Research on Big Data Professional Curriculum Construction Based on OBE Education Concept——Take Hadoop System Application as an example

Huatao Zhou[1,a], Fang Tu[1,b], Ju Chen[1,c], Jinfang Weng[2,*]

Zhouhuatao@163.com[a], 89722772@qq.com[b], 646027719@qq.com[c], 382935224@qq.com[*]

College of Information Engineering, Wuchang Institute of Technology, No. 110, Baishazhou Avenue, Hongshan District, Wuhan City, Hubei Province 430065, China[1]
Fujian Polytechnic of Information Technology, No. 106, Fufei South Road, Gulou District, Fuzhou, Fujian 350003, China[2]

Abstract: Today global economic and social big data is widely and deeply applied. The big data professional courses in colleges and universities are new in theory, strong in practice, and widely applied. The task of course construction is arduous and imminent. This paper takes Hadoop System Application as an example, takes OBE education concept as the guiding ideology, takes students' graduation requirements as the result orientation, takes students as the center, and uses online and offline hybrid teaching as the teaching means Some innovations have been made in the teaching mode and assessment method, and the next research direction of strengthening the construction of teaching resources and optimizing the construction of experimental environment has been proposed. Some useful exploration and attempts have been made in the course construction goals, teaching design, research ideas, course content construction and other aspects, hoping to have some enlightenment and help to the big data professional construction courses of similar universities.

Keywords: OBE; Results oriented; Big data; Curriculum construction; Hadoop

1 Introduction

OBE education concept refers to outcome based education (OBE for short, also known as ability oriented education, goal oriented education or demand oriented education). As an advanced education concept, it was put forward by Spady et al. in 1981, and soon gained people's attention and recognition, and has become the mainstream concept of education reform in the United States, Britain, Canada and other countries. The American Engineering Education Accreditation Association (ABET) has fully accepted the concept of OBE and put it through the engineering education accreditation standards [1]. In June 2013, China joined the Washington Agreement and became a signatory of the agreement, which marks the beginning of the certification of engineering education with international substantive equivalence in China. The professional certification of engineering education follows three basic concepts: achievement oriented, student-centered, and continuous improvement. These ideas are crucial to guide and promote specialty construction and teaching reform, and ensure and improve the

ICMEIM 2023, September 08-10, Wuhan, People's Republic of China

DOI 10.4108/eai.8-9-2023.2340078

quality of engineering education personnel training [2]. Achievement oriented education has become the mainstream concept of education reform in the United States, Britain, Canada and other countries, and has been fully adopted by the professional certification of engineering education. It is of practical significance to guide the reform of engineering education in China with the concept of achievement oriented education [3].

With the need of national development strategy, Hubei regional economic construction, big data industry development and the construction and development of our disciplines and specialties, the country and society are in urgent need of big data technology talents [4], so our school has added Hadoop System Application as the core professional course in the 2019 talent training program, with 3 credits, 48 class hours (32 theoretical class hours, 16 experimental class hours), and 4 weekly class hours, The computer science and technology major (big data technology direction) is scheduled to teach in the fifth semester of the third academic year, and the computer science and technology major (Sinosoft International Order Class) is scheduled to teach in the sixth semester of the third academic year. In September 2021, this course began its first teaching for 2019 computer science and technology majors (big data technology direction), and combined with the superstar learning platform for online and offline hybrid teaching. The teaching effect was good, and the students were unanimously praised.

2 Significance of course construction

2.1 Theoretical significance

This research project is to reconstruct and upgrade the course construction of Hadoop System Application under the guidance of OBE education concept and the standard of engineering education certification. As an application-oriented undergraduate university focusing on engineering, our in-depth research and practical exploration of this project is conducive to recommending engineering education reform, improving the quality of talent training, establishing an engineering education certification system connected with the registered engineer system [5], building a mechanism for linking engineering education with the business community, enhancing the adaptability of talent training to industrial development, and playing a demonstration role in the professional construction of the school, Comprehensively improve the level of professional construction, promote mutual recognition of domestic and international specialties, and enhance international competitiveness [6].

2.2 Practical significance

1. Renew the concept of engineering education. Reconstruct the open professional education system for students, teachers, industries and enterprises, and guide the curriculum construction and talent training of Hadoop System Application according to the student ability oriented education ideology [7].

2. Build a new teaching model. From teacher centered teaching to student-centered learning, all teaching work is student-centered, and teachers guide and help [8]; Transform ability from knowledge system as the center to achieve the goal, carry out teaching design and teaching

implementation with the goal of improving students' practical ability, and enable students to have the ability to implement big data engineering applications when they graduate [9].

3. Form a quality assurance mechanism. The quality standard of engineering education is the core of certification and the basis of mutual recognition of academic qualifications. It needs the support of the comprehensive teaching work of the school. At the same time, it can form an internal and external quality management system mechanism in the school, and establish a management mechanism and quality culture of continuous improvement [10].

3 Course construction objectives

The school running orientation of the university is an application-oriented undergraduate university. The training goal of the university is to cultivate application-oriented talents with reasonable knowledge structure, strong practical ability and high comprehensive quality [11]. Hadoop System Application, as the core professional course of computer science and technology in the talent training program, is the basic course of big data technology application ability. Through the study of this course, the following goals should be achieved:

3.1 Knowledge objectives

1. Master basic theories such as concepts and principles related to big data;

2. Understand and master various tools of Hadoop ecology and their relationships.

3.2 Capability objectives

1. Master the installation, configuration and deployment methods of Hadoop cluster, and be able to quickly build a big data processing platform;

2. Be familiar with the use of HDFS, MapReduce, Zookeeper, Hive and other tools.

3.3 Literacy objectives

1. Abide by academic norms, intellectual property rights, respect privacy, and have information ethics;

2. Establish information security awareness and manage network virtual identity;

3. Have inquiry and critical thinking on information, and develop the habit of lifelong learning.

4 Teaching design and research ideas

4.1 Instructional Design

This course adopts the online and offline hybrid teaching mode, with 16 hours online and 32 hours offline, 48 hours in total, and the online and offline courses alternate. The online classroom mainly realizes the knowledge goal. Teachers establish more than ten learning task lists according to the teaching content, and students learn teaching videos and course materials according to the task list. After that, they complete classroom assignments, question

discussions and questionnaires, and teachers answer questions online and have interactive discussions [12]. The offline classroom is mainly designed to achieve the ability goal from three stages: before class, during class and after class, and guide students to complete online learning before class; In the class, it mainly includes practical operation, task inquiry, student interaction, teacher comments and other activities; After class, students complete homework exercises, note presentation, and teachers answer questions about the course [13].

The online class grasps theoretical knowledge, the offline class cultivates operational application ability and information moral awareness, and the online and offline combination will ultimately achieve the goal of improving students' Hadoop system application ability.

4.2 Research ideas

Guided by the OBE education concept, guided by learning achievements, guided by teachers, student-centered, using case teaching methods and flipped classroom teaching methods, combined with the latest information technology means, supported by Superstar learning platform, and with teaching videos, special videos, cases, exercises, test papers, etc. as teaching resources, students are guided to learn, think and practice independently, Analyze and solve problems independently to achieve the expected teaching effect and teaching objectives [14], as shown in Figure 1.

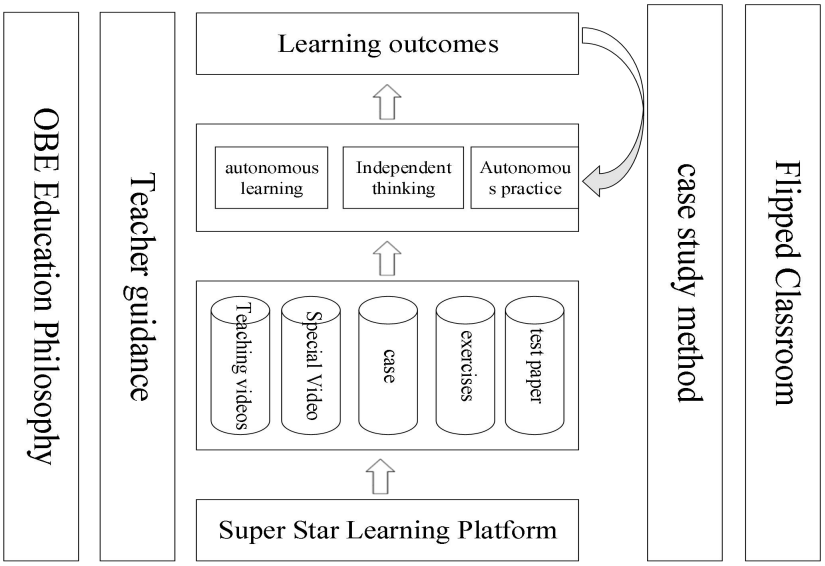

Figure 1 Research ideas

5 Course content construction

5.1 Teaching content construction

The teaching content of the course is divided into Hadoop overview, building Hadoop clusters, HDFS distributed file system, MapReduce distributed computing system, Zookeeper

distributed coordination service, new features of Hadoop2.0, Hive data warehouse. Hadoop overview is the theoretical basis, building Hadoop clusters is the practical basis, HDFS and MapReduce are the two core contents of Hadoop, HDFS is used for data storage, MapReduce is used for data calculation, Zookeeper is the foundation of Hadoop 2.0, which ensures the high availability of Hadoop clusters. Hadoop 2.0 is an upgrade and improvement of Hadoop 1.0, which is now a stable version of Hadoop. Hive is a data warehouse tool built on Hadoop high availability clusters, which can easily call MapReduce in Hadoop clusters for distributed computing using HQL language, providing convenience for data warehouse administrators to operate and improving work efficiency. The teaching content of the course, from theory to practice, from simple to deep, from the outside to the inside, from practical basis to tool application, is suitable for guiding students to gradually understand knowledge, master application, deepen understanding, and integrate knowledge with practice.

5.2 Organization and implementation

The curriculum organization is student-centered and guided by teachers. It actively uses information-based teaching methods to explore online and offline hybrid teaching mode. It uses Superstar's "one level and three ends" learning platform to release curriculum materials, manage classroom activities, and arrange a total of 16 class hours of teaching content of basic concepts easy to understand and master in the curriculum core standards for students to learn online independently, The problems encountered in autonomous learning were recorded, and the teachers were consulted and discussed in the offline learning class. The teachers arranged 32 hours of teaching content, including key and difficult teaching content and experimental teaching content, for offline theoretical teaching and experimental teaching. The online and offline hybrid teaching played the subjective initiative of students, and reformed "I want to learn" to "I want to learn", which achieved good teaching results.

Before class, students preview the 5-10 minute micro video corresponding to the knowledge points of each class to be familiar with the concepts and basic theoretical knowledge of each chapter of the course. In class, through learning through classroom interactive teaching, we can visualize the learning situation of each student, better realize targeted teaching, and make the teaching process more transparent and accurate. In the classroom, teachers sort out the knowledge points, focus on explaining the problems found in students' preview, draw inferences from one instance, systematize the content of knowledge units, and let students master the key and difficult points of each lesson. Teachers design teaching activities, organize classroom discussions, selection of people, preemptive answers, classroom exercises, unit tests and other teaching links. In the experimental link, teachers guide students to experiment, think independently, analyze and solve problems independently, and teachers demonstrate and guide them face to face. It can also improve the efficiency of classroom teaching by flipping the classroom, enhancing the interaction between teachers and students, students and students, giving full play to the active role of students, and cultivating students' thinking ability. After class, students can use computers or mobile phones to study on the learning platform, continue to watch teaching videos to consolidate knowledge, and summarize the contents of each class, as shown in Figure 2.

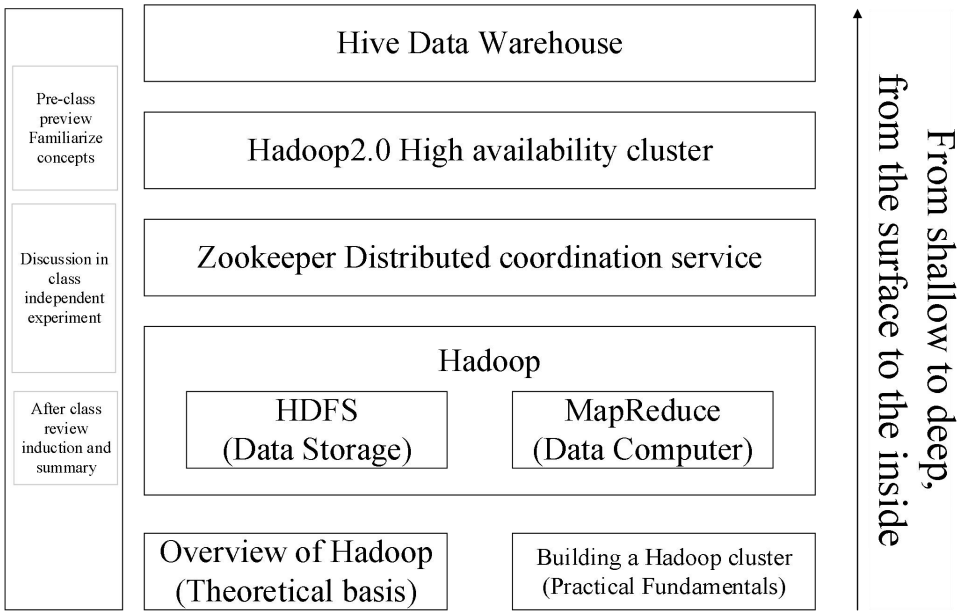

Figure 2 Teaching content and organization implementation

5.3 Course resource construction

The textbook is Hadoop Big Data Technology Principles and Applications (2nd Edition), written by the black horse programmer and published by Tsinghua University Press in July 2023. As an introductory textbook for Hadoop big data technology, this textbook aims to cultivate students' ability to develop big data, and simplifies some complex and difficult ideas and problems so that beginners can easily understand and quickly master them. This textbook has carried out in-depth analysis on each knowledge point, and carefully designed relevant cases for each knowledge point, and then simulated the application of these knowledge points in the actual work, so that the explanation of knowledge can really go from simple to deep, from easy to difficult. The core curriculum standards, overall teaching design, teaching schedule, teaching plans, courseware and other curriculum resources of this course have been completed in combination with the construction of this textbook. Online resource construction is based on the superstar learning platform （ https://mooc1.chaoxing.com/course-ans/ps/222795535 ）. First, the course contents are arranged according to the course standards, and then the teaching courseware and video (the teaching video of this textbook in NetEase open class) are uploaded to the corresponding chapters of each knowledge point. Students can preview and review before and after class in combination with the textbooks, teaching materials and teaching videos, and add a "chapter test" link after each chapter to test the effect of students' online autonomous learning, Finally, upload the core curriculum standards, overall teaching design, teaching schedule, teaching plans, courseware and installation software to the materials on the Superstar Learning Connect platform for students to use online autonomous learning, and improve the efficiency of students' online autonomous learning.

5.4 Course performance evaluation method

The assessment of this course adopts a combination of process assessment and final assessment. The process assessment accounts for 60% of the total score, and the final assessment accounts for 40% of the total score.

The process assessment focuses on assessing students' mastery of big data knowledge and algorithm design and professional quality in the learning process. The process assessment score of this course is recorded and exported through the Chaoxing Fanya platform system. The usual score consists of eight parts: sign in, homework, course audio and video, chapter test, resource access, discussion, classroom interaction, and experiment report.

The final assessment focuses on the assessment of students' knowledge mastery and application, including the concept and characteristics of big data, the system architecture, working principle, and operating mechanism of Hadoop, HDFS, MapReduce, Zookeeper, and Hive. The final assessment of this course is conducted by means of examination papers, with a full score of 100 points.

Assessment content requirements: the examination paper method shall be used for assessment, and the assessment content shall cover 90% of the knowledge points in the outline. The system architecture and working principle of Hadoop, HDFS, Hive, etc. pay attention to the knowledge understanding questions. Hive table operation, HDFS file reading and writing, MapReduce distributed computing, etc. pay attention to comprehensive application assessment. It can not only test students' mastery of basic concepts and knowledge of data structure, but also exercise their understanding of Hadoop system application, cultivate scientific thinking and engineering design methods, and have good scientific literacy and engineering awareness.

6 Curriculum construction innovation

This course is characterized by new content, high starting point and comprehensive requirements. The new content means that the Hadoop big data technology in this course has only been available for more than 16 years since 2006, and the relevant theories and practices are a new field for teachers and students. High starting point means that Hadoop system application developers must have basic computer knowledge, Linux foundation and object-oriented programming (Java) design ideas, which are not competent for ordinary programmers. The requirements all refer to the higher requirements for teachers and students at the same time. We should keep pace with the times in today's big data era and not follow the beaten track. We should master the download, installation, configuration, deployment and testing skills of Hadoop system, as well as the ability of Java API programming, so as to cultivate more and better application-oriented undergraduate talents.

6.1 Innovation of experimental environment.

Since this course requires students to master the installation, configuration, deployment and testing of Hadoop clusters under the Linux operating system, it has high requirements for computer configuration, especially memory configuration, and the combination of theory and

practice is very close. The existing laboratory environment cannot meet the above requirements. Therefore, this course relies on students' laptops for teaching in the whole computer room, no longer distinguishing between theoretical courses and experimental courses. Students carry their laptops into the classroom to carry out experiments. Teachers teach while doing, and students listen while doing. When students encounter problems in class, they can ask teachers to help solve problems at any time, effectively solving the complex problems of the experimental environment. After the pilot project in the class of 2002 of Planning (Big Data) last semester, The effect is very good, and has been praised by all the students in the class.

6.2 Teaching mode innovation.

On the basis of the innovation of the experimental environment, theoretical courses and experimental courses are no longer distinguished. Guided by OBE's advanced educational philosophy, online and offline integrated teaching mode is adopted. Teachers provide micro lesson videos as the "online" resources for students to learn. Usually, after students complete online knowledge point learning and test questions, teachers will guide students to further master knowledge points offline by focusing on the "flipped classroom" mode [15], and taking students as the center The knowledge used in project design and implementation will improve project development skills and promote students' independent thinking and practical ability.

6.3 Innovative assessment methods.

This course focuses on the cultivation of practical ability according to the assessment of "theory/experiment" integration and the combination of "online/offline" integration of teaching, emphasizing the process assessment. The total score of the course is composed of process assessment (60%) and final assessment (40%).

The process assessment focuses on assessing students' mastery of big data knowledge and algorithm design and professional quality in the learning process. The process assessment score of this course is recorded and exported through the Chaoxing Fanya platform system. The usual score consists of eight parts: sign in, homework, course audio and video, chapter test, resource access, discussion, classroom interaction, and experiment report.

The final assessment focuses on the assessment of students' knowledge mastery and application, including the concept and characteristics of big data, the system architecture, working principle, and operating mechanism of Hadoop, HDFS, Zookeeper, and Hive. The final assessment of this course is conducted through examination papers.

7 Next research direction

7.1 Strengthen the construction of teaching resources

Both theory and practice should be considered in the selection of course materials. Neither theory alone nor explanation of the operation process should be allowed, so that students can know more about it. Due to the late opening of big data courses, there are not many mature supporting resources, so it is necessary to collect a large amount of data, build and sort out characteristic digital teaching resources, such as teaching courseware, pre class guidance

micro tasks, knowledge point micro videos, operation demonstration videos, cases, exercise banks, discussion topics, questionnaires, expansion resources and other diversified digital resources [16]. Big data is closely linked with current events, providing more expansion papers, current events videos, news links, improving students' interest in learning, and letting students know how to apply what they have learned. Organize these resources according to teaching needs and processes, and provide them to students step by step. Students can study repeatedly before and after class according to their own conditions until they master.

7.2 Optimize the construction of experimental environment

Computer experiment is an important part of course teaching. The students have mastered the theoretical knowledge, further strengthen the memory of relevant knowledge points through computer operation, verify the knowledge learned, integrate theory with practice, and achieve the purpose of improving students' practical ability. The big data experiment faces many challenges. The construction of the big data experiment environment is complex, and students have limited time in the computer room. If one step is not completed, the next step may be delayed. Moreover, if you need to do experiments after class and carry out computer practice on your own computer, first, there are certain requirements for students' computer configuration; second, if the experimental environment is damaged, it may be difficult for you to repair it. In order to meet the diverse learning needs of students, the diversified practice environment is built by building laboratory environment and online experimental platform.

8 Conclusions

Carry out the construction of big data professional curriculum system based on the requirements of students' graduation ability and reverse thinking, carry out the teaching design with students as the center, give full play to the students' subjective initiative and enthusiasm, combine the content and key points of Hadoop System Application, try and explore the online and offline hybrid teaching mode of big data professional curriculum based on OBE education concept, and cultivate new engineering application-oriented undergraduate talents, Provide construction ideas and design models for other professional courses of big data, and make certain contributions and efforts for the construction of big data specialty in colleges and universities.

Fund project. The guiding project of philosophy and Social Sciences Research of Hubei Provincial Department of education in 2021, "Research on classroom teaching evaluation based on artificial intelligence" (21G155), and the project "Research on key technologies and applications of face recognition" (2021BSJ02) supported by the doctoral research start fund of Wuchang Institute of technology, Research Achievements of the Collaborative Education Project of the Ministry of Education on Industry and Education "Course Construction of Hadoop System Application Based on Engineering Education Professional Certification"

References

[1] Yang, Z. (2023) Exploration on the Construction of Network Ideological and Political Courses for Food Specialty in the Age of Big Data. Food and Machinery, 39 (03): 251

[2] Qian L.F., Ma J., Mi C.M. (2023) Research on the construction of big data courses for information management and information system majors of new liberal arts. Information Theory and Practice, 46 (03): 83-89. DOI: 10.16353/j.cnki.1000-7490.2023.03.012
[3] Ge X.Q., Yang Y.F. (2022) On the construction of online open courses for social students in higher vocational education under the background of enrollment expansion based on the OBE concept. Vocational and Technical Education, 43 (26): 38-43
[4] Zhang M., Wang Y.P., Li A. (2022) Integrated curriculum design in the context of new liberal arts construction -- taking the "big data audit" course as an example. Monthly Journal of Finance and Accounting, (15): 104-109. DOI: 10.19641/j.cnki.42-1290/f.2022.15.014
[5] Tanner L.R.,Collins E.L. (2021) Using Publicly Available Long-Term Climate Records in Undergraduate Interdisciplinary Big Data Curriculum. Journal of College Science Teaching,51(2)
[6] Wang L.L., He Y., Yang J. (2021) Research on online and offline hybrid teaching reform and innovation of college ecological environment courses in the context of big data -- Review of Introduction to Ecological Civilization Construction in the New Era. Education Development Research, 41 (20): 85. DOI: 10.14121/j.cnki.1008-3855.2021.20.013
[7] Cui X.L., Zhang M., Zhang L. (2021) Research and construction of applied big data talent training curriculum cluster under the background of new engineering. Experimental Technology and Management,38 (02): 213-218. DOI: 10.16791/j.cnki.sjg.2021.02.046
[8] S. J.G., Wilfredo A.,F. E.C. (2021) A Data Analytics/Big Data Framework for Advanced Metering Infrastructure Data. Sensors,21(16).
[9] Tang Y.Q. (2020) Discussion on the construction of information courses for criminal science and technology majors in the era of big data -- Taking the People's Public Security University of China as an example. Laboratory Research and Exploration,39 (10): 155-158
[10] Wu H.J., Rao Y.H. (2019) Big Data Curriculum Construction for New Engineering. China University Teaching, (04): 34-37
[11] Wang G.R., Jin F.S., Liu C. (2018) Construction of the professional curriculum system of data science and big data for internationalization. China University Teaching, (12): 43-45+51
[12] Huang Y.M. (2018) Construction of undergraduate curriculum system for information management in the era of big data: based on the needs of college students. Library and Information Work,62 (03): 45-54. DOI: 10.13266/j.issn.0252-3116.2018.03.006
[13] He W.W., Liu G.M. (2017) Exploration and Research on the Construction of Core Curriculum of Data Science and Big Data Technology. Education Review, (11): 31-35
[14] Gao T. (2017) Research on the construction of sports curriculum teaching case base from the perspective of big data thinking: take Capital University of Physical Education as an example. Journal of Capital University of Physical Education,29(06): 514-516. DOI: 10.14036/j.cnki.cn11-4513.2017.06.009
[15] Wang X.W., Guo Y., Shi J. (2015) A comparative study on the construction of the curriculum system of postgraduates majoring in library and information between China and the United States in the era of big data. Library and Information Work, 59 (23): 30-37. DOI: 10.13266/j.issn.0252-3116.201523.005
[16] Zhao J. (2015) Course resource construction based on big data: trend, value and direction. Course. Textbook. Teaching Method,35 (04): 18-23. DOI: 10.19877/j.cnki.kcjcjf.2015.04.004