of GIS technology, can accurately analyze the relevant data, positioning. This function can be fully played in the compilation of writers' chronology. On the basis of intelligent retrieval of the relevant ancient books database, the computer arranges all the materials of the writer according to the order of the first event and then the work, and becomes a chronology with certain credibility. On this basis, with the necessary artificial compilation, it is possible to give a quick and accurate combing of the life story of a writer or a group of writers. In the same way, it can also be applied to the writing of chronicle history\([5]\).
The study of classical literature, a subject with a long history, has developed to this day, and many problems have been recognized by academic circles. Modern academic is becoming more and more refined. It is important to put forward new viewpoints and conclusions, and to explain the process of derivation of some existing conclusions. For example, the scholars recognized that the southern migration of the scholars in the Eastern Jin Dynasty had an important impact on the literary development of the six dynasties, and the relevant literary history also elaborated on it in detail. However, if we measure it from the standard of modern academic development, we undoubtedly need to make a more detailed analysis of this historical phenomenon. By means of GIS technology, we can accurately locate which people migrate from the north to the south and what the specific literary creation of these people and their descendants is. In this way, we can have a very clear grasp of the cultural and literary center of gravity caused by population migration and its later evolution process. Under the premise that the micro research is very fine, the macro research must be more persuasive. Similar research methods can be extended and applied to many aspects of literary history research, ranging from the study of regional literature in successive dynasties to the official, roaming, demotion, exile and related communication and literary creation of a certain writer. Although traditional literary history research can solve these problems to some extent, the scope of traditional literary history research is still concentrated in a few large and middle writers for many years. It is difficult to obtain fundamental changes in a short period of time. This is due to the inevitable shortcomings caused by relying solely on manpower, and the GIS system based on accurate data retrieval can methodically locate all the writers of certain size in the corresponding data model, which can be compared vertically or horizontally. If literary phenomena are used to study a large range or a long historical span, the advantages of auxiliary research means such as GIS system will be revealed.

2.3. Aesthetic appreciation of literary works and analysis of the development law of literary history

At this level, based on the current level of intelligence that computers can achieve, it is not possible to participate substantially in these creative work that requires more human mental thinking activities, showing the limitations of computers. But the work of computer data analysis at the above two levels has laid a solid foundation for the third level of work. For example, if we need to study the history of Chinese women's literature, with the support of enough massive ancient books database, we can arrange all women's works in detail according to nationality, time and genre since the pre-Qin Dynasty. Such a kind of data display itself is a powerful explanation of the law of the development of female literature. In addition, the statistical analysis of word frequency of literary works by computer should also help to grasp the style of works. Of course, computer technology is applicable to the field of classical literature research far more than these aspects, in the process of applying these technologies to traditional literary and historical disciplines, there will be such twists and turns and mistakes. The function of computer will not change with the development of technology. With the development of technology and the emergence of more and more digital texts of ancient books, it is very convenient to collect relevant information in a far larger range than in the past. This provides a necessary prerequisite for the study of writers who have never been noticed in the past. Scholars' horizons will be further expanded, and the group of writers and writers will be further expanded. The object of study, the way of research and the new research topics derived from it will all appear or change large or small. The revolution of computer in the field of classical literature will be long and profound.

3. Conclusion

Human exploration of science is endless, only decades of computer technology has brought about a world-shaking change in our lives and work. The development trend of computer in the future and its powerful function, we can not accurately predict. But, just as anything should not be absolute, computers can help us deal with a lot of things, but we should not suffer from computer dependence, everything to the computer to complete. The application of computer technology to classical literature
is naturally more so. For such a subject that pays more attention to humanistic spirit and aesthetic analysis, it is more like the taste of semi-finished products and an incomplete work. It is a successful work to reprocess it and to put creative thinking and fresh thinking into practice. In the modern and future period, the combination of computer accurate calculation and human mind and aesthetic ability is the future academic development direction. Based on the increasingly advanced IT achievements, we have reason to believe that the study of classical literature will make a brilliant future in the future of the increasingly frequent interaction between the mind and technology of scholars.

References
[1] Xu Jinliang. Application of Excellent Traditional Culture in the Teaching of Classical Literature [J]. Education Drama House, 2019, No.331, 183.
[2] Li Wei. Exploration of the Application of Classical Literature in Landscape Design —— Review of "The Mysteries of Chinese Classical Garden Art "[ J].].1 Chinese Edible Fungi ,2019, v.38; No.224, 158-159.
[3] Lin Xijin. Application of Communicative Translation Theory in the Translation of Chinese Classical Literature [J]. A Journal of Hubei Second Normal University, v.36; No.260, 101-104., 2019.
[4] Mao Jianjun; Zhang Sanxi. Historical Map GIS and Classical Literature Research in Central China, v.9, 72-77., 2017.
[5] Shan Chengbin. Proceedings of the Third International Symposium on Digitalization of Chinese Ancient Books.
[6] Li duo; Wang Yi. The Open Space — of the Age of Data Analysis and the Study of Classical Literature and the Questioner [J.] on the Interaction between Information Engineering and Classical Literature Chinese Cultural Studies, 2006, 101-109.