English Mixed Teaching Database Based on Big Data Technology

Xiaoqin Zhang

Yichun Early Childhood Teachers College,Yichun, Jiangxi, 336000, China
*Corresponding author's e-mail:59711001@qq.com

Abstract

The focus of English majors is to communicate fluently with people from English-speaking countries, so translation teaching in English majors is very important. Now that we have entered the era of big data, the way of education has been updated with the development of science and technology. Big data technology can bring a new way to English professional translation teaching. The establishment of the English translation database can provide reference opinions for English teaching, provide more abundant and standardized teaching cases for English translation teaching, and improve the efficiency and quality of teaching. Today's databases can be built with the SQL Server database management system. The English database under the SQL Server database management system has a large capacity and stable operation, which is of great help to the establishment of an English database.

Keywords: *Big data technology; English; translation; teaching*

1. INTRODUCTION

Big data technology can help teaching innovation and reform, and the teaching platform built with big data technology can provide students with integrated, personalized, customized and socialized learning resources. Nowadays, in the teaching of English majors, the combination of online and offline teaching has become the mainstream. In addition to receiving education in the classroom, students also need to preview, review and submit homework through the teaching platform after class. After many experiments in colleges and universities in recent years, we can find that blended teaching can greatly improve the quality of teaching. Blended teaching is a kind of student-centered, targeted teaching, through which students can develop their own learning ability and interest in learning.

2. ADVANTAGES OF USING BIG DATA TECHNOLOGY IN ENGLISH TEACHING

With the popularization of information technology, people's work and life have been affected by the Internet [9]. Every Internet behavior of people produces data, the data collection, the data will be huge.Big data is a collection of data that greatly exceeds the scope of traditional database software tools in terms of acquisition, storage, management and analysis.Big data has four characteristics: massive data scale, fast data flow, diverse data types and low value density.Big data technology includes a series of technologies for the big data processing, including big data collection, big data storage, big data management, big data mining and analysis, big data display and so on [6]. Big data has a wide range of applications. Nowadays, wherever there is data information, big data technology can be used to mine the value information in the data information.Because of the value of big data technology in today's era, many scientists call the 21st century the era of big data [7].

The development of information technology and big data technology has put forward higher and higher teaching requirements for the teaching of English majors [8]. Nowadays, computers have been able to provide people with basic language translation. Therefore, the translation ability of English majors needs to be more professional and accurate to achieve translation work. "Cinda Ya" requirements. In the translation teaching of English majors in the past, teachers were in the dominant position in the classroom, and students only passively received knowledge. This teaching method could not give full play to students' subjective initiative, and had a bad influence on the cultivation of students' learning ability. With the development of today's information technology and big

data technology, English majors should combine online and offline courses, and encourage students to inquire about the learning resources they need through the Internet, so that students can cultivate autonomous learning ability and improve their English proficiency. The combination of offline classroom and offline classroom can make learning break through the limitations of time and space, and has great potential [9].

Big data technology can also help in English teaching to set up exams.Big data technology can collect a large amount of data information, and mine and analyze these data information.In the traditional English test, the English questions are set by the teacher, which consumes the teacher's energy, and the questions the teacher chooses may be written before.In order to improve the fairness of English, English exam questions set and difficulty set accurate grasp, can use big data technology in English test, big data technology can continue to mining, in the class of students 'information collection and sorting, from the data information understand students' word memory, grammar understanding and the application of complex sentence patterns and so on.Big data technology will give questions according to students' actual level and weaknesses, and such a test content can let students learn deeper knowledge [10].

In the United States, the United Kingdom and other English-speaking countries, have a local development background of a lot of slang, the slang in the local use scope and frequency is very wide, but in the English professional classroom teaching, the application of slang is ignored, in this case, students to find information on the Internet, independent learning, to improve their learning system, improve the practicability of English ability [11].

3. PRINCIPLES OF ENGLISH TRANSLATION DATABASE CONSTRUCTION

The construction of the English translation database is to respond to the changes in the demand for English translation talents and cultivate high-quality English translation talents. Therefore, the establishment of the English translation database must adhere to the principle of practicality [14]. The English translation database must meet the needs of the international development trend of all walks of life in my country. Nowadays, my country's economy is developing rapidly, and many industries need English translators as a way to communicate with other countries. Under such an economic background, English translators need to learn a lot of words to meet the needs of enterprises [15].

The goal of English translation database construction is to enable students to fully understand the process,

skills and principles of translation in the data, and to master the learning and training methods of translation. The construction of the English translation database should comply with the needs of students, combine the teaching objectives with the construction objectives of the case database, provide students with a large number of translation training opportunities, strengthen the students' skill level, and improve the translation's language organization and analysis capabilities [13].

There are numerous resources and learning cases of English translation available in the English translation database. The database covers a wide range of rich content and can meet many needs of students. English translation requires students to accumulate a large number of English words, phrases and grammar. Only when students have a solid basic ability can they fully use Chinese to show the situation, effect and connotation of English. The construction of English translation database should adhere to the principle of content diversification. The content of the English translation database cannot be static, and English will change with the development of time. The database should be updated and improved in time to make the English translation meet the development requirem0ents of the times.
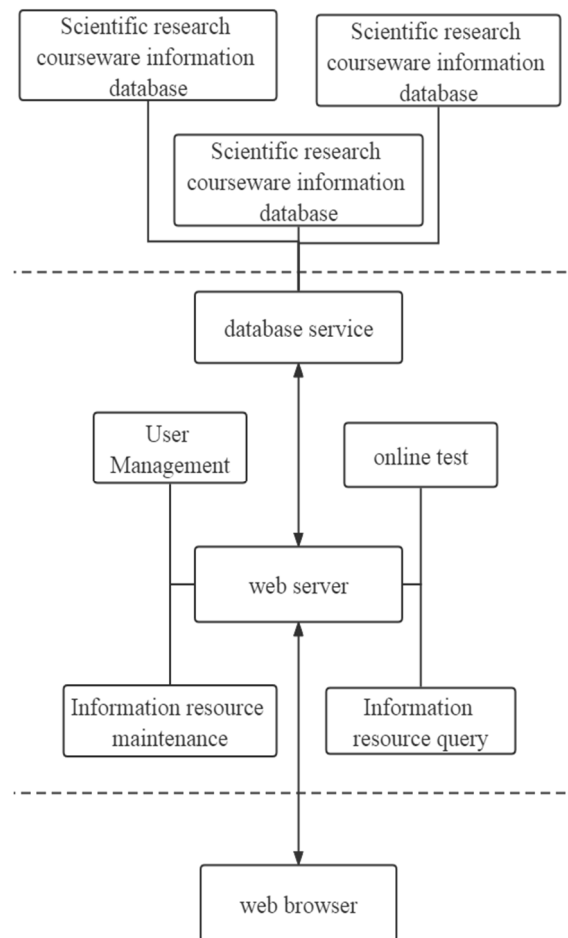

Figure 1: Database structure

4.   ENGLISH TRANSLATION DATABASE CONSTRUCTION

4.1.   Overall structure

The English translation database should focus on the information integration of translation theory and translation cases, and match translation concepts, theories, methods, procedures, literal translation, variable translation, full translation, domestication, foreignization and other concepts with the corresponding cases. In the translation database, we should also pay attention to the retrieval of translation keywords, and mark relevant information so that students can search for materials.

Application course network information transmission technology, big data technology and multimedia technology in English translation database. The establishment of the database uses the SQL Server database management system developed by Microsoft. Database construction must first search and organize relevant Chinese and English materials according to the teaching requirements of English translation, and import the information into the SQL Server system based on a large number of materials. The database relies on the Visual Studio 2019 platform to design the form interface of the database system. Connect the database with the form interface, and you can run a simple English translation case resource database.

The interface design of the English translation database should meet the requirements of the English translation teaching objectives, and the overall structure design concept and content categories should be deeply analyzed. The interface of the case library system is mainly divided into two parts: information query and information update. Each part has modules such as economy, culture, politics, science and technology, and education. Students can query according to their own needs. Each translation case in the database should be added with the case-specific number, editor information, title and other information.
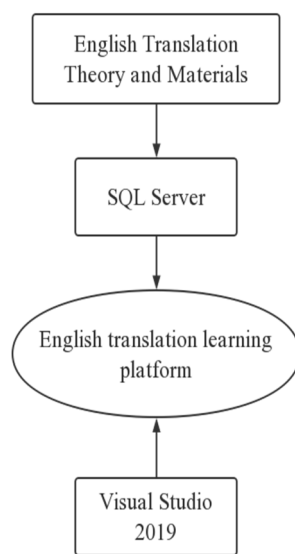

Figure 2: The process of building the database

4.2.   Data processing

When inputting database data, a lot of data is knowledge from books, so it is necessary to preprocess the information, transfer the paper information to the computer, and enter it into the database. There are two ways to turn paper information into electronic documents. The first is manual input, which is slow in efficiency and has a high error rate. Therefore, in the data preprocessing of this database, the second method is used. way, using scanning to enter text data. Colleges and universities can use the easy-to-operate and powerful text editor EmEditor to eliminate impurities in the corpus, and store it in the unified format of TXT and XML for database management (Comell 2018).

The data information of the database needs to be marked in detail, such as the title, keywords, time, etc. of the data. These hyper-language information can use the general markup language (SGML format) for the purpose of document exchange. In the middle of the brackets, use "<...>""</...>" to represent the start tag and end tag respectively. "<Text_head>...</Text_head> (annotation of data information)" "<Title>...</Title> (title of text)" "<Author>...</Author> (name of text author)"" </Source>…</Source> (the subject matter of the text)""<Difficulty>…<Difficulty> (the difficulty index of the text)", etc. In the database, language information such as part of speech, syntax, and picture structure of the learning materials also need to be marked. The database in this paper is mainly based on

tagging parts of speech and paragraphs. The Chinese part-of-speech coding software ICTCLAS version 3.0 and the English part-of-speech coding software CLAW4 can automatically process the parts of speech of Chinese and English corpora. The EmEditor software can automatically assign codes to paragraph marks, and mark the start and end of natural paragraphs in the corpus.

4.3.   Retrieval Module

Retrieval is the core module of the database, which mainly includes functions such as keyword or multiple keyword retrieval, span retrieval, and wildcard retrieval for keywords. The retrieval modules should be divided according to the user's authority. The specific modules include ordinary retrieval and advanced retrieval, which have different applicability and different roles. In the general retrieval module, users can retrieve the first 100 records that match the keywords. The main retrieval process is to first retrieve keywords in the database, summarize and save the specific lock cylinder information in the object array, and then determine the string range of the context before and after. Advanced search can retrieve eligible content. The database will replace the wildcards in the keywords with the symbols in the regular expression, extract the id number of the article to which the keyword belongs, store the position information of all words and indexes in the object array in order, and finally output the content.

4.4.   Density analysis

The emphasis on high-frequency vocabulary in the English translation resource library is necessary. Density analysis can find important words in a large amount of data information. Density analysis is the number of single words or word blocks in statistical articles or English materials, and the number is compared with the vocabulary of the whole study factor [13]. Density analysis function can be implemented based on accurate search and regular search. Density analysis can make statistics based on the results of a certain vocabulary or matched regular expression. The proportion of a single vocabulary or word block in the research factor or corpus can intuitively see the importance of vocabulary in the article.If significant, then vocabulary must be repeated in the article or research factors, indicating that vocabulary is very important. Densitometric analysis can also be performed in combination with the retrieval.The retrieval module plus density analysis can retrieve a single vocabulary, which can be density display for word blocks or regular expressions.

Based on the density analysis, the black and white position maps can be generated.Each matched word or regular expression result was sorted, and colleagues

count the number of all matched words in allCount.Use the CSS style to display the white background black bar.A black bar per percent position ends with a black and white position map.The black and white position map is implemented in the front-end drawing, where all the data is retrieved from the Services [2].
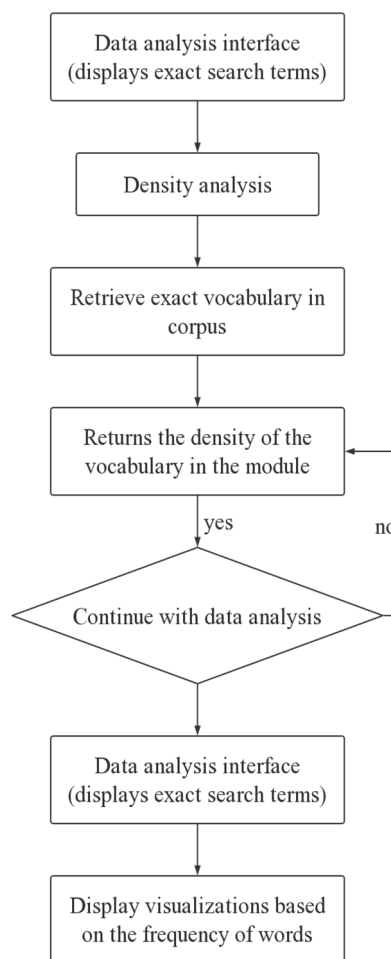

Figure 3: Density Analysis Process

5.   CONCLUSION

In the context of today's era, English translation requires the joint innovation of schools and teachers, improving the teaching mode and cultivating better talents. The development of big data technology provides sufficient realistic conditions for the construction of English translation database. This paper builds a database for English learners based on big data technology. Students can check English grammar, words, articles, slang and other information in the database to help students improve their learning ability.

REFERENCES

[1]   Adair Cornell. Research on Classification and Redundant Information Filtering of Massive Data

in Big Database[J]. Computer Informatization and Mechanical System,2018,1(3):

[2] Ali Akbar Khansir,Fatemeh Ghani Dehkordi,Mostafa Mirzaei. Learning Strategies and English Language Teaching[J]. Theory and Practice in Language Studies,2021,11(6):

[3] Bouamrane Karim,Matallah Houcine,Belalem Ghalem. Comparative Study Between the MySQL Relational Database and the MongoDB NoSQL Database[J]. International Journal of Software Science and Computational Intelligence (IJSSCI),2021,13(3):

[4] Cihan Küçükkeçeci,Adnan Yazıcı. Big Data Model Simulation on a Graph Database for Surveillance in Wireless Multimedia Sensor Networks[J]. Big Data Research,2018,11:

[5] Create a Database of Go-To Experts[J]. Nonprofit Communications Report,2021,19(8):

[6] E. F. E, F. Codd. Derivability, redundancy and consistency of relations stored in large data banks[J]. ACM SIGMOD Record,2009,38(1):

[7] Ford Bridget,Chilton Katherine,Endy Christopher,Henderson Michael,Jones Brad A.,Son Ji Y.. Beyond Big Data: Teaching Introductory U.S. History in the Age of Student Success[J]. Journal of American History,2020,106(4).

[8] Grisham William,Brumberg Joshua C,Gilbert Terri,Lanyon Linda,Williams Robert W,Olivo Richard. Teaching with Big Data: Report from the 2016 Society for Neuroscience Teaching Workshop.[J]. Journal of undergraduate neuroscience education : JUNE : a publication of FUN, Faculty for Undergraduate Neuroscience,2017,16(1).

[9] Kamal Mona Ibrahim,Zubanova Svetlana,Isaeva Anastasia,Movchun Vasily. Distance learning impact on the English language teaching during COVID-19.[J]. Education and information technologies,2021,26(6):

[10] Misra Biswapriya B. New software tools, databases, and resources in metabolomics: updates from 2020.[J]. Metabolomics : Official journal of the Metabolomic Society,2021,17(5):

[11] Rahmati Jafar,Izadpanah Siros,Shahnavaz Ali. A meta-analysis on educational technology in English language teaching[J]. Language Testing in Asia,2021,11(1):

[12] Sheik Yousuf T.,Devi M. Indra. FREQUENT PATTERN SUB-SPACE CLUSTERING OPTIMIZATION (FPSSCO) ALGORITHM FOR DATAMINING FROM LARGE DATA BASE[J]. International Journal of Business Intelligence and Data Mining,2017,12(3/4):

[13] T. Cheng,X. Zhou,Y. Jia,G. Yang,J. Bai. STUDY ON BIG DATABASE CONSTRUCTION AND ITS APPLICATION OF SAMPLE DATA COLLECTED IN CHINA'S FIRST NATIONAL GEOGRAPHIC CONDITIONS CENSUS BASED ON REMOTE SENSING IMAGES[J]. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,2018,XLII-3:

[14] Tomas Lang. A two-level architecture for a large data base[J]. ACM SIGIR Forum,1976,10(4):

[15] Vaalavuo Y.,Antila A.,Ahola R.,Rinta-Kiikka I.,Vornanen M.,Sand J.,Laukkarinen J.. 10 year survival and characteristics of operated IPMNs in the Finnish nationwide database[J]. Pancreatology,2021,21(S1):