

Research on Literature Data Analysis of Ideological and Political Education in Curriculum

Taizhi Lv, Enze Wu, Yujuan Zou

College of Information Engineering, Jiangsu Maritime Institute, Jiangsu Nanjing, 211170, China

Abstract

The quantitative study of literature on ideological and political education in curriculum helps to gain a deeper understanding of the research trends and progress in its construction, clarify research priorities, scopes, and theoretical frameworks, and provides guidance for enhancing the quality of talent cultivation. This paper employs the Scrapy crawling framework of Python to specifically crawl data on ideological and political education in curriculum literature, storing it in a MySQL database to construct a comprehensive and accurate literature database. The research utilizes ECharts for quantitative data visualization analysis, intuitively displaying research hotspots, trends, and shortcomings through visual charts, thus providing data support for researchers. Finally, a detailed comparative analysis is conducted based on the quantitative results, revealing key issues and potential opportunities, clarifying future research directions, and promoting the improvement of the theoretical system and the deepening of practical applications of ideological and political education in curriculum.

Keywords

Ideological and Political Education in Curriculum, Data Analysis, Web Crawling, Data Visualization.

1. Introduction

Ideological and political education in curriculum is an idea that takes the curriculum as the carrier and integrates ideological and political education throughout the entire teaching process [1-2]. This educational approach aims to cultivate students' ideological and moral qualities and promote their overall development. In recent years, research on ideological and political education in curriculum in Chinese universities has increasingly garnered widespread attention, becoming a research hotspot in academia and the education sector.

In recent years, with the continuous development of big data and artificial intelligence technologies, the application of web crawling technology in the field of education has become increasingly widespread [3]. Especially in the area of ideological and political education in curriculum, through quantitative research on literature data, one can gain insights into research hotspots, trends, and existing issues in this field, providing a scientific basis for the further development of ideological and political education in curriculum. To deeply study the application of web crawling technology in ideological and political education in curriculum, this paper employs the ECharts tool for quantitative data visualization analysis, aiming to clearly present hot topics, development trends, and existing shortcomings in the research field through vivid and intuitive chart forms, thereby providing researchers with solid data support and insights.

2. Related Technologies

2.1. CNKI Literature Crawling Technology Based on Scrapy Framework

The Scrapy framework consists of multiple components such as an engine, scheduler, downloader, and spider [4]. These components work together to complete the task of crawling CNKI literature. The engine, serving as the heart of the entire framework, handles data flow and triggers various events; the scheduler manages the request queue to ensure orderly execution of requests. The downloader simulates browser behavior to download webpage resources from the CNKI website, while the spider extracts the required literature information from the downloaded webpage based on preset rules. In the process of crawling CNKI literature, the flexibility and scalability of the Scrapy framework are fully demonstrated. Researchers can customize crawling rules according to actual needs to accurately capture target literature information. Additionally, Scrapy provides rich middleware and pipeline mechanisms, making data cleaning, validation, and saving more convenient. Through the Scrapy framework, researchers can efficiently acquire CNKI literature resources for quantitative data visualization analysis, intuitively displaying research hotspots, trends, and shortcomings, thereby providing strong data support for academic research.

2.2. CNKI Literature Mining Technology Based on Python

As an important literature resource platform, CNKI contains abundant academic information. To efficiently mine these valuable resources, this paper adopts a CNKI literature mining technology based on Python. After obtaining CNKI literature data, this paper utilizes data processing libraries such as Pandas and NumPy in Python to clean, organize, and analyze the data. By constructing data models, conducting statistical analysis, drawing visualization charts, and other means, this paper deeply mines academic hotspots, research trends, and potential research gaps in CNKI literature. The CNKI literature mining technology based on Python not only improves the efficiency and accuracy of literature acquisition but also provides strong data support and insights for in-depth academic research.

2.3. Literature Data Storage Based on MySQL

Literature data storage based on MySQL is an important part of CNKI literature mining. Firstly, based on the characteristics of literature data, a reasonable database table structure is designed, clarifying the field types and constraints. Subsequently, a dedicated database is created in MySQL, and data tables are set up to store detailed literature information, such as literature titles, authors, publication dates, etc. MySQL's powerful storage and management capabilities ensure the security, integrity, and queryability of literature data. By constructing indexes and optimizing query statements, rapid retrieval and analysis of massive literature data can be achieved, providing solid data support for subsequent academic research and knowledge mining. This storage solution based on MySQL not only improves the processing efficiency of literature data but also greatly promotes the sharing and utilization of academic resources.

2.4. Literature Data Visualization Based on ECharts

This paper utilizes the powerful ECharts data visualization library to present literature data in intuitive and vivid chart forms [5]. Through bar charts showing the distribution of research hotspots, line charts revealing research trend changes, and pie charts presenting the proportions of different research fields, complex literature data becomes easier to understand and analyze. The interactivity of ECharts also allows users to deeply explore the data, discovering hidden patterns and correlations. This visualization method not only enhances the interpretation efficiency of literature data but also provides academic researchers with intuitive data support, helping them gain insights into academic frontiers and grasp research dynamics.

3. Quantitative Research on Literature Data

3.1. Basic Analysis of Literature

The number of publications is one of the important indicators to measure the degree of attention received by a research field, and the annual analysis of the number of publications can reflect changes in the degree of attention received by the research field [6]. As shown in Figure 1, since 2016, the number of related literature has shown a significant and continuous upward trend. This robust growth is not only a powerful testament to the activity level of the research field but also specifically reflects that ideological and political education in curriculum, as a current hot research topic in the field of educational research, still maintains a high level of research popularity and widespread academic attention. This trend not only indicates that the important position of ideological and political education in curriculum in educational practice has been widely recognized but also predicts that this field will continue to attract more scholars for in-depth exploration in the future, contributing new wisdom and strength to promoting educational reform and development.

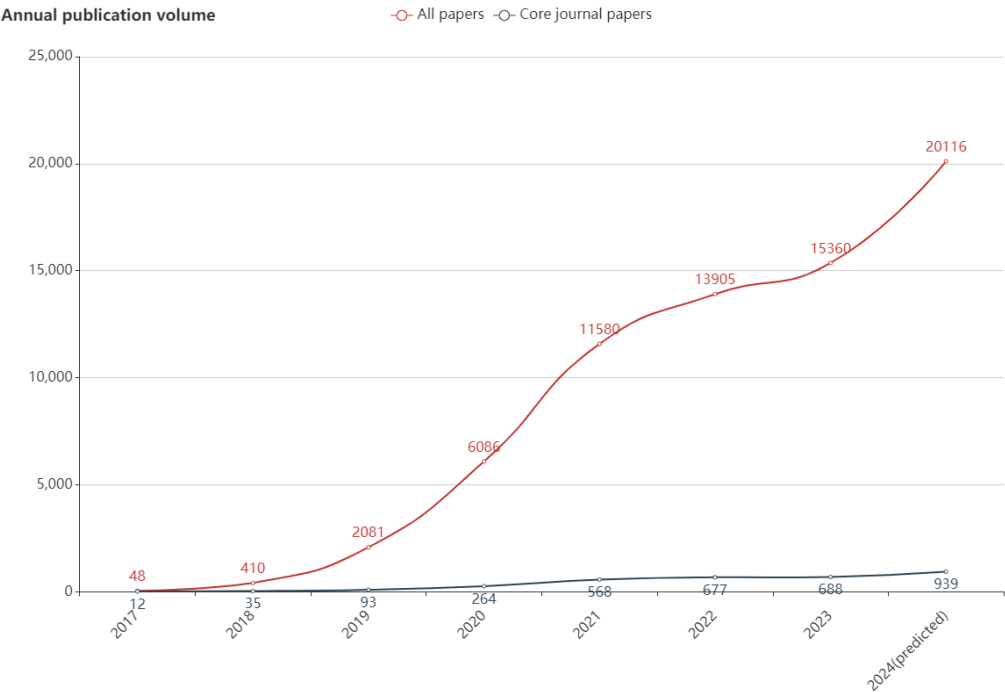


Figure 1: Annual Publication of Literature

Statistical analysis of the journals that publish the literature is beneficial for understanding the publication directions of the journals and assisting authors in submitting their papers. According to statistics, the journal "Education and Teaching Forum" has published the most papers, with 1,657 papers. Figure 2 displays statistical information on the top 10 journals in terms of publication volume.

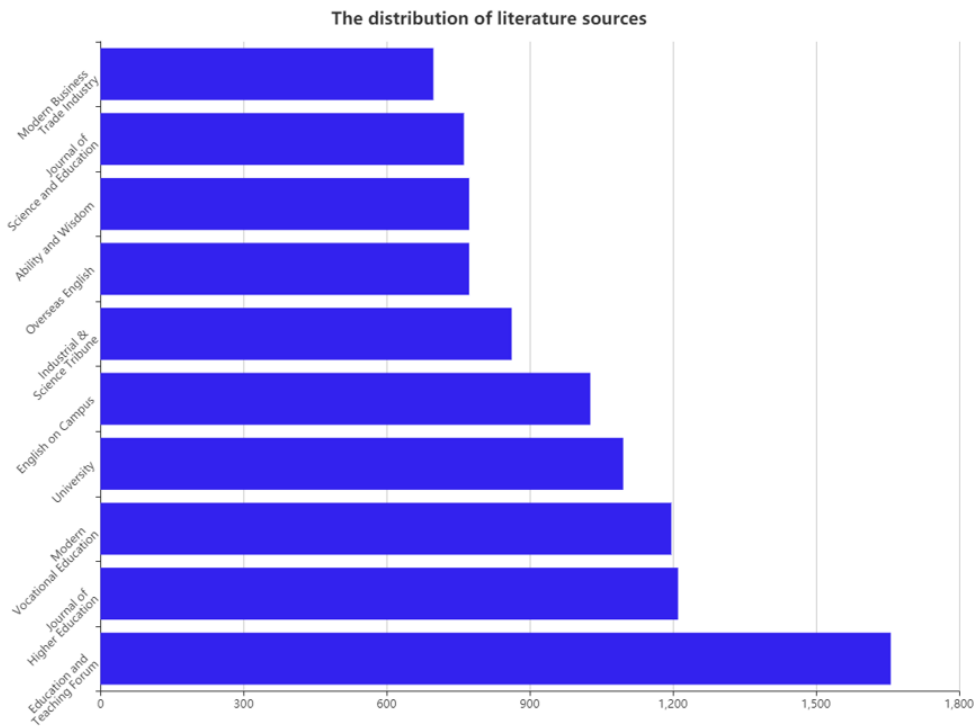


Figure 2: Literature distribution from different sources

Statistical analysis of the institutions affiliated with the authors is beneficial for understanding the identities of the authors and grasping the distribution of research forces. As shown in Figure 3, the publication volume of the top 11 institutions is above 150 papers, indicating that these institutions have strong potential and strength in the research on ideological and political education in curriculum.

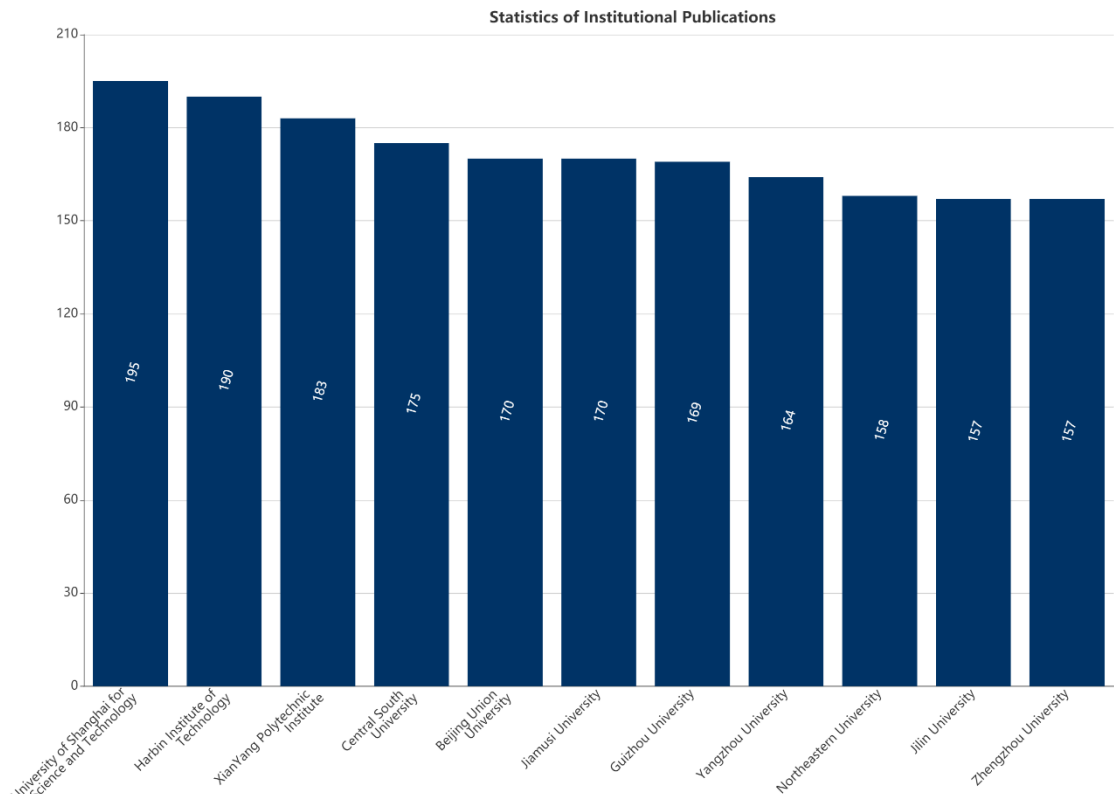


Figure 3: Literature distribution from different organizations

3.2. Analysis of Literature Keywords

Keywords are an important part of academic papers. Just a few words can reflect the theme of the article and express the content features and research areas of the article. Through analysis of high-frequency keywords, research hotspots in the discipline can be identified. Figure 4 displays a word cloud of research keywords on vocational education groups.



Figure 4: Word cloud of keywords

3.3. Citation Analysis of Literature

Citation frequency is an important indicator to measure the quality of a literature piece, reflecting both the degree of attention received by the literature and the hotspots of academic research over a period. This paper has counted the number of papers with different citation frequencies. There are 2 papers with a citation frequency of 100 or more. In terms of content, these papers mainly discuss the connotation, types, and functions of vocational education groups. The Figures 5 shows that most papers have low citation frequencies, with nearly half of the papers having zero citations, indicating that the content of many studies is not understood or recognized by most scholars.

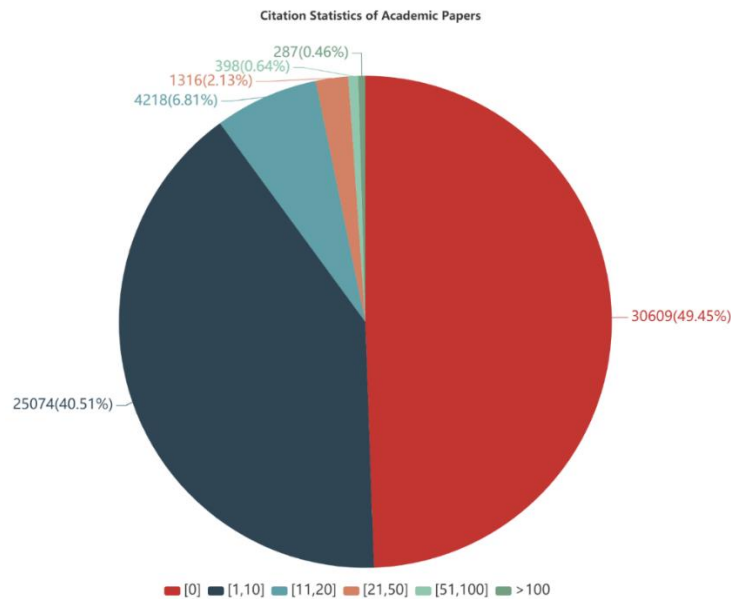


Figure 5: Citation frequency statistics of literatures

4. Conclusion

To deeply quantify the number of literatures on ideological and political education in curriculum and explore its connotations, this paper firstly adopts advanced web crawling technology, targeting authoritative academic platforms such as CNKI, and carefully designs efficient and stable crawler programs to comprehensively and accurately crawl data and information related to literatures on ideological and political education in curriculum. After obtaining the data, this paper stores it securely in a MySQL database, ensuring data integrity and queryability. Subsequently, this paper utilizes quantitative tool libraries such as Pandas and NumPy in Python to conduct in-depth quantitative analysis of the literature data, including multiple dimensions such as the number of literatures, publication time distribution, and research institution distribution, providing a solid data foundation for comprehensively summarizing the research status of ideological and political education in curriculum. Furthermore, this study innovatively adopts a content analysis method to conduct detailed statistics and in-depth analysis of relevant literatures on precise teaching included in CNKI, aiming to explore the internal connections and laws between ideological and political education in curriculum and precise teaching. Through this series of scientific and rigorous research methods, the study not only quantifies the number of literatures on ideological and political education in curriculum but also reveals the academic trends and research hotspots behind them, providing strong data support for in-depth research and practice of ideological and political education in curriculum.

Acknowledgements

This work was financially supported by the funding of the Philosophy and Social Science Research Project of the Jiangsu Higher Education Institutions of China (2022SJYB0804), the Fundamental Computer Education and Teaching Research Project of the Association of Fundamental Computing Education of Chinese Universities (2024-AFCEC-416), and the Excellent Teaching Team for QingLan Project of the Jiangsu Higher Education Institutions of China (Big Data Technology Teaching Team with Shipping Characteristic).

References

- [1] Liu, Xu, Zhao Xiantong, and Hugh Starkey. "Ideological and political education in Chinese Universities: structures and practices." *Asia Pacific Journal of Education* 43.2 (2023): 586-598.
- [2] Gao, Hui-Wen. "Innovation and development of ideological and political education in colleges and universities in the network era." *International Journal of Electrical Engineering & Education* 60.2_suppl (2023): 489-499.
- [3] Khder, Moaiad Ahmad. "Web scraping or web crawling: State of art, techniques, approaches and application." *International Journal of Advances in Soft Computing & Its Applications* 13.3 (2021).
- [4] Kaiying, Deng, Chen Senpeng, and Deng Jingwei. "On optimisation of web crawler system on Scrapy framework." *International Journal of Wireless and Mobile Computing* 18.4 (2020): 332-338.
- [5] Fan, Luqiao, Banxiang Duan, and Jie Gao. "Data Visualization Analysis of JD Mobile Phone Based on Python and ECharts and Bootstrap." *Proceedings of the 2nd International Academic Conference on Blockchain, Information Technology and Smart Finance (ICBIS 2023)*. Atlantis Press, 2023.
- [6] Moresi, Eduardo Amadeu Dutra, Isabel Pinho, and António Pedro Costa. "How to operate literature review through qualitative and quantitative analysis integration?." *World Conference on Qualitative Research*. Cham: Springer International Publishing, 2022: 194-210.