Journal of Contemporary Educational Research, 2024, Volume 8, Issue 2
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

Exploration and Practice of Big Data Introductory Courses for Big Data Management and Application Majors

Tinghui Huang [1], Junchao Dong [2], Liang Min [3]*

[1]School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, Guangxi Zhuang Autonomous Region, China
[2]School of Computer Engineering, Guilin University of Electronic Technology, Beihai 536000, Guangxi Zhuang Autonomous Region, China
[3]The Youth Innovation Team of Shaanxi Universities, Xi'an Jiaotong University City College, Xi'an 710018, Shaanxi Province, China

***Corresponding author:** Liang Min, minliang@xjtucc.edu.cn

Abstract: As an introductory course for the emerging major of big data management and application, "Introduction to Big Data" has not yet formed a curriculum standard and implementation plan that is widely accepted and used by everyone. To this end, we discuss some of our explorations and attempts in the construction and teaching process of big data courses for the major of big data management and application from the perspective of course planning, course implementation, and course summary. After interviews with students and feedback from questionnaires, students are highly satisfied with some of the teaching measures and programs currently adopted.

Keywords: Big data management and application; "Introduction to Big Data"; Teaching reform; Curriculum exploration

Online publication: February 25, 2024

1. Introduction

At present, China has entered the era of digital intelligence, data has become a novel production factor of new quality productivity, and big data has become the backbone of driving Chinese modernization [1]. In recent years, our country has introduced a series of policies, charting the course for the healthy development of big data in China. On December 28, 2021, China officially released the "14th Five-Year Plan National Informatization Plan." Through the search, it was found that the keyword "big data" appeared as many as 54 times. In December 2022, the State Council issued the "Twenty Data Articles," proposing basic data systems in many aspects such as data property rights, circulation transactions, income distribution, and security governance, which is of milestone significance in promoting data sharing and activating data potential. In December 2023, the State Council issued the "Overall Layout Plan for the Construction of Digital China" to the whole country [2]. The plan emphasized the importance of building Digital China and clarified the overall

construction and development framework. It will promote China's digitalization process and improve national governance capabilities and governance levels. It is of great strategic significance to further promote Chinese-style modernization and other aspects.

Big data management and application major is one of the few undergraduate majors directly related to big data in China [3]. In order to respond to the national big data strategy, cultivate big data talents, and make early arrangements in the fields of big data and artificial intelligence to seize opportunities, China established a professional formation group for big data and management and application in 2015. After two years of investigation and research, the major was officially established in 2017. In the same year, five universities including Xi'an Jiaotong University were selected as the first batch of pilot universities. Since then, the emerging major of big data management and application has been welcomed by most universities, and students are actively applying for this emerging major. As of 2023, the number of colleges and universities across the country that have established undergraduate majors in big data management and application has reached 221. Guilin University of Electronic Technology (abbreviated as Guidian), a local engineering college with distinctive electronic information characteristics in Guangxi Zhuang Autonomous Region, actively applied for big data in 2019 by virtue of its experience in the construction of information management and information systems majors and its interdisciplinary advantages in the field of electronic information management and application major. Guidian was successfully approved for this major in 2020 and officially enrolled students in 2021. It has become the first local university in Guangxi as well as in the country to offer this major. The major of big data management and application is to cultivate big data talents with comprehensive development of skills and literacy, with a foundation in various disciplines such as economics and management, computer science and mathematics, and who can manage and analyze data. It also aims to cultivate qualified data analysis, big data development, data mining, and other engineers for China in many fields such as finance and media.

"Introduction to Big Data" is the first professional required course for big data management and application majors with freshmen as the teaching target. Since the big data management and application major is an emerging major that was established not long ago, a widely accepted curriculum system and standards have not yet been formed. At present, Guidian Big Data Management and Application major has enrolled students three times, and the "Introduction to Big Data" course has also been offered three times. This article mainly introduces the preliminary exploration and practical effects of the big data introductory course for big data management and application majors.

2. Status quo of big data introductory courses

In order to fully understand the current situation of big data introductory courses and better teach the courses, the course objectives, course plans, and teaching syllabus of the current big data introductory course were sorted out and studied through various means such as literature review, online questionnaires, and on-site interviews and discussions. It was found that there are several problems in the current big data introductory course:

(1) Inadequate focus on the introductory course on big data: Since big data is a cutting-edge subject and the big data management and application major is an emerging major, teachers and teaching resources are relatively scarce compared to traditional majors. Therefore, some universities choose not to offer the big data introductory course in the big data management and application major or teach introductory computer content under the introductory course.

(2) Single teaching methods: At present, introductory courses on big data mostly use teachers' PowerPoint explanations and students' passive acceptance. Students lack interest in the course and the teaching effect is poor.

(3) Inappropriate curriculum: The course focuses too much on the explanation of theoretical knowledge and lacks the cultivation of students' big data awareness and thinking. It cannot play the essential roles that this course should play. After taking such a course, students will not be able to have a complete and macro understanding of the entire ecology of big data.

(4) Unscientific teaching evaluation method: The assessment of introductory courses in most schools mainly adopts the closed-book assessment model of "one paper at the end of the semester." Before the exam, students will memorize the learning content in order to cope with the exam, turning a course that should focus on literacy and quality cultivation into a course where students can get higher marks purely by relying on memory. In addition, a small number of colleges and universities have also joined the process assessment template, but their process assessment scores account for a low proportion and the existing process assessment scores mainly consist of three parts: attendance, homework, and course performance, which essentially does not change the emphasis on knowledge over competency assessment.

In response to the four existing problems in the big data introductory course, Guidian has made some new attempts and practices when offering such a course. This article will introduce specific methods from the three main links of course preparation, course implementation, and course summary.

3. Preparation for the "Introduction to Big Data" course

The course preparation stage mainly includes course syllabus formulation, course plan writing, and teaching material selection.

The syllabus is a programmatic document for teaching implementation. It consists of several professional parts: course introduction, course objectives, teaching content and requirements, teaching strategies and concepts, and assessment and evaluation methods. A good syllabus should have clear objectives, focused content, and an established and complete structure. New data literacy goals and requirements were added when writing the syllabus. Data literacy is the ability to understand, manage, process, analyze, and reflect on data, covering the entire life cycle of the big data ecosystem.

The "Introduction to Big Data" course is a professional basic course with 32 credit hours arranged in the first semester of freshman year. When formulating the course plan, only four classes are arranged in a week so that students can digest the knowledge explained in class and explore the mysteries of big data independently. The specific course plan is shown in **Table 1**.

Table 1. "Introduction to Big Data" course plan

Chapter number	Course content	Class schedule
1	Big data basics	2
2	Data collection and preprocessing	5
3	Data storage and management	6
4	Data processing and analysis	6
5	Data visualization	4
6	Dig data applications	4
7	Big data frontier lecture	2
8	Visit, exchange, and learn	2

The main material of the big data introductory course is "Introduction to Big Data" written by Professor Ziyu Lin of Xiamen University. Professor Ziyu Lin is one of the earliest teachers in China to compile a series of big data textbooks. He has written multiple textbooks in the field of big data and established a public service platform for big data teaching. The big data introductory textbook written by Professor Ziyu Lin has been selected by more than 300 universities in China. This textbook covers all dimensions required in data literacy, with particular emphasis on cultivating students' data thinking and checking data ethics, and is suitable for the current actual needs of cultivating big data composite talents. In addition, Professor Lin has developed an online development course corresponding to this textbook. The course quality is high and it was approved as a national-level quality course in 2018, which can meet the needs of students for self-study and course preview after class.

4. Implementation of big data introductory course

"Introduction to Big Data" is a comprehensive introductory course covering the concepts, characteristics, technologies, applications, and other aspects of big data. However, the "Introduction to Big Data" course only has 32 credit hours, so the selection of teaching methods must be able to achieve classroom efficiency. In the implementation of the big data introductory course, we mainly organize teaching based on data literacy theory and OBE (outcome-based education) concepts.

An efficient classroom refers to a classroom in which teachers can make full use of classroom time, effectively impart knowledge, stimulate students' interests and thinking, cultivate students' abilities and literacy, and achieve expected teaching goals [4]. To organize and implement efficient classrooms, we should begin from four aspects. Firstly, teachers have clear teaching goals and plans, and can flexibly adjust teaching strategies and methods according to students' actual situations and needs, so that classroom teaching can proceed in an orderly manner. Secondly, teachers have rich teaching content and forms and can apply a variety of teaching methods and media to make classroom teaching content rich and colorful, with diverse teaching forms to attract students' attention and participation. Thirdly, teachers have efficient teaching processes and methods and can adopt various teaching models such as heuristic, inquiry, and cooperative to make the classroom teaching process logical, coherent, hierarchical, and scientific, and effectively promote students' cognition, emotional, and behavioral development. Fourthly, teachers have good teaching evaluation and feedback, can evaluate and feedback on students' learning situations and effects in a timely manner, guide students to conduct self-evaluation and reflection, encourage students to carry out independent learning and innovative practices, and improve students' self-confidence and self-esteem.

Data literacy refers to people's ability to recognize, understand, and apply data in the information society. Data literacy theory includes data acquisition, processing, analysis, evaluation, application, and dissemination [5]. Data literacy theory is a comprehensive ability that not only requires students to master basic data-related knowledge and skills, but also requires data thinking, data values, and data ethics. Data literacy theory is an important part of information literacy theory and one of the basic qualities of citizens in modern society.

The OBE educational philosophy emphasizes an educational process that takes students' abilities and qualities as the core, students' learning outcomes as the guide, students' personality development as the goal, and students' lifelong learning as the basis [6]. The OBE concept believes that the essence of education is to cultivate learners with the knowledge, skills, attitudes, and values required for future social life and work. The OBE concept requires teachers to design appropriate teaching objectives, content, methods, evaluation, and feedback based on the characteristics and needs of learners to improve teaching effects and quality. The OBE concept also requires learners to actively participate in learning activities, independently explore problems, collaborate and exchange experiences, reflect on and evaluate their own learning processes and results, and

constantly improve their abilities and qualities.

Data literacy theory and the OBE concept are two important aspects of the education reform of big data management and application majors, both emphasizing students' skills cultivation and lifelong learning awareness. Data literacy theory believes that data is a basic cultural literacy and an essential ability for people to participate in social life, solve problems, and create value in the information society. The OBE concept is an education model that is learner-centered, oriented by learning outcomes, supported by process evaluation, and driven by continuous improvement. Data literacy theory and OBE concepts complement each other and can provide new ideas and methods for education.

"Introduction to Big Data" is a course that introduces the basic concepts, technologies, and applications of big data. It plays an important role in cultivating students' data thinking and analytical skills. In order to implement efficient classrooms, we carry out reforms in the following four aspects.

Firstly, multimedia and network resources are utilized to make the classroom more interactive and interesting. For example, videos, animations, charts, and other forms can be used to display the principles and cases of big data to stimulate students' interest and curiosity; the Qingruan U+ online learning platform can be used to design interesting small tests, questionnaires, discussions, and other activities to allow students to participate in the classroom and improve learning initiative and enthusiasm.

Combined with practical problems and cases, students are guided to explore the value and significance of big data. For example, some social network analysis, e-commerce recommendation systems, smart cities, and other issues and cases related to student life or majors are selected to let students understand the application and impact of big data in different fields and cultivate students' problem awareness and innovative thinking.

In addition, students are encouraged to join the visit to large-scale engineering workshops on campus or comprehensive off-campus internship bases to achieve the purpose of applying what they have learned. For example, students are arranged to visit or participate in some actual big data analysis projects, allowing students to personally experience the analysis process and methods of big data, exercise students' programming and logical thinking skills, encourage students to communicate and share their ideas and experiences in groups, and promote mutual learning and cooperation among students.

In order to make the course organization more efficient, we set a class duration of 90 minutes and designed a course organization model of "course introduction, course explanation, case analysis, class test and discussion, course summary and outlook, and after-class development." Designing an effective big data introductory course requires taking into account the overall structure of the course, teaching methods, and student engagement. Specific teaching time arrangements will be adjusted based on student feedback and progress. Throughout the entire process, maintaining interaction with students and guiding students to proactively think and participate will help improve the effectiveness of the course.

In the introduction of the course, we first introduce the theme and objectives of today's course to stimulate students' interest and let them understand the importance of learning this topic. Teachers explain the relevant knowledge of this course mainly through classroom explanation and share some practical industry applications and scientific research projects using case analysis. Teachers use big data application cases to emphasize the big data technology used in the case and the results achieved, allowing students to better understand the connection between theory and practice. Class tests and discussions Additionally, teachers distribute exercises or design small practical tasks to allow students to apply the knowledge they have learned to solve problems. Students are supervised as they complete exercises in class and immediate feedback is provided. Group discussions or question-and-answer sessions are conducted to allow students to participate in discussions and share their understanding. Teachers also answer the questions raised by students and resolve possible doubts. The summary and outlook section mainly summarizes the key contents of this lesson and emphasizes key concepts. Teachers

can briefly preview the content that will be covered in the next lesson to stimulate students' interest in future learning. Lastly, some additional reading materials, online resources, or practical projects are recommended in the after-class session to help students who are interested in in-depth learning further carry out independent learning activities.

Curriculum ideological and political education refers to integrating the concepts and requirements of ideological and political education into the course design and teaching implementation process and guiding them to establish a correct outlook on life, values, and worldview by cultivating students' ideological awareness and political literacy. The ideological and political course is an overall thought that aims to enable students to better understand social phenomena and social systems while learning subject knowledge, and to cultivate a sense of social responsibility and patriotic sentiment. **Table 2** shows the ideological and political elements of the courses related to the introduction to big data that we mined.

Table 2. Ideological and political elements of big data introduction

Ideological and political elements	Explanation
Data privacy and ethical issues	Guiding students to think about privacy and ethical issues faced in the era of big data, and discussing ethical considerations in data abuse, personal privacy protection, etc. Analyzing the regulations and policies on data privacy in different countries and regions, and emphasizing the ethical responsibilities that big data practitioners should have for data use.
Social equity and data bias	Discussing the possible bias and inequality issues of big data in social decision-making, and guiding students to think about how data affects social justice and equality. Analyzing potential biases in algorithms and discussing how to address them through technical means and institutional reforms.
Information transparency and public participation	Guiding students to think about the role of big data in government decision-making and social management, promote information transparency, and encourage public participation in social affairs. Analyzing the application of big data technology in the public domain, allowing students to understand the importance of data to people's livelihood and social governance.
National security and technological development	Analyzing the application of big data in the field of national security and guiding students to think about the relationship between technological development and national security. Discussing issues such as data security and network attacks, and cultivating students' sense of responsibility for information security.
Sustainable development and environmental protection	Emphasizing the application of big data in environmental monitoring and resource management, and guiding students to think about the relationship between technological development and sustainable development. Discussing the impact of digitalization on the environment and advocating the use of scientific and technological means to solve environmental problems.
Global vision and international cooperation	Cultivating students to have a global perspective and understand the application and cooperation of big data at the international level. Promoting students to participate in international scientific and technological innovation and cooperation projects and develop international competitiveness.

5. Big data introductory course summary

In order to comprehensively assess students' understanding of big data concepts, principles, and applications, and promote students' practical operation and problem-solving skills, the "Introduction to Big Data" course uses a combination of final open-book exams and comprehensive case reports. The final exam accounts for 60% and the comprehensive case report accounts for 40%. A comprehensive final exam that covers the main concepts, technical frameworks, case studies, etc., is designed in the course. Students are guided to analyze actual cases and write case analysis reports to analyze big data problems and solutions encountered in actual work. Students' practical application ability of theoretical knowledge is tested through actual case analysis.

6. Conclusion and future work

As professional teachers of "Introduction to Big Data," we have conducted a series of explorations and research on the construction of teaching courses for this course in the past three years. In the course preparation stage, we clarified the training objectives in the three dimensions of knowledge, skills, and literacy, and introduced common big data management systems and related algorithms. The teaching methods are rich, ideological and political elements are integrated into the whole teaching process, and students' participation is high. In the future, we plan to introduce more practical cases and projects to strengthen students' practical skills in technology application. Interactive links are added to stimulate students' interest in learning and meet the learning needs of different students. Regular feedback to students is enhanced to help them better understand and apply course content. Ideological and political elements are further integrated to guide students to pay attention to the relationship between technological development and social responsibility and cultivate their sense of social responsibility. More specific and practical course objectives are designed to enable students to better understand and apply big data technology based on academic knowledge.

Funding

(1) The 11th batch of education reform projects of Xi'an Jiaotong University City College (111001)
(2) Regular topics of Shaanxi Provincial Sports Bureau in 2023 (2023220, 2023161)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Jiang Q, 2023, Digital China: Big Data and Government Managerial Decision, Springer Nature, Singapore.

[2] Zhao X, Keane M, 2023, Sino-Futurism and Alternative Imaginaries of Digital China. Media International Australia, 2023. https://doi.org/10.1177/1329878X231215108

[3] Huang Y, Chen Z, Yu T, et al., 2018, Agricultural Remote Sensing Big Data: Management and Applications. Journal of Integrative Agriculture, 17(9): 1915–1931.

[4] Tang T, Abuhmaid A M, Olaimat M, et al., 2023, Efficiency of Flipped Classroom with Online-Based Teaching Under COVID-19. Interactive Learning Environments, 31(2): 1077–1088.

[5] Robertson J, Tisdall EKM, 2020, The Importance of Consulting Children and Young People About Data Literacy. Journal of Media Literacy Education, 12(3): 58–74.

[6] Yang S, 2023, Research on Curriculum Reform in Teacher Education Courses Based on the Outcomes-Based Education (OBE) Philosophy. Curriculum and Teaching Methodology, 6(19): 1–7.