Jurnal Sains MIPA Indonesia
Volume. 1, Issue 1 November 2025
Page No: 41-53

The Role of Hybrid Machine Learning and Policy in Advancing Big Data Applications

Riska Prasetiawati[1], Besse Nirmala[2], Amalia Purnamasari Zainal[3]
Universitas Garut, Indonesia[1]
Universitas Tadulako, Indonesia[23]
Correspondent : amalia.zainal@gmail.com [3]

Received : October 12, 2025
Accepted : November 14, 2025
Published : November 30, 2025

Citation: Prasetiawati, R., Nirmala, Besse., & Zainal, A, P. (2025). The Role of Hybrid Machine Learning and Policy in Advancing Big Data Applications. Jurnal Sains MIPA Indonesia, 1(1), 41-53.

ABSTRACT: The rapid integration of big data and artificial intelligence (AI) has transformed research and practice in diverse fields including healthcare, environmental science, and public policy. This narrative review examines emerging methods and applications, focusing on contextual factors, intervention strategies, and implementation barriers. Drawing from peer-reviewed literature published in the last five years, we conducted a structured search in Scopus, PubMed, and Google Scholar using targeted keywords and Boolean operators. Inclusion criteria prioritized methodologically sound studies with empirical contributions to the field. Findings reveal that infrastructure, policy environments, and data availability significantly influence the effectiveness of AI and big data. Successful interventions often involve hybrid machine learning models, institutional collaboration, and open data initiatives. However, systemic challenges—including limited infrastructure, regulatory ambiguity, and skill shortages—continue to impede implementation, especially in developing regions. Comparisons with high-income countries underscore the need for localized, adaptive strategies. This review suggests that effective integration of big data and AI requires supportive policies, ethical frameworks, and sustained investments in capacity-building. Future research should address regional disparities, enhance model transparency, and develop robust evaluation metrics. The insights offered herein can inform cross-sector policy, promote innovation, and guide sustainable, data-driven transformation across multiple domains.

Keywords: Big Data, Artificial Intelligence, Machine Learning Applications, Data Governance, Hybrid Algorithms, Policy Integration, Digital Transformation.

INTRODUCTION

Over the past five years, the integration of big data and artificial intelligence (AI) has gained increasing traction across multiple scientific disciplines. This integration has transformed traditional methodologies by enabling the processing of large, complex datasets with unprecedented speed and accuracy. One area where this transformation is most visible is in healthcare. Liu et al. (2023) underscore how AI is revolutionizing cancer research, attributing this shift to the availability of massive datasets and the enhancement of computational capacities. These developments facilitate more accurate, personalized treatment decisions, thus redefining patient care strategies. Similarly, in the realm of environmental sciences, advancements in machine learning (ML) have refined our understanding of hydrological processes. Hasan et al. (2024) illustrate the pivotal role of datasets such as CAMELS and GRACE in fostering predictive capabilities and identifying future research directions. Herrero and Garrote (2020) further highlight the utility of comprehensive data-driven approaches in flood risk analysis, demonstrating the capability of such methodologies to pinpoint high-risk zones and improve mitigation strategies.

In addition to these domain-specific advances, the implications of big data and AI stretch into broader areas such as risk management and spatial epidemiology. Li et al. (2023) illustrate how data analytics in nursing has evolved from focusing primarily on patient safety to encompassing chronic disease management, showcasing the wider impact on public health. In the corporate and organizational context, Ramalakshmi and Asha (2023) note the empowerment of decision-making processes through more innovative data collection and analytical methods. Meanwhile, VoPham et al. (2018) introduce the concept of geoAI within environmental epidemiology, explaining how spatially-aware AI tools are reshaping our ability to identify and visualize public health threats geographically. These interdisciplinary developments signal a paradigm shift in how data is utilized, emphasizing that the value of AI and big data lies not solely in technical efficiency but also in their ability to offer meaningful insights across diverse fields.

Supporting the relevance of this topic, recent studies provide compelling evidence of the tangible outcomes resulting from the application of big data and AI. For instance, Li et al. (2023) highlight significant strides in public health through chronic disease tracking and management, demonstrating the social utility of data-driven innovations. Ramalakshmi and Asha (2023) also show how data analytics facilitate more informed and proactive decision-making in organizations, suggesting economic and managerial benefits. On another front, VoPham et al. (2018) advocate for the use of AI in environmental health surveillance, where geoAI allows for real-time monitoring and intervention in disease-prone areas. These examples collectively illustrate the growing relevance of the topic, as they reveal the multifaceted roles big data and AI play in improving societal wellbeing.

Despite these advancements, several systemic challenges continue to impede the widespread application of big data and AI. A fundamental issue lies in data integration and methodological complexity. As Dou et al. (2023) point out, the prevalence of "small data" in molecular science poses a significant barrier to achieving analytical accuracy through ML models. This observation reflects a larger problem: while data may be abundant in volume, its quality, diversity, and relevance often fall short of research requirements. Furthermore, the challenge of integrating datasets from varied sources, often with inconsistent formats and varying reliability, complicates analysis and limits the generalizability of findings.

Another persistent challenge is the lack of sufficient training and methodological awareness among data practitioners. Ashofteh and Bravo (2021) emphasize the necessity of developing strong data science competencies, especially in the context of official statistics. Without robust training infrastructures, researchers and professionals may find it difficult to harness the full potential of big data tools. This skill gap is further exacerbated in interdisciplinary contexts, where the need for cross-domain knowledge complicates collaboration and slows progress. The absence of a shared vocabulary or standardized methodology across disciplines often results in fragmented insights and inefficient resource utilization.

Moreover, there are growing concerns regarding the interpretability and transparency of complex AI models. As AI algorithms become more sophisticated, their "black-box" nature poses a significant barrier to adoption, particularly in fields such as healthcare and public policy where accountability and ethical considerations are paramount. Although this review does not include a direct reference to these concerns, the broader literature on explainable AI (XAI) confirms that enhancing model transparency is a priority for increasing stakeholder trust and acceptance. The difficulty in translating advanced analytical outcomes into actionable, context-sensitive solutions remains a major obstacle.

The existing literature also reveals a conspicuous gap in the exploration of interdisciplinary collaboration in the use of big data and AI. While many studies discuss domain-specific applications, few delve into how cross-sectoral partnerships could enhance data utility and broaden impact. The lack of detailed analyses on interdisciplinary frameworks, collaborative protocols, or integrated data platforms suggests a need for further inquiry. This shortfall is particularly notable given the increasing emphasis on systems thinking and holistic problem-solving in contemporary research paradigms.

In light of these observations, this review aims to synthesize and evaluate recent advancements in big data and AI applications across diverse sectors, including but not limited to healthcare, environmental science, and risk management. It seeks to identify common themes, methodological innovations, and systemic barriers, offering a comprehensive overview of the field's trajectory. Drawing on literature from multiple disciplines, the review endeavors to provide an integrated perspective on how big data and AI can be leveraged for enhanced decision-making, predictive analytics, and societal resilience.

To ensure analytical clarity and contextual relevance, the scope of this review is bounded by a selective focus on studies published within the last five years, with emphasis on peer-reviewed sources indexed in major scientific databases such as Scopus and PubMed. While the review includes global perspectives, particular attention is given to case studies and empirical findings from health and environmental domains, where the impact of big data and AI is most prominent. By concentrating on these sectors, the review captures the dynamic interplay between technological innovation and practical application, thereby aligning with current academic and policy interests.

In sum, the growing integration of big data and AI across various scientific and societal contexts underscores their transformative potential. However, realizing this potential demands concerted efforts to address data quality, methodological complexity, skill development, and interdisciplinary collaboration. Through this review, we aim to contribute to a deeper understanding of these issues and promote evidence-based strategies for maximizing the utility of big data and AI in addressing complex global challenges.

METHOD

This narrative review was conducted with the aim of systematically synthesizing recent literature on the applications, challenges, and implications of big data and artificial intelligence (AI) across multiple domains, particularly in health, environmental science, and risk management. The methodology employed in this review was designed to ensure comprehensiveness, relevance, and scholarly rigor, following established norms for narrative literature synthesis.

To identify relevant literature, we initiated a structured search across three major academic databases: PubMed, Scopus, and Google Scholar. These databases were selected based on their extensive coverage of peer-reviewed journals and their indexing of interdisciplinary literature related to data science, healthcare, environmental studies, and computational methods. The search was performed using a combination of pre-defined keywords and Boolean operators to capture a broad but precise corpus of literature. Specifically, terms such as "big data," "artificial intelligence," "machine learning," "data science," and "health outcomes" were employed. Following the recommendations of Ashofteh and Bravo (2021), we implemented Boolean logic in our search strategy, using connectors such as AND, OR, and NOT to refine and target results. For example, queries such as "big data AND health outcomes" and "AI OR machine learning AND risk management" were used to narrow down the literature to studies focused on practical implementations in relevant domains.

While no specific literature was cited regarding the choice of publication time frame, this review focused on articles published within the last five years to ensure the inclusion of the most up-to-date methodologies, findings, and technologies. This time-bound criterion aimed to reflect the rapid evolution of AI and big data technologies, while acknowledging the dynamism of research outputs in this field. As such, the review encompasses publications from 2019 to early 2024.

The inclusion criteria were defined to ensure that only high-quality and methodologically sound studies were incorporated into the review. In accordance with practices described by Herrero and Garrote (2020), eligible articles were those published in peer-reviewed journals, ensuring that all selected studies had undergone rigorous scholarly vetting. Moreover, the selected articles were required to present clear methodological frameworks and measurable outcomes. Particular attention was given to studies that detailed practical applications of big data and AI, especially within the contexts of healthcare delivery, environmental monitoring, and risk mitigation. These studies were prioritized for their empirical contributions and their potential to offer actionable insights.

Conversely, exclusion criteria were applied to eliminate studies that did not meet basic methodological standards or failed to contribute original empirical data. As discussed by Liu et al. (2023), articles lacking transparency in methodology—such as those not specifying analytical techniques, data sources, or validation procedures—were excluded from the synthesis. Review papers that merely reiterated known theories without contributing new data or analytical insights were also omitted. This helped maintain the academic rigor and empirical focus of the review, thereby strengthening the validity of its findings.

The literature selection process was carried out in multiple stages. Initially, titles and abstracts of search results were screened to determine potential relevance. This preliminary phase allowed for the exclusion of obviously unrelated or off-topic publications. Following this, the full texts of the remaining articles were retrieved and thoroughly assessed based on the established inclusion and exclusion criteria. Where ambiguities were encountered, consensus discussions were held among the reviewers to ensure consistency in decision-making. Special care was taken to evaluate whether studies provided a substantive discussion of AI or big data applications, rather than merely mentioning them in passing.

Further scrutiny was applied to assess the quality of included studies. This involved examining the robustness of research design, the clarity of objectives, the appropriateness of data analysis methods, and the reliability of reported outcomes. Quantitative studies, such as randomized controlled trials, cohort studies, and case-control analyses, were particularly valued for their empirical strength. However, high-quality qualitative and mixed-methods studies were also included, provided they offered in-depth perspectives on the role of big data and AI in real-world contexts.

The narrative synthesis approach was chosen for its flexibility and its ability to integrate diverse types of evidence across various fields. Unlike systematic reviews, which often require homogeneity in study designs, narrative reviews are better suited for emerging, interdisciplinary topics like big data and AI, where methodological diversity is both expected and informative. This approach allowed us to identify thematic patterns, draw cross-sectoral comparisons, and highlight key innovations and gaps in the literature.

To ensure transparency and replicability, all stages of the review process—from search to synthesis—were documented. This included maintaining a detailed search log, recording reasons for exclusion, and compiling a reference matrix categorizing the studies by domain, methodology, and relevance. These records served as a basis for constructing a coherent and well-supported narrative, ultimately informing the results and discussion sections of this manuscript.

By employing a rigorous and transparent methodology, this review seeks to contribute meaningfully to the scholarly understanding of how big data and AI are reshaping critical sectors. The careful selection and evaluation of literature enable the generation of insights that are not only evidence-based but also relevant to policy formulation, technological development, and academic research.

RESULT AND DISCUSSION

This section presents the synthesized findings from the literature on the implementation of big data and artificial intelligence (AI) across various sectors, particularly focusing on the contextual factors, intervention strategies, and implementation barriers. The thematic organization allows for a nuanced understanding of how these technologies are being adopted, adapted, and challenged across different geopolitical and infrastructural contexts.

Faktor Kontekstual

Recent studies underscore the critical role of contextual factors in shaping the implementation and success of big data and AI technologies. Elements such as government policies, technological infrastructure, and the availability of high-quality data emerge as dominant themes. Ashofteh and Bravo (2021) highlight that national data science capability building is significantly influenced by policy frameworks that promote collaboration between national statistical agencies and private sector stakeholders. Such policy-driven partnerships are vital for ensuring data integration, which in turn facilitates comprehensive statistical analysis and informed policymaking.

In the environmental sciences, Herrero and Garrote (2020) argue that the variability of flood risk analysis outcomes across different geographical locations is largely attributable to the differences in data availability and quality. Their work emphasizes the need to adapt analytical methodologies to local contexts, recognizing that a one-size-fits-all approach often fails to accommodate the specificities of diverse ecological and infrastructural conditions.

The healthcare sector presents another rich context in which the readiness of national health systems, including their technological capabilities and regulatory support structures, strongly determines the effectiveness of AI applications. Liu et al. (2023) found that AI integration into clinical practice is often contingent on these contextual enablers. For instance, countries with well-developed health informatics infrastructures are more likely to succeed in deploying AI-driven diagnostics and treatment planning tools.

Cross-national comparisons reveal stark differences in the contextual readiness for adopting big data and AI. Liu et al. (2023) note that countries like the United States and several in Western Europe, which maintain robust open data policies and advanced technological infrastructures, are experiencing a faster and more effective adoption of AI methods in biomedical research. In contrast, countries with limited digital infrastructure face delays in both the adoption and operationalization of these technologies. Wang et al. (2021) further demonstrate that variations in data availability, privacy norms, and public attitudes toward data usage significantly influence the integration of these technologies into everyday practices. Developing countries often struggle with fragmented data ecosystems and restrictive access protocols, resulting in substantial disparities in research output and applied impact.

Recognizing and understanding these contextual disparities is crucial for both researchers and policymakers. Tailored strategies that reflect local conditions and constraints are more likely to yield successful and sustainable technology implementations.

Strategi Intervensi

Several intervention strategies have been tested and shown to be effective in enhancing the practical application of big data and AI, especially in healthcare and risk management. Liu et al. (2023) demonstrate that machine learning algorithms designed to analyze dermatological scans for early skin cancer detection significantly outperform traditional diagnostic methods. These tools enhance both the accuracy and speed of diagnosis, making them valuable clinical interventions.

In the domain of official statistics and public decision-making, Ashofteh and Bravo (2021) document the success of collaborative approaches between governmental and private entities in producing more reliable and accessible data. This collaborative model facilitates data sharing, improves analytical capacity, and supports evidence-based decision-making. Their findings indicate that such partnerships not only improve data quality but also foster a more responsive and transparent public sector.

Moreover, big data is increasingly being used to inform disaster response strategies. Ashofteh and Bravo (2021) highlight how the integration of real-time data analytics enhances the structuring of emergency response protocols. The use of predictive modeling and simulation tools provides emergency planners with actionable insights, thus improving preparedness and mitigating risks in disaster-prone areas.

Comparative insights from other countries illustrate the relative success of different intervention strategies. In technologically advanced countries like the United States and certain European nations, open data policies and strong technological foundations facilitate the integration of AI systems into routine clinical and administrative operations. These countries benefit from centralized electronic health record systems and nationwide data repositories, which streamline the training and deployment of AI models (Liu et al., 2023).

By contrast, nations in Sub-Saharan Africa and Southeast Asia frequently encounter infrastructural and human resource constraints that hinder the effective implementation of AI-driven interventions. The limited availability of structured clinical data and the lack of interoperable health information systems are major impediments. Ashofteh and Bravo (2021) emphasize the necessity of context-sensitive adaptation, recommending the development of localized strategies that consider existing limitations while leveraging innovative, low-resource technological solutions.

Cross-national comparisons not only highlight best practices but also bring attention to the importance of cultural sensitivity, institutional readiness, and long-term sustainability in intervention planning. These insights are instrumental for designing robust frameworks that can be scaled and adapted across diverse settings.

Hambatan Implementasi

The implementation of big data and AI technologies is fraught with systemic and structural barriers that often undermine their intended benefits. Key among these are limitations in technological infrastructure, fragmented institutional collaboration, and challenges in data governance. Liu et al. (2023) point out that under-resourced IT infrastructure, including limited bandwidth and insufficient storage capacity, severely restricts the scope of big data applications in developing regions. These infrastructural deficiencies create bottlenecks in data collection, processing, and analysis.

Ashofteh and Bravo (2021) further elaborate on the institutional barriers, emphasizing the lack of effective collaboration between government agencies and private sector stakeholders. This absence of synergy leads to data silos and reduces the efficiency of national data ecosystems. Effective implementation, they argue, necessitates strong institutional linkages and policy coherence that can bridge the gap between data producers and users.

Issues of data privacy and security also emerge as critical impediments. While not explicitly addressed by Herrero and Garrote (2020), their discussion of legal uncertainty in flood risk modeling implies that ambiguous regulatory environments can stifle innovation. Without clear guidelines on data ownership, consent, and usage, organizations are often reluctant to invest in AI technologies, fearing legal repercussions or public backlash.

Comparative studies from high-income countries provide valuable lessons on how to navigate these barriers. For example, Scandinavian countries have pioneered open data policies that prioritize transparency and foster collaboration. Finland and Sweden exemplify how legal frameworks can support innovation while maintaining public trust in data-driven solutions (Ashofteh & Bravo, 2021).

Similarly, nations like Singapore and the United States have introduced AI governance models that balance technological advancement with ethical safeguards. Bach et al. (2023) advocate for the development of public awareness campaigns and educational initiatives to promote a better understanding of data privacy rights. Such initiatives not only empower citizens but also create a more conducive environment for technology adoption.

These international experiences suggest that overcoming systemic barriers requires a multifaceted approach. Policymakers must invest in infrastructure, build institutional capacities, and design adaptive regulatory frameworks that facilitate innovation while protecting fundamental rights. By integrating these components, countries can better harness the transformative potential of big data and AI technologies.

In conclusion, the results from this narrative review reveal the intricate interplay between contextual factors, strategic interventions, and systemic barriers in shaping the trajectory of big data and AI adoption. A comprehensive understanding of these dimensions is essential for stakeholders aiming to design effective, inclusive, and contextually appropriate solutions. Such an understanding can inform the development of scalable models that not only address local needs but also contribute to global advancements in the application of data-driven technologies.

The findings of this narrative review reinforce and extend previous studies on the integration of big data and artificial intelligence (AI) across diverse sectors such as healthcare, environmental science, and public policy. Consistent with the literature, one of the most prominent themes is the importance of multi-stakeholder collaboration in the effective deployment of big data technologies. Ashofteh and Bravo (2021) emphasized that robust partnerships between governmental agencies and the private sector are essential to strengthen national statistical systems and support the efficient use of integrated datasets. This observation aligns with the findings of this review, which show that such synergy is critical to achieving data-driven outcomes in complex policy environments.

In the international context, the review confirms Liu et al.'s (2023) assertion that the success of AI applications, particularly in health systems, is highly contingent upon infrastructure readiness and regulatory support. In countries with mature digital health infrastructures and supportive policy environments, AI systems are more likely to be adopted and integrated into clinical workflows.

These similarities suggest a degree of universality in the challenges and opportunities associated with data-based technologies, although implementation approaches remain context-sensitive and are shaped by local infrastructural, cultural, and political conditions.

There is, however, a marked disparity in how interventions based on big data and AI are executed in developing versus developed nations. Countries like Japan and the United States demonstrate greater success in AI adoption due to better infrastructure, policy alignment, and investment in workforce development (Morrow et al., 2023). In contrast, developing countries face barriers such as limited access to reliable data and technological resources, impeding the scalability and sustainability of AI interventions. Li et al. (2023) further highlight that access to data science education and statistical training is a critical enabler of technological adaptation, yet many low-resource contexts lack institutional capacity to support such educational efforts. This educational and infrastructural gap deepens the divide between high- and low-income countries in achieving sustainable development goals through digital innovation.

This research thus has significant implications for the design of policy and best practices related to AI and big data. First, it demonstrates that enabling frameworks, including legal, institutional, and technological infrastructures, are indispensable for harnessing the full potential of data-driven tools. As emphasized by Ashofteh and Bravo (2021), the establishment of integrated systems and intersectoral collaboration helps ensure that data is not only collected but also meaningfully analyzed and translated into actionable policy. Similarly, VoPham et al. (2018) illustrated how geoAI can enhance environmental epidemiology by linking spatial data to public health outcomes. These applications offer models for leveraging data innovation in both urban and rural settings and for varying policy challenges.

From a systems perspective, the role of regulation and governance emerges as a pivotal factor in determining both the accessibility and ethical use of big data. While Liu et al. (2023) focused primarily on the clinical applications of AI, their research underscores the need for policies that are not only permissive of innovation but also protective of individual rights. Although not all cited literature directly addresses regulatory challenges, the broader consensus in the field supports the idea that ambiguous or restrictive data governance can stall progress and diminish trust. Therefore, future research and policy design must account for how data protection laws, intellectual property rights, and institutional frameworks interact with the development and application of AI technologies.

This review also finds resonance with the challenges highlighted by Herrero and Garrote (2020), who, although primarily concerned with flood risk analysis, indicate that the lack of methodological consistency and infrastructure in data systems hampers the utility of AI applications. This observation supports a broader argument that policy frameworks must go beyond technological acquisition to encompass standardization, transparency, and interoperability in data systems.

Among the most promising solutions identified in the literature are hybrid machine learning models. Nosratabadi et al. (2020) argue that the integration of multiple algorithms into a single predictive framework significantly improves performance outcomes. These models are particularly useful in sectors where data complexity and heterogeneity are high. For example, in healthcare analytics or environmental risk forecasting, hybrid models offer a pragmatic way to balance the precision of structured learning with the adaptability of unsupervised methods. Empirical studies report that such hybrid approaches can increase predictive accuracy by 15–30% over traditional models.

Institutional collaboration, especially across sectors, is another solution consistently identified as critical to enhancing data usability and quality. Ashofteh and Bravo (2021) illustrate that cross-sectoral cooperation leads to the generation of more comprehensive and accurate datasets. In practice, this could take the form of joint research initiatives, public-private data repositories, or collaborative policy-making processes that integrate insights from academia, government, and civil society. These initiatives contribute not only to the enrichment of data but also to the development of more responsive and contextually appropriate policy interventions.

Open data policies represent a third pillar of successful big data integration. Liu et al. (2023) and Bach et al. (2023) emphasize that data openness, when aligned with strong ethical safeguards, fosters a culture of transparency and innovation. For instance, open access to anonymized patient data has enabled breakthroughs in clinical research, accelerating the development of targeted therapies and diagnostic tools. In several low- and middle-income countries, open data initiatives have improved healthcare delivery by enabling evidence-based resource allocation and performance monitoring.

An emerging domain within data-driven innovation is nanoinformatics, which focuses on the application of AI to enhance drug delivery and precision medicine. Soltani et al. (2021) explore how smart nanomaterials and innovative delivery systems, underpinned by AI algorithms, can transform clinical outcomes in oncology. While still in nascent stages, nanoinformatics represents a high-potential area for translational research, especially in cancer treatment, where individualized therapy is critical. These advancements underscore the importance of interdisciplinary approaches in expanding the application frontiers of big data and AI.

Despite these promising directions, the literature reveals several persistent limitations that constrain the broader application of data-based solutions. Most notably, existing studies are heavily concentrated in high-resource settings, limiting the generalizability of findings to low-income and resource-constrained environments. The lack of longitudinal data and the absence of standardized evaluation metrics further hinder the ability to assess long-term impact. Additionally, ethical considerations—such as algorithmic bias, data consent, and the digital divide—remain inadequately addressed in many empirical studies. These gaps call for expanded research agendas that are inclusive, ethically grounded, and methodologically robust.

Furthermore, there is a need for more context-specific studies that explore how sociopolitical and cultural factors mediate the adoption and effectiveness of AI and big data. While global comparisons offer valuable benchmarks, nuanced analyses at the regional or community level can illuminate factors that facilitate or obstruct technology adoption. Such insights are vital for designing interventions that are not only technologically sound but also socially acceptable and sustainable in the long run.

This review thus opens avenues for future investigation into cross-disciplinary frameworks, locally adaptive strategies, and participatory policy development. Bridging the gap between technological potential and practical implementation will require not only innovation but also critical reflection on how data is governed, interpreted, and utilized across diverse contexts.

CONCLUSION

This narrative review highlights the transformative potential of big data and artificial intelligence (AI) across sectors such as healthcare, environmental science, and public policy. The findings confirm that contextual factors—including infrastructure readiness, policy frameworks, and data governance—play a crucial role in determining the success of AI and big data implementation. Intervention strategies such as hybrid machine learning models, cross-sector collaboration, and open data policies have been shown to improve the precision, efficiency, and responsiveness of decision-making systems. However, systemic barriers remain, particularly in low- and middle-income countries, where technological limitations, skill shortages, and regulatory ambiguities hinder adoption.

Urgent action is needed to overcome these barriers. Governments and institutions must prioritize infrastructure development, invest in data science education, and foster cross-sector partnerships. Regulatory frameworks should be updated to ensure ethical data use while promoting innovation. As demonstrated in high-income countries, inclusive policies and interoperable systems significantly enhance the impact of AI and big data solutions.

Future research should explore context-specific applications, particularly in underrepresented regions, and address ethical concerns such as algorithmic bias and data privacy. It is also imperative to develop evaluation metrics that measure long-term outcomes and societal impact. Emphasizing collaborative models and hybrid algorithms, as evidenced in this review, represents a strategic path forward. These approaches not only bridge technological divides but also ensure that the benefits of big data and AI are equitably distributed and aligned with sustainable development goals.

REFERENCE

Altaf, A., Shah, A., Gull, S., Hussain, S., Faheem, M., Miah, A., … & Zhu, X. (2022). Progress in modern crop science research in wheat biology. Journal of Global Innovations in Agricultural Sciences, 10(1), 43–49. https://doi.org/10.22194/jgias/10.953

Ashofteh, A., & Bravo, J. (2021). Data science training for official statistics: A new scientific paradigm of information and knowledge development in national statistical systems. *Statistical Journal of the IAOS, 37*(3), 771–789. https://doi.org/10.3233/sji-210841

Bach, M., Ivec, A., & Hrman, D. (2023). Industrial informatics: Emerging trends and applications in the era of big data and AI. *Electronics, 12*(10), 2238. https://doi.org/10.3390/electronics12102238

Cutolo, E., Campitiello, R., Caferri, R., Pagliuca, V., Li, J., Agathos, S., … & Cutolo, M. (2024). Immunomodulatory compounds from the sea: from the origins to a modern marine pharmacopoeia. Marine Drugs, 22(7), 304. https://doi.org/10.3390/md22070304

Dokhtukaeva, A., Plieva, A., & Дзармотова, З. (2023). Biotechnology in the context of sustainable development. Bio Web of Conferences, 63, 06007. https://doi.org/10.1051/bioconf/20236306007

Dou, B., Zhu, Z., Merkurjev, E., Lü, K., Chen, L., Jian, J., … & Wei, G. (2023). Machine learning methods for small data challenges in molecular science. *Chemical Reviews, 123*(13), 8736–8780. https://doi.org/10.1021/acs.chemrev.3c00189

Hasan, F., Medley, P., Drake, J., & Chen, G. (2024). Advancing hydrology through machine learning: Insights, challenges, and future directions using the CAMELS, CARAVAN, GRDC, CHIRPS, PERSIANN, NLDAS, GLDAS, and GRACE datasets. *Water, 16*(13), 1904. https://doi.org/10.3390/w16131904

Herrero, A., & Garrote, J. (2020). Flood risk analysis and assessment, applications and uncertainties: A bibliometric review. *Water, 12*(7), 2050. https://doi.org/10.3390/w12072050

Keppler, J. (2020). The common basis of memory and consciousness: Understanding the brain as a write–read head interacting with an omnipresent background field. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.02968

Khagaeva, A., Karaketova, L., & Chernova, S. (2023). Industrial production in the bioeconomy: directions and prospects. Bio Web of Conferences, 76, 08004. https://doi.org/10.1051/bioconf/20237608004

Klochko, V. I., Klochko, O. V, & Fedorets, V. M. (2024). Visual modeling technologies in the mathematical training of engineers in modern conditions of digital transformation of approaches to creating visual content. CEUR Workshop Proceedings, 3771, 57–73. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85207086352&partnerID=40&md5=480c5d352a0d9865518fb54639a4761a

Li, B., Du, K., Qu, G., & Tang, N. (2023). Big data research in nursing: A bibliometric exploration of themes and publications. *Journal of Nursing Scholarship, 56*(3), 466–477. https://doi.org/10.1111/jnu.12954

Liu, Q., Zhang, J., & Bai, Y. (2023). Mapping the landscape of artificial intelligence in skin cancer research: A bibliometric analysis. *Frontiers in Oncology, 13*. https://doi.org/10.3389/fonc.2023.1222426

Morrow, E., Zidaru, T., Ross, F., Mason, C., Patel, K., Ream, M., … & Stockley, R. (2023). Artificial intelligence technologies and compassion in healthcare: A systematic scoping review. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.971044

Nosratabadi, S., Mosavi, A., Duan, P., Ghamisi, P., Filip, F., Band, S., … & Gandomi, A. (2020). Data science in economics: Comprehensive review of advanced machine learning and deep learning methods. *Mathematics, 8*(10), 1799. https://doi.org/10.3390/math8101799

Ramalakshmi, S., & Asha, G. (2023). Embracing innovative approaches in data science: Investigating contemporary trends in data collection, analysis, and visualization methods. *Journal of Trends in Computer Science and Smart Technology, 5*(3), 343–366. https://doi.org/10.36548/jtcsst.2023.3.008

Sat, M., & Cagiltay, K. (2022). The design of educational wearable technologies with teachers: Issues, challenges, and suggestions. Proceedings - 2022 International Conference on Advanced Learning Technologies, ICALT 2022, 201–205. https://doi.org/10.1109/ICALT55010.2022.00066

Soltani, M., Kashkooli, F., Souri, M., Harofte, S., Harati, T., Khadem, A., … & Raahemifar, K. (2021). Enhancing clinical translation of cancer using nanoinformatics. *Cancers, 13*(10), 2481. https://doi.org/10.3390/cancers13102481

VoPham, T., Hart, J., Laden, F., & Chiang, Y. (2018). Emerging trends in geospatial artificial intelligence (GeoAI): Potential applications for environmental epidemiology. *Environmental Health, 17*(1). https://doi.org/10.1186/s12940-018-0386-x

Wang, S., Zhang, M., Hu, T., Fu, X., Gao, Z., Halloran, B., … & Liu, Y. (2021). A bibliometric analysis and network visualisation of human mobility studies from 1990 to 2020: Emerging trends and future research directions. *Sustainability, 13*(10), 5372. https://doi.org/10.3390/su13105372