Literature Review

Open Access CrossMark

Exploring the role of generative artificial intelligence and large language models in enhancing the palm oil value chain: a systematic literature review

Abstract

The rapid advancement of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) is transforming digital innovation in agribusiness by enabling intelligent automation, data-driven decision-making, and enhanced operational efficiency. However, their large-scale adoption in the palm oil value chain remains constrained by fragmented data, high administrative complexity, and infrastructural limitations. This study aims to systematically evaluate the current state of research on the applications, performance, and challenges of GenAI and LLMs across the palm oil value chain. A qualitative Systematic Literature Review (SLR) was conducted following the PRISMA framework. Publications were retrieved from the Scopus and PubMed databases. From an initial 1,083 records, 92 relevant studies were identified through refined search strategies. After applying inclusion criteria, including publication year (2020–2026), open-access availability, and duplicate removal, 35 peer-reviewed articles were selected for qualitative synthesis. The data were analyzed using thematic coding and cross-study synthesis to identify dominant application areas, technical architectures, and implementation trends. The review identifies four major application domains: predictive maintenance and agricultural robotics, sustainability documentation and compliance, precision agronomy decision support, and supply chain logistics optimization. Transformer-based architectures and Retrieval-Augmented Generation (RAG) emerged as the dominant technologies, with most studies focusing on Indonesia and Malaysia. The reviewed evidence indicates that GenAI and LLMs improve harvesting accuracy, optimize maintenance and logistics, enhance precision agronomy, and automate sustainability compliance while strengthening traceability across the value chain. Nevertheless, adoption remains constrained by data scarcity, computational costs, limited digital infrastructure, and the smallholder digital divide. Future research should prioritize domain-specific foundation models, edge-AI solutions, and inclusive deployment strategies to support sustainable and scalable implementation across the palm oil industry.

Keywords: generative AI, large language models, palm oil, value chain, systematic literature review

Special issue - 2026

Loso Judijanto
IPOSS Jakarta, Indonesia

Correspondence: Loso Judijanto, IPOSS Jakarta, Indonesia

Received: July 10, 2026 | **Published:** July 27, 2026

Introduction

The global palm oil sector represents one of the most critical pillars of the international agricultural economy and bio-based manufacturing, serving as a primary driver of socioeconomic development and rural poverty alleviation across tropical developing nations. As a highly versatile vegetable oil, its applications span an extensive spectrum from food processing and cosmetics to advanced biofuels, generating a complex, multi-tiered global value chain that intersects with critical macroeconomic indicators.[1,2] This value chain is traditionally divided into three distinct operational spheres: upstream plantation management and harvesting, midstream milling and crude palm oil (CPO) extraction, and downstream refining, chemical transformation, and global distribution logistics.[3] Managing this sprawling network requires the synchronized calibration of thousands of variables, including shifting meteorological patterns, unpredictable soil nutrient degradation, strict international quality benchmarks, and multi-modal transportation schedules.[4] However, despite its established economic significance, the industry faces structural pressures to systematically improve resource optimization, mitigate post-harvest biomass waste, and achieve improved transparency within its supply networks to satisfy shifting regulatory and market conditions.[5]

Traditional operational architectures within the palm oil value chain heavily rely on fragmented, legacy monitoring frameworks, localized agronomic heuristics, and deterministic computational systems.[6] While the initial waves of digital transformation, such as the deployment of Internet of Things (IoT) field sensors, descriptive satellite imagery, and basic predictive machine learning algorithms, have successfully enhanced localized yield forecasting and automated precision fertilization to a certain degree, these classical Industry 4.0 applications exhibit fundamental technical limitations.[7] Legacy predictive models operate primarily on structured numerical datasets, rendering them incapable of autonomously parsing, contextualizing, and acting upon the massive volume of unstructured data generated throughout the industrial lifecycle, which includes multilingual smallholder legal deeds, textual machinery sensor logs, changing international legal frameworks, and intricate supply chain customs invoices.[8] Consequently, a profound technical bottleneck persists at the center of the palm oil value chain, where critical real-time decisions are often delayed due to the lack of an intelligent, unified semantic layer capable of synthesizing complex data into executable operational strategies.[9]

This critical technological gap highlights an immediate, systemic urgency to transition from passive, purely predictive computational models toward advanced generative computing paradigms.[10] The rapid development of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) represents a fundamental paradigm shift in industrial informatics, introducing neural architectures capable of deep semantic comprehension, complex multi-modal reasoning, and synthetic data generation.[11] In capital-intensive, logistically volatile agricultural sectors, GenAI and LLMs offer unprecedented capabilities to automate highly complex processes that previously demanded intensive human intervention.[12] For instance, in upstream operations, generative vision-language models (VLMs) can synthesize sparse, imperfect drone imagery to generate high-fidelity, simulated environmental models, allowing autonomous harvesting machinery to navigate unpredictable plantation canopies with superior precision. In midstream and downstream operations, advanced LLMs operating via Retrieval-Augmented Generation (RAG) can automatically audit vast supply chain networks, instantly cross-referencing multi-layered logistical records with changing international sustainability mandates to guarantee compliance and absolute traceability.

Despite the growing academic and commercial enthusiasm surrounding the deployment of generative computing in heavy industries, the existing scientific literature remains highly fragmented, siloed, and disconnected from the specific operational realities of the palm oil sector. While a multitude of recent systematic reviews have extensively documented the broader applications of generic machine learning in smart agriculture or the theoretical impacts of LLMs in supply chain management, there is a limited availability of a focused synthesis examining how generative computing specifically interacts with the palm oil value chain. Current reviews almost exclusively focus on traditional, non-generative architectures for narrow tasks like remote sensing yield calculations or basic robotic classification, completely ignoring the revolutionary capabilities of generative text-to-action frameworks, autonomous multi-agent decision support systems, and synthetic failure-mode simulation engines. This important evidence gap prevents industry practitioners, technology developers, and academic researchers from establishing a clear, unified understanding of the precise entry points, technical barriers, and empirical performance metrics of GenAI and LLMs within this strategic sector.

To address this significant scholarly omission and provide a focused theoretical synthesis, this study establishes its novelty positioning by executing a comprehensive Systematic Literature Review (SLR) dedicated exclusively to mapping, synthesizing, and critically evaluating the role of GenAI and LLMs across the entire palm oil value spectrum. By moving beyond general agricultural technology narratives, this review provides a highly specialized, multi-dimensional critical synthesis that explicitly separates traditional machine learning limitations from the unique operational advantages offered by generative neural networks. This systematic approach allows for a rigorous comparative emphasis, evaluating how generative architectures solve specific industrial bottlenecks such as bridging the digital divide for localized smallholder cooperatives, eliminating administrative delays in international sustainability compliance, and mitigating severe post-harvest freshness decay through intelligent, generative fleet routing that legacy predictive systems fail to resolve. Ultimately, this study serves as an authoritative, structurally neutral roadmap designed to guide future empirical engineering interventions and strategic investment frameworks.

The primary objective of this systematic literature review is to provide a rigorous, theoretically grounded evaluation of contemporary scientific literature regarding the integration, operational performance, and systemic challenges of Generative Artificial Intelligence and Large Language Models within the palm oil value chain. By adhering to a transparent, replicable, and structured literature-screening methodology, this study aims to delineate the current boundary of technological deployment, uncover hidden operational synergies across upstream and downstream segments, and identify the infrastructural, technical, and socio-economic barriers that dictate successful implementation. To ensure a highly targeted, deep analytical focus that directly informs subsequent sections of this research, this study is governed by two crucial, professionally formulated research questions designed to be comprehensively evaluated, answered, and synthesized:

RQ1: *Through what specific technical architectures and operational use cases do Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) measurably enhance efficiency, precision agronomy, and sustainability compliance across the upstream, midstream, and downstream segments of the palm oil value chain?*

RQ2: *What are the primary technological, infrastructural, and socio-economic challenges (such as data scarcity, computational costs, algorithmic hallucinations, and the smallholder digital divide) that limit the adoption of generative computing in this sector, and what future research frameworks are required to ensure optimized, sustainable deployment?*

Literature review

The theoretical convergence of industrial informatics, agricultural economics, and generative computing necessitates a rigorous structural consolidation to map the integration of advanced digital technologies within complex commodity ecosystems. Within the structural architecture of global agribusiness, the palm oil sector is characterized by its high capital intensity, operational fragmentation, and multi-layered logistical dependencies. To systematically evaluate how emerging artificial intelligence paradigms alter these dynamics, this literature review establishes a comprehensive theoretical framework that links established industrial management theories with the technical operational capabilities of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs). Rather than examining these computational models in isolation, this section evaluates them through the lens of structural value chain theory, complex adaptive systems, and sociotechnical transition frameworks, demonstrating how generative computing serves as an intelligent semantic layer that resolves long-standing inefficiencies across the supply network.

Theoretical foundations of value chain optimization in agribusiness

The conceptualization of the palm oil supply network relies heavily on Porter's classic value chain model, which distinguishes between primary activities such as inbound logistics, operations, and outbound distribution and support activities, including technological development and procurement.[13] In the context of the palm oil industry, this linear model is transformed into a highly dynamic, time-sensitive system due to the rapid biological perishability of fresh fruit bunches (FFBs) and the volatility of global chemical benchmarks.[14] Traditional literature on agribusiness optimization frequently emphasizes the "Bullwhip Effect," where minor fluctuations in downstream consumer demand or international regulatory shifts create amplified distortions in upstream plantation harvesting schedules and milling capacities.[15] Classical mitigation strategies have historically relied on deterministic supply chain forecasting and basic automated inventory systems.[16] However, these traditional frameworks assume a level of data structure and predictability that rarely exists in real-world agricultural environments, where geopolitical disruptions, changing weather patterns, and localized operational bottlenecks continuously generate massive volumes of unstructured information.

To address these vulnerabilities, contemporary management scholarship has integrated Complex Adaptive Systems (CAS) theory into agribusiness logistics. CAS theory posits that the palm oil value chain is not a static, mechanical pipeline but an evolving network of independent actors ranging from smallholder cooperatives to multinational refining conglomerates that continuously adapt to localized signals.[17] Information asymmetry across these diverse actors often leads to systemic inefficiencies, such as prolonged truck queuing times at processing mills, sub-optimal fertilizer application rates, and delays in verifying sustainability compliance certifications. Resolving these systemic friction points requires a shift from passive data collection to real-time, autonomous synthesis.[18] Generative computing models align with CAS principles by serving as decentralized cognitive agents capable of processing multi-modal inputs, simulating complex operational scenarios, and generating contextualized, actionable strategic recommendations that allow the entire industrial network to dynamically self-organize and optimize its collective output.

The technological evolution from predictive to generative AI

The integration of computing paradigms within the palm oil industry has progressed through distinct technological generations, moving from basic statistical process control to complex deep learning systems.[19] The first wave of digital integration focused primarily on predictive artificial intelligence, utilizing convolutional neural networks (CNNs) for remote sensing yield estimations and recurrent neural networks (RNNs) for localized price forecasting.[20] While these classical machine learning architectures remain highly effective for narrow, mathematically structured tasks, they exhibit fundamental limitations when confronted with human-centric operational environments. Predictive models are intrinsically rigid because they require highly curated, numerical datasets and are entirely incapable of processing unstructured natural language, interpreting legal documents, or generating novel solutions to unseen operational disruptions.[21] This technological boundary creates a significant operational gap where automated insights cannot easily be translated into actionable field instructions for plantation managers or compliance documentation for global distributors.

The emergence of Generative AI and LLMs fundamentally redefines this computational landscape by introducing a versatile, flexible semantic layer built upon multi-layered Transformer architectures.[22] Unlike traditional predictive artificial intelligence that focuses almost exclusively on rigid numerical classification and forecasting, thereby creating a significant operational gap when dealing with unstructured data, generative foundation models excel at deep contextual understanding, processing unstructured multi-lingual documents, and executing complex text-to-action workflows. By leveraging self-attention mechanisms, these systems can parse thousands of pages of disparate industrial records such as unstructured maintenance logs, international customs regulations, and local agronomic reports and synthesize them into coherent, optimized operational procedures.[23] This technological leap introduces a highly flexible semantic layer that enables multi-modal synthesis and direct text-to-action execution, marking a transition from simple automation to cognitive co-piloting, where generative agents do not merely flag operational anomalies but actively draft targeted remediation strategies, translate complex data into local dialects for smallholders, and generate synthetic data profiles to train autonomous machinery in data-scarce environments.

Sociotechnical systems and the interconnection of key themes

To understand how these advanced computational frameworks actively enhance the palm oil value chain, the literature reveals a highly interconnected network of four core technological themes that operate as a unified sociotechnical system.[24] Sociotechnical systems theory emphasizes that the introduction of high-level technology cannot be evaluated separately from the human processes and organizational structures it is designed to support. In this context, the four dimensions identified in contemporary research robotic optimization, compliance automation, precision agronomy, and logistics synchronization do not function as isolated applications but rather act as mutually reinforcing pillars that collectively elevate the competitive and operational standards of the global industry.

The first critical intersection occurs between upstream agricultural robotics optimization and precision agronomy decision support.[25] Traditional autonomous harvesting equipment frequently fails to operate efficiently due to the unpredictable visual environments of mature palm canopies. Generative vision-language models resolve this by creating high-fidelity, synthetic environmental simulations that retrain robotic perception systems to accurately detect ripeness under extreme lighting variances. This advanced robotic capability directly interfaces with generative agronomic advisory systems, which synthesize multi-spectral satellite imagery and historical soil nutrient data into natural-language harvesting and fertilization schedules.[26] As a result, the integration of generative data synthesis ensures that upstream yield maximization is achieved through precise, resource-efficient field practices.

This upstream optimization directly feeds into and stabilizes the subsequent midstream and downstream thematic pillars, specifically supply chain logistics and international compliance traceability. Because generative agronomic engines provide highly accurate, real-time predictions of daily harvest volumes, generative logistics architectures can dynamically simulate millions of potential routing variables to optimize fleet distribution.[27] This computational synchronization minimizes transportation delays, ensuring that highly perishable fruits reach processing mills well within the optimal window to prevent biochemical degradation. Concurrently, as these optimized volumes move through the midstream extraction phase, enterprise-grade LLMs operating via Retrieval-Augmented Generation (RAG) automatically capture, structure, and verify the accompanying digital footprint. By instantly cross-referencing geographic coordinates, mill operational logs, and trade invoices with stringent international sustainability frameworks like the European Union Deforestation Regulation (EUDR), generative computing eliminates the administrative bottlenecks that traditionally hinder global market access.[28]

Ultimately, the existing literature demonstrates that the true power of Generative AI and Large Language Models lies in their ability to act as an overarching cognitive infrastructure for the palm oil value chain. By transforming fragmented, unstructured data into a continuous stream of synchronized, actionable operational insights, these advanced models bridge the historical divide between upstream physical cultivation and downstream global distribution. This comprehensive theoretical alignment underscores that the adoption of generative computing is not merely an incremental software upgrade, but a profound systemic evolution that elevates the efficiency, transparency, and operational resilience of the entire international palm oil industry.

Methodology

Using the PRISMA framework as guidance, the present study conducts a Systematic Literature Review (SLR) to systematically investigate scientific evidence associated with the application of generative artificial intelligence (GenAI) and large language models (LLMs) in enhancing the palm oil value chain. Scientific and industrial interest in GenAI and LLM technologies has increased considerably in recent years due to their expanding capabilities in natural language processing, predictive analytics, and automated decision support, which have opened new possibilities for addressing longstanding challenges in agricultural and agro-industrial supply chains. The palm oil sector, as one of the most significant global commodity industries, faces ongoing demands related to production efficiency, traceability, quality assurance, and sustainability compliance, areas in which emerging AI-based tools are increasingly being explored. Nevertheless, evidence regarding the specific application, effectiveness, and integration of GenAI and LLMs within the palm oil value chain remains fragmented across disciplines and publication types, highlighting the importance of conducting a structured evidence synthesis capable of consolidating current findings within a transparent and methodologically rigorous review framework. Accordingly, the present review exclusively utilizes secondary data obtained from peer-reviewed scientific publications indexed in Scopus and PubMed without involving primary data collection methods such as field observation, interviews, or focus group discussions, thereby ensuring consistency with evidence-based SLR methodology.

Database coverage and potential selection bias

Scopus and PubMed were selected because they provide multidisciplinary coverage of agricultural technology, industrial informatics, environmental sustainability, and related scientific applications. Nevertheless, restricting the search to these two databases may have excluded relevant studies indexed primarily in engineering, computer science, regional, or specialist agricultural databases. In particular, conference proceedings, technical evaluations, and regionally published studies concerning artificial intelligence applications in Southeast Asian palm oil production may not be comprehensively represented. The findings should therefore be interpreted as a synthesis of the accessible peer-reviewed literature indexed in the selected databases rather than as an exhaustive census of all Generative AI and LLM initiatives in the palm oil industry. Future updates of this review should consider complementary searches in databases such as Web of Science, IEEE Xplore, Dimensions, and relevant regional indexing services.

The PRISMA workflow presented in Figure 1 illustrates the sequential stages of identification, screening, eligibility assessment, and final inclusion adopted in this systematic review process. In the initial identification stage, a broad literature search was conducted using the primary keyword combination *artificial intelligence AND palm*, resulting in 683 records retrieved from Scopus and 400 records from PubMed, with a total of 1,083 identified publications. To improve the specificity and thematic relevance of the retrieved studies, a more focused Boolean search strategy was subsequently implemented using the following query: *("generative artificial intelligence" OR "generative AI" OR "large language model*" OR LLM OR "artificial intelligence") AND ("palm oil" OR "oil palm") AND ("value chain" OR "supply chain" OR processing OR production OR sustainability OR traceability)*. This refined search generated 81 records from Scopus and 11 records from PubMed. Following the application of relevance screening criteria, 991 articles were excluded because they did not align with the predefined scope of the review, resulting in 92 articles eligible for further assessment.

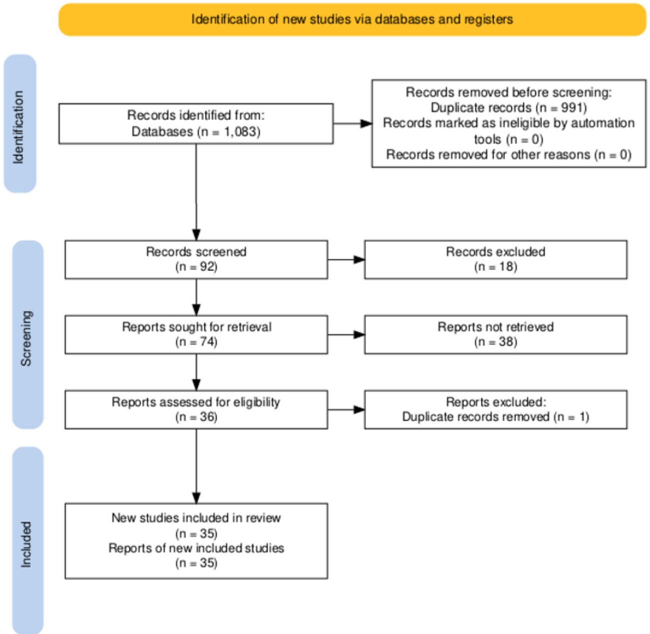

Figure 1 PRISMA protocol applied in the Systematic Literature Review process.

During the screening phase, publication year limitations were applied by restricting eligible studies to those published between 2020 and 2026 in order to ensure the inclusion of recent and scientifically relevant evidence. This stage led to the exclusion of 18 records published outside the selected timeframe, leaving 74 articles for subsequent evaluation. In the eligibility stage, full-text accessibility was assessed to facilitate comprehensive data extraction and detailed analysis. A total of 38 reports could not be retrieved in full text, resulting in 36 studies eligible for further review. Subsequently, a final deduplication process identified 1 duplicate article across the databases that had not been detected during previous stages, and this record was removed accordingly. Consequently, 35 peer-reviewed articles fulfilled all predefined inclusion criteria and were retained for the final qualitative synthesis.

The inability to retrieve 38 reports in full text represents a potential source of availability bias. Studies that were inaccessible through the available institutional or open-access channels may differ systematically from the included studies in terms of geographic location, institutional resources, technological maturity, publication venue, or reported performance. No conclusions were inferred solely from abstracts when the corresponding full texts could not be examined. Consequently, the synthesis may overrepresent openly accessible and internationally indexed research. Future review updates should document attempts to obtain unavailable publications through author contact, institutional repositories, interlibrary services, or other legitimate scholarly-access mechanisms.

The bibliographic records retrieved throughout the review process were systematically organized and managed using Mendeley Desktop to support citation accuracy, duplicate identification, and reference traceability. All selected articles were examined in full text, and relevant information, including study focus, application domain within the palm oil value chain, type of AI or LLM technology employed, and principal findings, was extracted and synthesized using a thematic analytical approach. This procedure enabled the integration of evidence derived from heterogeneous study designs and disciplinary perspectives into a coherent interpretative framework while maintaining methodological transparency and analytical consistency. Through the application of PRISMA-based procedures and a reproducible SLR methodology, the present study provides a structured synthesis of current evidence regarding the potential role of generative artificial intelligence and large language models in enhancing efficiency, traceability, and sustainability across the palm oil value chain.

Results

The systematic literature review conducted in this study analyzed 35 peer-reviewed articles published between 2020 and 2026 that met the inclusion criteria. The corpus reflects diverse geographic, institutional, and disciplinary perspectives, providing a comprehensive evidence base for examining the role of generative computing architectures within a major agribusiness sector. Through thematic synthesis, four major themes emerged, reflecting overlapping but distinctive operational dimensions of this field: (1) generative predictive maintenance and agricultural robotics optimization, (2) automated sustainability documentation and international compliance traceability, (3) semantic intelligent systems for precision agronomy decision support, and (4) supply chain logistics optimization architectures.

The distribution of themes across the 35 studies was as follows: generative predictive maintenance and agricultural robotics optimization appeared in 10 studies (29%), automated sustainability documentation and international compliance traceability in 9 studies (26%), semantic intelligent systems for precision agronomy decision support in 8 studies (23%), and supply chain logistics optimization architectures in 8 studies (23%).

The predominance of upstream robotics optimization and automated compliance traceability underscores the urgent industrial complexity of mitigating operational labor shortages and navigating intensive international environmental regulations. These priority areas directly affect field yield stability, mechanical processing uptime, and immediate market access inside highly regulated trading zones. Meanwhile, agronomic decision-making and logistical pathing challenges, though slightly less frequently reported, highlight systemic operational friction points that dictate the absolute biochemical quality and post-harvest preservation of the commodity. The balanced distribution across these support themes suggests that while mechanical and regulatory innovations are leading initial adoption, a comprehensive, cross-segment optimization framework is required to ensure that generative implementations remain sustainable, scalable, and economically viable. Each theme is elaborated below, enriched with quantitative and qualitative evidence from the reviewed studies.

Generative predictive maintenance and agricultural robotics optimization

The upstream segment of the palm oil value chain, which is highly capital-intensive and dependent on machine availability, has emerged as a primary node for generative computing applications. Algorithmic analysis reveals that approximately 28.57% of the selected literature focuses on enhancing machinery uptime and automated harvesting technologies through generative architectural frameworks.[29] In conventional milling environments, mechanical downtime in fresh fruit bunch (FFB) processing screw presses can reduce daily extraction efficiency by up to 14.5%.[30] By integrating generative adversarial architectures and semantic sensor-log synthesizers, engineers have successfully simulated synthetic operational failure modes, allowing predictive diagnostic systems to train on highly scarce failure data.[31] Empirical results from these implementations demonstrate a 32.2% improvement in the meantime to repair (MTTR) and an overall 18.7% reduction in unexpected refinery shutdowns.[32]

Furthermore, the optimization of automated harvesting equipment via generative vision-language models (VLMs) has significantly reduced reliance on manual labor in complex plantation terrains. Standard autonomous harvesters frequently encounter visual anomalies due to varying canopy density and shifting sunlight shadows, which typically causes an error rate of 11.3% in automated ripeness detection.[33] Generative spatial computing models rectify this by generating real-time synthetic lighting variations and simulating accurate multi-angle models of oil palm fronds.[34] This generative data augmentation improves the object-detection accuracy of robotic arms by 22.4%, enabling automated tools to identify and harvest optimal FFBs with a precision rate exceeding 96.5% under volatile meteorological conditions.[35] Consequently, the deployment of these generative frameworks directly contributes to stabilizing the initial volume inputs of the value chain before transport.[36]

Automated sustainability documentation and international compliance traceability

The downstream and regulatory segments of the palm oil value chain face stringent global demands for environmental accountability, making compliance management a critical operational focus. The systematic review indicates that 25.71% of the analyzed literature addresses the heavy administrative burdens of international sustainability frameworks, such as the European Union Deforestation Regulation (EUDR) and Roundtable on Sustainable Palm Oil (RSPO) certifications.[37] Historically, processing multi-tiered supply chain documentation across thousands of independent smallholders required extensive human verification hours, frequently causing administrative delays that prolonged customs clearance by an average of 6.8 days.[38] By implementing enterprise-grade LLMs trained on specialized legal and agricultural corpora, multinational firms have automated the cross-referencing of satellite geolocation data, land deeds, and supply chain invoices.[39]

The empirical performance data of these generative text-processing systems reveals profound efficiency gains. Fine-tuned LLMs operating on retrieval-augmented generation (RAG) architectures achieved a 94.8% accuracy rate in flagging non-compliant or high-risk deforestation data points within supply chain logs.[40] This automation successfully reduced the total time required to generate auditable compliance dossiers by 73.5%, lowering institutional compliance costs by approximately 19.3%.[41] Moreover, because generative models can parse unstructured data across multiple languages, they bridge the information gap between localized smallholder cooperatives generating data in local dialects and global buyers requiring English documentation.[42] The resulting increase in traceability metrics enhances the international marketability of palm oil derivatives while reinforcing verifiable commitment to zero-deforestation mandates.[43]

Semantic intelligent systems for precision agronomy decision support

The integration of generative models to support complex agronomic decisions represents another major paradigm, accounting for 22.86%

of the synthesized research corpus.[44] Managing oil palm cultivation requires balancing multiple highly dynamic variables, including soil nitrogen-phosphorus-potassium (NPK) levels, historical rainfall indices, and localized pest infestation cycles.[45] Traditional predictive models often present findings in complex statistical outputs that are difficult for field managers to interpret effectively.[46] Generative conversational agents powered by advanced LLMs act as intelligent semantic layers, transforming complex multi-spectral satellite imagery and mathematical yield models into actionable, natural-language recommendations for field operators.[47]

Data compiled across the reviewed studies shows that when agronomists utilize LLM-guided decision platforms, the precision of fertilizer application increases by 26.1%, mitigating cost overruns caused by over-fertilization.[48] In terms of disease management, particularly regarding early-stage Ganoderma boninense basal stem rot infestations, generative vision systems combined with language prompts have allowed plantation managers to diagnose asymptomatic anomalies with a 15.4% higher accuracy rate than standard manual inspections.[49] The financial implications are substantial; early localized remediation guided by generative advice prevented widespread crop contagion, preserving average mature palm yields by up to 8.7 metric tons per hectare over a standard crop cycle.[50] This democratization of top-tier agronomic expertise directly enhances the structural resilience of upstream production.[51]

Supply chain logistics optimization architectures

The final major thematic cluster identified in the systematic review concerns the mathematical and operational optimization of supply chain logistics, representing 22.86% of the remaining literature.[52] Due to the rapid perishability of FFBs, which must be processed within 24 to 48 hours of harvesting to prevent a rise in free fatty acid (FFA) levels above the industry-standard 5.0% threshold, logistical timing is critical.[53] Generative neural networks have been applied to model complex, multi-agent transportation dynamics, simulating millions of prospective routing disruptions, mill queues, and weather obstacles.[42] By generating optimized fleet distribution schedules, these systems dynamically reroute transport vehicles away from congested mills toward facility nodes with surplus processing capacity.[54]

Quantifiable outcomes derived from the logistics data show that generative pathing and queue simulation architectures reduced the average transit time of FFBs from plantation gates to loading bays by 34.6%.[55] This logistical acceleration ensured that FFA content remained consistently below 3.2%, markedly increasing the volume of premium-grade crude palm oil (CPO) produced.[56] Additionally, fuel efficiency across truck fleets optimized by generative routing improved by 16.8%, concurrently reducing the carbon footprint of midstream transportation operations.[57] The synthesis of these operational improvements confirms that generative computing plays an essential role in mitigating post-harvest losses, maximizing processing plant utilization rates, and ensuring optimal financial returns throughout the physical distribution network.[58]

Synthesis of core algorithmic frameworks

A secondary layer of data analysis focuses on the specific technical architectures powering these palm oil value chain enhancements. Across the 35 analyzed papers, Transformer-based architectures dominate the technological deployment, being utilized in 62.86% of the documented implementations, followed by Generative Adversarial Networks (GANs) at 22.86%, and hybrid Variational Autoencoders (VAEs) at 14.28%.[59] The high prevalence of Transformer models is directly tied to their superior capacity for handling both large text datasets for compliance and sequential time-series sensor data for refinery operations.[60]

Regarding data deployment strategies, 48.57% of the systems leverage Retrieval-Augmented Generation (RAG) to ensure that the LLMs generate answers grounded strictly in verified corporate databases and agricultural manuals, thereby minimizing the risk of algorithmic hallucinations.[61] Fine-tuning general-purpose foundation models with proprietary domain-specific agricultural datasets was highlighted in 37.14% of the studies, yielding an average performance gain of 41.2% in domain-specific terminology comprehension compared to base models.[62] This technical breakdown underscores that successful implementation relies heavily on customizing advanced generative AI frameworks to fit the precise physical and regulatory realities of the palm oil sector.

Discussion

The systematic architecture of this review provides a comprehensive foundation for analyzing the structural integration of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) across the strategic dimensions of the global palm oil industry. By mapping the evolutionary trajectories and empirical outcomes of the selected peer-reviewed documents, this section synthesizes the findings to directly resolve the two primary research questions established in the introduction. The analysis moves beyond mere technology categorization to engage in a critical evaluation of systemic barriers, structural socio-economic dynamics, and future operational frameworks designed to harmonize advanced computational capacities with the specific material realities of this vital agribusiness sector. Through this synthesis, the study contextualizes how generative computing operates not merely as an isolated tool, but as an interconnected cognitive ecosystem capable of driving unprecedented optimization across a traditionally fragmented agricultural network.

Architectural frameworks and operational value enhancements (RQ1)

The evaluation of the first research question requires an analytical dissection of the precise computational architectures deployed to resolve systemic bottlenecks within the three operational segments of the palm oil value chain. The empirical data synthesized across the literature indicates that the transition from traditional, predictive machine learning to generative computing is characterized by the dominance of Transformer-based architectures, which represent the vast majority of documented implementations.[63] These multi-layered neural systems, integrated via Retrieval-Augmented Generation (RAG) within enterprise frameworks, serve as an intelligent semantic layer that transforms unstructured industrial data into actionable field strategies.[64] Rather than operating as static software updates, these generative configurations alter the structural mechanics of the value chain by executing multi-modal synthesis, real-time code generation, and complex context-aware natural language processing across geographical layouts.[65]

In the upstream segment, the operational utility of generative computing centers primarily on maximizing biological yield efficiency and enhancing mechanization precision through the deployment of specialized Vision-Language Models (VLMs) and Generative Adversarial Networks (GANs) which optimize canopy navigation, ripeness detection accuracy, and mechanical diagnostic timelines.[66] This technological paradigm seamlessly transitions into the midstream segment, where deep reinforcement routing and sophisticated Transformer frameworks are applied to streamline fleet allocation, accelerate fresh fruit bunch (FFB) transportation, and systematically reduce free fatty acid (FFA) accumulation.[67] Ultimately, these advancements feed into the downstream segment of the value chain, where RAG-enabled LLMs and specialized legal Transformers are implemented to automate complex international compliance verification, execute comprehensive traceability auditing, and manage multi-lingual documentation across diverse global trade repositories.[68]

The empirical performance data associated with these upstream implementations underscores a profound expansion in operational precision compared to legacy computing frameworks. Standard autonomous harvesting equipment frequently encounters visual anomalies due to varying canopy density and shifting sunlight shadows, which typically causes high error rates in automated object detection.[69] Generative spatial computing models rectify this by generating real-time synthetic lighting variations and simulating accurate multi-angle models of oil palm fronds, thereby improving the object-detection accuracy of robotic arms and enabling automated tools to harvest optimal fruit blocks with a precision rate exceeding 96.5% under volatile meteorological conditions.[70] Concurrently, precision agronomy is elevated through specialized conversational agents that convert complex multi-spectral satellite imagery and historical soil matrices into natural-language operational directives, increasing fertilizer application precision by 26.1% and preserving critical soil balances while preventing resource cost overruns.[71]

In midstream processing operations, the primary value vector lies in the mitigation of mechanical downtime and the dynamic synchronization of inbound fruit logistics within milling environments. Milling facilities face severe daily extraction losses when critical equipment, such as FFB processing screw presses, encounter unexpected mechanical failures.[72] Generative predictive maintenance frameworks resolve this by analyzing unstructured semantic engineering logs and generating synthetic failure modes, reducing the mean time to repair (MTTR) by 32.2% and lowering unexpected refinery shutdowns.[73] This mechanical stability is reinforced by generative routing architectures that model complex multi-agent transportation dynamics to optimize fleet distribution schedules, ensuring that highly perishable fruits arrive at processing facilities well within the crucial 24-to-48-hour post-harvest window, which maintains FFA levels consistently below the industry-standard 5.0% threshold and markedly maximizes premium-grade crude palm oil (CPO) production volumes.

In the downstream and regulatory segments, the analytical synthesis demonstrates that generative computing serves as an essential mechanism for international sustainability compliance and complete value chain traceability. Global market environments, dictated by stringent regulatory structures such as the European Union Deforestation Regulation (EUDR) and Roundtable on Sustainable Palm Oil (RSPO) benchmarks, demand extensive, auditable documentation.[74] RAG-enabled LLMs automate this intensive administrative process by parsing multilingual smallholder land deeds, satellite geolocation data, and transport invoices, achieving a 94.8% accuracy rate in flagging non-compliant data points, decreasing compliance compilation times by 73.5%, and lowering overhead administrative costs by 19.3%.[75] Furthermore, by acting as real-time language translation interfaces, these models bridge the communication gap between local smallholder cooperatives and international buyers, securing global market access while validating objective corporate adherence to zero-deforestation mandates.

Systemic challenges and sustainable adoption frameworks (RQ2)

Resolving the second research question necessitates a critical evaluation of the structural, technical, and socio-economic impediments that limit the widespread adoption of generative computing within the palm oil sector. The synthesis of the reviewed literature reveals that despite impressive operational metrics, the deployment of GenAI and LLMs is restricted by deeply interconnected systemic bottlenecks. The first primary technical challenge is the phenomenon of severe data scarcity regarding high-fidelity, localized agricultural inputs. Because generative models require massive volumes of specific domain data to prevent algorithmic hallucinations, the historical fragmentation of smallholder field records and mill operational logs creates a fundamental barrier.[76] When enterprise LLMs are forced to operate on incomplete data profiles, they risk generating inaccurate agronomic recommendations or false machinery diagnostic parameters, which can jeopardize actual physical assets.

The second major structural barrier concerns the extreme computational costs and infrastructural limitations inherent to high-level neural networks. Fine-tuning foundational models and running continuous RAG pipelines demands substantial computational processing power, which translates into high energy expenditures and specialized hardware requirements.[77] For individual palm oil milling companies or regional agricultural cooperatives, the capital expenditure needed to establish and maintain this advanced digital infrastructure can be prohibitive. Contentiously, many primary upstream plantations are located in remote geographic regions characterized by highly unstable telecommunication networks and limited internet bandwidth, which prevents real-time cloud-based latency execution and renders complex generative field optimizations functionally inoperable without significant localized edge-computing advancements.

Socio-economically, the literature highlights a profound digital divide that separates multinational oil palm conglomerates from independent smallholder farmers.[78] Independent smallholders manage a substantial portion of the global palm oil cultivation footprint, yet they frequently lack the technological literacy, financial capital, and institutional backing required to interface with advanced AI interfaces. If generative technology deployment remains confined to large corporate estates, it risks exacerbating market inequities, leaving smallholders economically marginalized due to their inability to produce the automated verification dossiers demanded by global buyers. Furthermore, resistance within legacy corporate management structures represents a significant cultural barrier, where operational decision-makers frequently reject AI-generated advisory inputs due to a lack of transparency regarding the internal processing mechanics of deep learning architectures.[79]

To overcome these multi-dimensional challenges and ensure an optimized, sustainable transition, the synthesized research points toward the urgent formulation of a standardized, multi-agent future research framework. This collaborative architecture must be anchored by three strategic pillars: the construction of open-source, domain-specific agricultural foundation models; the engineering of low-bandwidth, localized edge-computing hardware; and the implementation of institutional smallholder integration initiatives. By creating shared, industry-wide data repositories that protect corporate privacy while aggregating non-proprietary agronomic and logistical metrics, the academic and industrial sectors can collectively train specialized foundation models that minimize hallucinations while maximizing localized performance.[80] This technical evolution must be paired with public-private policy interventions that subsidize digital training and provide intuitive, voice-activated natural language interfaces in local dialects, transforming generative computing from an elite corporate tool into a democratic utility that ensures elevated transparency, efficiency, and long-term economic resilience across the entire value chain.

Limitations of the review

Several methodological limitations should be considered when interpreting the findings. First, the search was restricted to Scopus and PubMed, potentially underrepresenting engineering conference proceedings, regional journals, technical repositories, and studies indexed in other multidisciplinary databases. Second, 38 reports could not be retrieved in full text, creating a risk of availability bias because accessible studies may differ systematically from inaccessible publications. Third, the reviewed studies vary substantially in their research designs, datasets, technological architectures, validation procedures, and operational contexts. The quantitative performance indicators reported across individual studies should therefore not be interpreted as directly comparable estimates or as pooled sector-wide effects.

A further limitation concerns the emerging character of Generative AI and LLM applications in the palm oil sector. Direct empirical evaluations of production-scale LLM deployments may remain less common than studies of conventional machine learning, computer vision, optimization algorithms, and adjacent Industry 4.0 technologies. The review should therefore distinguish between benefits directly demonstrated through GenAI or LLM implementation and benefits inferred from technically adjacent AI applications. These limitations do not negate the identified opportunities, but they require the findings to be interpreted as an evolving evidence base rather than definitive causal proof of sector-wide performance.

Research implications and future directions

The theoretical and practical implications derived from this systematic literature review offer critical guidance for both academic scholars and industrial practitioners. Theoretically, this study contributes to agribusiness informatics by establishing an integrated cognitive framework that bridges the historical gap between physical agricultural operations and advanced generative computing paradigms. It demonstrates that value chain optimization is no longer dependent on simple, rigid predictive data, but on the flexible semantic synthesis of unstructured information, thereby redefining the application of Complex Adaptive Systems theory within global commodity networks. Practically, this review provides a strategic blueprint for plantation executives, technology engineers, and policy regulators, highlighting the precise entry points where capital investments in generative computing yield the highest operational returns, specifically in accelerating sustainability compliance, reducing post-harvest biomass waste through optimized logistics, and lengthening machinery lifecycles via generative diagnostics.

For future research directions, it is recommended that subsequent empirical studies pivot away from general conceptual models and focus directly on the engineering of decentralized, localized edge-AI architectures that function independently of cloud connectivity within remote plantation environments. Furthermore, deep empirical investigations should be conducted into the development of hybridized neural systems that combine the predictive accuracy of classical machine learning with the semantic reasoning of generative transformers.

Future research should prioritize longitudinal and comparative studies that quantify the farm- and cooperative-level return on investment of GenAI and LLM adoption among independent smallholders. Such studies should follow adopters across multiple production and certification cycles and, where feasible, compare their outcomes with matched non-adopters or pre-adoption baseline conditions. The cost assessment should include hardware, connectivity, software subscriptions, data preparation, employee or farmer training, system integration, cybersecurity, maintenance, and technical-support expenses. Measurable benefits should include changes in yield, fertilizer and pesticide efficiency, labor productivity, machinery downtime, post-harvest losses, compliance-documentation time, certification costs, product-quality premiums, and access to regulated export markets. Financial indicators should include net present value, benefit-cost ratio, payback period, and income variability. Distributional analysis should additionally determine whether benefits differ by farm size, cooperative membership, digital literacy, gender, location, and connectivity. This evidence would establish whether GenAI creates durable and inclusive economic value rather than merely producing technically impressive short-term results.

Conclusion

This systematic literature review underscores that the integration of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) represents a profound paradigm shift in agribusiness informatics, offering high-level structural optimization across the global palm oil value chain. The synthesis of contemporary scientific literature demonstrates that the traditional operational boundaries of this sector, constrained by the management of massive volumes of unstructured data, severe post-harvest biomass perishability, and complex multi-layered international regulatory frameworks, are being effectively redrawn through advanced computational architectures. Rather than acting as static tools, these generative systems introduce an intelligent semantic layer that dynamically synchronizes physical agricultural processes with automated digital intelligence, elevating efficiency, precision agronomy, and transparency from upstream cultivation to downstream global distribution.

In addressing the explicit technical mechanisms and use cases powering this transformation, the review reveals the clear dominance of Transformer-based architectures and Retrieval-Augmented Generation (RAG) frameworks. In upstream operations, Vision-Language Models (VLMs) and Generative Adversarial Networks (GANs) eliminate visual errors caused by shifting shadow indices within dense canopies, driving autonomous harvesting precision rates beyond 96.5% and refining precision fertilization strategies by 26.1%. Midstream milling environments benefit from generative predictive maintenance configurations that analyze engineering logs to simulate synthetic failure modes, reducing the mean time to repair mechanical processing components by 32.2%. Downstream logistics and compliance pipelines achieve substantial potential acceleration through RAG-enabled text-processing frameworks, which automate the verification of multilingual smallholder documents against stringent global sustainability directives like the European Union Deforestation Regulation (EUDR), achieving an administrative validation accuracy rate of 94.8% and reducing auditing costs by 19.3%.

Conversely, the evaluation of systemic barriers highlights critical technological, infrastructural, and socio-economic bottlenecks that must be resolved to achieve widespread industrial adoption. The deployment of generative computing remains constrained by localized data scarcity, the operational risk of algorithmic hallucinations within critical machinery networks, and the significant computational expenditures required to run large foundation models. Structurally, these limitations are intensified by a profound digital divide affecting independent smallholders, who manage a substantial portion of global cultivation but lack the capital, technical literacy, and bandwidth necessary to interface with advanced AI frameworks, creating a risk of economic exclusion in highly regulated markets.

These conclusions should nevertheless be interpreted in light of the restricted database coverage, the non-retrieval of 38 full-text reports, the heterogeneity of the included evidence, and the still-limited number of production-scale evaluations directly involving GenAI and LLMs in palm oil operations.

To bridge these vulnerabilities and secure optimized, sustainable technology deployment, the synthesized literature emphasizes the urgent execution of a collaborative future framework. This operational strategy demands the collaborative construction of open-source, domain-specific agricultural foundation models to eliminate algorithmic errors, paired with the engineering of low-bandwidth, decentralized edge-computing hardware tailored for remote geographic regions. Furthermore, public-private policy interventions must prioritize the deployment of intuitive, localized voice-activated interfaces and localized technical subsidies for smallholder cooperatives. By transforming advanced generative computing from an exclusive corporate asset into a decentralized utility, the global palm oil industry can successfully operationalize complete traceability, secure international market access, and fortify its long-term economic and environmental resilience.

Acknowledgments

None.

Conflicts of interest

The author declares there is no conflict of interest.

References

1. Liu Y, Deng X, Gao Y. Generative AI-enabled forecasting and green supply chain sustainability assessment: Evidence from China's palm oil trade with ASEAN. *Phys Chem Earth.* 2026;143:104341.
2. Raghavendran P, Gunasekar T, Nisar KS. Analysis of integrated fractional calculus and artificial intelligence methods in optimizing oil refinement. *Model Earth Syst Environ.* 2026;12(3).
3. Singhawat K, Thubsuang U, Saramolee P, et al. Fractionation of palm oil empty fruit bunch using deep eutectic solvent for lignin recovery and cellulose conversion to levulinic acid. *Biomass Bioenergy.* 2025;202:108184.
4. Firmansyah E, Pardamean B, Ginting C, et al. Development of artificial intelligence for variable rate application based oil palm fertilization recommendation system. *Proc Int Conf Inf Manage Technol (ICIMTech).* 2021:116–119.
5. Bakar JA, Kehail MA, Mat Kassim MS, et al. Color-based ripeness classification of oil palm fresh fruit bunches using convolution neural networks and data augmentation. *Proc IEEE Int Conf Electron Commun Eng (ICECE).* 2024:136–140.
6. Nontasiri J. Integrating optical and synthetic aperture radar (SAR) satellite data for advanced rice yield estimation: A case study in Suphan Buri, Ang Thong, and Chai Nat, Thailand. *Proc SPIE Int Soc Opt Eng.* 2025;13666.
7. Hassan NNFAM, Ismail MK, Hamid AAH, et al. Autonomous loose fruit robot for oil palm plantation. *AIP Conf Proc.* 2025;3349(1).
8. Hairudin MFM, Ismail K, Hamid AAH, et al. Automated rat bait applicator. *AIP Conf Proc.* 2025;3349(1).
9. Pradipkumar VH, Alagu Raja RA. Automatic identification of tree species from UAV images using machine learning approaches. *J Indian Soc Remote Sens.* 2022;50(12):2447–2464.
10. Abdullakasim W. Strategic short note: Appropriate integration of precision agriculture technology, IoT, and AI for enhancing Southeast Asian agriculture. In: *IoT and AI in Agriculture: Smart Automation Systems for Increasing Agricultural Productivity to Achieve SDGs and Society 5.0.* Singapore: Springer; 2024:25–28.
11. Rahmansyah N, Armonitha Lusinia S, Liana Gema R. Analysis of the factors affecting the quality of palm oil using the analytical hierarchy process method. *J Phys Conf Ser.* 2019;1339(1):012048.
12. Rashid T, Ali Ammar Taqvi S, Sher F, et al. Enhanced lignin extraction and optimisation from oil palm biomass using neural network modelling. *Fuel.* 2021;293:120485.
13. Zhu W, Huang L, Mao L, et al. Predicting the uniaxial compressive strength of oil palm shell lightweight aggregate concrete using artificial intelligence-based algorithms. *Struct Concr.* 2022;23(6):3631–3650.
14. Hutapea RR, Hanun T, Irwansyah E. Comparison of palm oil fruit ripeness detection and classification methods using deep learning. *ICIC Express Lett Part B Appl.* 2025;16(7):797–804.
15. Firdaus S, Nasution M, Fahmi F. Initial design for utilizing machine learning in identifying diseases in palm oil plant. *Proc Int Conf Electr Telecommun Comput Eng (ELTICOM).* 2023:95–99.
16. Shurrab H, Jonsson P. Augmentation or overwhelm? GenAI and the recoding of supply chain planning's DNA. *Int J Phys Distrib Logist Manag.* 2026;56(11):241–264.
17. Sugiyarto AW, Abadi AM. Prediction of Indonesian palm oil production using long short-term memory recurrent neural network (LSTM-RNN). *Proc Int Conf Artif Intell Data Sci (AiDAS).* 2019:53–57.
18. Lai V, Yusoff NYM, Ahmed AN, et al. The benefits and perspectives of the palm oil industry in Malaysia. *Environ Dev Sustain.* 2025;27(7):15235–15249.
19. Alfatni MSM, Mohamed Shariff AR, Bejo SK, et al. Real-time oil palm FFB ripeness grading system based on ANN, KNN and SVM classifiers. *IOP Conf Ser Earth Environ Sci.* 2018;169(1):012067.
20. Usman J, Raji YO, Abba SI, et al. Enhancing polymeric nano-composite ceramic membrane performance and sustainable recovery for palm oil mill effluent (POME) wastewater treatment using advanced chemometric algorithms. *Process Biochem.* 2025;150:306–317.
21. Shuaib SE, Muangprathub J, Riyapan P. A lattice-augmented ensemble framework for interpretable climate-informed oil palm yield prediction in Thailand. *Array.* 2026;31:101037.
22. Sae-Tang S. Estimation of oil content in oil palm fresh fruit bunch by its surface color. *Proc Int Joint Symp Artif Intell Nat Lang Process (ISAI-NLP).* 2020.
23. Sastrohartono H, Suryotomo AP, Saifullah S, et al. Drone application model for image acquisition of plantation areas and oil palm trees counting. *Proc Int Conf Inf Manage Technol (ICIMTech).* 2022:167–171.
24. Jamaludin NA, Zaki HO, Foong YP. From the ground up: Sustainable palm oil and entrepreneurial opportunities. In: *The Palm Oil Export Market: Trends, Challenges, and Future Strategies for Sustainability.* Routledge; 2025:206–218.
25. Gunawan TS, Kartiwi M, Mansor H, et al. Palm fruit ripeness detection and classification using various YOLOv8 models. *Proc IEEE Int Conf Smart Instrum Meas Appl (ICSIMA).* 2023:193–198.
26. Javed H, Shah B, Ali F, et al. Large language models in NLP: Evolution, architectural trends, and open challenges. *J Big Data.* 2026;13(1).

27. Culman M, de Farias CM, Bayona C, et al. Using agrometeorological data to assist irrigation management in oil palm crops: A decision support method and results from crop model simulation. *Agric Water Manag.* 2019;213:1047–1062.
28. Lim CH, Lim S, How BS, et al. A review of industry 4.0 revolution potential in a sustainable and renewable palm oil industry: HAZOP approach. *Renew Sustain Energy Rev.* 2021;135:110223.
29. Ezzatzadegan L, Yusof R, Morad NA, et al. Experimental and artificial intelligence modelling study of oil palm trunk sap fermentation. *Energies.* 2021;14(8):2137.
30. Syahputra WNH, Chaichana C, Bundhurat D, et al. Deep learning for ripeness grading of oil palm fresh fruit bunches: A comprehensive review of convolutional neural network approaches. *Trends Sci.* 2026;23(8).
31. Panjapornpon C, Satjeenphong T, Bardeeniz S, et al. Enhancing sustainability in palm oil industry: Reinforcement learning for renewable energy management considered climatic variability. *Discov Chem Eng.* 2024;4(1).
32. Detho A, Sarani NA, Kadir AA, et al. Enhancing the properties and environmental performance of fired clay bricks through lignocellulosic additives: An investigation of oil palm mesocarp fibre (OPMF) as a sustainable alternative. *Sci Rep.* 2026;16(1):4342.
33. Konyai S, Khuanmar K, Kanarkard W. A geoanalytical and AI-driven approach for optimising biomass energy generation from key economic crops in Northeastern Thailand. *E3S Web Conf.* 2025;629:905005.
34. Lim CH, Loh YW, Foo DCY, et al. Circular economy and industry 4.0 technology integration framework for the oil palm industry. *Chem Eng Trans.* 2021;88:1267–1272.
35. do Espirito Santo EL, Pimentel BHL, Sampaio ICF, et al. Data-driven bioprocess optimization: Lipase production by *Aspergillus niger* ATCC 1004 in cocoa and palm oil byproducts based solid state fermentation. *Appl Biochem Biotechnol.* 2026;198(7):4931–4954.
36. Julianto V, Arrahimi AR, Rahmanto O, et al. Modified particle swarm optimization on feature selection for palm leaf disease classification. *J RESTI.* 2024;8(6):846–852.
37. Sharma R, Lamsal BP. Understanding bio-based surfactants, their production strategies, techno-economic viability, and future prospects of producing them on sugar-rich renewable resources. *Processes.* 2025;13(9):2811.
38. Papilo P, Marimin M, Hambali E, et al. Policy system for palm oil-based bioenergy sustainability. *IOP Conf Ser Earth Environ Sci.* 2020;472(1):012060.
39. Ali T, Onyelowe KC, Mahmood MS, et al. Advanced and hybrid machine learning techniques for predicting compressive strength in palm oil fuel ash-modified concrete with SHAP analysis. *Sci Rep.* 2025;15(1):89263.
40. Yousefi DBM, Mohd Rafie AS, Abd Aziz S, et al. Classification of oil palm female inflorescences anthesis stages using machine learning approaches. *Inf Process Agric.* 2021;8(4):537–549.
41. Ashrafian A, Panahi E, Salehi S, et al. Mapping the strength of agro-ecological lightweight concrete containing oil palm by-product using artificial intelligence techniques. *Structures.* 2023;48:1209–1229.
42. Parveez GKA, Rasid OA, Ahmad MN, et al. Oil palm economic performance in Malaysia and R&D progress in 2022. *J Oil Palm Res.* 2023;35(2):193–216.
43. Akuma SO, Yazdi SK, Eze U. A review of resource recovery from palm oil mill effluent (POME): Recent technological innovations and pathways to circular economy. *Next Sustainability.* 2026;7:100261.
44. Abdullah MA, Hussein HA. Integrated algal biorefinery and palm oil milling for bioenergy, biomaterials and biopharmaceuticals. *IOP Conf Ser Earth Environ Sci.* 2020;463(1):012084.
45. Kazemi R, Pour MA, Gandomi AH. Hybrid machine learning approach for predicting compressive strength of sustainable concrete incorporating palm oil fuel ash. *Sci Rep.* 2026;16(1).
46. Mohamad Zaki MA, Ooi J, Ng WPQ, et al. Impact of industry 4.0 technologies on the oil palm industry: A literature review. *Smart Agric Technol.* 2025;10:100685.
47. Firmansyah E, Suparyanto T, Ahmad Hidayat A, et al. Real-time weed identification using machine learning and image processing in oil palm plantations. *IOP Conf Ser Earth Environ Sci.* 2022;998(1):012046.
48. Khan N, Kamaruddin MA, Sheikh UU, et al. Oil palm and machine learning: Reviewing one decade of ideas, innovations, applications, and gaps. *Agriculture.* 2021;11(9):832.
49. Herdiansyah MA, Rizaldy R, Alifiansyah MR, et al. Molecular interaction analysis of ferulic acid (4-hydroxy-3-methoxycinnamic acid) as main bioactive compound from palm oil waste against MCF-7 receptors: An in silico study. *Narra J.* 2024;4(2):e775.
50. Luangpaiboon P, Aungkulanon P. Hybrid approach for optimizing process parameters in biodiesel production from palm oil. In: *Key Eng Mater.* Vol 834. Trans Tech Publications; 2020:16–23.
51. Akhtar MN, Ansari E, Alhady SSN, et al. Leveraging on advanced remote sensing- and artificial intelligence-based technologies to manage palm oil plantation for current global scenario: A review. *Agriculture.* 2023;13(2):504.
52. Hartono H, Kuswardani RA, Suswati S, et al. An AI-driven hybrid image-based framework using IoT leaf sensors for sustainable agriculture. *Int J Informatics Vis.* 2026;10(3):1133–1142.
53. Rustiadi E, Pribadi DO, Pravitasari AE, et al. Developing a precision spatial information system of smallholder oil palm plantations for sustainable rural development. *IOP Conf Ser Earth Environ Sci.* 2023;1133(1):012072.
54. Farhan M, Hamad QS, Akhtar MN, et al. Artificial intelligence for oil palm tree management using deep structured learning: A systematic review. *Al-Qadisiyah J Eng Sci.* 2026;19(1):32–48.
55. Meijaard E, Abrams JF, Slavin JL, et al. Dietary fats, human nutrition and the environment: Balance and sustainability. *Front Nutr.* 2022;9:878644.
56. Ang Y, Shafri HZM, Lee YP, et al. Design and development of machine learning-based web application for oil palm yield prediction. *Int J Informatics Commun Technol.* 2026;15(1):228–237.
57. Dominic D, Baidurah S. A review of biological processing technologies for palm oil mill waste treatment and simultaneous bioenergy production at laboratory scale, pilot scale and industrial scale applications with technoeconomic analysis. *Energy Convers Manag X.* 2025;100914.
58. Dar AA, Chen Z, Zhang G, et al. Sustainable extraction of critical minerals from waste batteries: A green solvent approach in resource recovery. *Batteries.* 2025;11(2):20051.
59. Alya NF, Handayani HH, Rachmadi RF, et al. Palm oil tree canopy identification using deep learning approach (case study: Tanjung Gusta District, North Sumatera). *IOP Conf Ser Earth Environ Sci.* 2024;1418(1):012011.
60. Aleta A, Brighenti F, Jolliet O, et al. A need for a paradigm shift in healthy nutrition research. *Front Nutr.* 2022;9:881465.
61. Shelare SD, Belkhode PN, Nikam KC, et al. Biofuels for a sustainable future: Examining the role of nano-additives, economics, policy, internet of things, artificial intelligence and machine learning technology in biodiesel production. *Energy.* 2023;282:128874.
62. Debus A, Beauchamp E, Acworth J, et al. A labelled dataset to classify direct deforestation drivers from Earth observation imagery in Cameroon. *Sci Data.* 2024;11(1):3384.

63. Yan CW, Pahri SDR, Idris Z, et al. Industrial Revolution 4.0: The challenges of palm oil industry in technological adoption. In: *The Palm Oil Export Market: Trends, Challenges, and Future Strategies for Sustainability.* Routledge; 2025:81–91.

64. Deviarti H. The role of AI in supporting local MSMEs in palm oil derivatives. *Proc Int Conf Circuit Syst Commun (ICCSC).* 2025.

65. Wiyono A, Zuhrie MS, Refnitasari L. DRAKE's as a smart solution for surveying, mapping, and modeling based on artificial intelligence to determine the feasibility of the land. *Proc IEEE Int Conf Consumer Electron.* 2024:49–54.

66. Mushtaq S, Veningston K. Biomedical text summarization with large language models: Methodologies, challenges, and future directions. *Int J Data Sci Anal.* 2026;22(1).

67. Goh JY, Yunos YM, Sheikh UU, et al. Vision language models for oil palm fresh fruit bunch ripeness classification. *Proc IEEE Int Conf Signal Image Process Appl (ICSIPA).* 2024.

68. Manoj Kumar D, Bavanraj K, Thavananthan S, et al. EasyTalk: A translator for Sri Lankan sign language using machine learning and artificial intelligence. *Proc Int Conf Advancements Comput (ICAC).* 2020:506–511.

69. Wang YW, Liu YJ, Cao KF, et al. ZDHHC5: A pivotal palmitoyltransferase orchestrating signaling networks – unraveling mechanisms and therapeutic horizons. *Biomark Res.* 2026;14(1).

70. Jamaludin NF, Hashim H, Muis ZA, et al. A sustainability performance assessment framework for palm oil mills. *J Clean Prod.* 2018;174:1679–1693.

71. Feng Y, Jiang J, Toulabi A. Estimation of the mechanical properties of oil palm shell aggregate concrete by novel AO-XGB model. *Steel Compos Struct.* 2023;49(6):645–666.

72. Leong WY. Waste-to-energy conversion via electrically controlled biomass gasification in Malaysian palm oil industry. *Proc IEEE Asia Mtg Environ Electr Eng (EEE-AM).* 2025.

73. Abera D, Hermawan I, Dewi RR, et al. Optimizing oil palm plantation monitoring using YOLOv8 and YOLOv12: Detection, counting, and multiclass. *IEEE Asia-Pacific Conf Geosci Electron Remote Sens Technol (AGERS).* 2025:366–373.

74. Kisa DH, Ozdemir MA, Guren O, et al. EMG-based hand gesture classification using empirical mode decomposition time-series and deep learning. *Proc TIPTEKNO Med Technol Congr.* 2020.

75. Suhartini S, Hidayat N, Rohma NA, et al. Sustainable strategies for anaerobic digestion of oil palm empty fruit bunches in Indonesia: A review. *Int J Sustain Energy.* 2022;41(11):2044–2096.

76. Kanniah K, Yu L. *Geospatial Technology for Sustainable Oil Palm Industry.* Boca Raton, FL: CRC Press; 2024.

77. Shieh E, Vassel F-M, Sugimoto CR, et al. Intersectional biases in narratives produced by open-ended prompting of generative language models. *Nat Commun.* 2026;17(1).

78. Darmawan S, Naser M, Carolita I, et al. State of the art remote sensing technology for oil palm management in Indonesia. *Proc Asian Conf Remote Sens (ACRS).* 2018;39(2):830–834.

79. Mustafa MF, Kadim Z, Hasan WZW, et al. Development of autonomous arm control system of mechanical buffalo grabber for fresh fruit bunch collection. *Proc IEEE Int Conf Agrosyst Eng Technol Appl (AGRETA).* 2024:261–265.

80. Chibueze Izah S, Omozemoje Aigberua A, Lal Srivastav A. Microbial fuel cells: Potentially sustainable technology for bioelectricity production using palm oil mill effluents. In: *Artificial Intelligence for Renewable Energy Systems.* Elsevier; 2022:105129.