DOI: 10.53469/wjimt.2025.08(09).01

The Current Status and Development Trends of Big Data Processing Technology

Junquan Xiao, Xuwei Chen

Zhejiang Bangsheng Technology Co., Ltd. Hangzhou 310000, Zhejiang

Abstract: ***With the rapid development of information technology, big data has become an important force driving social progress and economic development. This article aims to explore the current status and development trends of big data processing technology, and conduct in-depth analysis from multiple dimensions such as technical architecture, application fields, security and privacy protection, and policy support, in order to provide reference and inspiration for the future development of big data technology.***

Keywords: Big data processing technology; Present situation; Development trend.

1. INTRODUCTION

The rise of big data technology marks a new stage of development in the information age. Big data technology provides strong support for government decision-making, business innovation, social services, etc. by collecting, storing, processing, and analyzing massive amounts of data. This article will start from the current situation, analyze the key elements of big data processing technology, and look forward to its future development trends. In the realm of privacy-preserving technologies, Li, Lin, and Zhang (2025) developed a framework combining federated learning and differential privacy for advertising personalization[1]. System optimization approaches include Tu's (2025) modeling-driven neural architecture search for smart regression detection[2], Xie and Liu's (2025) multimodal sentiment analysis for recruitment processing[3], and Zhu's (2025) LLM-based backbone for enhancing small business platform stability[4]. Zhang Yuhan (2025) further contributed to business applications through reinforcement learning for automated ad campaign optimization[5]. Industry-specific AI applications are extensively explored, with Tan (2024) analyzing AI trends in automotive production[6], Zhuang (2025) examining digital transformation in real estate marketing[7], and Han and Dou (2025) proposing a hierarchical graph attention network for user recommendation[8]. Advanced learning techniques are represented by Yang et al.'s (2025) RLHF fine-tuning for conversational recommenders[9] and Zhang Jingbo et al.'s (2025) AI-driven sales forecasting in gaming[10]. Yang Yifan (2025) focused on web performance improvement through component-based architecture[11], while Cheng et al. (2025) investigated the relationship between executive human capital and stock volatility[12]. Computer vision research includes Chen et al.'s (2022) gaze-estimation based object referring[13] and Tong et al.'s (2024) hybrid framework for credit approval prediction[14]. Tian et al. (2025) introduced cross-attention multi-task learning for digital advertising[15], and Chen Yinda et al. (2023) developed vision-language pretraining for medical segmentation[16]. Financial and environmental applications feature Zhang Zongzhen et al.'s (2025) deep learning approach for carbon market forecasting[17]. The domain generalization field is advanced by Peng, Zheng, and Chen's (2024) dual-augmentor framework for 3D pose estimation[18], Pinyoanuntapong et al.'s (2023) self-aligned domain adaptation for gait recognition[19], and Zheng et al.'s (2025) motion-aware diffusion framework for human mesh recovery[20].

2. THE CURRENT STATUS OF BIG DATA PROCESSING TECHNOLOGIES

2.1 Evolution of Technical Architecture

Currently, the architecture of big data processing technology has undergone significant evolution and is gradually maturing. This technological architecture covers multiple key aspects such as data collection, storage, processing, and analysis, providing a comprehensive solution for processing large-scale and highly complex data. In terms of data collection, various efficient data capture, transmission, and preprocessing techniques are widely used to ensure the integrity and accuracy of data. In the field of data storage, distributed file systems such as HDFS (Hadoop Distributed File System) and columnar storage databases such as HBase, Cassandra, and other NoSQL databases have been widely used, which can efficiently store and process massive amounts of data. Data processing is the core component of big data technology architecture. At present, distributed processing frameworks such as Hadoop MapReduce, Apache Spark, and stream processing frameworks such as Apache

Storm have become industry standard. These frameworks provide powerful data processing capabilities, capable of handling various complex data processing tasks such as batch processing, real-time stream processing, etc. In terms of data analysis, various data mining, machine learning, and artificial intelligence algorithms are widely used in big data processing. These technologies can extract valuable information from massive amounts of data, providing powerful data support for business decision-making and strategic planning of enterprises. The current architecture of big data processing technology already has relatively complete functions and performance, which can efficiently process and analyze large-scale and high complexity data. The maturity and development of this technological architecture provide a solid foundation and support for the application of big data technology in various fields.

2.2 Wide Expansion of Application Fields

Big data technology has been widely applied in multiple fields, including government services, commercial applications, healthcare, and smart cities, bringing significant changes and improvements to these areas. In the field of government services, big data technology plays a crucial role. Through smart transportation projects, the government can analyze real-time traffic flow, road conditions, and other information, accurately schedule vehicles and traffic signals, and effectively alleviate traffic congestion problems. At the same time, the smart healthcare project utilizes big data technology to optimize the allocation of medical resources and intelligentize medical services, improving medical efficiency and service quality. In the field of commercial applications, big data technology has also demonstrated enormous potential. By utilizing big data technology, enterprises can deeply mine and analyze massive amounts of data, thereby achieving precise marketing. By analyzing consumer behavior, preferences, and other data, companies can more accurately target their audience, develop personalized marketing strategies, and enhance market competitiveness. In addition, big data technology plays an important role in risk assessment, helping enterprises identify potential market and business risks and providing strong support for their stable development. In the field of healthcare, big data technology has brought innovation to medical research and clinical services. Through the analysis and mining of massive medical data, researchers can discover information such as the incidence patterns and treatment effects of diseases, providing scientific basis for new drug development and disease treatment. Meanwhile, big data technology also plays an important role in clinical services, such as assisting doctors in disease diagnosis and developing personalized treatment plans, improving the accuracy and effectiveness of medical services. In the field of smart cities, big data technology also plays an important role. By integrating and analyzing various data in urban operation, such as transportation, environment, energy, etc., smart cities can achieve optimized allocation and efficient utilization of resources, improve the intelligence level of urban management, and provide citizens with a more convenient and comfortable living environment.

2.3 Challenges of Data Security and Privacy Protection

With the rapid increase in data volume, data security and privacy protection issues have become increasingly prominent, becoming key issues that urgently need to be addressed in the development of big data technology. At present, big data technology has made certain progress in encryption technology, access control, and other aspects, providing a certain degree of guarantee for data security. However, in the face of increasingly complex data environments and constantly evolving security threats, it is still necessary to further strengthen data governance and compliance construction to ensure the secure use and privacy protection of data. In terms of encryption technology, big data technology adopts various encryption algorithms and protocols, such as AES, RSA, etc., to encrypt storage and transmission of sensitive data, effectively preventing data leakage and illegal access. Meanwhile, access control technology has also been widely applied, ensuring that only authorized users can access and operate specific data by setting permissions and access rules. However, despite these advances, big data security still faces many challenges. On the one hand, the surge in data volume has made data management and protection more complex and difficult. On the other hand, frequent security incidents such as hacker attacks and data breaches pose serious threats to data security. Therefore, further strengthening data governance and compliance construction is particularly important. To address these challenges, it is necessary to establish a sound data governance system, clarify the ownership, usage rights, and responsibilities of data, and ensure the legal and compliant use of data. At the same time, it is necessary to strengthen the research and application of data security technologies, such as data anonymization and differential privacy, in order to protect data privacy while achieving effective utilization of data. In addition, it is necessary to strengthen the formulation and implementation of laws and regulations, strictly supervise and manage data security, and ensure the safe use and privacy protection of data.

3. DEVELOPMENT TRENDS OF BIG DATA PROCESSING TECHNOLOGIES

3.1 Deep integration of artificial intelligence and big data

With the continuous advancement of technology, the deep integration of artificial intelligence and big data is becoming an important trend in the development of big data processing technology. This integration will further promote the intelligence and automation of big data processing, providing enterprises with more efficient and accurate data analysis and decision support. Natural language processing (NLP), as an important component of artificial intelligence technology, will play a crucial role in big data processing. Through NLP technology, enterprises can automatically extract and analyze key information from textual data, such as customer feedback, market trends, etc., in order to gain a more comprehensive understanding of market dynamics and customer needs. This will greatly improve the efficiency and accuracy of data processing, providing strong support for enterprises to formulate more accurate market strategies. Machine learning technology will also play an important role in big data processing. Through machine learning algorithms, enterprises can automatically classify, cluster, and predict massive amounts of data, discovering hidden patterns and associations in the data. This will help companies gain a deeper understanding of business operations, identify potential market opportunities and risks, and provide scientific basis for strategic planning and business decision-making. The deep integration of artificial intelligence and big data will further promote the development of data governance and compliance. With the help of artificial intelligence technology, enterprises can manage and protect data more efficiently, ensuring the legal and compliant use of data. At the same time, artificial intelligence technology can also help enterprises detect anomalies and violations in data, improving the security and credibility of data.

3.2 Improvement of Data Governance System

With the continuous deepening and extensive expansion of big data applications, the construction and improvement of data governance systems have become an important task that enterprises cannot ignore. In the face of increasingly complex data environments and growing data processing demands, building a sound and efficient data governance system is of great significance for enhancing enterprise data management capabilities, ensuring data security, and promoting data sharing and utilization. In the future, data governance will place greater emphasis on balancing data quality, data security, and data sharing and utilization efficiency. In terms of data quality, enterprises will strive to improve the accuracy, completeness, and consistency of data to ensure the reliability and effectiveness of data analysis results. To this end, enterprises will establish a comprehensive data quality monitoring mechanism, regularly inspect and clean data, and promptly identify and correct data issues. In terms of data security, enterprises will strengthen security measures such as data encryption and access control to prevent data leakage and illegal access. At the same time, companies will also pay attention to data privacy protection and ensure that personal data and sensitive information are handled properly. In terms of data sharing and utilization efficiency, enterprises will promote the construction and improvement of data sharing platforms, and facilitate the flow and sharing of data between different departments and businesses. By optimizing the data sharing process and improving data sharing efficiency, enterprises will be able to better utilize data resources to support business decision-making and innovative development.

3.3 Cross domain data management and sharing

Cross domain data management is gradually becoming an important trend in the field of big data processing. With the continuous increase in data volume and the diversification of data sources, traditional single domain or single system data management methods are no longer able to meet current needs. Cross domain data management aims to break down data silos, achieve data sharing and collaboration between different fields and systems, and maximize the value of data. Implementing cross domain data management can bring many benefits to enterprises and organizations. Firstly, it can facilitate the flow and sharing of data between different departments and teams, improving data utilization efficiency. Secondly, cross domain data management helps integrate data from different sources to form a more comprehensive and accurate dataset, providing stronger support for data analysis and decision-making. Finally, through cross domain data management, enterprises and organizations can better respond to market changes and business challenges, enhance overall competitiveness and innovation capabilities. However, cross domain data management also faces some challenges. Among them, network latency is an important issue. In cross domain data management, data needs to be transmitted and shared in different network environments, which may lead to increased network latency and affect the real-time and accuracy of data. In

addition, data heterogeneity is also a problem that needs to be addressed. Data from different fields and systems may have differences in format, standards, quality, and other aspects, which can pose difficulties for data integration and analysis. To address these challenges, continuous technological innovation is needed. For example, more efficient data transmission protocols and technologies can be developed to reduce the impact of network latency. At the same time, data integration and cleaning tools can also be developed to address data heterogeneity issues and improve data quality and consistency. Cross domain data management is an important trend in the field of big data processing. Although it faces some challenges, through technological innovation and continuous improvement, it is expected to maximize the value of data and bring more opportunities and development space to enterprises and organizations.

3.4 Development of High Performance Computing and Storage Technologies

Faced with the demands of large-scale and highly complex data processing, high-performance computing and storage technologies are constantly evolving to meet the growing challenges. In this context, new hardware such as GPU (graphics processing unit), TPU (tensor processor) and other accelerators have been widely used in big data processing systems, aiming to significantly improve data processing performance. Accelerators such as GPUs and TPUs perform well in parallel computing, capable of processing large-scale datasets and executing complex mathematical and logical operations. Integrating these accelerators into big data processing systems can greatly accelerate the data processing and analysis process, and improve the overall performance of the system. For example, in applications such as machine learning and deep learning, GPUs and TPUs can significantly accelerate the speed of model training and inference, thereby supporting more complex and accurate data analysis. At the same time, the development of storage technology has also attracted much attention. To ensure the persistence, reliability, and scalability of data, storage systems need to constantly innovate. On the one hand, new storage media such as flash memory and phase-change memory are gradually replacing traditional hard drives, providing higher read and write speeds and better data persistence. On the other hand, the rise of distributed storage systems and cloud storage technology enables redundant storage of data across multiple nodes and geographic locations, improving data reliability and availability. In addition, storage technology is also focusing on the scalability of data. With the continuous growth of data volume, storage systems need to be able to seamlessly expand capacity to meet ever-changing storage requirements. This requires storage systems to have high scalability and flexibility, capable of supporting different data types and access modes [4]. The development of high-performance computing and storage technology is an important component of the evolution of big data processing technology. By continuously introducing new hardware and storage technologies, big data processing systems will be able to better respond to large-scale and highly complex data processing needs, providing enterprises and organizations with more efficient and reliable data processing and analysis capabilities.

3.5 Maturity of Industrial Ecology

With the continuous maturity of big data technology and the continuous expansion of application fields, the big data industry ecosystem is gradually moving towards perfection. In this process, various links in the industrial chain will strengthen cooperation and collaboration, jointly building a complete and collaborative ecosystem. In the big data industry ecosystem, various links such as data generation, collection, storage, processing, analysis, and application are closely connected, forming an organic whole. In order to ensure the smooth flow and efficient utilization of data, enterprises in the industrial chain need to strengthen cooperation, jointly develop data exchange standards, establish data sharing mechanisms, and optimize data processing processes. This close collaborative relationship will help enhance the competitiveness and innovation capabilities of the entire industry chain. At the same time, the position of the open source community in the construction of the big data software and hardware ecosystem will continue to strengthen. As an important driving force for technological innovation, the open source community gathers the wisdom and strength of numerous developers and enterprises. In the field of big data, open source communities not only provide rich software and hardware resources, but also promote technological innovation and standardization processes. Through collaboration and sharing in open source communities, enterprises can quickly access the latest technological achievements, reduce research and development costs, and accelerate product iteration and upgrading. In addition, with the maturity of the big data industry ecosystem, more professional service providers and solution providers will emerge. These enterprises will focus on the research and development, application, and promotion of big data technology, providing customized big data solutions for various industries. Their joining will further enrich the big data industry ecosystem and promote the popularization and application of big data technology. The maturity of the big data industry ecosystem is an important trend in the development of big data technology. By strengthening industry chain collaboration,

leveraging the role of open source communities, and cultivating professional service providers, the big data industry ecosystem will continue to grow and improve, injecting new vitality into economic and social development.

4. CONCLUSION

Big data processing technology is developing at an unprecedented speed, with its application areas constantly expanding and its technical architecture becoming increasingly mature. In the future, with the promotion of trends such as real-time data analysis, integration of artificial intelligence, and improvement of data governance systems, big data technology will inject new vitality into economic and social development. At the same time, policy support and the maturity of industrial ecology will also provide strong guarantees for the continuous innovation and development of big data technology.

REFERENCES

[1] Li, X., Lin, Y., & Zhang, Y. (2025). A Privacy-Preserving Framework for Advertising Personalization Incorporating Federated Learning and Differential Privacy. arXiv preprint arXiv:2507.12098.
[2] Tu, Tongwei. "ProtoMind: Modeling Driven NAS and SIP Message Sequence Modeling for Smart Regression Detection." (2025).
[3] Xie, Minhui, and Boyan Liu. "EvalNet: Sentiment Analysis and Multimodal Data Fusion for Recruitment Interview Processing." (2025).
[4] Zhu, Bingxin. "ReliBridge: Scalable LLM-Based Backbone for Small Business Platform Stability." (2025).
[5] Zhang, Yuhan. "Learning to Advertise: Reinforcement Learning for Automated Ad Campaign Optimization for Small Businesses." (2025).
[6] Tan, C. (2024). The Application and Development Trends of Artificial Intelligence Technology in Automotive Production. Artificial Intelligence Technology Research, 2(5).
[7] Zhuang, R. (2025). Evolutionary Logic and Theoretical Construction of Real Estate Marketing Strategies under Digital Transformation. Economics and Management Innovation, 2(2), 117-124.
[8] Han, X., & Dou, X. (2025). User recommendation method integrating hierarchical graph attention network with multimodal knowledge graph. Frontiers in Neurorobotics, 19, 1587973.
[9] Yang, Zhongheng, Aijia Sun, Yushang Zhao, Yinuo Yang, Dannier Li, and Chengrui Zhou. "RLHF Fine-Tuning of LLMs for Alignment with Implicit User Feedback in Conversational Recommenders." arXiv preprint arXiv:2508.05289 (2025).
[10] Zhang, Jingbo, et al. "AI-Driven Sales Forecasting in the Gaming Industry: Machine Learning-Based Advertising Market Trend Analysis and Key Feature Mining." (2025).
[11] Yang, Yifan. "Web Front-End Application Performance Improvement Method Based on Component-Based Architecture." International Journal of Engineering Advances 2.2 (2025): 24-30.
[12] Cheng, Ying, et al. "Executive Human Capital Premium and Corporate Stock Price Volatility." Finance Research Letters (2025): 108278.
[13] Chen, J., Zhang, X., Wu, Y., Ghosh, S., Natarajan, P., Chang, S. F., & Allebach, J. (2022). One-stage object referring with gaze estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5021-5030).
[14] Tong, Kejian, et al. "An Integrated Machine Learning and Deep Learning Framework for Credit Card Approval Prediction." 2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS). IEEE, 2024.
[15] Q. Tian, D. Zou, Y. Han and X. Li, "A Business Intelligence Innovative Approach to Ad Recall: Cross-Attention Multi-Task Learning for Digital Advertising," 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Shenzhen, China, 2025, pp. 1249-1253, doi: 10.1109/AINIT65432.2025.11035473.
[16] Chen, Yinda, et al. "Generative text-guided 3d vision-language pretraining for unified medical image segmentation." arXiv preprint arXiv:2306.04811 (2023).
[17] Zhang, Zongzhen, Qianwei Li, and Runlong Li. "Leveraging Deep Learning for Carbon Market Price Forecasting and Risk Evaluation in Green Finance Under Climate Change." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-27.
[18] Peng, Qucheng, Ce Zheng, and Chen Chen. "A Dual-Augmentor Framework for Domain Generalization in 3D Human Pose Estimation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

[19] Pinyoanuntapong, Ekkasit, et al. "Gaitsada: Self-aligned domain adaptation for mmwave gait recognition." 2023 IEEE 20th International Conference on Mobile Ad Hoc and Smart Systems (MASS). IEEE, 2023.
[20] Zheng, Ce, et al. "Diffmesh: A motion-aware diffusion framework for human mesh recovery from videos." 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). IEEE, 2025.