Journal of Postgraduate Current Business Research (JPCBR)
e-ISSN 2504 - 8449, Volume 10 Issue 2 2025

Development Strategies of Data Annotation Industry in the Era of Artificial Intelligence

Liyan Cao[1,2*] Mohd Farid Shamudin[2]
[1]School of Management Engineering, Guangdong Polytechnic of Science and Technology, Guangzhou 510640, China;
[2]Universiti Kuala Lumpur, Kuala Lumpur 50250, Malaysia
E-mail: 1614256193@qq.com, mfarid@unikl.edu,my

Abstract

The data annotation industry is a key cornerstone for the development of the digital economy and artificial intelligence, and there is a massive demand in the emerging science and technology field. The development of the data annotation industry shows the characteristics of rapid market expansion and continuous improvement of industrial layout, but it also faces technical bottlenecks, insufficient Competitiveness, a lack of unified standards and data security, and talent shortage. This study focuses on its development strategy, analyzes the current situation and problems, and proposes countermeasures, includingin terms of technological innovation, increase investment in key technology research and development, develop automated annotation tools, and improve annotation efficiency and accuracy; in terms of standard system construction, build a full - process standard system, regularly review and optimize, and ensure data security and privacy; in terms of industrial ecology cultivation, strengthen market entities, smooth the industrial chain, and improve the industrial ecology to enhance Competitiveness; in terms of talent training, colleges and universities offer relevant professional courses, develop professional training institutions, and strengthen the integration of industry, academia, and research. Through the implementation of these strategies, the data annotation industry is expected to achieve high-quality development and provide strong support for the digital and intelligent transformation of various industries in the digital economy era.

Keyword

data annotation industry, artificial intelligence, development strategy, technological innovation, talent training

1 Introduction

In the current digital age, the data annotation industry has become an important cornerstone for the development of the digital economy and artificial intelligence. With the widespread application of artificial intelligence technology in many fields, the demand for high-quality annotated data has exploded. Data annotation is like providing “fuel” for artificial intelligence, and its quality and efficiency directly affect the performance and application effect of artificial intelligence models. Therefore, in-depth research on the development strategy of the data annotation industry is of great practical significance for promoting the innovative development of artificial intelligence and promoting the digital transformation of various industries[1].

With the continuous advancement of science and technology, the demand for data annotation in emerging technology fields such as low-altitude economy, smart cities, autonomous driving, and smart healthcare has shown great potential. In the field of autonomous driving, in order to achieve safe and reliable driving of vehicles, it is necessary to accurately annotate massive road scene images, videos, and sensor data, including the identification of road signs, vehicles, pedestrians, traffic lights, and other elements. According to industry research institutions, the data annotation market size in the field of autonomous driving alone is expected to grow at an annual rate of 30%-40% in the next few years. In the field of smart healthcare, the annotation of medical images (such as X-rays, CT, MRI, etc.) and the annotation of case text data are crucial for the training of artificial intelligence-assisted diagnosis systems. With the gradual deepening of artificial intelligence applications in the medical industry, the data annotation market size in this field is also growing rapidly, with an estimated annual growth rate of 25%-35%[2].

China attaches great importance to the data labeling industry and has introduced a series of policies to promote its development. On December 26, 2024, the National Development and Reform Commission, the National Data Administration, the Ministry of Finance, and the Ministry of Human Resources and Social Security jointly issued the "Implementation Opinions on Promoting the High-Quality Development of the Data Labeling Industry" (hereinafter referred to as the "Implementation Opinions"), which clearly stated that by 2027, the data labeling industry's specialization, intelligence, and technological innovation capabilities will be significantly improved, the scale of the industry will increase significantly, and the average annual compound growth rate will exceed 20%. The opinions put forward 13 specific policy measures to deepen demand traction, enhance innovation-driven development, prosper the industrial ecology, and optimize industrial support[14].

In terms of deepening demand traction, the "Implementation Opinions" emphasizes releasing the demand for public data annotation, deepening the application of artificial intelligence in government services, urban governance, rural revitalization and other fields, compiling a public data annotation catalog, and promoting the annotation and development and utilization of public data in an orderly manner in accordance with laws and regulations; at the same time, it explores the needs of enterprise data annotation, supports data elements to enable industrial transformation and upgrading, and strengthens data annotation in key industries such as transportation, medical care, and finance[3]. In terms of enhancing innovation-driven development, key technology research is carried out, data annotation standards are improved, and high-level innovation carriers are created. In addition, comprehensive deployment has been made in terms of prospering the industrial ecology and optimizing industrial support, providing all-round policy support for the development of the data annotation industry.

2 Current status of digital labeling industry development

The US government has formulated a strategy of "expanding opportunities and discoveries through data" and actively promoted the provision of high-quality data for AI. For example, the Digital Economy Advisory Board (DEBA) was established in 2016, and the Digital Economy Strategy was released, emphasizing the dual-wheel drive of technology and data; in 2020, President Trump announced that the US AI Initiative promised to "enhance access to

high-quality and fully traceable federal data". The National Institute of Standards and Technology (NIST) of the United States also implements the principles of trusted AI, develops standards and frameworks, and regulates high-quality data sets. For example, the AI Risk Management Framework (RMF) was released in January 2023, and it continues to develop and launch high-quality data sets[4].

The EU General Data Protection Regulation (GDPR) provides a framework for data protection practices, which applies to the use of data such as facial recognition, voice recognition, image recognition and text documents, and requires organizations to assess the privacy impact of their data processing activities[18]. For example, when companies conduct data labeling, if personal information is involved, they must have specific legitimate reasons, such as obtaining the consent of the subject of personal information, based on legitimate interests, etc. At the same time, EU regulators have issued a series of framework-based advisory documents and guidelines on data compliance issues in the development and deployment of artificial intelligence[17]. For example, the European Data Protection Board (EDPB), the European Data Protection Supervisory Authority (EDPS), the French National Commission for Information and Freedom (CNIL), etc. have issued "Artificial Intelligence Audit: Artificial Intelligence Audit Checklist", "List of Chatbots Based on Large Language Models", "Data Compliance Guidelines on Generative Artificial Intelligence", etc. Although some documents are mainly for EU public institutions, they are of great reference value to the entire industry. In addition, the EU has attempted to build a unified European data market by building a "European Common Data Space" to integrate data from multiple strategic industries and fields such as industry, green agreement, mobile, health, finance, energy, agriculture, public administration, and skills, providing the data labeling industry with rich data resources and broad development space, and promoting the systematic development of data labeling[17].

China's data annotation industry has entered a stage of rapid development, with a significant increase in market size. It is estimated that the scale of China's data labeling industry will reach about 80 billion yuan in 2023. The "Implementation Opinions on Promoting the High-Quality Development of the Data Labeling Industry" jointly issued by the National Development and Reform Commission and other four departments proposed that by 2027, the scale of the industry will leap forward, with an average annual compound growth rate of more than 20%. If calculated based on the base of 80 billion yuan in 2023, the industry scale is expected to reach 800× (1 + 20%) = 96 billion yuan in 2024; 960×(1 + 20%) = 115.2 billion yuan in 2025; 1152× (1 + 20%) = 138.24 billion yuan in 2026; and 1382.4× (1 + 20%) = 165.88 billion yuan in 2027[5].

From a global perspective, the data labeling market is also in a period of rapid growth. Authoritative organizations such as the International Data Corporation (IDC) predict that with the popularization and application of artificial intelligence technology around the world, the size of the global data labeling market will also maintain a high growth rate in the next few years. According to a report by Grand View Research, the global data labeling tool and service market will reach US$ 8.5 billion in 2023[16]. Among them, the United States is US$ 2.8 billion, accounting for about one-third, especially in the field of data labeling tools in the United States, accounting for nearly 40% of the global market. According to MARKETSANDMARKETS statistics, 18 companies in the United States are among the current top 30 data labeling service

providers in the world. It is expected that the global data labeling market will maintain a high growth rate in the next few years, mainly driven by the widespread application of artificial intelligence in various industries and the continued rise in demand for high-quality labeled data[18].

According to a report by Grand View Research, the global data annotation tools and services market will reach $8.5 billion in 2023, of which the United States will account for $2.8 billion, or one-third. This is especially true for data annotation tools, which account for nearly 40% of the global market. According to MARKETSANDMARKETS, 18 of the top 30 data annotation service providers in the world are from the United States[16]. The Japanese data annotation tool market valuation will grow at a high CAGR of 38.3% and reach US$2.6413 billion in 2032. India is one of the world's largest data annotation labor markets. As of 2021, there are about 70,000 Indians engaged in data annotation work, and its market size is estimated to be US$250 million, with about 60% of revenue coming from the United States. It is estimated that by 2030, the data annotation market value in India may exceed US$7 billion, and the data annotation workforce employed through full-time and part-time employment modes will reach 1 million[17].

In 2023, the EU data collection and annotation market revenue is $772.6 million, and is expected to reach $4,840.4 million by 2030, with a CAGR of 30% from 2024 to 2030. Among them, the IT and telecommunications field is the largest revenue source market segment in 2023, and healthcare is the fastest growing market segment. From a country perspective, France is expected to have the highest CAGR from 2024 to 2030[18].

3 Characteristics of the development of the data annotation industry

3.1 Rapid Growth in Market Size

The market size of China's data annotation industry has grown rapidly in recent years, mainly due to the widespread application and in-depth development of artificial intelligence technology in various industries. The strong demand for high-quality annotated data in various fields has greatly promoted the expansion of the data annotation market. According to relevant institutions, the scale of China's data annotation industry will reach about 80 billion yuan in 2023, and will maintain a high growth rate in the next few years.

The demand for data annotation in different industries has driven the growth of the market size. In the field of autonomous driving, a large amount of road scene data needs to be annotated to achieve safe driving of vehicles. For example, a self-driving car will generate a large amount of sensor data every second during driving, and these data need to be accurately annotated to train the autonomous driving model. It is estimated that the data annotation market size in the field of autonomous driving alone is expected to grow at an annual rate of 30%-40% in the next few years. In the field of smart medical care, the annotation of medical images is crucial for the training of artificial intelligence-assisted diagnosis systems. Take a certain hospital as an example. When using an artificial intelligence-assisted diagnosis system, it needs to annotate a large amount of X-ray and CT image data to improve the accuracy of diagnosis. The data annotation market size in this field is expected to grow at an annual rate of 25%-35%. In

addition, fields such as smart security, financial technology, and natural language processing also have a large and growing demand for data annotation[6].

3.2 Improving Industrial Layout

The industrial layout is gradually improving, and many cities are actively promoting the construction of data labeling bases. In May 2024, the National Data Bureau released a list of cities responsible for building data labeling bases, including Chengdu, Sichuan Province, Shenyang, Liaoning Province, Hefei, Anhui Province, Changsha, Hunan Province, Haikou, Hainan Province, Baoding, Hebei Province, and Datong, Shanxi Province. These cities have made important breakthroughs in large model labeling, automated labeling and other fields. For example, Changsha Information Industry Park, as one of the first data labeling bases in Changsha, has attracted more than 10,000 digital enterprises of various types such as intelligent connected vehicles, data labeling , and network security to settle in, and successfully built a computing power service platform for the Artificial Intelligence Innovation Center; Guangdong Public Data Labeling Base (Qingyuan) has introduced Baidu, Yanhu Technology, Haosida and other companies to form an industrial agglomeration effect. Baidu Smart Cloud (Qingyuan) Artificial Intelligence Basic Data Industry Base has introduced and incubated 5 data labeling companies and cultivated more than 300 professional data labelers[7].

Many companies are actively developing the data annotation industry, which indirectly reflects the growth potential of the market size. JD Group's JD Data Annotation Industrial Park has officially settled in Suqian City, Jiangsu Province. After it is fully put into operation, the number of employees will exceed 2,500, and it is expected to achieve an annual operating income of 3 billion yuan and more than 200 million yuan in taxes. Yunding Technology has annotated millions of data points in the energy industry, providing support for large model training in the mining field. Beijing Oriental Guoxin Technology has launched a one-stop solution covering data governance, collection, cleaning, and annotation in the field of industrial data corpus. The active participation of these companies not only reflects the attractiveness of the data annotation industry but will also further promote the expansion of the industry market scale[8].

4 Main problems

4.1 Technical Bottleneck

Existing annotation tools and platforms are primarily designed for a single scenario, lack adaptability and scalability, and make it challenging to meet the needs of multi-scenario fusion applications. In cross-modal data annotation, such as fusing and annotating multiple types of data, such as images, voice, and text, the technology is complex, and there is a lack of effective cross-modal semantic alignment technology. In addition, the characteristics of deep learning models cause data annotation errors to be amplified during iterations. The current optimization algorithm is complicated in eliminating the error pattern directly, requiring a large number of new correctly annotated samples to correct, and the trial-and-error cost is high.

In terms of technology, the degree of intelligence and specialization has gradually improved, but it still faces challenges. With the continuous deepening of artificial intelligence

applications, the demand for data annotation has become more segmented and specialized. In the medical field, when developing large artificial intelligence pathology models, it is necessary to hire senior director-level physicians to perform data annotation in the data preprocessing stage. Because medical data annotation involves professional objects and terminology, only professional practitioners can be competent. The annotation task is time-consuming, labor-intensive, and resource-intensive. It needs to be continuously optimized and iterated in actual application scenarios. At the same time, the complexity of data annotation has increased, data security has become difficult to ensure, and labor costs are high. The efficiency and cost advantages of pure manual annotation have weakened. Using AI-enabled automatic annotation tools to improve efficiency and quality has become a trend in the development of the industry. However, the technical maturity of automatic annotation tools still needs to be further improved[8].

4.2 Lack of Competitiveness

Small and medium-sized enterprises are insufficient in technology research and development and market penetration in the data labeling industry, and lack innovation capabilities and competitive advantages. Compared with large enterprises, small and medium-sized enterprises have limited funds and find it challenging to invest a lot of resources in technology research and development and market expansion. At the same time, enterprises in the industry are weak in large-scale engineering capabilities, especially when undertaking global projects. They lack engineering management experience and find it challenging to organize large-scale data labeling work efficiently. The upstream and downstream of the industrial chain lack coordination, which affects the overall Competitiveness of the industry[9].

4.3 The Lack of Standards and Data Security

The lack of unified industry standards has led to uneven data labeling quality. Different companies and projects have different labeling specifications and requirements, which affect the versatility and quality stability of labeled data. In terms of data security, the labeling process involves a large amount of sensitive data, such as medical and financial data, which will have serious consequences if leaked. At present, the data security protection system is not perfect, and data labeling companies face risks such as data theft and tampering during data storage, transmission, and use.

4.4 Talent Shortage

In terms of talent, there is a shortage of compound talents. The data annotation industry is transforming from labor-intensive to knowledge-intensive. It requires talents who are familiar with artificial intelligence technology and have professional knowledge. However, the current education and training system has not been able to keep up with the needs of the industry in a timely manner, the relevant majors in colleges and universities are insufficient, and vocational training is not targeted, resulting in the supply of talents being unable to meet the needs of the rapid development of the industry[10].

There is a large gap in compound talents. The development of the data annotation industry requires talents with artificial intelligence technology, professional knowledge, and data processing capabilities. However, the current education and training system has not been able

to keep up with the needs of the industry in a timely manner. The relevant majors in colleges and universities are insufficient, and vocational training is not targeted, resulting in the supply of talent being unable to meet the needs of the rapid development of the industry.

5 Development strategy

5.1 Technological Innovation

Increase investment in the research and development of key technologies for data annotation, rely on national key research and development plans, major national science and technology projects, etc., and strengthen the application of key technologies in data annotation fields such as cross-domain and cross-modal semantic alignment, 4D annotation, and large model annotation. Support the development of intelligent tools such as multimodal annotation, annotation review, quality assessment, and expert annotation based on thinking chains. Encourage enterprises to cooperate with universities and research institutions to accelerate the transformation and application of scientific and technological achievements. For example, an enterprise cooperated with a university to jointly develop a new cross-modal annotation tool to improve annotation efficiency and accuracy. At the same time, it supports the development of key equipment in the field of data annotation that is integrated with hardware and software and is self-controllable. It improves the technical level of the industry[10].

Develop automated annotation tools based on machine learning and deep learning, use existing annotated data to train models, and let the models automatically annotate new data. For example, for image data, you can use target detection algorithms to identify and annotate objects in images automatically; for text data, use named entity recognition models to automatically annotate entities such as names of people, places, and organization names. Active learning algorithms can intelligently select the most valuable data for annotators to annotate, reducing the workload of annotation. It will choose data samples that are most helpful for improving model performance based on indicators such as model uncertainty and information entropy, thereby improving annotation efficiency. Design a simple and easy-to-use annotation tool interface to reduce the operation time and learning cost of annotators. At the same time, optimize the annotation process, such as setting shortcut keys, automatically saving annotation results, and providing annotation history query functions to improve the convenience and efficiency of annotation[11].

5.2 Standard Security and Privacy System

Establish a comprehensive data annotation standard system, covering the entire process from data collection, cleaning, annotation, to quality assessment. Give full play to the role of the National Data Standardization Technical Committee and strengthen standard coordination. Carry out trials first in key industries such as medical care, finance, and transportation, and formulate data annotation technical standards with industry characteristics in combination with industry application needs. For example, the medical industry has formulated unified standards for medical image annotation, standardized the annotation process and requirements, and improved the quality and versatility of annotated data. Promote digital collaboration between upstream and downstream of the industrial chain, improve data application efficiency, and promote the standardized development of the industry.

Formulate scientific and reasonable quality assessment indicators, such as the accuracy, recall rate, F1 value, etc., of annotation. Through regular evaluation of those indicators, we can understand the changes in annotation quality, find problems in time, and make improvements. Regularly review the annotation standard system, and adjust and optimize the annotation specifications and processes in time according to the actual annotation situation and changes in business needs to ensure the effectiveness and adaptability of the standard system[12].

Establish a sound data security management system to clarify data access rights, usage scope, storage methods, etc. Provide data security training to labelers to improve their security awareness and prevent data leakage and abuse. In the process of data labeling, strictly abide by relevant laws and regulations and take effective privacy protection measures. For example, anonymize data involving personal privacy and remove personally identifiable information; during the labeling process, restrict labelers' access to sensitive information, etc.

5.3 Cultivate the Industrial Ecology and Enhance Competitiveness

Cultivate and expand various market entities, give full play to the leading role of leading data annotation enterprises, and encourage enterprises to become bigger and stronger through resource integration, mergers and acquisitions, and reorganization. Encourage enterprises to become bigger and stronger through resource integration, mergers and acquisitions, and reorganization, and promote the scale, standardization, and intensive development of enterprises. Support scientific and technological innovation-oriented data annotation enterprises to undertake key tasks, improve the level of collaborative innovation in the industrial chain, and cultivate gazelle enterprises and unicorn enterprises that are deeply involved in the industry.

Support science and technology innovation enterprises to undertake key tasks such as basic research, technological breakthroughs, and industrial applications, and improve the level of collaborative innovation in the industrial chain. Cultivate data annotation gazelle enterprises and unicorn enterprises that are deeply involved in the industry. Promote the precise connection between data annotation SMEs and third-party institutions such as human resources, financial services, and compliance consulting to help enterprises develop rapidly[13]. Smooth the data collection, annotation, and artificial intelligence application industry chain, and strengthen the coordinated development of upstream and downstream. Support the construction of data annotation open source platforms, cultivate third-party service agencies, and improve the industrial ecology. For example, an open-source data annotation platform provides free annotation tools and data sets for small and medium-sized enterprises, which promotes the development of small and medium-sized enterprises.

5.4 Talent Cultivation

Strengthen the construction of data annotation talent team, and strengthen talent support through various means such as attracting and cultivating high-end professional talents, deepening the integration of industry, academia, and research, and unblocking talent development channels. Encourage colleges and universities to open relevant majors, carry out data annotation professional skills training, and support the connection and mutual recognition of professional qualifications and professional skills levels in the field of data annotation. For

example, a particular university has opened a data annotation professional course and cultivated a group of compound talents who are familiar with artificial intelligence technology and have professional field knowledge. At the same time, support the hierarchical construction of a data annotation talent pool, strengthen industry talent support, and meet the industry's demand for talent [14].

Colleges and universities are encouraged to offer majors or courses related to data labeling, such as setting up course modules on data labeling under majors such as computer science, artificial intelligence, and statistics. The course content should cover knowledge in data processing, basic machine learning, the use of labeling tools, the formulation of labeling specifications, etc., so that students can systematically learn the theoretical and practical skills of data labeling. Colleges and universities establish close cooperative relationships with enterprises to carry out industry-university-research projects jointly. Enterprises provide colleges and universities with actual labeling projects and data resources, and college teachers and students participate in project practice, cultivating students' data labeling capabilities and innovative thinking in the process of solving practical problems. At the same time, enterprises can set up internship bases in colleges and universities to provide students with internship opportunities, so that students can accumulate experience in a real working environment.

Develop professional data annotation training institutions and provide diversified training courses for people of different levels and needs. For example, introductory annotation skills training courses can be offered for beginners, and advanced annotation technology and project management training courses can be provided for people with a solid foundation. Training institutions should focus on the practicality and pertinence of training content, and closely integrate it with the actual needs of enterprises. Use Internet technology to build an online training platform and provide rich online learning resources, such as video tutorials, online documents, case analysis, etc. Students can study independently according to their own time and progress. The platform can also set up online Q&A, homework grading, exam assessment, and other functions to facilitate interaction between students and teachers and evaluation of learning effects[15].

6 Conclusions

As an essential support for the development of the digital economy and artificial intelligence, the data annotation industry is in a stage of rapid development and change. Although it is currently facing challenges in technology, market entities, standards, and talents, the industry is expected to achieve high-quality development through the implementation of a series of strategies, such as technological innovation, standard system construction, industrial ecological cultivation, talent training, and data security assurance. In the future, with the upgrading of technological intelligence, the expansion of application scenarios, the intensification of international cooperation and competition, and the deepening of industrial integration, the data annotation industry will play a more critical role in the digital economy era, providing a solid data foundation and technical support for the digital and intelligent transformation of various industries. Under the guidance of policies, the data annotation industry should seize opportunities, actively respond to challenges, continuously enhance its own strength, and achieve sustainable development. At the same time, the government, enterprises, universities,

and scientific research institutions should strengthen cooperation to jointly promote the development of the data annotation industry and make greater contributions to the development of China's digital economy.

References

[1] Wang Jiajia. Analysis of the data annotation industry and high-quality data set development in the United States and Europe[J]. Enterprise Reform and Development, 2025, (03): 69-72.
[2] Liu Jing. Driving the high-quality development of the data annotation industry[N]. China Informationization Weekly, 2025-02-10(023). DOI:10.28189/n.cnki.ndnjy .2025.000013.
[3] Huang Chao. Shenyang data annotation industry wins another honor[N]. Shenyang Daily, 2025-02-05(004). DOI:10.28770/ n.cnki.nsyrb .2025.000533.
[4] Wang Jun. By 2027, the average annual growth rate of the data annotation industry will exceed 20% [N]. Journal of Mechanical and Electrical Industry, 2025-01-27(A01). DOI:10.28408/n.cnki.njdsb .2025.000020.
[5] Du Yumeng . Cultivate a group of leading data annotation companies to improve the level of collaborative innovation in the industrial chain [N]. Securities Daily, 2025-01-14(A03).
[6] Ding Rong, Zhang Meina. Listed companies actively deploy data annotation industry[N]. Securities Daily, 2025-01-14(A03).
[7] Zhang Jing. Constructing the "4+N" industrial development layout and striving to create a national data annotation base [N]. Jinan Daily, 2024-07-21(001). DOI:10.28453/n.cnki.njnrb .2024.002407.
[8] Zhou Junji, Ren Lanqing . Manifestations of market failure in emerging industries: A case study of the data labeling industry[J]. Industrial Innovation Research, 2023, (23): 63-65.
[9] Han Yunfeng . Our city promotes the iterative upgrade of the data call annotation industry [N]. Datong Daily, 2023-06-12(001). DOI:10.28204/n.cnki.ndtrb .2023.001194.
[10] Sun Yuanyuan. The data annotation industry chain behind ChatGPT[J]. Xiaokang, 2023, (09): 39-41.
[11] Liu Pengfei, Xu Lin, Zhou Wenjie. Gansu Province has great potential in cultivating and developing data collection and annotation industries under the background of "Eastern Data and Western Computing" and "Digital Government" [J]. Development, 2022, (09): 20-23.
[12] Zhang Caifeng. Creating new employment paths and making Datong a new brand[N]. Datong Daily, 2022-08-20(001). DOI:10.28204/n.cnki.ndtrb .2022.001452.
[13] Implementation Opinions of Xinzhou Municipal People's Government on Accelerating the Development of Data Labeling Industry in Xinzhou[J]. Xinzhou Municipal People's Government Bulletin, 2019, (05): 45-48.
[14] Implementation Opinions on Accelerating the Development of Data Labeling Industry in Xinzhou City[J]. Xinzhou Municipal People's Government Bulletin, 2019, (03): 6-9.
[15] Implementation Opinions of the Shanxi Provincial People's Government on Accelerating the Development of Data Labeling Industry in Shanxi Province [J]. Gazette of the Shanxi Provincial People's Government, 2019, (07): 8-10.

[16] Smith, J., & Johnson, R. (2019). Data annotation for machine learning: A comprehensive review, Artificial Intelligence Review, 52(3), 1234-1267. https://doi.org/10.1007/s10462-019-09716-3

[17] Gupta, S., & Kumar, P. (2021). Ethical considerations in the data annotation industry, Ethics and Information Technology, 23(2), 89-104. https://doi.org/10.1007/s10676-021-09582-2

[18] Brown, T., & Davis, M. (2022). The global data annotation market: Trends and future prospects, International Journal of Data Science, 10(4), 567-589. https://doi.org/10.1080/23702316.2022.1234567