Research on the Training Model for Big Data Major

Shuo Wen, Wei Du, Lixuan Zhao, Kangyuan He, Xinghua Liu, Qi Wu
Guangdong Police College, Guangzhou, Guangdong, China

Abstract: With the continuous development of information technology, the new development model led by big data has become the trend of informatization reform across various industries. This paper, based on the recent background of the big data strategy, studies and analyzes the thinking abilities and professional skills required for big data major, considering the challenges faced in the training of a big data workforce. Adhering to the education philosophy of "practice-oriented, student-centered," and focusing on "enhancing data awareness and data literacy among students in the new era," this paper proposes a big data talent cultivation model that emphasizes student development and aligns with the needs of practical units.

Keywords: Big Data; Talent Training Program Construction; Specialty Construction

1. Introduction

With the rapid development of computer information technology, data has become an important resource in the new era and has continuously penetrated all aspects of social life and economic development[1]. In order to promote the high-quality development of big data technology, existing research has conducted in-depth research from multiple aspects such as theory[2], policy[3], information security[4], laws and regulations[5, 6]. Big data technology and its industry have developed steadily, and a report pointed out that the market size of the big data industry has continued to expand and is expected to reach 1.9 trillion in 2023[7]. As the big data industry matures, the quantity and quality of big data talents are also increasing. This requires all kinds of universities to strengthen professional talent training, innovate talent training models, and establish and improve a multi-level and multi-type big data talent training system.

As key institutions for social talent output, universities bear the important responsibility of cultivating high-quality reserve talents for society and providing intellectual and technical support across various industries. Current research on optimizing the cultivation of big data talent in universities primarily focuses on aspects such as curriculum system construction and practical teaching. For example, Li et al.[8] analyzed the curriculum system of big data management and application in 84 higher education institutions in China, used LDA topic modeling to cluster the curriculum settings of different types of institutions, identified commonalities, characteristics, and existing issues, and proposed suggestions for optimizing the curriculum system in the context of the "new liberal arts". Cui et al.[9] established a curriculum system that spans from basic to comprehensive, from theory to practice, designed case-based and task-driven teaching methods, and used modular experimental design to enhance students' mastery of big data technology and practical application skills. Under the background of the big data era, Tang[10] proposed suggestions for curriculum reform in response to the current situation of information-related course construction and the new changes in talent demand, aiming to promote the deep intersection and integration of disciplines and computer science, and strengthen students' information application capabilities. It is evident that existing research has explored big data talent cultivation models across various fields such as new liberal arts, new engineering, and public security technology, targeting different disciplinary needs. However, as many industries now require big data capabilities to enable informatization reforms, further exploration is needed on how various disciplines can integrate big data skills into their educational models.

To fulfill the mission of universities in training and delivering professional talents to society, it

is essential to perfect the new operational model of "Discipline + Mechanism + Big data." This involves accelerating the construction of an integrated "intelligence-command-body" operation mechanism, where new structures, new mechanisms, and new methods contribute to the generation and enhancement of new social development capabilities. Under the requirements of national higher education reform, universities in the new era must carefully study the principles of big data education and teaching in the context of the new situation. They should also consider the practical applications of rapidly developing science and technology in various industries and strengthen practical, diversified teaching to cultivate students' innovative and practical abilities. This will enhance students' professional qualities in multiple ways. Therefore, this study, building on previous research, proposes a big data professional talent training model that aligns with the demands of society and addresses the challenges faced by universities in the information age.

2. The Challenges in Big Data Talent Cultivation

As an important export place of talent training for social development, colleges and universities shoulder the important responsibility of cultivating high-quality reserve talents for social development and progress, and also provide intellectual support and technical support for the development of social industries. Under the requirements of national higher education reform, colleges and universities in the new era should seriously study the law of big data major education and teaching under the new situation, combined with the use of rapidly developing science and technology in the industry in today's era, strengthen practical and diversified teaching in teaching, cultivate students 'innovative practice ability, and improve students 'professional quality in many aspects, which is of great significance.

Currently, according to various professional training programs, students are required to master fundamental theoretical knowledge and the current state of practice in their field. They should also be familiar with relevant knowledge in the natural sciences, social sciences, technology, and humanities. However, in the context of big data development, universities must adapt their talent cultivation strategies to meet new demands. Big data major need sufficient skills and expertise to contribute to significant advancements in areas such as data collection, data processing and integration, data analysis, and data mining after entering the workforce. This requires a forward-looking approach and in-depth research to drive quality, efficiency, and innovation in their professional fields.

3. The Establishment of a Big Data Talent Cultivation Model

Based on the societal demand for big data major and considering the realities of higher education, this paper proposes a talent cultivation model centered on "one direction." The goal is to ensure the delivery of highly skilled and innovative major who meet the needs of societal development. The model introduces a curriculum system that integrates "two major categories of courses," allowing big data major to advance in both big data business and technology. It also suggests "three types of customized student training," where different training plans are developed for different types of students. Finally, the model advocates "four combined tactics" as a strategy for talent cultivation, providing clear guidance and ensuring its implementation. The specific framework is illustrated in Figure 1.

3.1 The Talent Cultivation of “Based on a Single Concept”

The cultivation of big data major adheres to the "practice-oriented, student-centered" educational philosophy. This approach emphasizes enhancing "data awareness and data literacy" among college students in the new era, while organically integrating practical elements of big data into the curriculum. This method subtly fosters students' spirit of innovation, professional ethics, and pride in the field of big data. By maintaining a focus on "student development" and "meeting the needs of practical units," this model continuously improves teaching quality and ensures the delivery of high-level, highly-skilled major who are well-prepared to contribute to society.

3.2 The Teaching Model of "Integration of Two Major Approaches"

This model moves away from traditional technology-centered teaching methods by incorporating the "new engineering" mindset. It promotes the deep integration of modern information technologies, such as big data online platforms, internet-based teaching platforms, and interactive teaching software, with education. The model extends learning from the "first classroom" to the "second classroom," effectively merging with practical enterprise units to create a hybrid teaching model that combines online and offline, school-enterprise integration.

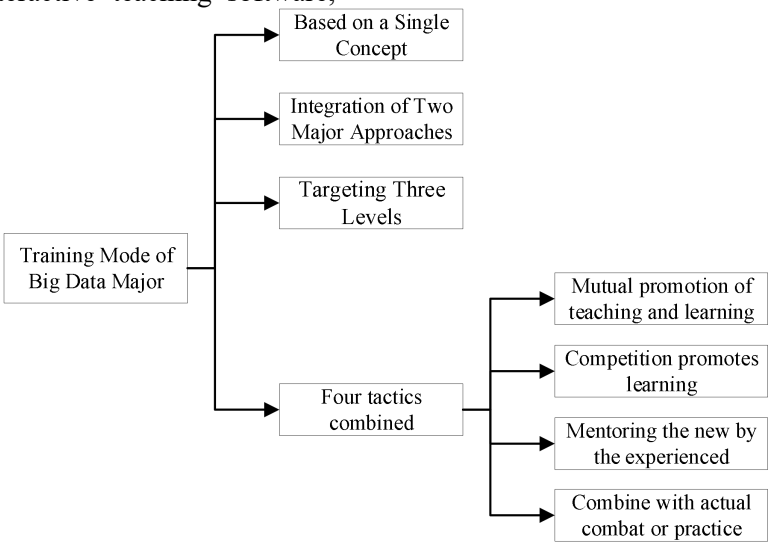

Figure 1. Big Data Talent Cultivation Model

To closely align with the practical needs of big data operations, this model involves in-depth cooperation with the developers of big data platforms commonly used across industries. It meticulously crafts online training courses that serve as resource carriers for conducting deeply integrated hybrid teaching, combining online and offline components. These resources include comprehensive online teaching videos, big data business practice platforms, and instructional training data, covering areas such as data processing, data analysis, common visualization techniques, data modeling and implementation, and data analysis report preparation.

Based on these resources, the model fully utilizes informational teaching tools to create a blended teaching approach that merges online and offline education with practical enterprise collaboration. This approach is deeply integrated with real-world practices, fostering the development of a "data thinking in-class + practical thinking" model alongside "after-class online autonomous learning." Students not only master data technology but also expand their data thinking and recognize the importance of data thinking in real-world applications. Moreover, by engaging with custom online training courses that accompany various big data platforms, students seamlessly connect with future big data business practices.

Before class, students engage in self-directed learning through online platforms, acquiring a foundational understanding of the material and performing basic practical exercises. During class, interactive, discussion-based teaching methods are used, supported by modern information technology tools like teaching platforms and interactive software. Teachers focus on building a comprehensive knowledge framework, clarifying data thinking, and cultivating students' sensitivity to data as big data major. Group discussions use case-based teaching to stimulate students' further exploration. By analyzing given cases and discovering related examples from a big data perspective, students are encouraged to develop their data thinking and data security awareness. Additionally, industry experts are invited to give 1-2 lectures, providing students with direct exposure to practical work in the big data industry. After class, preparatory assignments and major projects linked to online courses are assigned to deepen students' learning and spark innovative thinking. Extra-credit assignments, such as research paper reviews or explorations of the latest technologies, are also provided to challenge and engage advanced students, guiding their

future learning.

3.3 The Curriculum Structure of "Targeting Three Levels"

Under the background of the construction of the "new engineering" and big data-driven new policing model, the teaching plan is formulated according to three levels of students: "students without a computer science background", "ordinary students with a computer science background", and "top students with a computer science background". The teaching content is arranged as needed to meet the requirements of subsequent courses. Each category of teaching content is designed into three levels of difficulty and presented through different teaching methods. The integrated plan can better meet the needs of talent cultivation with different computer science teaching backgrounds, ensuring the integrity of the knowledge system and achieving the teaching objectives. It creates a learning environment of "teaching, learning, practicing, and engaging in real-world scenarios". It pays close attention to student needs, organizes interactive teaching sessions that motivate high enthusiasm and a strong sense of participation, monitors classroom feedback and the actual needs of real-world units, and promptly adjusts teaching strategies. The focus is on the effectiveness of students' knowledge acquisition, emphasizing the cultivation of their self-study abilities and active learning of "advanced, specialized, and in-depth" knowledge. Efforts are made to enhance the practical applicability of teaching, enabling both teachers and students to improve simultaneously.

3.3.1 The students without computer science background

Students with a non-computer science background are often overlooked in big data talent cultivation. In reality, the need for big data major extends beyond just computer-related industries across various sectors of society. Enterprises, institutions, catering services, manufacturing, and countless other industries all require the participation of big data talents in their development. The presence of a professional with big data skills can significantly alleviate work pressures within a position. Although these students lack a computer science background and programming fundamentals, through a well-structured curriculum, they can still master the three technical levels of data preprocessing, data analysis, and data visualization. To ensure a seamless learning experience for students without programming knowledge, the content taught focuses primarily on industry-mature software, including advanced applications of Microsoft software, BI software, and police practical platform software.

3.3.2 Ordinary students with a computer science background

Ordinary students with a computer science background constitute the backbone of big data major. Through their studies, they have acquired fundamental computer programming skills, knowledge of databases, Linux operating systems, and the ability to use code to address the entire process of data preprocessing, analysis, and visualization. For these students, beyond teaching basic principles and code writing, emphasis should be placed on cultivating their ability to automate the entire data analysis process using code. In practical scenarios, numerous data analysis processes remain disconnected, and various industries often need to obtain data from different sources, manually cleaning, integrating, analyzing, and visualizing it, leading to low efficiency and a high error rate. Fully automated integration programs can significantly enhance the efficiency of various industries, allowing staff to focus on more crucial tasks and reducing the likelihood of errors.

3.3.3 Outstanding students with a computer science background

Building upon the foundation of ordinary computer science students, specialized courses tailored to the unique needs and potential of outstanding students are essential. These courses encompass more challenging content such as advanced algorithm analysis, satisfying the eagerness of these exceptional students for challenges and in-depth learning. Simultaneously, the school adopts an individualized "tailor-made" training approach, providing more precise and intensive guidance and training for students interested in research and big data practical work. This training model aims to keep big data majors at the forefront of information technology development. By exposing them to the latest technologies and theories, it ignites their

curiosity and sustained research interest in information technology, big data, and artificial intelligence, fully tapping into and unleashing their remarkable potential.

3.4 Talent Cultivation Strategy "Combining Four Tactics"

The talent cultivation strategy "combining four tactics" refers to the implementation of a model that integrates "teaching by mutual growth, learning through competition, mentoring through experience, and combining with practical application." By aligning this strategy with the "practice-oriented philosophy" and the "integration of two major courses" curriculum, the four tactics create a closed-loop for big data professional talent cultivation. This model aims to develop major who not only master big data technology but also know how to apply it, possess strong problem-solving abilities across industries, and stay aligned with the development trends of the big data era. We endeavor to enhance the practical applicability of teaching, enabling both teachers and students to improve simultaneously.

3.4.1 Teaching by mutual growth

"Teaching by mutual growth" has two aspects: first, teachers, while proficiently teaching the fundamental knowledge of their courses, should also uphold the spirit of scientific inquiry and embrace the concept of lifelong learning. They should be willing to explore key courses related to their teaching, even if they are unfamiliar, to broaden their knowledge and update their understanding, ensuring continuity in the learning of specialized courses. By deeply integrating course content, the traditional isolated and compartmentalized structure of the curriculum is broken, allowing students to truly appreciate the importance of what they learn and understand how to transfer and apply knowledge, thereby promoting "learning through teaching." Second, after students grasp the foundational knowledge, teachers can experiment with "flipped classroom" formats, encouraging students to independently gather information and present their findings, thus enhancing their understanding of the material and their communication skills, promoting "teaching through learning."

3.4.2 Learning through competition

The main focus of "learning through competition" is on the students, who test their team's skills and overall quality by participating in relevant big data skills competitions. This approach helps to consolidate and reinforce students' professional knowledge, improving the alignment of their knowledge system with big data analysis and application skills. The competitive atmosphere closely mirrors actual professional work, providing an opportunity for students to strengthen their psychological resilience, remain calm under pressure, and tackle challenges head-on. Additionally, participating in competitions reinforces practical skills, fostering a positive cycle where theory informs practice, and practice consolidates theory. After competitions, teachers organize student teams to review and analyze the competition process and outcomes, identifying areas of improvement in teaching and learning, recognizing successful strategies, and maintaining them. This process helps students identify knowledge gaps, address them, and summarize their findings, improving their ability to identify, analyze, solve, and summarize problems. This approach further enhances students' self-learning abilities and innovative thinking.

Simultaneously, this strategy efficiently manages resources, facilitates the sharing of big data practice resources within the network, and meets students' practical needs. It establishes a comprehensive teaching curriculum system, complete teaching resources, and corresponding practical training scenarios, solving the issues related to the construction and configuration of experimental and competition platforms, thereby achieving "learning through competition." As shown in Figure 2.

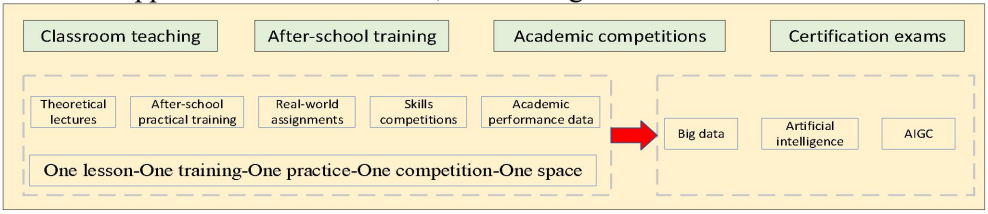

Figure 2. Learning through Competition

3.4.3 Mentorship strategy

The "mentorship strategy" involves experienced students with a solid knowledge foundation and practical training guiding those who have not yet developed a comprehensive knowledge framework or lack hands-on experience. Through peer-led assistance and mutual exchanges, this approach promotes overall collective advancement and steady improvement, providing sustainable momentum for the development of big data professional teams.

In recent years, the rapid development of the internet and information industries has led to continuous updates in big data-related knowledge. As a discipline with a fast pace of knowledge growth and renewal, younger students who have established a strong foundation are better positioned to stay abreast of the latest developments in the big data era. Their accumulated learning outcomes and practical experiences serve as valuable resources to foster the growth of big data professional teams, establishing a sustainable development mechanism.

3.4.4 Integration with practical experience

The "integration with practical experience" strategy focuses on strategic cooperation between universities and big data industry practice units. To bridge the gap between students' capabilities and industry needs, and to help students adapt more quickly and easily to relevant fields after graduation, collaboration with big data industry units is crucial. Specific measures include inviting experienced industry experts to conduct lectures or take on practical course teaching roles, and having these experts provide input on course development based on current industry trends. These initiatives aim to enhance classroom case studies and improve students' practical skills.

At the same time, before graduation, students are arranged for internships at carefully selected organizations that are closely related to big data technologies. This enables them to truly apply what they have learned and closely integrate it with practical work, where they can contribute their efforts. By applying the knowledge acquired in the classroom to actual work situations, students personally engage with and utilize big data platforms for modeling, running, and obtaining results. They analyze and mine data programming in real-world tasks, accomplishing practical assignments such as querying associated personnel information across various systems, participating in epidemic prevention and control data verification, and data cleaning. In the process of actively exploring and optimizing work methodologies and strategies, students significantly enhance their abilities to independently think, solve problems hands-on, and their proficiency in big data technologies and thinking.

4. Conclusion

Currently, China is actively advancing big data and intelligent applications, positioning big data as a key driver of innovation across industries and a new growth point for generating competitiveness. Big data major are indispensable in the information age. To meet the urgent demand for big data talent in society, universities must update and adjust their talent cultivation models. This paper proposes a big data talent cultivation model that aligns with the "practice-oriented, student-centered" educational philosophy, enabling students of different levels to develop unique knowledge systems and innovative thinking through learning and practice. This model aims to cultivate new major proficient in big data skills, providing a fresh approach to both big data-driven industry transformation and educational reform in higher institutions.

Acknowledgments

This paper is supported by The "New Generation Information Technology Innovation Project" of the China University Industry, University and Research Innovation Fund(No. 2022IT171), Youth scientific research project of Guangdong police college (No. 2022-QN04), Cooperative education project of industry university cooperation between Ministry of education and 360 enterprise security group -- Web Attack and Defense Technology (web security and operation) (No. 201802007006), 2023 Campus level Teaching Quality and Teaching Reform Project of Guangdong Police College (Research on the Practical Teaching Reform of Police Big Data Course Based on the Integration of "Teaching, Studying, Training, Combating and Competition"), 2022 Campus level Teaching Quality and Teaching Reform Project of Guangdong Police College

(Teaching Reform and Research on Police Big Data Course for Practical Use), Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds, No. pdjh2021b0356), Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds, No. pdjh2022b0373).

References

[1] BAOPING R, WEI X. Transformation of China's new normal economic growth path and policy in the era of big data. Journal of Humanities, 2015, (04): 29-35.

[2] HUI L. Theoretical mechanism, practical basis and policy choice of big data promoting high-quality economic development in my country. Economist, 2019, (03): 52-9.

[3] TAO Z, HAIQUN M. Analysis of my country's big data policy themes and development trends. Intelligence Theory and Practice, 2022, 45(03): 72-80.

[4] National Data Bureau and 17 Departments. Three-Year Action Plan for "Data Element ×" (2024-2026). Beijing, 2024.

[5] YU Y, FEI Y. Overcoming personal information protection challenges involving real-world data to support public health efforts in China. FRONTIERS IN PUBLIC HEALTH, 2023, 11.

[6] ZHANG H, JAYAKAR K, TANG Z. From Lofty Guarantees to Basic Protections: Benchmarking Privacy and Data Protection Standards in Chinese Government Programs. TPRC 46: The 46th Research Conference on Communication, Information and Internet Policy 2018. 2018: 1-22

[7] China Business Industry Research Institute. 2024 China Big Data Industry Market Outlook Research Report (Simplified Edition). Beijing, 2024.

[8] LANLAN L, JIANLING J, RANZHONG Y. Research on the Construction of Big Data Management and Application Curriculum System under the Background of New Liberal Arts. Library Science Research, 2023, (10): 2-12.

[9] XIAOLONG C, MIN Z, LEI Z. Research and Construction of Big Data Talent Training Curriculum Group under the Background of New Engineering. Experimental Technology and Management, 2021, 38(2): 213-8.

[10] YUNQI T. Discussion on the Construction of Information Courses in Criminal Science and Technology Specialty in the Big Data Era: A Case Study of the People's Public Security University of China. Laboratory Research and Exploration, 2020, 39(10): 155-8.

