English Sociological Terminology Based On Functional Equivalence Theory

Cuiping Xu, M.Sc
Cuiping Xu, Lecturer, School of Foreign Languages, Henan University of Science & Technology, Luoyang , Henan , 471000
Correspondence LecXu;mqpqtf@163.com

Objectives: In the practice of translating English, there are often situations that may lead to missing words. In this case, a computer technology is needed to improve the translation studies of sociological terms in English. This time, based on the characteristics of Internet network data, intelligent robot information is extracted. Methods: According to the knowledge ontology constructed this time, based on the functional equivalence theory, a method based on the automatic construction of the ontology library in the party building domain is proposed. Results:In order to verify the proposed method algorithm, the example study of some sociological terms conceptual terms above the interactive encyclopedia is studied by the ontology created by encyclopedia resources, such as father/sub-relationship, class and instance relationship and attribute relationship, and a total of 72474 relationships are obtained through the final statistical study.Conclusion: From the overall analysis, it can be seen that the sociological terminology research of English computer network based on functional equivalence theory can achieve a good classification effect.

Keywords:functional equivalence; theory; english; sociological terms
Tob Regul Sci.™ 2021;7(5-2): 4818-4828
DOI: doi.org/10.18001/TRS.7.5.2.46

In the practice of translation, the author finds that the absence of equivalence in the translation of legal terms into Chinese is frequent and unavoidable. With the development of computer technology, it is more and more necessary to apply computational technology to the study of English social terminology. Distance education has become an assistant means of education in many schools and training institutions, and has been applied more and more widely [1]. While distance education provides excellent service and experience, it has higher data processing capabilities. Therefore, based on functional equivalence theory, distance education technology is applied to English sociological terminology research [2].

In the research methods, according to the characteristics of network data, the network data is extracted by using the technology of information extraction of network intelligent robot, information extraction based on CRF template and deep web information extraction technology of filling in forms. At the same time, the knowledge ontology is constructed for encyclopedia resources by combining the ontology construction technology. In view of the characteristics of information in the field of party building, a method based on the automatic construction of the ontological library in the field of Party building is proposed, and the construction of the ontological library in the field of Party building is completed[3].

The research has certain innovations. Firstly, the Deepweb knowledge is elaborated to prepare for the deep acquisition of data, and the data collection is improved to enrich the ontology library. The current situation and latest development of ontology creation are studied. Ontology is used to describe knowledge points and realize automatic ontology creation. Research and implement an ontology-based distance education platform, the goal is to establish a distance education platform with better flexibility,

scalability and maintainability.

The research is divided into three parts. The first part is literature review, the second part is based on functional equivalence theory, and the third part is tested and verified.

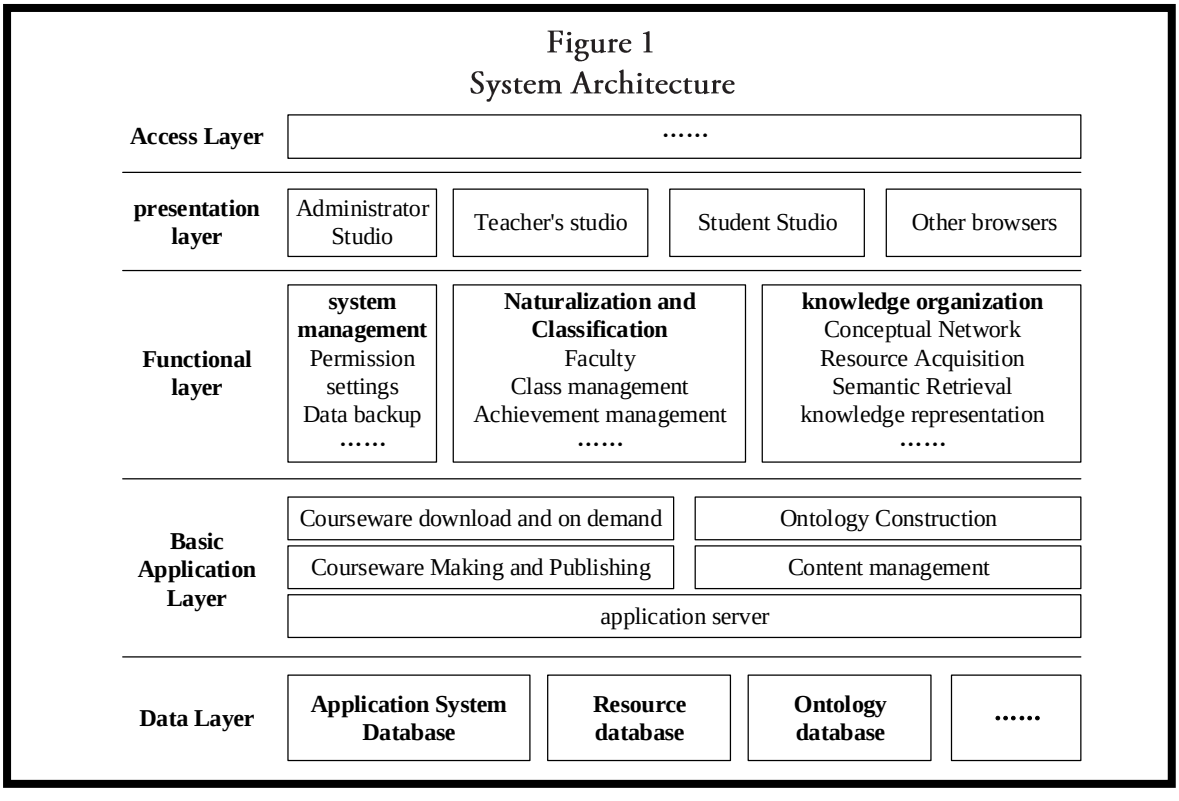
It is not easy to achieve full functional equivalence between the target language and the source language in the translation of English sociological academic terms into Chinese, because no one can completely convey the concepts and ideas in a text in another language. For centuries of sociological, political and literary use, words and structures always carry cultural traces, which cannot be fully transmitted [4]. When translators translate the source culture into the target culture, they find that there are some areas that can well correspond to each other, but some do not correspond to each other. This means that there are some factors in the source culture that do not exist in the target culture. The linguistic expressions of these factors are "gap" or "void" in the target language. There is no exact equivalence or unequal phenomenon in the translation of English and Chinese sociological terms[5]. In short, the main reasons for this kind of peer-to-peer are: differences in sociological systems; differences in cultural background; differences in social systems; differences in language. Obviously, when translating the terminology of sociology, the translator faces such a problem: on the one hand, sociological translation requires the equivalence of linguistic functions and the equivalence of sociological functions; on the other hand, there are often no exact equivalents in translation. So, how should the translator solve this contradiction and thus guarantee the quality of the translation of the term? Nida is a famous translation theorist in the United States. He proposed the translation principle of "dynamic equivalence" in the "Exploration of Translation Science" published in 1964 [6]. The so-called dynamic peer-to-peer translation refers to the reproduction of primitive information from the semantics to the style in the most similar and natural equivalents in the target language. In the translation of dynamic equivalence, the translator should grasp the meaning and spirit of the original text, and not constrain the form correspondence [7]. However, Nida's dynamic peer-to-peer translation is different from the traditi onal "free translation". The former requires the translation to reproduce the original meaning as completely as possible in different language structures, while the latter is often the translator's unrestrained freedom [8]. In summary, sociological translation requires equivalence in language function and equivalence in sociological function. However, in translation practice, the authors found that there are often no exact equivalents in translation. So, how should the translator solve this contradiction and thus guarantee the quality of the translation of the term? That is to say, this is studied.

METHODS

Through a large number of research and analysis, the work mainly has the following difficulties: The strategy of data acquisition includes two parts: Firstly, which acquisition strategy is used to access the Web, so that the data acquisition robot can accurately judge whether a web page is related to the subject, extract as many relevant pages as possible, and visit as few irrelevant pages as possible to ensure the quality of the web page. Secondly, according to the characteristics of domain deep network data, it is a main problem to design a strategy to effectively collect and rationally utilize deep network data to improve the coverage of data acquisition. The creation of ontology: At present, there are four main methods of ontology construction: manual construction, reuse of existing ontology, semi-automatic construction of ontology and automatic construction of ontology. However, each of the four methods has its own drawbacks. How to effectively create domain ontology is the key problem to be solved in this paper. Ontology construction of learning resources in learning platform: How to build an effective resource ontology and what kind of construction model to adopt can not only solve the sharing and reuse of learning resources, but also solve the problem of machine understanding, so the ontology construction of internal resources of learning platform is also one of the problems to be solved [9]. First of all, to solve the problem of data ollection, it requires a better recall rate and precision rate and pays attention to domain switching and transplantation, including two main parts. Firstly, data acquisition in surface network domain is

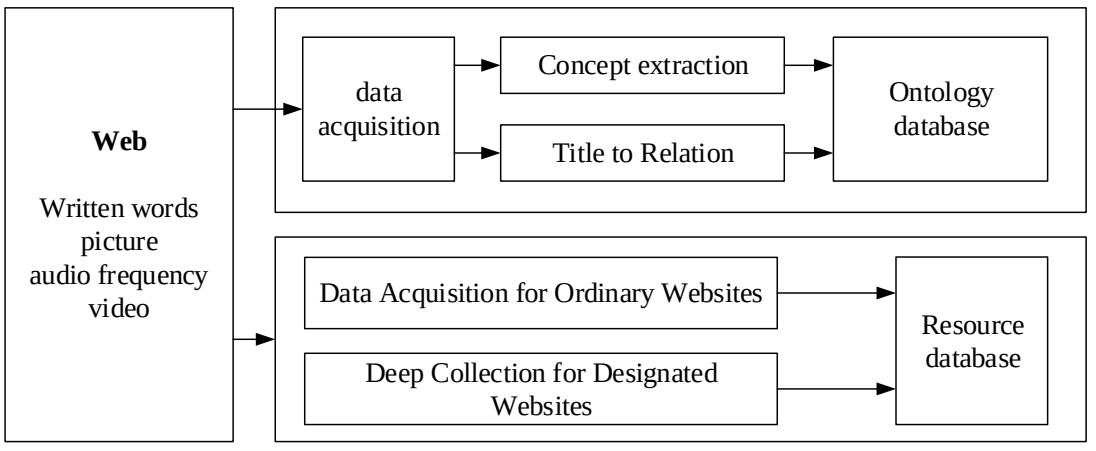
carried out. According to the characteristics of data distribution, data acquisition strategy is set up. Subject correlation is judged by the existing SVM classification algorithm which has better effect on text classification. A surface network domain classification robot is designed and implemented. Secondly, the collection of deep-net data, using the domain characteristics of deep-net distribution to design the information integration strategy of deep-net, find an efficient and feasible method to design and implement the deep-net information extraction robot, reuse the deep-net data, and improve the search coverage [10]. Secondly, find an efficient method to automatically create domain ontology. Through the previous research and study, reading the related articles on the use of Chinese Wikipedia to create domain ontology is deeply inspired. A lot of research has been done on the problem of automatic creation of Encyclopedia resources, and domain switching is taken into account. A method of automatic creation of ontology using encyclopedia resources is proposed. Taking Party building knowledge as an example, domain ontology of Party building is automatically created to support text follow-up work [11]. Finally, a platform for distance education with rich content and comprehensive knowledge is designed and implemented. According to the design principles and basic ideas of distance education platform, the overall logical framework of the platform is shown in Figure 1.

Figure 1
System Architecture



In particular, the resource data and ontology database in the data layer are the key databases and key technology applications. The resource database and ontology database are created by data acquisition and ontology automatic construction. The construction process is shown in Figure 2.

Figure 2
Resource and Ontology Database Construction Flow Chart



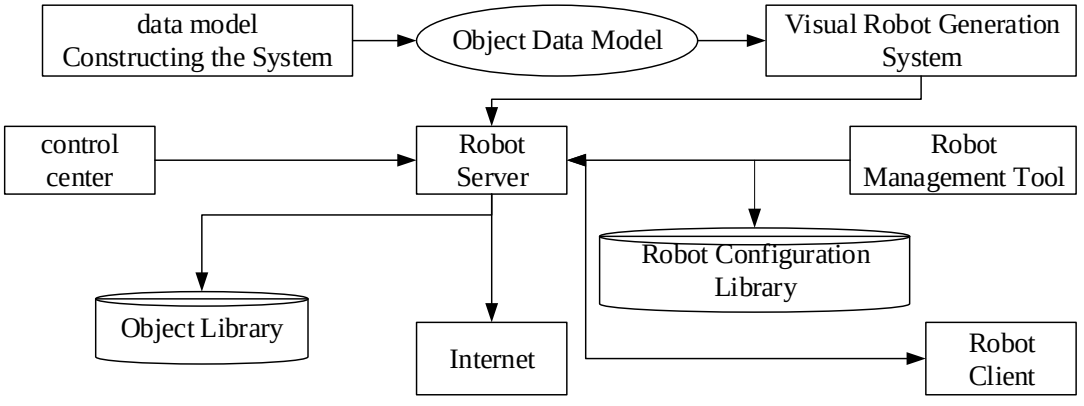
An Analysis of Data Characteristics of English Sociological Terms

Through the previous research and analysis, we found that the information in specific areas of the network has the following characteristics: information distribution is not centralized and deep-seated. It is not difficult to find that most of the information in a certain field on the network is divided into several categories and distributed throughout the network. Each category often exists in the form of an independent website. Some of the information exists in the secondary pages of some websites, and a small part of the information scatters irregularly in the network [12]. In addition, domain information is becoming more and more deeply distributed in the network, many of which are "hidden" under the query interface (form) of the website. Distribution characteristics of information are conducive to collecting most of the data centrally, mainly from the theme website or website containing a large amount of domain information, which can achieve efficient collection of information. The deepening of information makes traditional web crawlers unable to effectively collect deep information resources, and the deep web information collection technology based on domain is a feasible idea [13]. The breadth of information coverage varies from site to site: some areas are clear, such as forestry, agriculture and medicine, which is conducive to finding feature sets to describe; for areas with broad coverage and unclear classification, it is difficult to determine feature sets to describe, which increases the difficulty of data collection [14]. The ontology database to be built in the field of Party building is more complex. Firstly, there are many kinds of Party building information, such as characters, events, meetings, documents, culture, films, music and so on. These classifications exist in the categories of characters, history, society and art, but the contents of these classifications do not belong to the field of Party building, which poses new difficulties for the accuracy of data collection [15].

A solution to this problem is proposed. Firstly, since distance education platform provides learners with not only knowledge in a certain field, but also knowledge covering all fields, it is necessary to collect all data under the big classification. In view of the limited time relationship and ability, only four kinds of information related to the field of Party building, namely, personage, history, society and art, have been collected. Second, the information requirements in the field of party building must be accurate. However, the amount of network information is huge, and it is inevitable that there will be flaws. Therefore, an authoritative media website must be found as the source of information extraction, which can improve the accuracy of the data. Due to the characteristics of network information, the following two collection

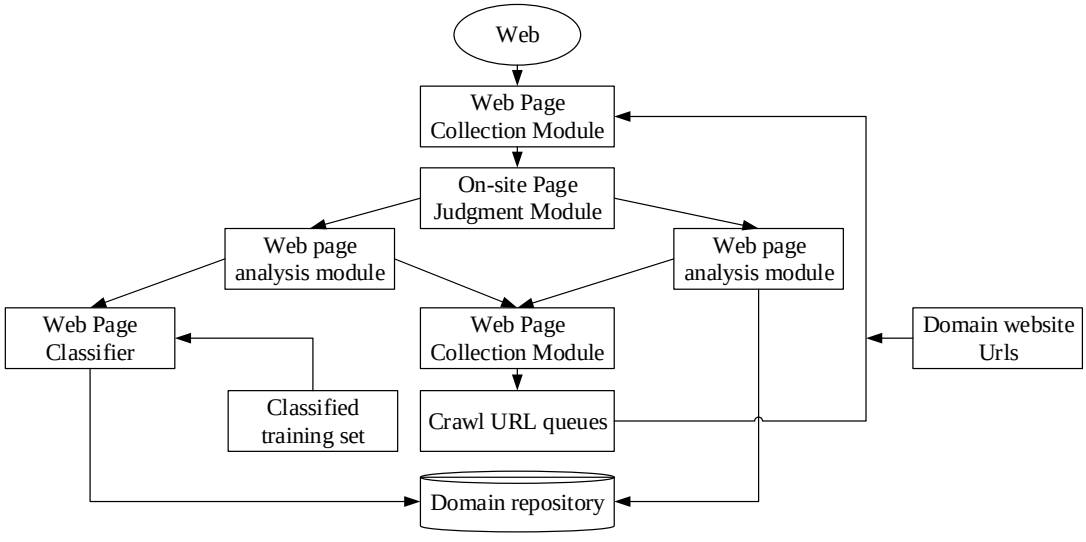
strategies will be adopted: data collection for ordinary websites and deep data collection for designated websites. According to the above analysis, in view of the distribution characteristics of domain information, for ordinary websites, the information is collected by a crawler strategy which combines manual intervention with first acquisition and then filtering. Network intelligent robot VWIE is an information extraction technology based on Wrapper which has been implemented by our team. The technology provides a graphical interface, built-in a web browser, just click the mouse to choose, not only can flexibly define any content to be collected in the web page (such as text, pictures), but also can fully simulate the user to browse the operation of all the steps of the web page (such as entering registration information, looping, turning pages, etc.), and automatically form a network intelligent robot, to achieve the automatic collection of web information. The architecture of the network's intelligent robot VWIE is shown in Figure 3.

Figure 3
VWIE System Architecture



Similarly, the network intelligent robot VWIE can also be used to create a topic classification robot to improve the accuracy and effectiveness of information classification. Figure 4 is a diagram of the classification robot architecture.

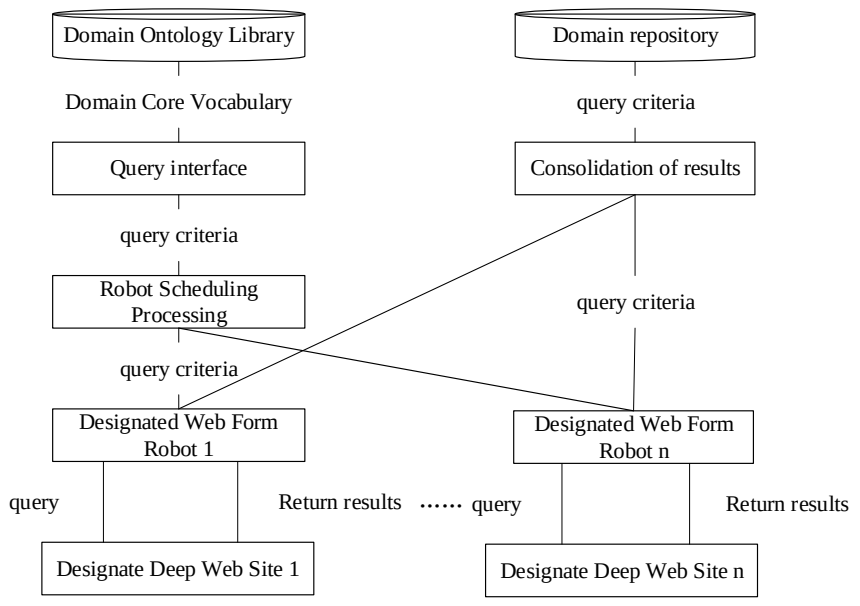
Figure 4
Classified Robot Architecture



Deep Collection of English Sociological Terminology Data

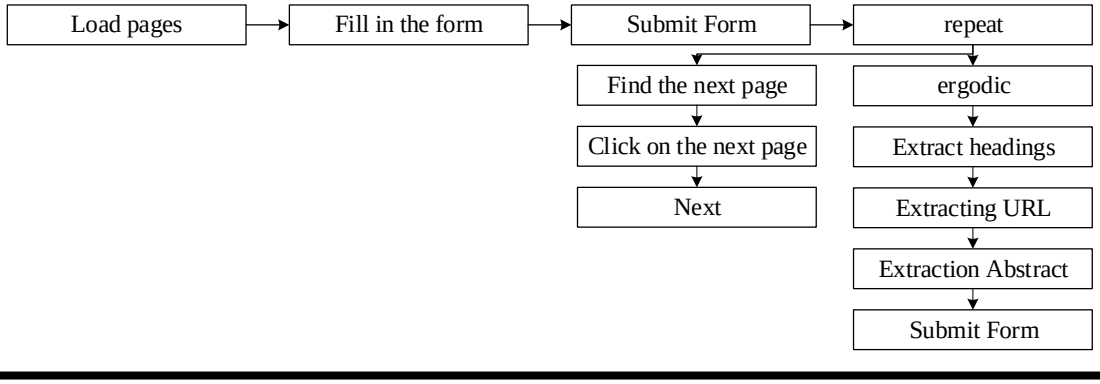
The distribution of information in the network domain is deepened. Based on this feature, a form-filling technology is designed to form a deep-seated information collection scheme based on form filling for a specified website. The data collection amount is increased, and the data resources of the deep network domain are searched more accurately and comprehensively. The deep network information collection scheme based on form filling is performed in two ways: continuous and real-time acquisition. Continuous collection method, make full use of the domain ontology, the core vocabulary of the field as a keyword input, the background continues to the designated deep-net website data integration, according to the URL re-merger results into the warehouse, to ensure that the field of deep-net data large and complete collection; Real-time way, according to the user's query criteria real-time query deep network website and the results combined feedback to the user, so that to a certain extent to make up for the search engine data update speed. Taking the field of party building as an example, manually select a certain number of query effect good deep-net party building website to make robots, query interface is provided, in order to enter query conditions and select the deep-net website to query, and then by the robot scheduling module according to the user query conditions call the designated robot. By filling out the deep net form in real time to query its online database, and finally the robot query results combined and returned, the deep net acquisition framework is shown in Figure 5 below:

Figure 5
Deep Web Information Integration Framework



Due to different websites have their own independence, unable to find a unified web page extraction templates for different site effectively extract the query results, data collection is mainly based on network intelligent robot VWIE completed. VWE are template information extraction tool, therefore need to use its extensibility, to add without a template information extraction of CRF components, in order to realize the automatic extraction of deep web site query results. Visualization is used to customize the software robot's filling and submitting actions for a single designated deep web site, so as to realize the real-time form query function of the designated deep web site. The robot that can fill and submit forms is called the designated deep web form robot. Its workflow is shown in Figure 6 below.

Figure 6
Specifies the Workflow of the Deep Web Form Robot



RESULTS

Test Environment

The system uses My Eclips as the main development platform, My SQL as the background database, resin server, JSP and other front-end technologies, and the platform is basically built. The Java-based language for development is based on Java simplicity, platform irrelevance, security, robustness, and comprehensive support for the Internet, in addition to the following considerations: Search engine system is a web-oriented application software for Web data, and Java is the programming language based on the Internet development, in the network programming has a unique advantage; Java has good stability and security, multi-threaded technology, its perfect memory allocation and release technology so that programmers do not have to worry about memory overflow, thread deadlock and other programming problems, but focus on the development of high-performance programs; There are many Java open source communities in the current network, and many excellent Java open source projects can be borrowed and referenced, such as Nutch, Apache HTTP Request Library, HTML Parser, Lucene, Paoding and other open source toolkits.

Test Results and Analysis

Interactive Encyclopedia (www.hudong.com), the world's largest Chinese encyclopedia website, is founded on July 18, 2005. Interactive Encyclopedia is dedicated to providing hundreds of millions of Chinese users with massive, comprehensive and timely encyclopedia information for free. Compared with other encyclopedia websites such as Chinese Wikipedia and Baidu Encyclopedia, Interactive Encyclopedia has richer vocabulary resources and more understandable knowledge structure. As of February 2012, Interactive Encyclopedia has 6.36 million entries, 6.4 billion words and 6.74 million pictures. Interactive encyclopedia is chosen as the main source of knowledge acquisition to construct ontology automatically. The first-level classification tree of Interactive Encyclopedia consists of 13 categories, which are classified as nature, culture, people, history, life, society , art, economy, science, sports, technology, geography and Hot words. Interactive encyclopedia is not only rich in classification, but also quite clear in classification. Based on this feature, it is feasible to use encyclopedia resources to create ontology. Due to the huge content of interactive encyclopedia, it is impossible to extract all the data. Based on the time relationship, the ontology of the extracted data is automatically constructed for the four categories of personages, history, society and art involved in the field of Party building. Unlike wiki, Interactive Encyclopedia does not have dump data. Therefore, it is necessary to collect the data of interactive encyclopedia, select four categories of people, history, society and art as the four categories of knowledge field, and then analyze and process them to get the ontology of science and technology field. The data acquisition here is still completed by using the network intelligent robot VWIE, which has existed in our research group. It lasts 5 months, and uses the width optimization step by step to traverse all kinds of data in interactive encyclopedia, which is stored in storage for subsequent processing.

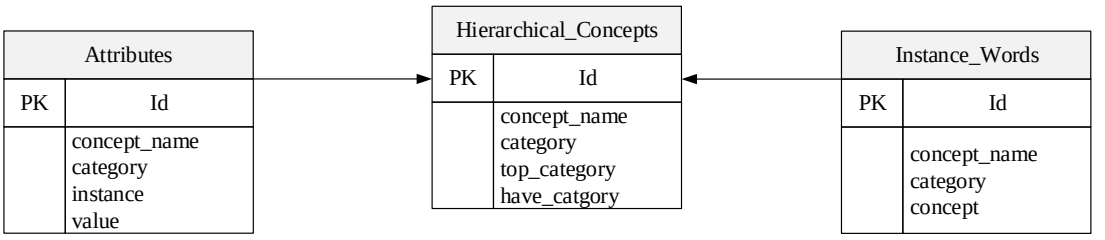
In the Interactive Encyclopedia page structure, there is a page that gives all the entries of the category. Using the natural structural advantage of Encyclopedia resources, it is easy to extract the relationship between classes and examples. According to the ontology inference rule, an instance belonging to a subclass must also belong to its parent class. The interactive encyclopedia page gives all the terms that belong to the class. There are duplicate entries in all the entries, which are duplicated. Here, you need to do the weight removal, and delete the entries in the parent class that belong to the subclass at the same time. For the extraction of relations, the semantic characteristics unique to the encyclopedia resources are still utilized, and the extraction methods are given for the classification relationship and the attribute relationship respectively. Think of a concept class as a class of domain ontology. In the data collection, this paper finds that the interactive encyclopedia has a good classification relationship structure, and each classification page will display the current classification, the superior classification and the sub-categories. It is considered that each concept class is a class, and its subclasses constitute a

parent/subclass relationship. With this obvious hierarchical relationship, it is easy to extract parent/child class relationships. It needs to be pointed out that not all the classes in the hierarchical relations in Interactive Encyclopedia can be regarded as conceptual classes. The most confusing one is examples, so we need to prune the classes in the hierarchical relations. The basic idea of pruning is to use some empirical rules to prune all non-conceptual classes in hierarchical relationships. The remaining conceptual classes and their hierarchical relationships constitute the classes of candidate domain ontology. The rule of thumb adopted is as follows: Any conceptual class that satisfies the following rules is considered to be filtered by non-conceptual classes: Rule 1: Class names that denote years and equal time; Rule 2: Class names that denote bibliographies with " 《》 "and Rule 3: Class names that have no subclasses. Through the analysis of interactive encyclopedia pages, this paper finds that although different terms are different, but the structure is generally the same, the left side generally has a content directory, if the term editing is perfect, the right side will have a basic information box and this kind of basic information. This feature is particularly prominent in military and character categories, which can be attributed to the same category by extracting the common part of the entries under the same subcategory. Attribute extraction is gradually improved from lower level to higher level.

At present, there are two main methods of ontology storage: text storage and data storage. Text storage is to store ontology libraries in ontology file system as files. If the ontology is stored in this way, the system needs to read the ontology library from the file into the memory before searching for the ontology, then operate the ontology in the memory, and finally write back the processed results to the file. Ontology can be stored in many file forms, such as XML, RDF, OWL, etc. Text file can preserve the semantic information of ontology completely, and it has good expansibility and flexibility. But every operation of ontology in this way involves the operation of the whole ontology file. Especially when the ontology is large, the system needs a lot of memory management, and the efficiency is difficult to be guaranteed. Ontology storage based on file system is usually only suitable for ontology editing and building, but not suitable for large-scale ontology data storage and query in practical applications. The way of database storage is to organize and store ontologies in database according to certain strategies. This method takes advantage of the advantages of database in managing and storing data, so it can improve the efficiency of ontology storage and query. Select data inventory as the storage domain ontology. In order to minimize the loss of semantic information caused by the ontology being stored in the database, the use of relational database as the storage background of the ontology needs to carefully design the storage mode of the ontology, and improve the efficiency of ontology storage and management without losing the semantic information of the ontology. A hybrid model combining vertical and decomposition modes is designed to store domain ontologies. Classes, instances, attributes and synonyms are stored in four tables respectively. Each class has a unique ID identifier. In order to maximize the retention of semantic information and minimize the space occupied by information, the ID identifier is skillfully designed. The semantic information of the parent/child class is hidden in the ID, so the parent/child class relationship does not need to be stored separately. Similarly, ID representations of instance tables also contain instance relationships, which will greatly save storage space. In order to further improve the query efficiency, the storage space is balanced. Column fields are added to each table to query aspects. Instance table, all instances in the body of the domain and the relationship between the class and the instance, named Instance_Words. Similarly, each instance word is assigned a unified ID to ensure the uniqueness of the ID, and each ID is designed to give each ED class an instance relationship. For the convenience of query, a category and a concept field are set to indicate the number of levels to which the current instance belongs and the concept class to which the current instance belongs. Attributes table, which stores all attribute relations and attribute values in domain ontology, is named Attributes. Also give each attribute relationship a unique id, other fields have attribute_name, concept, instance, value, and so on. Each tuple represents the value of an attribute in an instance of the

concept to which it belongs. Similarly, you can use SQL statements to easily find attribute relationships and attribute values. Figure 7 below is a schema diagram of ontology storage.

Figure 7
Ontology Storage Relational Schema Diagram



Based on the ontology created by the encyclopedia resources, the extracted relationships include parent/subclass relationships, class and instance relationships, and attribute relationships. The total relationship is 72474 relationships, as shown in Table 1.

Table.1
Extracts Relationships and Number Lists

relationship	Number
Parent/child classes	5704
Example	63592
attribute	3178
Total	72474

Encyclopedia resource is a semantic resource platform which integrates resource sharing and openness. It is a huge resource database maintained and managed by the vast number of netizens. The categories and concepts in different fields of encyclopedia resource are different. On the other hand, encyclopedia resources are rich in resources, but lack of professional participation in the field. There is still much room for improving the quality of the domain ontology.

DISCUSSION

The key technologies involved in distance education platform are deeply studied and their implementation methods are described. In the data acquisition part, according to the characteristics of data, data acquisition in Web domain is realized by using SVM classification technology, network intelligent robot information extraction technology, CRF template-free information extraction technology and deep-net information acquisition technology based on form filling. In the part of ontology creation, the automatic establishment of domain ontology is discussed by using the semantic advantage of encyclopedia resources, and the method of creating domain ontology is put forward. In order to verify the proposed research method, a search engine classification is built based on the development platform of My Eclips. Taking the classification of multiple entries on the Interactive Encyclopedia as an example, the ontology created in the encyclopedia resources is extracted. The extracted relationships include parent/child relationship, class-instance relationship and attribute relationship, totaling 72474 relationships. This study fills in a certain gap for similar English social terms and has very good practical significance and value.

Human Subjects Approval Statement

This paper did not include human subjects.

Conflict of Interest Disclosure Statement

None declared.

Reference:

1. Adikusuma F, Pederick, Daniel, McAninch, Dale, et al. Functional Equivalence of the SOX2 and SOX3 Transcription Factors in the Developing Mouse *Brain and Testes. Genetics,* 2017, 206(3):1495.
2. Chan A, Lew, Claire, Wang, Xiao Jun, et al. Psychometric properties and measurement equivalence of the Multidimensional Fatigue Syndrome Inventory- Short Form (MFSI-SF) amongst breast cancer and lymphoma patients in Singapore. *Health & Quality of Life Outcomes,* 2018, 16(1):20.
3. Enfield, Katey S. S, Martin, Spencer D, Marshall, Erin A, et al. Hyperspectral cell sociology reveals spatial tumor-immune cell interactions associated with lung cancer recurrence. *Journal for ImmunoTherapy of Cancer,* 2019, 7(1):13.
4. Gordon E M, Laumann, Timothy O, Adeyemo, Babatunde, et al. Individual Variability of the System-Level Organization of the Human Brain. *Cerebral Cortex,* 2017, 27(1):386.
5. Hischier R, Inge Reichart. Multifunctional electronic media-traditional media. *International Journal of Life Cycle Assessment,* 2018, 8(4):201-208.
6. Obradovic-Wochnik J, Hayes S. re-visualising international relations: audio-visual projects and direct encounters with the political in security studies. *European Political Science,* 2017, 16(3):415-429.
7. Pesen A, Üzüm B. The Analysis of Self Efficacies of English Language Teachers in Terms of Creative Drama Use in Education. *Universal Journal of Educational Research,* 2017, 5(8):1378-1385.
8. Peter P, Scholten P. Assimilation of Ethnic-Religious Minorities in the Netherlands: A Historical-Sociological Analysis of Pre-World War II Jews and Contemporary Muslims. *Social Science History,* 2017, 41(3):477-504.
9. Singh B P. Sikh militants' terms of discourse: Religion, Khalistan/nation and violence. *Sikh Formations Religion Culture Theory,* 2017, 12(2-3):1-16.
10. Sotillos S, Aguilar-Aragon, Mario, Hombría, James Castelli-Gair. Functional analysis of the Drosophila RhoGAP Cv-c protein and its equivalence to the human DLC3 and DLC1 proteins. *Scientific Reports,* 2018, 8(1):4601.
11. Thiel A, Seiberth, Klaus. Migrant background, culture and ethnicity: sociological terms without explanatory value?!. *European Journal for Sport and Society,* 2017, 14(3):183-185.
12. Thiel A, Seiberth, Klaus. Migrant background, culture and ethnicity: sociological terms without explanatory value?!. *European Journal for Sport and Society,* 2017, 14(3):183-185.
13. Ushioda E, Dörnyei Z. Beyond global English : motivation to learn languages in a multicultural world. *Modern Language Journal,* 2017, 101(3):451-454.
14. Xing H P. On E-C Translation of Attributive Clauses in English for Science and Technology from Perspective of Functional Equivalence. *Overseas English,* 2018(1):150-151.
15. Yong L, Shuang L, Song H. Ultracontractivity and Functional Inequalities on Infinite Graphs. *Discrete & Computational Geometry,* 2018, 149(5):1-14.