Adult and Higher Education (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/aduhe.2024.060610
ISSN 2523-5826 Vol. 6 Num. 6

Exploration and Practice of Micro-Major Course Development in Big Data Management and Application under the New Liberal Arts Background

Zongjing Liang[1,a], Zhijie Li[1,b], Xiaobo Liang[2,c], Junyi Du[1,d], Yun Kuang[3,e,*]

[1]School of Economics and Management, Guangxi Normal University, Guilin, China
[2]Center for Network and Educational Technology, Guilin Normal College, Guilin, China
[3]Library, Guilin Normal College, Guilin, China
[a]2981893541@qq.com, [b]812811205@qq.com, [c]123070286@qq.com, [d]283853340@qq.com, [e]59937770@qq.com
[*]Corresponding author

Keywords: New Liberal Arts; Big Data Management and Application; Micro-major Construction

Abstract: This paper proposes the construction of a course development system for a "micro-major" in Big Data Management and Application under the prospects of the New Liberal Arts. The paper discusses the importance, necessity, and feasibility of building such a "micro-major." The research content specifically begins with an explanation of the theoretical foundation of this micro-major, followed by a description of the course construction status, and then outlines the specific steps for teaching implementation. Finally, a concrete case study is provided to demonstrate the feasibility and correctness of the proposed plan. The research outcomes can serve as a reference for the development of New Liberal Arts in higher education institutions.

1. Introduction

With the widespread application of emerging information technologies such as big data and artificial intelligence across various industries, there is an increasing demand for information technology skills among job seekers, which has, in turn, sparked a new wave of interest in cultivating interdisciplinary talents in higher education institutions. The cultivation of interdisciplinary talents is not limited to students in science and engineering fields but is also relevant for students in the humanities. Traditionally, humanities education has often focused on theoretical studies in the humanities and social sciences, lacking instruction in data-related knowledge. However, with the advent of the big data and AI era, data processing skills have become a fundamental requirement for many jobs. Therefore, integrating data processing knowledge into traditional humanities education has become a key issue in the current cultivation of interdisciplinary talents.

In response to this need, the concept of New Liberal Arts has emerged. The essence of implementing New Liberal Arts education reform in higher education is to integrate emerging information technologies into traditional humanities education, thereby creating a new model for

cultivating interdisciplinary talents. There are many approaches to New Liberal Arts education reform, one of which is the establishment of a "micro-major" in Big Data Management and Application. The construction of this "micro-major" involves offering core courses related to big data management, providing students with a specialized, short-term study opportunity to develop basic data processing skills.

This paper begins with an overview of the relevant theories on the construction of the "micro-major" in Big Data Management and Application, followed by a study of course development practices. It then delves into the practical aspects of course teaching, and finally, analysis the teaching outcomes through a specific case study, concluding with a summary of the entire paper.

2. Introduction to Theoretical Foundations

The construction of the "micro-major" in Big Data Management and Application is primarily aimed at humanities students. Since the academic background of humanities students differs from that of science and engineering students, the course structure is also distinct. To comprehensively describe the theoretical underpinnings and practical foundations of this micro-major, this section will provide a necessary introduction to the relevant theories.

2.1 Micro-major

The "micro-major" is an educational form that has gradually emerged in the field of higher education in recent years. The educational goal of a "micro-major" is to provide focused learning in a specific knowledge area over a relatively short period, aiming to cultivate particular skills. A "micro-major" can be seen as a supplement to traditional degree programs, helping students enhance their professional skills in specific fields[1,2]. The "micro-major" originated in the late 1990s and early 2000s, emerging against the backdrop of the development of the internet and digital technologies. As online education gradually became more widespread, this new educational model was born [3]. Subsequently, in the mid2010s, Massive Open Online Courses (MOOCs) related to "micro-majors" emerged. These courses were seen as one of the practical implementations of the "micro-major" concept. MOOCs offer many advantages compared to traditional offline education, such as flexibility in learning time and location, the ability to connect with likeminded learners worldwide, and the facilitation of shared learning and interaction[4]. Over time, the "micro-major" educational model has developed rapidly in recent years, with the variety and scope of micro-major courses expanding quickly [5]. Among them, "micro-majors" related to big data and artificial intelligence technologies have received widespread attention and have been increasingly applied in higher education institutions around the world.

2.2 Big Data Management and Application

The "Big Data Management and Application" major covers the entire process of big data, from data collection, processing, and analysis to its application. The core theories include the following:(1)Data Governance: This section covers topics such as data quality management, access control, and data privacy and security[6] . (2) Data Lifecycle Management: This involves a discussion of the various stages of the data lifecycle, the different tasks involved in managing data from various sources, and the application, challenges, conclusions, and future areas of focus for data lifecycle management in industries such as manufacturing and healthcare[7]. (3) Data Mining: This involves extracting valuable patterns and information from big data using various algorithms and models, including supervised learning, unsupervised learning, and deep learning algorithms. This knowledge is specifically applied to tasks such as classification, regression, clustering, and association analysis[8].

(4) Big Data Applications in Industries: The application of big data varies across different sectors such as finance, healthcare, retail, manufacturing, and government. Each industry has its specific use cases and requirements, such as financial risk management in finance, and diagnostic optimization in healthcare [9]. (5) Privacy Protection Theory: Protecting data privacy and security during data analysis and application is crucial. A comprehensive data protection system, including specific safeguards, is essential to compensate for the loss of individual control in data-intensive health research[10]. (6) Future Trends in Data Management and Application: Cross-Industry Data Sharing and Collaboration. Organizations from different sectors will increasingly collaborate to share data, generating greater value, particularly in areas such as public health, safety, and environmental protection[11].

3. Course Development Practice

The development of the "micro-major" courses in Big Data Management and Application for humanities student needs to retain the core content of big data technology while considering the actual background and abilities of students from the humanities. The focus should be on the application features of big data and the integration of interdisciplinary knowledge. The following outlines the specific course structure.

3.1 Foundational Courses

1) Introduction to Computers and Office Software Applications: This course covers the basic theoretical concepts of computers, including the components of a computer and an introduction to software and hardware. The main objective is for students to master basic data analysis functions in Excel.

2) Econometrics: Econometrics is an interdisciplinary subject that combines statistics, economics, and mathematical models. Its primary purpose is to quantitatively analysis the causal relationships and magnitudes between economic variables by building and estimating economic models, thereby providing a reference for policymaking.

3) Usage of Common Econometrics Software: This course focuses on the application of common econometrics software such as SPSS, Eviews, Oxmetrics, ArcGIS, and GeoDa. SPSS, Eviews, and Oxmetrics are mainly used for econometric data processing, while ArcGIS and GeoDa are primarily used for spatial data analysis.

3.2 Core Courses

1) Introduction to Big Data: This course provides an overview of the concepts, technologies, and applications of big data. It emphasizes the four main characteristics of big data: large volume, realtime processing, objectivity, and diversity. Big data is widely used in various industries, including business forecasting, healthcare, disease monitoring, traffic management, and financial risk control. However, the course also addresses potential issues that arise in its application, such as data privacy, data quality, and data ethics.

2) Text Analysis: This course covers the process of extracting information from text using automated techniques, which is a branch of natural language understanding within computer science. The course content includes text preprocessing, sentiment analysis, topic model identification, text classification, and text clustering. One of the most commonly used topic models is the Latent Dirichlet Allocation (LDA) model. LDA is a statistical model used for text topic model that can automatically classify and cluster text, making it one of the most widely applied models for topic extraction.

3) Introduction to Interdisciplinary Studies (e.g., Structural Equation Model): This course focuses on introducing Structural Equation Model (SEM), a multivariate statistical model. The key difference between SEM and regression analysis models is that SEM can simultaneously handle multiple

dependent variables (Y) and multiple independent variables (X). SEM allows for a comprehensive approach to factor analysis, path analysis, and regression analysis. It is widely applicable in fields such as sociology, psychology, and education.

4. Course Teaching Practice

Given the reality that humanities students lack a technical background, the teaching practice for the "micro-major" courses needs to emphasize interdisciplinary teaching. Methods such as analogies and case studies should be used to enhance students' understanding and achieve the final teaching objectives. Below is a description of the teaching process for the four main courses in the "micro-major."

4.1 Econometrics

Teaching Content: The content includes statistical description of data, correlation coefficients and regression analysis, stationarity tests, causal analysis, time series case analysis, and an introduction to nonlinear models. Additionally, the course covers the use of common econometrics software, such as SPSS and Eviews, including both introduction and practical application.

Teaching Methods: A combination of lecture-based teaching, case-based teaching, and blended online and offline teaching methods.

Teaching Evaluation: Evaluation methods include class performance and participation, homework assessments, and exams.

4.2 Spatial Econometrics

Teaching Content: The main topics include types of spatial data, spatial weight matrices, spatial autocorrelation, spatial lag models, spatial error models, spatial Durbin models, and geographically weighted regression. Additionally, the course covers the use of classic spatial data analysis software such as ArcGIS and GeoDa.

Teaching Methods: The teaching methods include theoretical instruction, case-based teaching, and experimental teaching.

Teaching Evaluation: Evaluation methods include class performance, homework assessments, and lab reports.

4.3 Introduction to Big Data

Teaching Content: The course covers the definition of big data, sources of big data, big data storage and management, data analysis tools, and big data visualization. It also introduces specific applications of big data and artificial intelligence in fields such as tourism management, infectious disease prediction, and online religious dissemination, with a focus on the current usage of generative AI.

Teaching Methods: A combination of theoretical teaching, experimental teaching, and project-based teaching methods.

Teaching Evaluation: Evaluation combines classroom performance, homework and lab reports, and course paper assessments.

4.4 Text Analysis

Teaching Content: The content includes the basics of text analysis, text classification techniques,

sentiment analysis, topic model, and text clustering and similarity analysis. A key focus is on learning the theory and practical application of the LDA topic model algorithm.

Teaching Methods: The teaching methods primarily involve a combination of theoretical classroom instruction, experimental teaching, project-based teaching, and flipped classroom techniques.

Teaching Evaluation: Evaluation methods include class performance, exams, and course paper assessments.

5. Analysis of Practical Outcomes

To assess the practical teaching effectiveness of the "micro-major" in Big Data Management and Application proposed in this paper, an actual teaching exploration was conducted within the Information Management and Information Systems program at a university in Guangxi. To evaluate the teaching outcomes, students were actively encouraged to participate in various academic competitions during the learning process. This participation aimed to train students' abilities in applying knowledge, organization, and collaboration.

The following is a case study of one learning group (comprising three undergraduate students) that participated in a national academic competition. By analyzing the group's research outcomes, the steps for applying big data management and application in practice are illustrated. The competition they participated in was the "National College Student Market Research and Analysis Competition," with their project titled "Research on the Use of AI Large Models and the Spread of Online Public Opinion." According to the competition's organization requirements, the team first needed to participate in the school's preliminary competition and, upon winning, advance to the provincial level. The team successfully passed the school selection and advanced to the provincial competition, ultimately winning second prize in the "Chia Tai Cup 14th National College Student Market Research and Analysis Competition, Guangxi Division, Undergraduate Group," organized by the Chinese Business Statistics Society in 2024.

The project, "Research on the Use of AI Large Models and the Spread of Online Public Opinion," was set against the backdrop of the rapid development of artificial intelligence technology, which has sparked a new generation of information revolution. The research focused on university students' use of generative AI technology, employing big data analysis methods for market research and public opinion dissemination studies. The results were presented in the form of a paper, which was divided into three parts:

1) **Part One**: Market Research on the Use of AI Large Models in Guangxi. This section utilized a questionnaire to gauge university students' satisfaction with using generative AI (AI large models). The research found that more than half of the respondents had used AI large models, with the majority using them once or twice per week. The data also shows that higher education level of the respondents, the more frequently they use AI models.

2) **Part Two**: Topic Mining on AI Large Models on Weibo. To better understand online users' willingness and sentiment towards using AI large models, a web crawler was used to collect data from Sina Weibo. Text analysis methods were then employed for topic mining. After data cleaning and preprocessing (including word segmentation), a word cloud was generated. Subsequently, LDA topic analysis was performed, and a "bubble" chart was drawn to visually display the results of text classification and clustering. The findings indicated that online users were primarily concerned with AI large model performance, application scenarios, and development trends. Regarding sentiment, the public generally held a positive attitude towards AI large models, though there were some concerns and skepticize.

3) **Part Three**: Study on Factors Influencing Global Public Attention to AI Large Models. This

section aimed to investigate global public opinion on AI large models by analyzing global news big data (GDELT data) and Google Trends data from December 25, 2022, to the present. An Autoregressive Distributed Lag (ARDL) model was used to study the impact of news big data on public attention. The results showed a good fit for the research model, indicating that the findings could serve as a reference for public opinion monitoring.

In summary, by the analysis the research outcomes of the "Research on the Use of AI Large Models and the Spread of Online Public Opinion," the steps involved in using big data management and application technologies can be understood. This research applied various techniques from the "micro-major," including time series data analysis, structural equation model, and text analysis. Due to the cutting-edge nature and systematic approach of the research, the project won second prize in the provincial academic competition, fully demonstrating the practicality and validity of the "micro-major" construction proposed in this paper.

6. Conclusion

This paper proposes the construction of a "micro-major" curriculum system in Big Data Management and Application under the framework of New Liberal Arts. It discusses the importance, necessity, and feasibility of developing such a micro-major. The research begins by outlining the theoretical foundations of the micro-major, followed by the course development plan, specific teaching implementation steps, and finally, a case study to demonstrate the feasibility and correctness of the proposed plan. The following insights can be drawn from this study:

1) The necessity of establishing a "micro-major" in Big Data Management and Application. As data has become a critical factor driving transformation across various industries, mastering skills in big data management and application has become a powerful tool for humanities students to gain a deeper understanding of social and economic phenomena. The "micro-major" allows for breaking the boundaries of traditional humanities education, fostering the organic integration of the humanities with information technology. This approach cultivates interdisciplinary talents with both humanistic qualities and data analysis skills, thus providing qualified professionals for the construction of society in the new era.

2) The effectiveness of teaching strategies as a key to successful "micro-major" implementation. This paper proposes a series of strategies, including making full use of existing educational resources, enhancing interdisciplinary collaboration, and flexibly utilizing online platforms. By pooling excellent teaching resources from inside and outside the institution and combining the technical strengths of science and engineering disciplines, humanities colleges can establish a specialized system suitable for big data education. At the same time, ensuring a gradual and systematic teaching process enables humanities students to master big data skills, providing a solid foundation for the successful implementation of the "micro-major."

3) Active participation in academic competitions as an effective means of evaluating teaching outcomes. On a practical level, this paper emphasizes the importance of academic competitions in integrating theory with practice. For example, the National College Student Market Research Competition is an excellent example, demonstrating how students apply big data technologies to solve real world problems. This not only deepens students' understanding of the course content but also significantly enhances their data analysis and problem-solving skills. These successful cases prove the practicality and application value of the "micro-major."

In conclusion, offering a "micro-major" in Big Data Management for humanities students aligns with the development trend of New Liberal Arts and effectively enhances the adaptability and competitiveness of humanities students in the era of big data. This "micro-major" not only broadens the academic and career prospects of humanities students but also provides a successful example of

the transformation of humanities education in the digital age. With the accumulation of practical experience and the continuous improvement of the curriculum system, this "micro-major" is expected to become a highlight of innovation in humanities education, fostering more talents with data literacy and interdisciplinary skills for society.

Acknowledgments

The work of this article is supported by the following funds: Guangxi Zhuang Autonomous Region Level New Liberal Arts Research and Practice Project (No.: XWK2022008): "Research and Practice on the Transformation and Enhancement of Economic Management Majors Based on Big Data Technology", The Ministry of Education's Supply and Demand Matching Employment Education Project(No.:20230113798): "Exploration of the Digital and Intelligent Upgrade of the Business Administration Curriculum System",2021 Guangxi Education Science Planning Key Funding Project 'Category A' (No.: 2021A004): "Research on the Current Situation and Improvement Strategies of Information Literacy Cultivation in Guangxi Vocational Colleges under the Smart Campus Environment", The 2023 Special Key Project of Innovation and Entrepreneurship Education in Colleges and Universities in Guangxi Education Science's "14th Five Year Plan" (No.: 2023ZJY1537): "Research on the practice system of innovation and entrepreneurship education in higher vocational colleges under the background of smart campus", The Guangxi Higher Education Undergraduate Teaching Reform Project (No.:2023JGA134): "Research on the Innovation of the 'One Core, Dual Threads, Three Dimensions, Four Orders' Hybrid Teaching Model Aimed at Deep Learning", Research and Practice Project of New Engineering and New Liberal Arts at Guangxi Normal University (No.: 2023JGXW02): "Optimization and Practical Research of the Business Administration Curriculum System in the Era of Digital Intelligence".

References

[1] Wang Yu, Luo Shufang, Fan Yizhou, et al. 2017 Review of Global MOOC Development [J]. China Distance Education, 2018, 9(1): 53-63.

[2] Zhou Yi. Discussion on the Practice of "Micro-major" Construction in Library, Information, and Archival Studies under the New Liberal Arts Background [J]. Information and Documentation Services, 2021, 5-12.

[3] BATES A W. Teaching in a digital age: Guidelines for designing teaching and learning [M]. BC campus, 2015.

[4] BATURAY M H. An overview of the world of MOOCs [J]. Procedia-Social and Behavioral Sciences, 2015, 174(1): 427-433.

[5] GALLAGHER S R. The future of university credentials: New developments at the intersection of higher education and hiring [M]. Harvard Education Press, 2022.

[6] International D. Dama-Dmbok: Data management body of knowledge [M]. Technics Publications, LLC, 2017.

[7] Rahul K, Banyal R K. Data life cycle management in big data analytics [J]. Procedia Computer Science, 2020, 173(1): 364-371.

[8] Mitchell T M, Learning M. Mcgraw-hill science [J]. Engineering/Math, 1997, 1(1): 27.

[9] Mayer-Schönberger V, Cukier K. Big data: A revolution that will transform how we live, work, and think [M]. Houghton Mifflin Harcourt, 2013.

[10] Mostert M, Bredenoord A L, Van Der Slootb B, et al. From privacy to data protection in the EU: Implications for big data health research [J]. European Journal of health law, 2018, 25(1): 43-55.

[11] Gawer A, Cusumano M A. Industry platforms and ecosystem innovation [J]. Journal of product innovation management, 2014, 31(3): 417-433.