ISSN : 0974 - 7435 Volume 8 Issue 2

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 8(2), 2013 [173-179]

The application and exploration of web3.0 technology in distance education system

Xiaogang Liu
ZheJiang Economic & Trade Polytechnic, Hangzhou, (CHINA)
E-mail: sylxg@163.com

ABSTRACT

This paper studies on Web 3.0 technologies and distance education system, point out that distance education is in great demand because of large area and large population in China. In the current distance education system, teaching resources are usually simply distributed with the poor share which has to be constructed repeatedly and the system can't meet the need of being enlarged with poor configurability. According to the characteristics of distance education and Web 3.0 technology, propose the building of the distance education system base on Web 3.0, the new system makes much progress on educational service, teaching resources share, intelligent and personalized, which can guarantee the large-scale teaching activities. In the paper the preprocessing mechanism and system architecture for the distance education system are designed which are based on Web 3.0,the workflow and the main function of the system is described, a theoretical basis for the specific implementation of the system is provided.

KEYWORDS

Web 3.0;
Distance education system;
Preprocess;
System architecture.

INTRODUCTION

With the development of internet, modern education has changed greatly. More and more people have become beneficiaries of distance education. As China has an enormous population with a large land area, the demands of distance education are huge. The characteristics of distance education are also different from those of overseas education. In 1999, China started to implement modern distance education project. Through ten years of development, the market scale of China's distance education market reached 44 billion Yuan in 2009, and 72.3 billion Yuan in 2012, forming the largest distance education system. After developing for many years, China's distance education development has accumulated large quantity of teaching materials, which are dispersed in all places and lack in management and share; The existing distance education has not formed standard, and the individuation and customization of system service cannot reach the comprehensive service which has intelligent recognition function. The characteristics of Web 3.0 technology are individual, interactive, and in-depth application services. The modern distance education system based on WEB3.0 technology can solve some problems in China's current distance education.

FULL PAPER

THE SIGNIFICANCE OF ESTABLISHMENT OFDISTANCE EDUCATION SYSTEM BASED ON WEB3.0

China has many schools which engage in distance education, and most schools have their own website servers, which store large quantity of resources. The schools publish the resources on internet though B/S structure for their teachers and students to use. The resources of other distance education schools can only be searched through search engine. However, the search engine generally targets at webpage document, but is not powerful in searching literature, courseware, and video; moreover, the search is based on key word search and full-text search mechanism, so the search results are enormous, which are mixed with numerous useless information. As a result, the teachers and students find it hard to choose. The distance education system of B/S structure designed by traditional mode is only applicable to the conditions that the range is small and learners are not many[1]. The distance education system based on WEB 3.0 can provide the powerful functions which the traditional system doesn't have, so it is very suitable for China's vast distance education market, and mainly includes the following several points:

(A) Beneficial to realizing the standardization of distanceeducation

Currently, China has not formulated the technical standard and text standard concerning distance education. The distance education system of each network school establishes its own system and the resource format is not unified, which cannot realize effective communication. The distance education system based on Web3.0 can give full play to the characteristics of Web 3.0, and plans to use the software and hardware resources of each place as a whole, and provides powerful network teaching abilities and resources. For example, in underdeveloped western part of China, the software and hardware construction lag generally exists; moreover, the problem cannot be solved within a short term. Through distance education system based on Web 3.0, it is available to integrate the software and hardware of east and west regions, and provide powerful service abilities uniformly, thus effectively avoiding repeated construction of resources.

(B) Realize high-efficient resource share and access

With regard to China's current distance education system, the network resources of other schools are mainly obtained through search engine, but it is very hard to find out the latestand most useful parts, which causes poor share of resources. After the establishment of distance education system based on Web 3.0, high-efficient resource share can be realized: the resource visitors need not to know which server the resource is located at, and the system coordinates to manage all resources and uses unified resource list to provide most reasonable services, which are not restricted by domain and time. For any resource visitor, it only requires him to log in for one time, and the system will automatically analyze his UP, define route, search the resource closest to him, and establish good connection, so as to provide the fastest and best access.

(C) Realize intelligence and individuality

In current distance education system, the provider of teaching course stores designed courseware in server and waits for the user to visit through browser. Such distance education lacks in guidance on user learning, doesn't fully reflect teachers' leading role, and cannot adjust study strategy intelligently according to learner's conditions. The distance education system based on Web3.0 has intelligent recognition function. For different users, the system will, in light of users' interests and hobbies etc. and through selection of contents, automatically provides users with the best study schemes and resources on the basis of users' behaviors, habits, and information; the student can participate in the learning of many groups, publish comment, write experience, hand over homework, participate in evaluation, systematically summarize the information for students and collect it for teacher to examine and appraise. One of the most outstanding characteristics of distance education based on Web 3.0 is that individualized education can be realized. Moreover, the data collected by intelligent system actively analyzes student's basis, ability, interest, hobby, etc., so as to automatically provide appropriate study schemes and advices.

At the same time, the distance education system based on Web 3.0 will ensure network study any is available everywhere. You can learn anywhere, any time,

and with any machine0You can also use various mobile equipment for access. The student will get good achievement of distance education if they strive to learn.

THE DESIGN OF DISTANCE EDUCATION SYSTEM BASED ON WEB3.0

(A) Data preprocessing mechanism

The Web pages distributed on internet not only have large quantity, but also are dynamic and without organization structure, so it is very difficult for user to seek accurate information on Web. The general method is to use search engine to search web site, download webpage, and create relevant sentence index. However, the search result includes much irrelevant information, and lost the meaning of semantics at the same time. We adopt the resource crawler and resource identification module based on semantics to collect information, so as to avoid the problems of search engine.

(1) The resource crawler based on semantics

During data preprocessing, it is necessary to construct a kind of resource crawler based on crawler. The crawler is a program which can grasp Web document automatically. It supports more effective information search and analysis, which can reduce returned irrelevant webpages. The crawler can collect data from mixed data source and create semantics database, and use knowledge configuration file point at resource, key word, search order, and inquiry progress schedule. The search progress is driven by semantics and time, which replaces problem stimulation and query-driven[2] in traditional Web search process. The core of resource crawler computation is to construct sequencing system of resource prior values. We adopt many methods to calculate the sequencing of prior values of source correlation degree, such as correlation degree analysis, webpage content analysis, linking analysis, reference substance filtration, resourcelinking prediction. The Figure 1 shows the module part of crawler. The following introduces the work principles of the resource crawler.

Firstly, the crawler grasps the Web page, which is not visited yet, for processing. Calculatelinking and content analysis on initial page. Eliminate webpage HTML mark and stopword. Then get pure text contents. After processing of participle and work characteristic mark, lay foundation for latter semantic analysis. After completing preprocessing, the linking structure and relevant example of content resource in visited webpage will be extracted and calculated.

In the second step, the resource linking prediction module will calculate resource prediction on linking information, and calculate resource linkingcorrelation degree; in order to realize "intelligence" and enable sys-

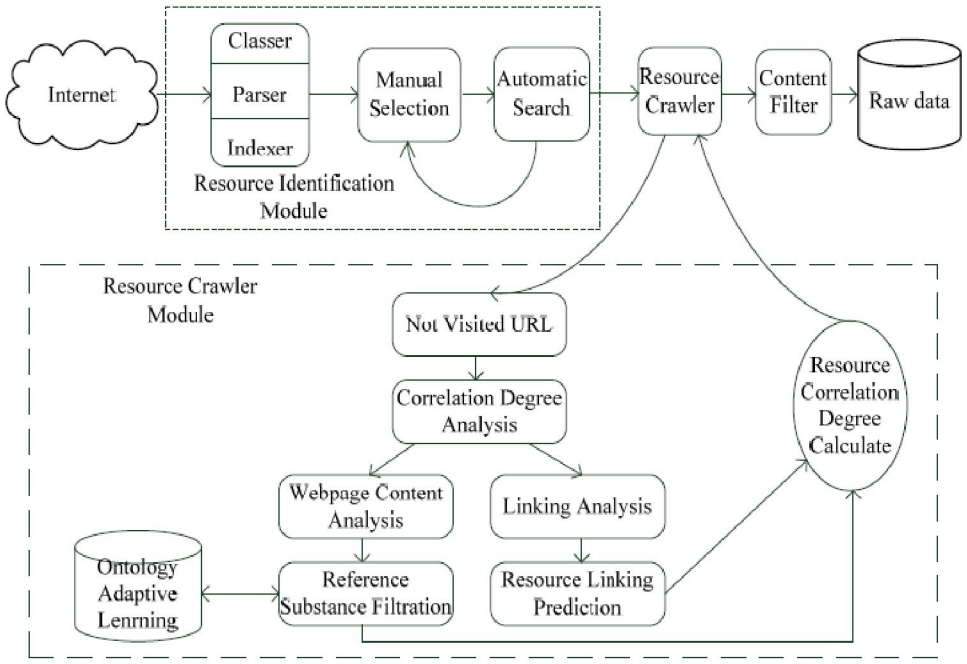

Figure 1 : Schematic Diagram of Preprocessing Working Mechanism.

tem to understand and integrate resource and user file, adopt appropriate tool to construct standard, formal, and operable adaptive resource ontology. The ontology can conduct adaptive learning of statistical principle, reach better screening effects. The model of webpage content analysis is based on the computation of resource correlation degree calculated by resource ontology on webpage content information.

In the third step, the prior value computation module of resource correlation degree integrates linking prior value and content prior value. In light of computation method of resource correlation degree prior value, get crawler to grasp prior value sequencing. Once good prior value sequencing is got, the resource collecting of crawling can go on successfully.

(2) Resource identification module

Data preprocessing firstly uses a resource identification module to position and collect relevant data resource, as shown in Figure 1. In data preparation stage of resource identification module, use data classifier to classify and collect data of different document types. The parser and indexer are used for analyzing data and constructing index and dictionary library during data-collecting process. In the latter stage, through ontology development, extract and collect list, key word advise, abstract and classification constitution, and use these data to create database of raw data. At last, depend on high-level semantic data collection, help user to locate information rapidly and accurately.

In combination with manual selection and automatic Web search, utilize recursive collection construction system to realize resource identification module, Firstly, through detailed and planned manual selection, identify several seed URL. Then use automatic Web searcher to search these URL and capture relevant Web sites and extract favorite page. These output favorite webpage linking list may point at other domain or irrelevant website. Then implement manual selection again. Use obtained URL as latest identification seed to repeat the above process to create and collect the result.

After identification and processing of resource, the document collection work is automatically implemented by resource crawler based on semantics. Most Web pages include the linking of other webpage. The crawler can almost craw from any place. In order to increase efficiency, define the webpage, which the crawler needs to seek and extract,is relevant with resource. The crawler is limited to special Hyperlink obtained from resource identification module, and handles data of different types and collects all data resource. In order to increase the efficiency of crawler, the work of crawler is parallel. The key words collected from webpage can be used as crawler's initial seed to start Web crawling.

In order to ensure the resource ontology reference has self-adaptability, based on method of ontology learning in ontology project, and under statistics and principles, calculate adaptive learning of resource ontology, which brings about better screening effect. The ontology learning module based on statistics mainly depends on the feedback of Web environment in the crawling process, and summarizes the feedback data from resource obtained. Then calculate weight learning and inheritance relationship evolution in light of domain ontology frame diagram.

In preprocessing module, with the continuous generation of crawler's analysis data, obtain analysis data of correlation degree of a series of resources. These data can be used as training dataset of ontology statistical learning and include one piece of important learning information: notional word hit rate. Record the frequency of the word and use it for learning of resource weight. From this, continuously learn and evolve reference substance of resource ontology, so as to ensure the crawling can be more scientific and effective.

The linking mar in webpage usually includes the resource characteristic information of the webpage. The text content around linking is the concentrated reflection of resource anchor. If the linking points to resource relevant page, the webpage also have certain resource correlation degree. Therefore, in light of analysis of resource linking relation, obtain another computation dimension of correlation degree of webpage resource, which is the main functions of resource linking analysis prediction module.

At last, use a content filter module to eliminate noise and junk data during data collection process. Then store the screening data in the database of raw data, thus completing the process of preprocessing.

(B) System architecture design

As shown in Figure 2, the entrance utilized by sys-

tem includes learning platform, teacher's teaching platform, and social media platform. Wherever and whenever, no matter what equipment is used, once the equipment is connected to internet, the student can learn by logging in student studying platform, and automatically receive the assessment and credit of learning. One-time logging in can enjoy all System resource; the teacher can log in teacher teaching platform for teaching anytime and anywhere, correct student's homework and assessment paper. Then send achievement and comment to corresponding student automatically. In the platform, it is available to see the student's all comment and learning experience published on forum, message board, SNS, etc. The teacher can appraise score at ordinary time through these contents, and acquire student's learning state at the same time; social media platform is provided for non-registered user, whose functions will be restricted. It mainly helps non-registered user to acquire distance education, application information, and advertisement.

There are two information collecting channels in the system: resource crawler and user submission. According to data preprocessing mechanism in Figure 1, the crawler acquires different series of key words in preprocessing, then crawl to collect information on Web according to key words. Some websites and forum sites of distance education are collected manually and used as the seed of crawler to increase efficiency. For specific search words, the resource crawler can collect comprehensive information about these words, including professional information, course information, course credit, student information, courseware material case, teaching audio and video, work and test questions etc. At the same time, students' activity during various social media platforms produces another type of information concerning students. The type of information can be treated as students' homework and used for score assessment of ordinary time, interest identification, and hot issues of search theme, and discussion hot issues.

Besides the data which are crawled, through user submission platform, the teacher can upload teaching resource and information, and student can submits homework and questions. In user submission platform, it is required to use unified predefined format easily applicable to adaptive ontology, which can transmit courseware, video, article, and comment, etc. The user submits unified module check and adjusts the contents submitted by user, which enables these content formats meet uniform standard and is convenient for the semantic analysis ofadaptiveontology. Then store it in the database of raw data. The system includes high-level and low-level databases. The low-level database stores the original data collected by crawler and submitted by user. The system extracts more specific information in low-level database through adaptiveontology, and stores it in high-level database. The elements of user files such as character, interest, etc. include high-efficient information to identify relevant data, can be used for the screening of document, and are important contents of the raw data database[3]. Adaptive ontology semantic analysis, detect and summarize various information of original data, such as metadata obtained from teaching article, including author name, Email address, contact, etc., teacher's individual address, schools related to teaching group and association, and addresses of distance education schools of various places, etc. Then store it in database of each high level. There are 6 high-level databases, as shown in Figure 2.

1)Professional information database

Includes basic information of various majors of distance education, including talent training scheme, course information of various majors, credits of each course, etc. Due to the particularity of distance education, student is only required to complete specified course, complete graduation thesis, get required credits, then he can graduate. The learning order of course is not required. This is the basic database of system, which only allows authorized teaching manager to modify. The database only undergoes regular update annually.

2) Studentdatabase

As the students of distance education are too many, it is necessary to establish special student database. The database includes students' name, student number, sex, date of birth, major name, class name, contact, address, and also includes learning sore, credits, rewarding, punishment, etc. The base only allows authorized teaching manager to modify. Besides annual upgrade, there may be daily student status change, academic records, such as major, suspension of schooling, credit reward, punishment record, etc.

FULL PAPER

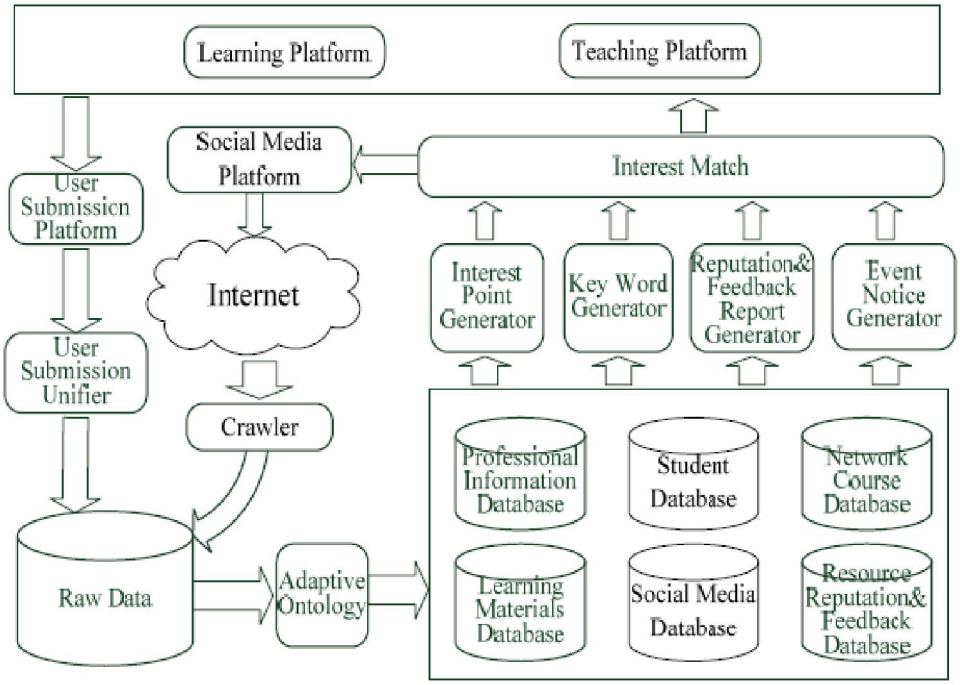

Figure 2 : System Architecture of Distance Education System Based on Web3.0.

3) Network course database

Includes audio, video courseware of various formats. In order to control quality of play and increase fluency of play, it is required to unify formats of audio and video. The audio uniformly adopts MP3 or WMA compressed formats, and video uniformly adopts RMVB or WMV formats.

4) Learning materials database

Includes various case and materials for teaching, and exercises of each course. The exercises adopt standardized format. The objective item is key point and subjective item is second-important, with standard answer. It can be used as daily homework, unit test, and test question.

5) Social media database

The users of social media group can share their resource in system and hold a meeting, discuss online, establish interest group, and calculate project cooperation, etc. After one-time registration, a user can visit system's Blog, Micro-Blog, SNS, and forum service. Store all the following information in the database: relevant articles and comments of social media, relevant discussion of organization and group, especially relevant article and news report collected from various distance education websites, and relevant comments collected from each forum, Blog, Micro-Blog, SNS and other media platform. The database not only stores users' important introduction, and also records users' behaviors.

6) Resource reputation and feedback database:

The database storage acquires resource reputation, user's feedback information, and interest point distribution through many methods such as network analysis and semantic analysis. Collect and store the following information in the database: feedback information collected from original article and association remark, other relevant learning information collected from media platform, and relevant comments and evaluations.

From contents of original data, extract users' interest points, such as research interest description in publisher's individual file, relevant introduction of website publisher, articles published by user's group, resource detail, resource relevant report, introduction, etc. The system generates statistical report of interest point through interest point generator. Manager will trace hot-spot resource according to these reports and timely adjust hot issues, and guide to pay attention to hot-spot resources. Through users' interest information, it is also available to extract their interested major field and research direction. The information is provided for intelligent service and also contributes to helping system to

constructing ontology.

The system generates appropriate key words through key word generator. These key words can be offered to users for search, so as to find out the resource quickly.

Adopt semantic method to generate the reputation feedback and summary report of distance education resource, and help resource developer to formulate development plan and development direction. After generating the report, match users' interest through user ontology interest matching module, and generate matching report and submit it to corresponding user platform. The teacher or student will get report from the platform. The system will also send report to appropriate social media platform to help users' to obtain latest information.

The function of event reminder generator is to remind people of important learning event. It lists latest various teaching meeting, video live broadcast, and other types of activities, including main theme, speaker, and time section, etc. The user can select interested event to participate in activities. For example, participate in video meeting, on-line discussion, etc.

CONCLUSION

In light of current situations and difficulties of China's distance education, the paper puts forward distance education system based on Web3.0 technology, so as to provide more powerful resource share, intelligent, and individualized abilities. The paper proposes preprocessing work mechanism and system Architecture based on distance education system, and describes System work flow and main functions. The future research emphasis is specific construction and realization method of platform.

REFERENCES

[1] Xiaogang Liu, Hongli Xu; A Study on Cloud Computing and Its Application to Distance Education System, Modern Distance Education, **131**, 64-67 Oct. **(2010)**.

[2] Wei Huang, Liyi Zhang; Research on Focused Merchandise Information Crawling Based on Semantic Crawler, New Technology of Library and Information Service, **188**, 3-8 Jan.**(2010)**.

[3] Guanpi Lai, Qingpeng Zhang, Ding Wen; A Prototype of the Next-Generation Journal System for ITS: Academic Social Networking and Media Based on Web 3.0, IEEE Trans. Intell. Transp. Syst., **13(3)**, 1078-1087 Sep. **(2012)**.