AIRA
ARTIFICIAL INTELLIGENCE RESEARCH AND APPLIED LEARNING

Vol. 1, No. 2, July 2022
AIRA: Artificial Intelligence Research and Applied Learning
https://journal.dinamikapublika.id/index.php/aira

OPTIMIZING BIG DATA PROCESSING THROUGH ARTIFICIAL INTELLIGENCE: A SYSTEMATIC LITERATURE REVIEW

Vannes Winata[1]

[1]Digital Business Department, Faculty of Social Sciences and Humanities, Universitas Bunda Mulia, Indonesia

E-mail: [1)] s35220112@student.ubm.ac.id

ABSTRACT

The exponential growth of big data presents significant challenges for traditional processing methods, driving the need for more efficient solutions. This systematic literature review explores the integration of artificial intelligence (AI) in big data processing, examining its transformative role in enhancing speed, scalability, and insight extraction. The review, conducted by following systematic guidelines, synthesizes findings from key academic databases such as IEEE Xplore, ACM Digital Library, and Scopus. We identify and discuss the primary advantages of this convergence, including the automation of data lifecycle, enhanced predictive analytics, and improved decision-making. The paper also highlights critical challenges such as data privacy, algorithmic bias, and resource requirements. By leveraging AI's capabilities, organizations can unlock unprecedented opportunities across various domains, from healthcare and finance to retail and smart cities. The review concludes that while the synergy between AI and big data offers immense potential, its full realization depends on addressing these challenges through ethical frameworks, transparent models, and strategic investments in governance and upskilling.

Keywords: ***Big Data, Artificial Intelligence, Data Processing, Machine Learning, Analytics***

1. INTRODUCTION

The relentless march of technological advancement has transformed the digital landscape, leading to an unprecedented explosion in the volume, velocity, and diversity of data being generated. This phenomenon, commonly referred to as "Big Data," has presented both challenges and opportunities for organizations and researchers across various domains.(Rahmani et al., 2021) A crucial aspect of effectively harnessing the power of Big Data is the application of Artificial Intelligence techniques, which have the potential to revolutionize the way we process, analyze, and derive insights from these massive datasets.(Leng et al., 2022)

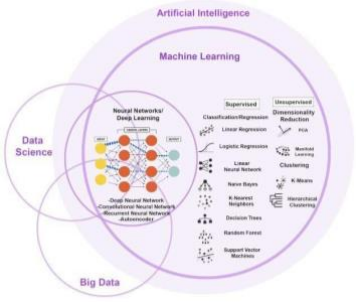

Image 1.1 *Venn diagram on the relationship artificial intelligence, data science, and big data. (https://www.researchgate.net)*

This systematic literature review aims to explore the current state of research on the intersection of Artificial Intelligence and Big Data processing, with a focus on identifying the key trends, opportunities, and challenges in this rapidly evolving field.(Wang, 2024). The integration of Artificial Intelligence with Big Data processing has emerged as a promising approach to address the complexities and demands of the modern data landscape.(Shi et al., 2022) This systematic literature review aims to explore the current state of research on the intersection of Artificial Intelligence and Big Data processing, with a focus on highlighting the key advancements, opportunities, and challenges in this rapidly evolving field.(Zaripova et al., 2023)

The symbiotic relationship between Artificial Intelligence and Big Data is undeniable, as each field has the potential to enhance and amplify the other.(Shi et al., 2022) Big Data provides the vast and diverse datasets required to train and refine AI models, while Artificial Intelligence offers the computational power, algorithms, and techniques necessary to extract meaningful insights from these large-scale, unstructured data sources.(Leng et al., 2022; Qi, 2024) For example, in e-commerce, platforms like Amazon utilize AI-powered recommendation systems to analyze user behavior data, providing personalized shopping experiences (Judijanto & Rolando, 2024; Rolando et al., 2024; Tanuwijaya et al., 2024). This mutually beneficial relationship has led to significant advancements in various domains, including natural language processing, computer vision, recommender systems, healthcare, finance, and beyond (Mulyono & Rolando, 2024; Nuraini et al., 2024; Wijaya et al., 2024).

2. RESEARCH METHOD

The literature review conducted for this study has delved deeper into the rich tapestry of research exploring the intersection of Artificial Intelligence and Big Data processing. These in-depth studies have explored the various applications, methodologies, and implications of leveraging AI techniques to optimize the processing and analysis of Big Data, providing a comprehensive understanding of this rapidly evolving field.

To ensure the rigor and comprehensiveness of the literature review, the research team has adhered to the systematic review guidelines proposed by Kitchenham et al. . The key steps involved in the review process are as follows: Identification of relevant literature: A comprehensive search was conducted across multiple academic databases, including IEEE Xplore, ACM Digital Library, and Scopus, using a combination of keywords such as "Artificial Intelligence," "Big Data," "Data Processing," "Machine Learning," and "Analytics." Screening and selection: The initial search yielded a large number of studies, which were then carefully screened and selected based on their relevance to the research topic, quality of content, and methodological rigor. Data extraction and synthesis: The selected studies were thoroughly reviewed, and key information such as research objectives, methodologies, findings, and implications were extracted and synthesized to identify the prevailing trends, opportunities, and challenges in the field.

Quality assessment: The research team employed a set of predetermined criteria to assess the quality and reliability of the included studies, ensuring the robustness of the literature review. Data analysis and interpretation: The extracted data was analyzed and interpreted to uncover the core themes, insights, and implications of the research on the intersection of Artificial Intelligence and Big Data processing. To ensure the rigor and comprehensiveness of the literature review, a systematic approach was adopted. This included identifying relevant literature from academic databases such as IEEE Xplore and Scopus, screening studies based on their relevance, and synthesizing key insights. A flowchart of the research process is provided to visualize these steps.

The systematic literature review process outlined above has enabled the research team to gather a comprehensive understanding of the current state of research in this field, paving the way for a holistic and evidence-based discussion of the key findings.

The seminal study by Mario D'Arco (Arco et al., 2019) highlights the transformative potential of Artificial Intelligence in the context of literature reviews, a process that often involves the synthesis and analysis of large volumes of unstructured data. The authors emphasize how AI-powered tools

and techniques can expedite the individual steps of the literature review process, from data collection and organization to the identification of key trends and meaningful insights that can drive research forward.

Similarly, another notable study by Muhammad Bilal(Bilal et al., 2024) examine the application of Artificial Intelligence and Machine Learning in library settings, where these cutting-edge technologies are being increasingly employed to enhance information retrieval, knowledge management, and user-centric services. The systematic review provides a thorough overview of the current state of research in this domain, underscoring the growing importance of AI-powered solutions in modern library operations and their potential to revolutionize the way information is accessed and utilized.

3. DISCUSSION

The systematic literature review conducted for this study underscores the transformative impact of Artificial Intelligence (AI) on Big Data processing (Maha et al., 2024; Rolando, 2024). AI-powered tools and techniques have emerged as critical enablers, allowing for efficient processing, analysis, and insight extraction from vast, complex, and heterogeneous datasets that characterize Big Data (Qi, 2024; Zaripova et al., 2023). One of the primary advantages of integrating AI into Big Data workflows is its ability to automate and streamline key steps in the data management lifecycle, such as data collection, preprocessing, analysis, and visualization (Rahardja et al., 2024; Rolando & Ingriana, 2024). AI-powered solutions enhance speed, accuracy, and scalability by automating tasks like data cleaning, feature engineering, and anomaly detection, which are otherwise resource-intensive and prone to error.(Rahmani et al., 2021)

The convergence of AI and Big Data has also unlocked unprecedented opportunities in predictive analytics, decision-making, and problem-solving. AI-driven algorithms uncover hidden correlations, trends, and insights that were previously difficult or impossible to discern using traditional analytical methods.(Bansal et al., 2024)These insights empower the creation of highly personalized, user-centric applications, such as recommendation systems, tailored healthcare solutions, and adaptive learning platforms.

Despite its potential, the adoption of AI in Big Data processing comes with significant challenges. Data privacy concerns, algorithmic bias, and the interpretability of AI-driven insights have emerged as critical issues that must be addressed to ensure ethical and responsible deployment (Hosen et al., 2023). For example, opaque "black box" AI models can hinder transparency, particularly in critical domains like healthcare and finance, where explainability is essential.(Badawy, 2023)

Additionally, the integration of AI with Big Data requires substantial investments in technological infrastructure, robust data governance frameworks, and the cultivation of specialized expertise within the workforce.(Mills et al., 2024) Organizations must prioritize the development of governance structures and upskilling initiatives to fully realize the potential of AI-driven Big Data processing while ensuring ethical compliance and inclusivity.(Chioma Ann Udeh et al., 2024)

The literature review revealed a rich array of research on the intersection of Artificial Intelligence (AI) and Big Data processing. Studies explored diverse applications, methodologies, and implications of using AI techniques to optimize Big Data analysis (Ingriana, Chondro, et al., 2024; Mulyono, Ingriana, et al., 2024; Wigayha et al., 2024).

3.1 Key Applications of AI in Big Data Processing

AI-powered tools such as IBM Watson Health analyze patient records to identify patterns for early disease diagnosis and personalized treatments. For instance, genomic data processed by AI models helps oncologists recommend targeted cancer therapies.

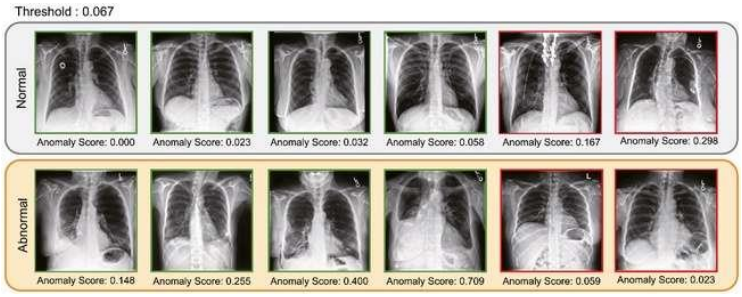

Image 3.1.1 *Annotated medical images showing AI detection of anomalies in genomic sequences. (https://www.cell.com/iscience/)*

Finance Machine learning algorithms detect fraudulent transactions by analyzing spending patterns. Visa's AI system, for example, identifies anomalies in real time, reducing fraud and saving billions annually (Ingriana, Gianina Prajitno, et al., 2024; Mulyono, Hartanti, et al., 2024).

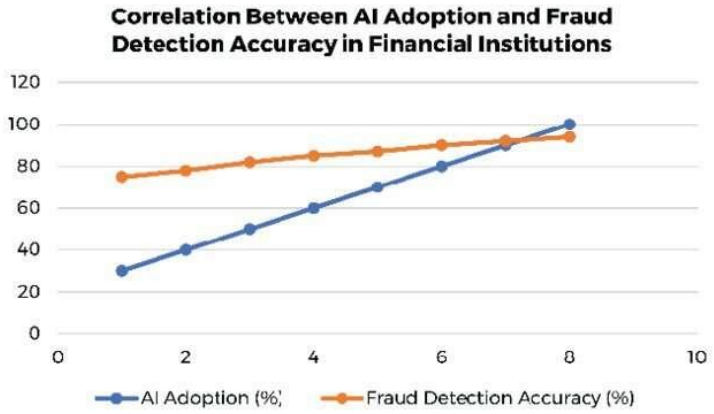

Image 3.1.2 *Graph illustrating improved fraud detection rates with AI adoption. (https://www.researchgate.net/)*

Retail and E-Commerce, AI enhances user engagement through recommendation engines. Platforms like Netflix and Spotify leverage user behavior data to predict preferences, offering tailored content suggestions.

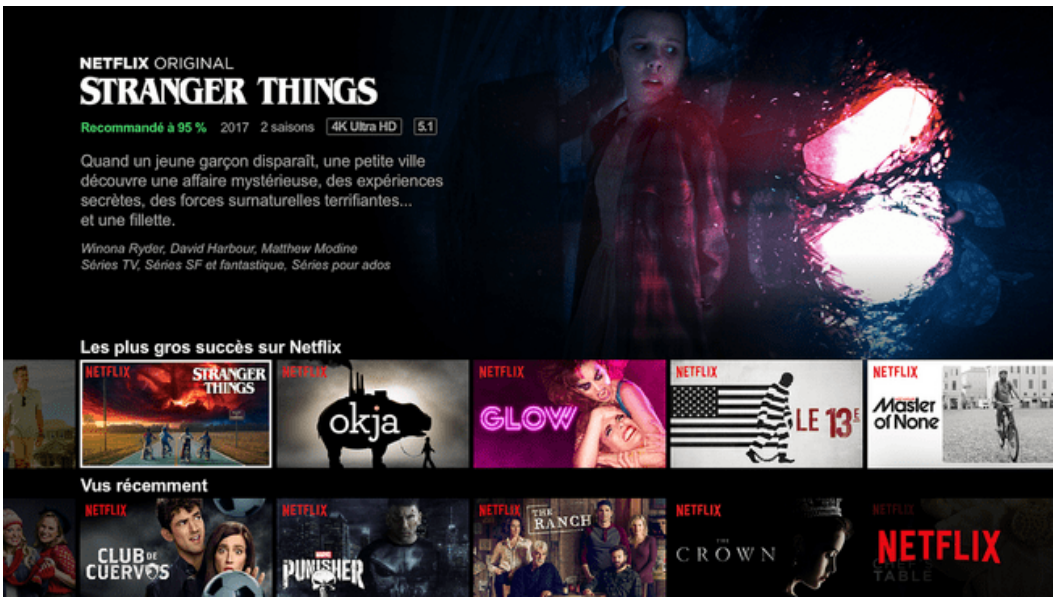

Image 3.1.3 *Netflix's personalized recommendation section. (https://gibsonbiddle.medium.com/)*

Smart Cities and Transportation, AI optimizes urban mobility by analyzing real-time traffic data. Singapore, for instance, employs AI to manage traffic lights, reducing congestion and improving travel efficiency.

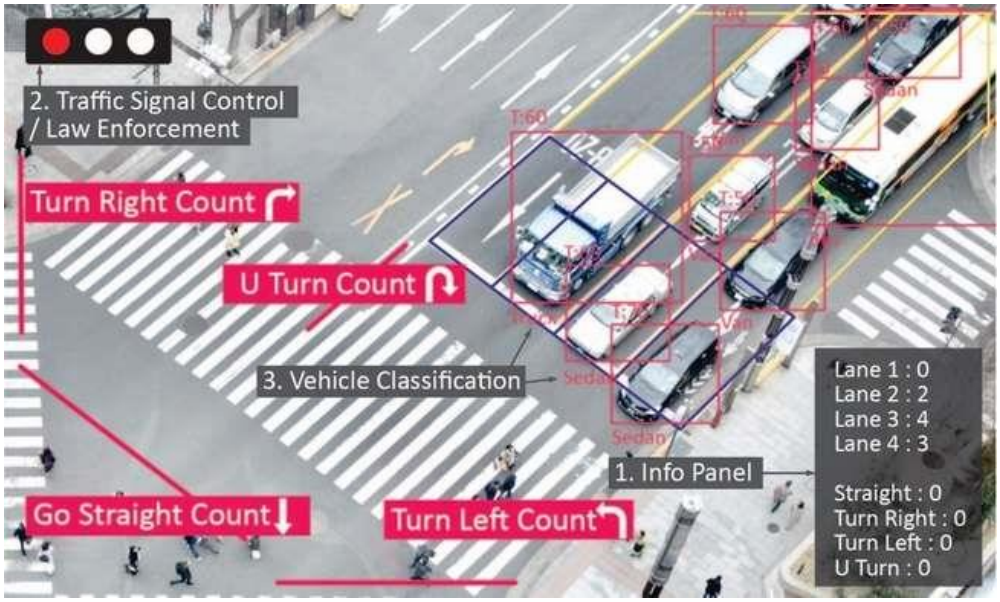

Image 3.1.4 *A dashboard showing live traffic flow managed by AI systems.(https://www.securityworldmarket.com/)*

Agriculture Precision farming solutions like John Deere's AI-driven tools process IoT sensor data to optimize irrigation, predict pest infestations, and improve crop yields.

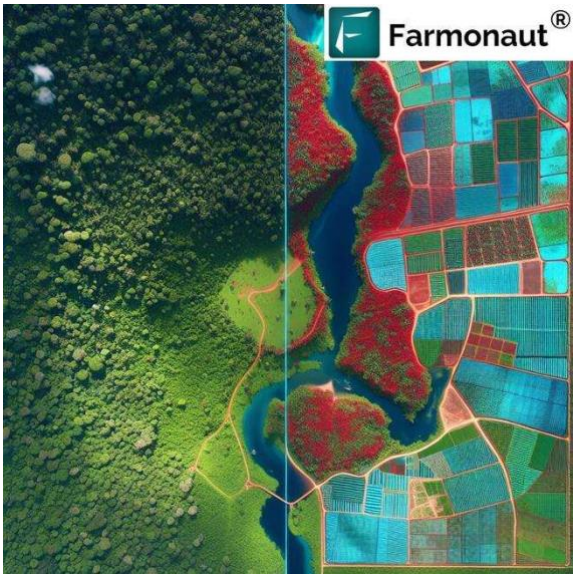

Image 3.1.5 *Satellite image highlighting AI-detected areas needing irrigation. (https://farmonaut.com/)*

3.2 Advantages of AI in Big Data Processing

Automation of Data Lifecycle: AI automates data collection, cleaning, preprocessing, and visualization, significantly enhancing speed and accuracy (Putri & Setiawan, 2025; Rolando et al., 2025). Example: Automated feature engineering in AI models for predictive analytics. AI processes vast and heterogeneous datasets efficiently, unlocking insights that traditional methods struggle to detect. Example: AI-based climate models analyzing satellite data to track deforestation. Enhanced Decision-Making: Predictive analytics driven by AI uncovers correlations and trends within Big Data, enabling informed decision-making (Widjaja, 2025). Example: AI predicting disease outbreaks during the COVID-19 pandemic based on mobility data.

3.3 Challenges in AI-Driven Big Data Processing

Data Privacy Concerns: Integrating AI in sensitive domains like healthcare raises questions about data security and compliance with privacy regulations (e.g., GDPR). Algorithmic Bias: Biased AI models in hiring or judicial systems can perpetuate inequity if not carefully monitored. Example: Gender bias in resume-screening algorithms (Ingriana, 2025; Wigayha et al., 2025; Zahran, 2025). Explainability: Black-box AI models pose challenges in critical areas where interpretability is essential, such as healthcare diagnostics. Resource Requirements: The integration of AI requires significant investment in infrastructure, skilled personnel, and governance frameworks (Tan & Alexia, 2025; Winata & Arma, 2025).

3.4 Emerging Trends

Deep Learning for Image Analysis: Convolutional Neural Networks (CNNs) improve tasks like medical imaging and facial recognition.

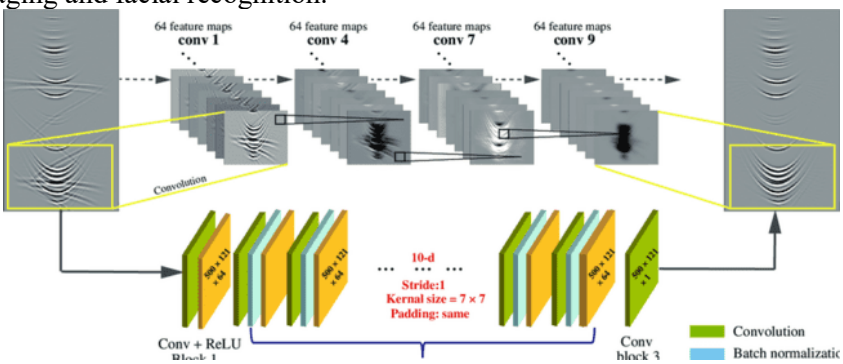

Image 3.4.1 *Simplified CNN architecture highlighting image feature detection. (https://www.researchgate.net/)*

Natural Language Processing (NLP): Advanced models like GPT-4 analyze vast text datasets, improving sentiment analysis, misinformation detection, and chatbots.

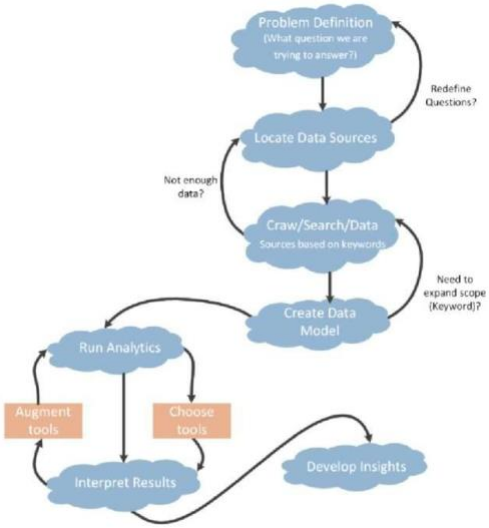

Image 3.4.2 *Visualization of AI analyzing social media text data. (https://www.researchgate.net/)*

4. CONCLUSION

AI-powered Big Data processing is reshaping industries by enabling smarter decision-making, better personalization, and enhanced efficiency. However, realizing its full potential requires addressing critical challenges such as algorithmic bias, data privacy, and infrastructure limitations. Future research should prioritize the development of ethical, interpretable, and scalable AI solutions to ensure the responsible and impactful deployment of these technologies.

The convergence of AI and Big Data has catalyzed innovation across sectors, from personalized healthcare to the optimization of urban environments. Notable advancements include the exploration of explainable AI models to improve transparency and trust, as well as emerging technologies like quantum computing, which promise to scale Big Data applications significantly. Additionally, federated learning offers a promising path to enhance data privacy while enabling

collaborative AI training across institutions, opening avenues for broader applications in sensitive domains like healthcare and finance.

By leveraging the synergies between AI and Big Data, organizations can unlock transformative opportunities, driving strategic decision-making, improving operational efficiency, and fostering growth and innovation. Stakeholders must also prioritize investments in governance frameworks, workforce upskilling, and sustainable technology infrastructure to ensure these technologies are adopted responsibly and inclusively.

REFERENCES

Arco, M. D., Presti, L. Lo, Marino, V., & Resciniti, R. (2019). Embracing AI and Big Data in customer journey mapping: From literature review to a theoretical framework. *Innovative Marketing*, *15*(4), 102–115. https://doi.org/10.21511/im.15(4).2019.09

Badawy, M. (2023). Integrating Artificial Intelligence and Big Data into Smart Healthcare Systems: A Comprehensive Review of Current Practices and Future Directions. *Artificial Intelligence Evolution*, *4*(2), 133–153. https://doi.org/10.37256/aie.4220232980

Bansal, R., Saini, S., Ngah, A. H., & Durga Prasad, T. (2024). Proselytizing the potential of influencer marketing via artificial intelligence: mapping the research trends through bibliometric analysis. *Cogent Business and Management*, *11*(1). https://doi.org/10.1080/23311975.2024.2372889

Bilal, M., Zhang, Y., Cai, S., Akram, U., & Halibas, A. (2024). Artificial intelligence is the magic wand making customer-centric a reality! An investigation into the relationship between consumer purchase intention and consumer engagement through affective attachment. *Journal of Retailing and Consumer Services*, *77*, 103674. https://doi.org/10.1016/J.JRETCONSER.2023.103674

Chioma Ann Udeh, Omamode Henry Orieno, Obinna Donald Daraojimba, Ndubuisi Leonard Ndubuisi, & Osato Itohan Oriekhoe. (2024). Big Data Analytics: a Review of Its Transformative Role in Modern Business Intelligence. *Computer Science & IT Research Journal*, *5*(1), 219–236. https://doi.org/10.51594/csitrj.v5i1.718

Ingriana, A. (2025). *THE INFLUENCE OF E-TRUST ON CONSUMER PURCHASING BEHAVIOR IN E-COMMERCE*. *1*(3). https://journal.dinamikapublika.id/index.php/Jumder

Ingriana, A., Chondro, J., & Rolando, B. (2024). *TRANSFORMASI DIGITAL MODEL BISNIS KREATIF: PERAN SENTRAL E-COMMERCE DAN INOVASI TEKNOLOGI DI INDONESIA* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/JUMDER

Ingriana, A., Gianina Prajitno, G., & Rolando, B. (2024). *THE UTILIZATION OF AI AND BIG DATA TECHNOLOGY FOR OPTIMIZING DIGITAL MARKETING STRATEGIES* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/IJEBS

Judijanto, L., & Rolando, B. (2024). EFFECTIVE STRATEGIES IN HUMAN RESOURCE MANAGEMENT. *International Journal Of Financial Economics (IJEFE)*, *1*(3), 747–754.

Leng, Z., Li, S., & Li, J. (2022). The Main Influencing Factors of Artificial Intelligence Big Data Sports News Coverage. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/7440337

Maha, V. A., Derian Hartono, S., Prajitno, G. G., & Hartanti, R. (2024). *E-COMMERCE LOKAL VS GLOBAL: ANALISIS MODEL BISNIS DAN PREFERENSI KONSUMEN* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/Jumder

Mills, N., Issadeen, Z., Matharaarachchi, A., Bandaragoda, T., De Silva, D., Jennings, A., & Manic, M. (2024). A cloud-based architecture for explainable Big Data analytics using self-structuring Artificial Intelligence. *Discover Artificial Intelligence*, *4*(1). https://doi.org/10.1007/s44163-024-00123-6

Mulyono, H., Hartanti, R., & Rolando, B. (2024). *SUARA KONSUMEN DI ERA DIGITAL: BAGAIMANA REVIEW ONLINE MEMBENTUK PERILAKU KONSUMEN DIGITAL* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/JUMDER

Mulyono, H., Ingriana, A., & Hartanti, R. (2024). *PERSUASIVE COMMUNICATION IN CONTEMPORARY MARKETING: EFFECTIVE APPROACHES AND BUSINESS RESULTS* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/IJEBS

Mulyono, H., & Rolando, B. (2024). Savoring The Success: Cultivating Innovation And Creativity For Indonesian Culinary MSMEs Growth. *Economics and Business Journal (ECBIS)*, *2*(4), 413–428.

Nuraini, I., Mulyono, H., Purnama, B., & Rolando, B. (2024). Pengaruh Lingkungan Kerja, Disiplin Kerja, dan Motivasi Kerja Terhadap Kinerja Karyawan Jasa Travel. *Journal of Trends Economics and Accounting Research*, *4*(4), 968–975.

Putri, L. W. B., & Setiawan, B. L. T. (2025). *ANALYZING THE STRATEGIC CONTRIBUTION OF SOCIAL MEDIA INFLUENCERS TO E-COMMERCE MARKETING EFFECTIVENESS*. *1*(2). https://journal.dinamikapublika.id/index.php/Jumder

Qi, X. (2024). Application analysis of big data and AI technology in new media communication channels. *Media and Communication Research*, *5*(2), 23–29. https://doi.org/10.23977/mediacr.2024.050204

Rahardja, B. V., Rolando, B., Chondro, J., & Laurensia, M. (2024). *MENDORONG PERTUMBUHAN E-COMMERCE: PENGARUH PEMASARAN MEDIA SOSIAL TERHADAP KINERJA PENJUALAN* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/JUMDER

Rahmani, A. M., Azhir, E., Ali, S., Mohammadi, M., Ahmed, O. H., Ghafour, M. Y., Ahmed, S. H., & Hosseinzadeh, M. (2021). Artificial intelligence approaches and mechanisms for big data analytics: a systematic study. *PeerJ Computer Science*, *7*, 1–28. https://doi.org/10.7717/peerj-cs.488

Rolando, B. (2024). *CULTURAL ADAPTATION AND AUTOMATED SYSTEMS IN E-COMMERCE COPYWRITING: OPTIMIZING CONVERSION RATES IN THE INDONESIAN MARKET* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/IJEBS

Rolando, B., Chandra, C. K., & Widjaja, A. F. (2025). *TECHNOLOGICAL ADVANCEMENTS AS KEY DRIVERS IN THE TRANSFORMATION OF MODERN E-COMMERCE ECOSYSTEMS*. *1*(2). https://journal.dinamikapublika.id/index.php/Jumder

Rolando, B., & Ingriana, A. (2024). *SUSTAINABLE BUSINESS MODELS IN THE GREEN ENERGY SECTOR: CREATING GREEN JOBS THROUGH RENEWABLE ENERGY TECHNOLOGY INNOVATION* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/IJEBS

Rolando, B., Nur Azizah, F., Karaniya Wigayha, C., Bangsa, D., Jl Jendral Sudirman, J., Jambi Selatan, K., & Jambi, K. (2024). *Pengaruh Viral Marketing Shopee Affiliate, Kualitas Produk, dan Harga Terhadap Minat Beli Konsumen Shopee*. https://doi.org/10.47065/arbitrase.v5i2.2167

Shi, H., Cao, G., Ma, G., Duan, J., Bai, J., & Meng, X. (2022). New Progress in Artificial Intelligence Algorithm Research Based on Big Data Processing of IOT Systems on Intelligent Production Lines. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/3283165

Tan, D. M., & Alexia, K. R. (2025). *THE INFLUENCE OF TIKTOK AFFILIATE CONTENT QUALITY AND CREDIBILITY ON PURCHASE DECISIONS VIA THE YELLOW BASKET FEATURE*. *1*(2). https://journal.dinamikapublika.id/index.php/Jumder

Tanuwijaya, M., Mulyono, H., Purnama, B., & Rolando, B. (2024). Pengaruh Kompensasi, Motivasi dan Disiplin Kerja Terhadap Kepuasan Kerja Karyawan. *Journal of Trends Economics and Accounting Research*, *4*(4), 956–967.

Wang, J. (2024). Research on Big Data Analysis Methods Based on Artificial Intelligence Technology. *Automation and Machine Learning*, *5*(1), 62–68. https://doi.org/10.23977/autml.2024.050108

Widjaja, A. F. (2025). *FACTORS INFLUENCING PURCHASE INTENTION IN E-COMMERCE: AN ANALYSIS OF BRAND IMAGE, PRODUCT QUALITY, AND PRICE*. *1*(3). https://journal.dinamikapublika.id/index.php/Jumder

Wigayha, C. K., Rolando, B., & Wijaya, A. J. (2024). *PELUANG BISNIS DALAM INDUSTRI HIJAU DAN ENERGI TERBARUKAN* (Vol. 1, Issue 1). https://journal.dinamikapublika.id/index.php/Jumder

Wigayha, C. K., Rolando, B., & Wijaya, A. J. (2025). *A DEMOGRAPHIC ANALYSIS OF CONSUMER BEHAVIORAL PATTERNS ON DIGITAL E-COMMERCE PLATFORMS*. *1*(2). https://journal.dinamikapublika.id/index.php/Jumder

Wijaya, F., Mulyono, H., Utami, F. N., & Rolando, B. (2024). Pengaruh Kualitas Pelayanan, Harga, dan Kepuasan Pelanggan Terhadap Loyalitas Pelanggan Motor. *Journal of Trends Economics and Accounting Research*, *4*(4), 976–984.

Winata, V., & Arma, O. (2025). *ANALYZING THE EFFECT OF E-WALLET USABILITY ON CUSTOMER RETENTION IN MOBILE PAYMENT APPS*. *1*(2). https://journal.dinamikapublika.id/index.php/Jumder

Zahran, A. M. (2025). *THE IMPACT OF MARKETING STRATEGIES ON THE SUCCESS OF THE FAST FASHION INDUSTRY: A SYSTEMATIC REVIEW*. *1*(3). https://journal.dinamikapublika.id/index.php/Jumder

Zaripova, R., Kosulin, V., Shkinderov, M., & Rakhmatullin, I. (2023). Unlocking the potential of artificial intelligence for big data analytics. *E3S Web of Conferences*, *460*, 1–9. https://doi.org/10.1051/e3sconf/202346004011