Journal of Deep Learning, Computer Vision and Digital Image Processing
E-ISSN: 2986-8939 P-ISSN: 2986-8920
Vol. 4, No. 2, June 2026
DOI. https://doi.org/10.61255/decoding.v4i2.1404

Design and Implementation of a Dual-LLM Prescriptive ESG Reporting System in the Indonesian Palm Oil Industry

Niko Firzi Anansyah
(Institut Teknologi Sawit Indonesia, Medan, Indonesia)

Ratu Mutiara Siregar ✉
(Institut Teknologi Sawit Indonesia, Medan, Indonesia)

Andi Prayogi
(Institut Teknologi Sawit Indonesia, Medan, Indonesia)

Muhammad Akbar Syahbana Pane
(Universitas Negeri Medan, Medan, Indonesia)

OPEN ACCESS

ARTICLE HISTORY
Received: April 29, 2026
Revised: June 5, 2026
Accepted: June 28, 2026

KEYWORDS
ESG;
Information system;
LLM;
Palm oil sustainability;
Prescriptive analytics

ABSTRACT

Purpose – This study aims to develop an automated Environmental, Social, and Governance (ESG) reporting information system based on Large Language Model (LLM) for the palm oil industry to overcome low efficiency, data inconsistencies, and analysis limitations inherent in manual reporting processes.
Methods – This applied research employs the Design Science Research (DSR) paradigm, encompassing needs analysis, system design, implementation, and black-box testing. The system was developed using the Laravel MVC framework and integrated a Dual-LLM API failover architecture (Groq Llama 3 as primary and Gemini as backup). The case study was conducted at PT Surya Mata Ie.
Findings – The developed system successfully automated ESG indicator extraction and prescriptive narrative generation. It utilizes a Strict Weighting Rule, a programmatic safeguard capping the ESG score at 50.0 (as a proof-of-concept testing constraint) if Ganoderma infection exceeds a 20% threshold (supported by agronomic research). During prototype evaluation, this rule intercepted an overly optimistic raw LLM score of 60.0 and corrected it to 50.0. This demonstrates the system's capability to function as a risk-control mechanism, mitigating potential hallucination-driven score inflation and supporting mathematically accountable outputs.
Research implications – The implementation of this system significantly accelerates reporting workflows and serves as an early warning instrument for environmental risks, thereby enhancing real-time managerial decision-making and corporate transparency in complying with global sustainability standards.
Originality – This study pioneers the integration of a Dual-LLM failover mechanism within a Laravel framework tailored for the palm oil sector. It introduces a novel programmatic constraint approach in JSON object parsing to maintain strict mathematical accountability in AI-generated ESG drafts.

Correspondence Author: ✉ratu_ms@itsi.ac.id
To cite this article: Anansyah, N. F., Siregar, R. M., Prayogi, A., & Pane, M. A. S. (2026). Design and implementation of a dual-LLM prescriptive ESG reporting system in the Indonesian palm oil industry. Journal of Deep Learning, Computer Vision and Digital Image Processing, 4(2), 176–191. https://doi.org/10.61255/decoding.v4i2.1404

This is an open access article under the CC BY-SA license

INTRODUCTION

The palm oil industry in Indonesia has long served as one of the most vital pillars supporting the national economy. As a leading global producer, it employs millions of workers across its extensive supply chain and contributes substantially to the country's foreign exchange reserves. Yet the continuous expansion and daily operations of this sector face intense global scrutiny, especially regarding the management of its social and environmental impacts. In today's business environment, palm oil companies must strictly comply with increasingly rigorous sustainability standards. A key requirement is the preparation of accurate, comprehensive, and transparent Environmental, Social, and Governance (ESG) reports. The ESG framework evaluates company performance beyond financial metrics by addressing three core dimensions. It serves as an essential guide to ensure sustainable operations and responsibility toward all stakeholders. In the palm oil sector, which is often linked to issues such as deforestation, land conflicts, and labor rights, transparent ESG reporting is crucial to demonstrate genuine commitment to sustainability [1].

Despite this urgent need, many companies still depend on conventional manual reporting systems. These processes are highly time-consuming and prone to recording errors. Such reliance creates major obstacles in meeting global demands for sustainability reporting, which are increasingly enforced by strict international regulations [2], [3]. As a result, manual ESG workflows frequently generate incomplete, inconsistent, and delayed reports, leading to high data redundancy. This inefficiency complicates timely decision-making and reduces the reliability of information provided to stakeholders [4], [5]. The large volume of field data that must be manually reviewed also makes real-time monitoring of operational sustainability nearly impossible, ultimately harming corporate reputation [6]. In an ideal setting, a modern information system should integrate human expertise with advanced technology. Systems such as Transaction Processing Systems (TPS) form the foundation for recording routine transactions and converting raw data into timely, accurate information for decision makers [6].

Extensive prior research has shown how advanced information technology can fundamentally transform data management and reporting processes. Artificial Intelligence (AI), particularly Large Language Models (LLM) and Natural Language Processing (NLP), has proven highly effective in automating large-scale administrative tasks. Large Language Models are advanced AI systems trained on vast linguistic datasets, enabling them to understand, extract, and generate natural language with impressive accuracy. Recent state-of-the-art studies indicate that NLP techniques and LLMs can automatically extract unstructured environmental factors to create accurate ESG alternative data for corporate sustainability monitoring [1]. In the plantation sector, the adoption of Industry 4.0 principles through Smart Agriculture approaches has supported sustainability efforts and enhanced productivity.

Current literature further highlights that LLMs excel at processing massive volumes of complex text data. These systems can automatically generate coherent report narratives even within heterogeneous supply chains, significantly reducing human analytical efforts and automating knowledge extraction from unstructured operational data, marking a critical transition toward prescriptive analytics in complex agricultural and agri-food domains [7], [8], . For unstructured data such as daily field reports from plantation foremen, modern NLP algorithms play a vital role. Techniques including entity extraction, text classification, and Cosine Similarity allow for more accurate analysis than manual methods [9]. The application of AI in agricultural management has consistently reduced workload while improving report quality [10]. To minimize risks such as hallucination, strict prompt engineering and inference parameter restrictions are essential [7]. Building upon these principles to ensure domain-specific precision in the palm oil industry, the proposed system implements deterministic JSON output controls and programmatic rule enforcement on the backend.

Although many studies have confirmed the success of Large Language Models and NLP in simplifying reporting within finance and healthcare, their application in the Indonesian palm oil industry remains very limited. The primary gap lies in the sector's continued heavy dependence on traditional, fragmented ESG reporting practices. Most data handling still occurs through manual mechanisms that lack integration within a unified system. This creates a wide disparity between the enormous daily operational data generated by plantations and the ability to transform it into comprehensive sustainability reports. Existing managerial information systems in the palm oil sector rarely utilize unstructured textual data effectively. Qualitative inputs such as daily field notes and incident descriptions remain largely underprocessed. Consequently, ESG indicator extraction and narrative drafting still depend almost entirely on human interpretation. This absence of automation leads to inconsistent standards, potential bias, redundant work, and prolonged reporting cycles. Ultimately, companies struggle to achieve real-time monitoring and respond quickly to evolving international sustainability requirements.

Given the operational challenges and clear technological gap identified, there is an urgent need for an innovative solution to streamline data management in palm oil plantations. Developing an integrated ESG reporting information system powered by Large Language Models represents a logical and strategic approach. AI offers unmatched capabilities in efficiently handling both qualitative and quantitative data at scale. By integrating LLM technology, repetitive administrative tasks can be significantly reduced, freeing the industry from over-reliance on manual processes that often cause errors. LLMs also adapt well to varied data formats, from narrative reports to structured metrics, and can synthesize them into consistent, high-quality sustainability narratives. Furthermore, the system enables predictive analytics to provide deeper insights into long-term sustainability performance. In a competitive global market, such automation not only eases operational burdens but also strengthens competitiveness by enhancing transparency and adaptability to international regulations.

This research seeks to address these systemic issues by designing, developing, and evaluating an automated ESG reporting information system based on Large Language Models tailored to the palm oil industry. Specifically, the study aims to create a robust system that digitizes and integrates the entire ESG reporting workflow. The technical core involves implementing NLP-LLM algorithms to extract ESG indicators from raw text and automatically generate accurate report narratives. In addition to system development, this research focuses on evaluating the functional reliability of the system as a working prototype. The study is conducted as a case study at PT Surya Mata Ie. To explicitly address the identified gaps, this study establishes several core research objectives. Primarily, it aims to design and develop an automated, LLM-based ESG reporting information system tailored for the palm oil industry that integrates both unstructured daily field notes and structured operational datasets. Furthermore, the study seeks to implement NLP and LLM algorithms for the automated extraction of ESG indicators and the generation of structured sustainability narratives. Finally, it evaluates the functional feasibility of the proposed system architecture using structured black-box testing methods. Successful implementation is expected to enhance corporate transparency, improve report quality, and support better management decision-making, thereby strengthening the sustainability of Indonesia's palm oil industry on the global stage.

METHOD

This study employed an applied research methodology adapting the system architecture design paradigm, which emphasizes systematic stages of analysis, design, implementation, and testing to create innovative farm management IT artifacts [11]. Ethical approval and institutional authorization were formally obtained from the management of PT Surya Mata Ie, granting permission to utilize historical operational datasets for research purposes. The population consists of datasets from the palm oil plantation at PT Surya Mata Ie. Purposive sampling was applied, focusing on two specific operational blocks during the April 2026 evaluation period to fulfill distinct testing objectives. Blok 1 was selected for routine performance evaluation and end-to-end system demonstration, specifically to validate the system's capability in synthesizing data and rendering a complete, formal ESG PDF report. Conversely, Blok 20 was specifically selected as a high-risk stress-test scenario due to its extreme operational variables, which is necessary to explicitly validate the activation of the programmatic safeguard (Strict Weighting Rule). The selection of these targeted test cases and the use of a lightweight SQLite database were deliberately chosen because this system is a functional prototype. SQLite provides a serverless environment ideal for rapid prototyping and validating the Dual-LLM API architectural logic, serving as a targeted proof-of-concept demonstration before scaling to a robust enterprise-level SQL server deployment.

Figure 1. Class diagram of the proposed MVC-based system architecture

The ESG reporting system was developed using a Model-View-Controller (MVC) architecture on the Laravel framework, with PHP as the programming language and SQLite as the database. This architecture is highly relevant for smart agricultural systems as it facilitates responsive data routing and secure database transactions [12], [13], [14]. This approach responds to the industry's need for digitized monitoring to replace manual harvest tracking, ensuring data integrity and mitigating duplication in operational environments [15], [16]. To systematically define these functional boundaries, a Class Diagram was modeled as presented in Figure 1. As noted by previous research, Unified Modeling Language (UML) modeling accelerates requirement validation and minimizes design flaws before implementation [15]. Figure 1 illustrates the static structure and modular responsibilities within the MVC architecture, highlighting the hierarchical relationships among the User classes (Admin, Mandor, and Manager) and the central Laporan (Report) entity. Crucially, the diagram depicts the LLMService class as the bridging component that handles entity extraction and JSON parsing requests, ensuring structured interactions between the local operational data and the external AI engines. To ensure high availability and mitigate single-point-of-failure issues during server downtimes, the system employs a Dual-LLM API failover architecture. The primary computational engine used was the Groq infrastructure with the Llama 3 model (llama-3.3-70b-versatile), while Google's Gemini model (gemini-1.5-flash) served as the secondary failover backup. LLM parameters were strictly configured to align with academic and industrial data standards. Natural Language Processing (NLP) inherently carries the risk of hallucination when interpreting numerical values [17]. Therefore, the temperature parameter was locked at a deterministic value of 0.1 to minimize AI hallucination when processing statistical numbers and to maintain logical output consistency [7].

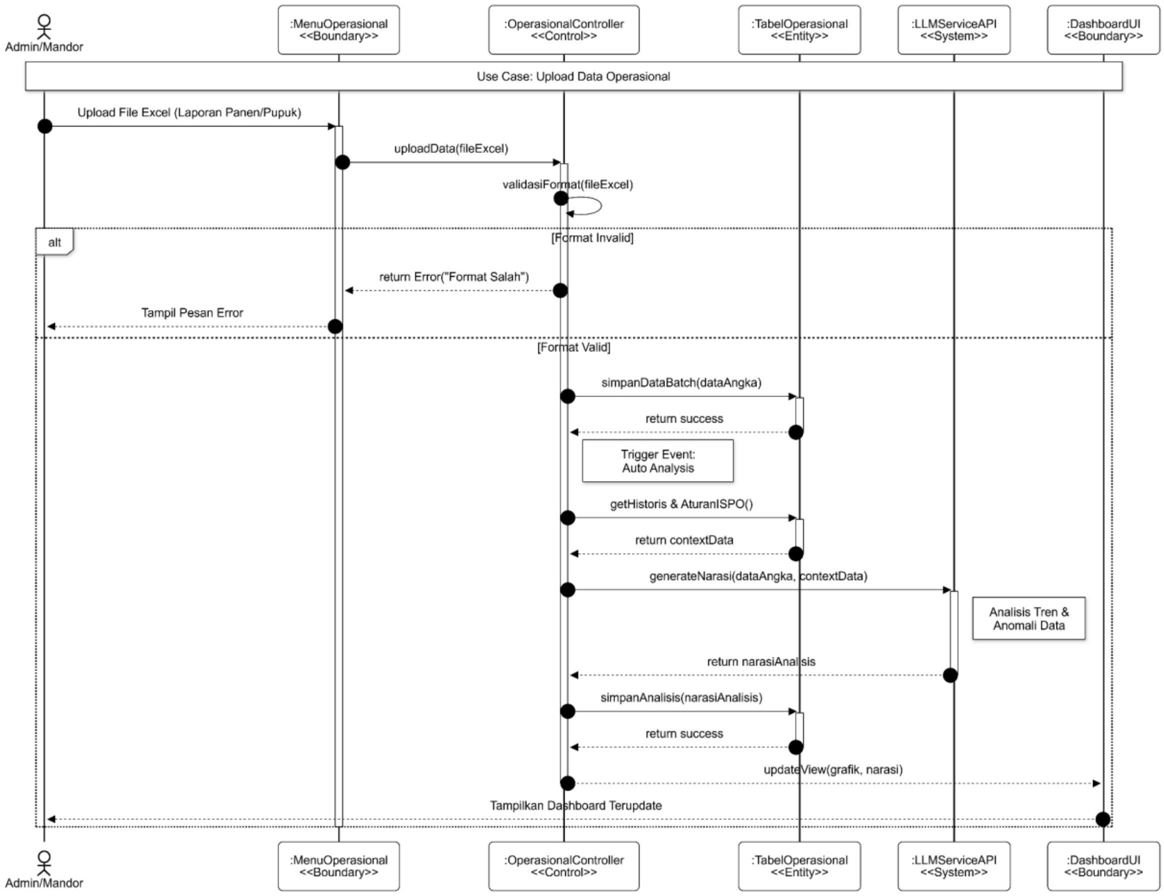


Figure 2. Sequence diagram of operational data submission and AI-driven ESG analysis

The research procedures were organized into a clear data processing flow. Figure 2 presents the Sequence Diagram illustrating the exact end-to-end procedure for transforming field data into a prompt context for the AI. The sequence begins when the user uploads operational data, triggering the application controller to validate the format and save the batch data into the local database. A backend event then automatically consolidates the newly added data with historical records and Indonesian Sustainable Palm Oil (ISPO) compliance rules to construct a specific prompt structure. This contextual prompt instructs the LLM to act as the "Primary Palm Oil ESG Auditor." The LLMServiceAPI subsequently performs trend analysis and anomaly detection, responding with pure JavaScript Object Notation (JSON) output containing the "ESG Risk Score," "Insight Narrative," and "Recommendations." This procedure included programmed integrity controls, for instance, if the Ganoderma infection rate exceeded 20 percent, the maximum ESG score was automatically capped at 50.0. This 20% threshold represents a critical epidemiological boundary where basal stem rot disease caused by Ganoderma boninense requires immediate intervention before reaching a mass mortality phase that severely limits plantation sustainability [18]. The 50.0 maximum score cap was established as a proof-of-concept programmatic constraint to demonstrate the system's capability to override and mitigate hallucination-driven score inflation during the software testing phase. This deterministic constraint, applied directly into the software's backend logic, bridges the gap between probabilistic generative text and exact mathematical thresholds [9]. Finally, the system updates the database, refreshes the manager's dashboard UI with the synthesized ESG insights, and renders the JSON format into a formal PDF document.

Instrument feasibility was evaluated exclusively using the black-box testing method. The main focus of this testing was to validate whether the system's functional outputs such as file upload, AI-generated JSON parsing, and report rendering met the specified requirements without examining the internal code logic [19]. Success was analyzed by comparing the strict weighting rules, such as score limitations during high infection rates with the final documents produced by the LLM. The structured test cases, input conditions, and functional evaluation results of this black-box testing are systematically documented in Table 1. The methodological limitations of this study lie in the absence of extensive usability testing with end users and the lack of inter-rater reliability assessment such as Fleiss' Kappa statistics [20] from a panel of human agronomy experts to measure the linguistic validity of the LLM narratives. The validity of the artificial intelligence performance in this study was measured purely based on its functional black-box capability to correctly interpret prompt syntax and complete the process without errors or timeouts.

Table 1. Black-box testing scenarios and functional evaluation results

Test Scenario & Repetitions	Test Input & Failure Condition	Expected Behavior	Actual Observation & Evidence	Status
Authentication & RBAC (Repetitions: 5x)	**Input:** User logs into the system using Admin, Mandor, and Manager accounts**. Failure Condition:** Unauthorized role attempting to access a restricted URL.	System redirects each user to the appropriate interface according to RBAC policies. Unauthorized access is blocked.	**Actual Output:** System successfully authenticated and redirected users based on role permissions. **Evidence:** User sessions initialized correctly without privilege escalation errors.	Valid
Data Input and Upload (Repetitions: 5x)	**Input:** Administrator uploads an operational .xlsx file. **Failure Condition:** Invalid file format or data type mismatch.	System parses the uploaded file and stores the extracted numeric records to the SQLite database.	**Actual Output:** File parsed successfully and data integrated into SQLite. **Evidence:** Data records verified in database tables without column mismatch errors.	Valid
LLM Analysis Execution (Repetitions: 5x)	**Input:** Manager clicks Generate AI Analysis for Blok 1. **Failure Condition:** LLM hallucinates format, returning non-JSON text.	Backend transforms data into prompt, enforces JSON parsing, and extracts variables accurately.	**Actual Output:** LLM returned valid JSON format. **Evidence:** parseJsonResponse() executed successfully without triggering Log::error('Gagal men-decode JSON response dari AI').	Valid
API Failover Validation (Repetitions: 3x)	**Input:** Generate AI Analysis request. **Failure Condition:** Network connection to the primary AI provider (Groq) is intentionally interrupted (timeout).	System's try-catch catches the exception and automatically switches to the secondary provider (Gemini 1.5 Flash).	**Actual Output:** Primary API failed. Fallback API executed seamlessly without UI crash. **Evidence Log:** Log::warning('groq gagal, mencoba failover', ['blok' =>	Valid

Test Scenario & Repetitions	Test Input & Failure Condition	Expected Behavior	Actual Observation & Evidence	Status
			'Blok 1']) recorded.	
Safeguard Rule (Strict Weighting Rule) (Repetitions: 5x)	**Input Data (Blok 20):** Ganoderma = 25.78%, Production = 41,083 kg. **Failure Condition:** Ganoderma > 20% critical threshold.	**Expected Score Behavior:** Backend controller must intercept overly optimistic AI scores and cap the ESG Risk Score at max 50.0.	**Raw LLM Score:** 60.0 **Corrected Score:** 50.0 **Final Result:** Score of 50.0 successfully stored. **Evidence:** AuditTrail::record() captured and stored the final adjusted score of 50.00.	Valid
Final Report Generation (Repetitions: 3x)	**Input:** User selects Print Sustainability Report (PDF). **Failure Condition:** Missing metrics or incomplete database retrieval.	System compiles operational statistics and AI narratives into a complete PDF report.	**Actual Output:** PDF rendered completely, including AI narratives and human-in-the-loop signature fields. **Evidence:** PDF generated flawlessly.	Valid

RESULTS AND DISCUSSION

Interface Modeling and System Management

Based on the software architecture engineering process, the Environmental, Social, and Governance (ESG) reporting information system was successfully developed using the Laravel framework with SQLite as the database management system. The implementation of the Model-View-Controller (MVC) approach produced an adaptive and well-structured operational interface, creating a robust separation between backend data processing and frontend presentation. This architecture includes role-based access interfaces such as secure login functionality, an admin panel for data ingestion, a foreman (Mandor) input page, and the central Managerial ESG & Productivity Dashboard.

The system has effectively digitized previously fragmented paper-based recording mechanisms into an actionable Transaction Processing System (TPS) for stakeholders. As explicitly illustrated in Figure 3, the main dashboard interface functions as a visual control tower that centralizes daily operational data and real-time field aggregation as a Decision Support System (DSS). The dashboard is interactively designed to display several critical components. First, it presents macro-level Key Performance Indicators (KPIs), providing executives with an instant overview of total area coverage, active tree populations, overall health indexes, and agronomy compliance rates. Second, it visualizes tree health stratification through intuitive graphical representations, accurately categorizing the plant population into normal, stunted (non-valuer), and dead trees. This granular visualization allows for the rapid detection of biological anomalies and critical environmental threats across different plantation blocks. Furthermore, the interface seamlessly integrates the 'AI Sustainability Insight' module, displaying the autonomously generated prescriptive narratives and ESG Risk Scores directly to the managers. By consolidating these multimodal datasets into a single responsive interface, the dashboard not only visualizes real-time field conditions but also significantly accelerates managerial evaluation, ensuring proactive interventions in plantation sustainability management.

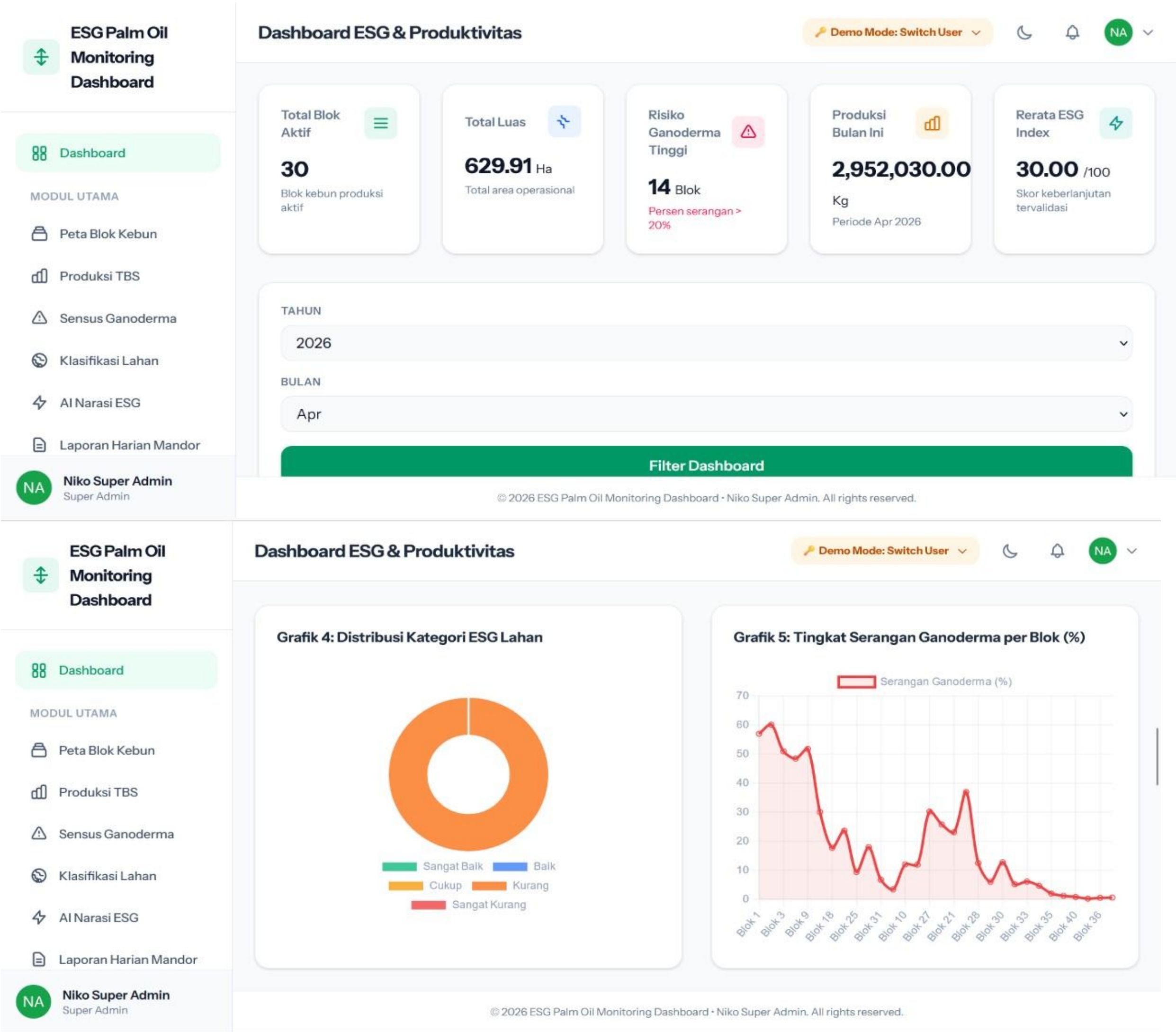


Figure 3. Dashboard interface for real-time monitoring of operational and ESG performance

Failover Mechanism and Dual-LLM Architecture

One of the most critical achievements of this system engineering is the successful implementation of a failure control algorithm, known as the Failover Sequence, within the AIOrchestrator class. In modern cloud-based system architectures, the dependence on external Application Programming Interfaces (APIs) for artificial intelligence computation introduces significant vulnerabilities, including unpredictable latency issues, strict daily rate limiting (token quotas), and unexpected server downtime [7]. To structurally mitigate these risks, the ESG reporting system was configured with a Dual-Large Language Model (LLM) ecosystem. This architecture integrates two primary AI providers, the Groq infrastructure powered by the Llama 3 model (llama-3.3-70b-versatile) as the primary computational engine, and Google's Gemini model (gemini-1.5-flash) serving as the secondary failover backup.

The operational logic of this architecture is managed dynamically through the ApiSetting database, which actively detects and verifies the availability of registered API keys before execution. During the generation of sustainability narratives, the AIOrchestrator applies a robust try-catch exception handling mechanism. If a data request transmitted to the primary service provider (e.g., Groq) encounters an exception, connection timeout, or service unavailability, the system does not halt or crash. Instead, the application controller immediately captures the error, records a detailed warning in the system log (Log::warning, including the error message and the specific plantation block name), and automatically triggers the failover sequence to route the prompt to the backup service provider.

Furthermore, to ensure complete traceability and system auditability, once the failover succeeds, the generated JSON response is securely parsed and saved into the ai_narratives table. This database transaction records not only the generated text but also crucial metadata, including the active ai_provider, ai_model, and token_digunakan (tokens used), allowing management to monitor API consumption efficiently. Internal software evaluation using black-box testing methodology confirmed that this failover mechanism demonstrates the capability to support high availability [19]. The integration within the Laravel MVC framework facilitates the complex sustainability analysis process to completes seamlessly in the background, minimizing potential interruptions or interface crashes for the end-users [14].

Prompt Engineering and Parameter Restriction Logic

This system does not simply forward raw data to the AI engine; rather, it applies highly structured prompt engineering techniques to ensure domain-specific accuracy. The application controller orchestrates a data fusion process, collecting a comprehensive set of variables from the database, including monthly production volume (kg), average bunch weight (BJR), Ganoderma infestation percentage, circle condition (kondisi piringan), and harvest rotation cycle. All these variables are dynamically compiled into a single contextual system instruction that explicitly directs the AI model to act as the “Primary Palm Oil ESG Auditor”. This role-playing technique is crucial in Natural Language Processing (NLP) to restrict the model's analytical scope and ensure that the generated narrative adheres to the terminology and standards of agricultural sustainability. Providing explicit contextual instructions ensures the model follows human intent more closely, akin to instruction-tuning methodologies [22], and effectively elicits step-by-step reasoning capabilities [23]. The detailed architecture for constructing this instruction and transmitting the data from the controller to the AI service provider is illustrated in Figure 4.

The computation results of this algorithm demonstrate that the LLM must be strictly regulated to mitigate the risks of AI hallucination, particularly when interpreting exact statistical figures [7]. Restricting the temperature parameter is a well-documented approach to managing the inherent probabilistic nature of LLMs, reducing fact-conflicting outputs [21], and maintaining expert-level precision in domain-specific tasks [24]. To achieve this deterministic behavior, the system locks the generation temperature parameter at an absolute value of 0.1. Furthermore, the system imposes a rigorous output format constraint. The AI is required to return its response in pure JavaScript Object Notation (JSON) format, specifically encapsulating the fields narasi_insight, rekomendasi, and skor_risiko_esg. This structured data extraction method ensures that the probabilistic generative text can be systematically parsed and safely injected into the relational database without causing data type mismatch errors [9]. Black-box testing confirmed that the parseJsonResponse function executed these extractions flawlessly without any JSON formatting errors [19].

Beyond linguistic controls, the system successfully implements a programmatic safeguard known as the "Strict Weighting Rule" to maintain the mathematical accountability of the ESG assessment. An absolute causal logic is injected directly into the back-end controller: if the field input detects a Ganoderma infestation rate exceeding 20 percent, the AI-generated ESG Risk Score is forcibly capped at a maximum of 50.0 (Categorized as Critical/Warning), regardless of how exceptionally high the production tonnage is. This hard-coded intervention is strategically designed to mitigate "greenwashing" and productivity bias, where massive harvest yields might otherwise mislead the AI into generating overly optimistic sustainability scores despite severe environmental degradation [3], [25]. By integrating this deterministic rule, the system successfully bridges the gap between the probabilistic nature of generative AI and the exact mathematical thresholds required for standardized ESG compliance.

```php
// Parse response JSON
$parsed = $this->parseJsonResponse($result['narasi']);

// Enforce Strict Weighting rule programmatically as a safeguard
$persenSerangan = (float) ($inputData['persen_serangan'] ?? 0);
if ($persenSerangan > 20 && $parsed['skor_risiko_esg'] > 50) {
    $parsed['skor_risiko_esg'] = 50.0;
}

// Simpan ke database
$narrative = AiNarrative::create([
    'blok_id'         => $blok->id,
    'periode'         => $periode,
    'narasi_insight'  => $parsed['narasi_insight'],
    'rekomendasi'     => $parsed['rekomendasi'],
    'skor_risiko_esg' => $parsed['skor_risiko_esg'],
    'status_validasi' => 'menunggu',
    'ai_provider'     => $result['provider'],
    'ai_model'        => $result['model'],
    'token_digunakan' => $result['tokens'],
    'raw_response'    => $result['raw_response'],
    'input_data'      => $inputData,
    'created_by'      => auth()->id(),
]);

// Catat audit trail
AuditTrail::record('ai_generate', $narrative, [], [
    'blok_id'  => $blok->id,
    'provider' => $result['provider'],
```

Figure 4. Prompt engineering formulation algorithm and API integration flow in AIOrchestrator

Actual Analysis Results and ESG Report Document Generation

The system's functional feasibility was demonstrated through direct Large Language Model (LLM) processing tests using historical field data from "Blok 1" during the April 2026 evaluation period. The operational dataset for Blok 1 presented a highly contrasting scenario: it recorded a massive production volume reaching 11,588.00 kg with a rapid harvest rotation of 4.6 days across a land area of 8.32 Ha containing 732 palm trees. However, alongside this peak productivity, the block suffered from an extreme Ganoderma boninense infestation rate of 57.02 percent. Using the llama-3.3-70b-versatile model deployed via the Groq infrastructure, the system's controller successfully orchestrated the data fusion and executed the instructional parameters flawlessly without any timeout errors [19].

The computational results demonstrated the AI's advanced inferential reasoning capabilities. Rather than being misled by the exceptionally high harvest yield, the AI accurately diagnosed the biological anomaly. The model formulated a highly prescriptive diagnostic narrative: "Block 1 has a high production of 11,588.00 kg, but with a very high Ganoderma infection rate of 57.02%. This current high production is a critical phase before a sharp decline due to mass tree mortality" (the original Indonesian output is retained in the system's database). Consequently, the AI naturally generated an ESG Risk Score of 30.0/100. Because this score was already below the critical threshold, the programmatic safeguard correctly remained dormant.

To explicitly demonstrate the activation of the Strict Weighting Rule, an additional evaluation from the system's operational database is presented using the case of Blok 20. In this scenario, Blok 20 exhibited an extreme Ganoderma infection rate of 25.78% based on the April 2026 evaluation period. As visually evidenced in Figure 5, during the initial LLM inference, the AI model generated an overly optimistic raw ESG Risk Score of 60.0 in its JSON response, prioritizing the exceptionally high production yield of 41,083.00 kg over the biological threat. However, the programmed safeguard mechanism functioned exactly as designed. The backend rule evaluated the Ganoderma infection rate (25.78%), and because it exceeded the established 20% critical threshold, the controller intercepted the JSON parsing process. Consequently, the raw AI score of 60.0 was programmatically corrected and capped at a maximum of 50.0. The final score of 50.00 was then successfully stored in the database and displayed on the ESG dashboard. This before-and-after demonstration indicates that the system not only possesses strong natural inferential capabilities but also maintains mathematical accountability through deterministic risk-control intervention when the AI hallucinates, aligning with recent state-of-the-art guidelines which emphasize that LLMs require strict task-specific inference constraints to ensure downstream analytical reliability and safety [25].

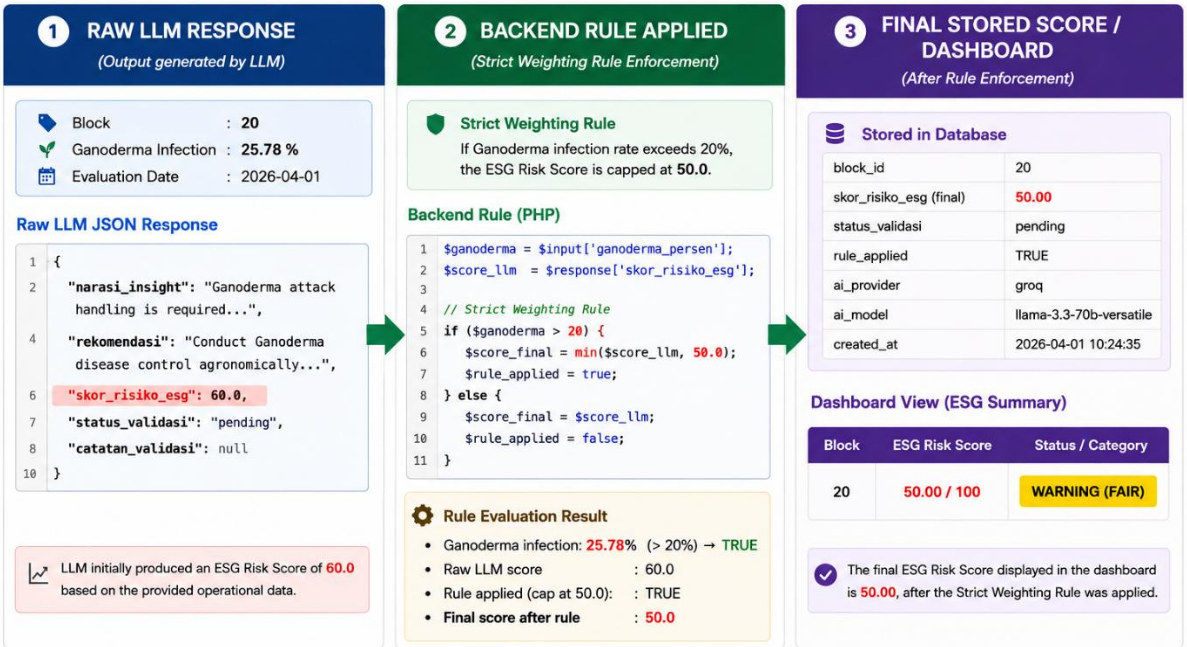
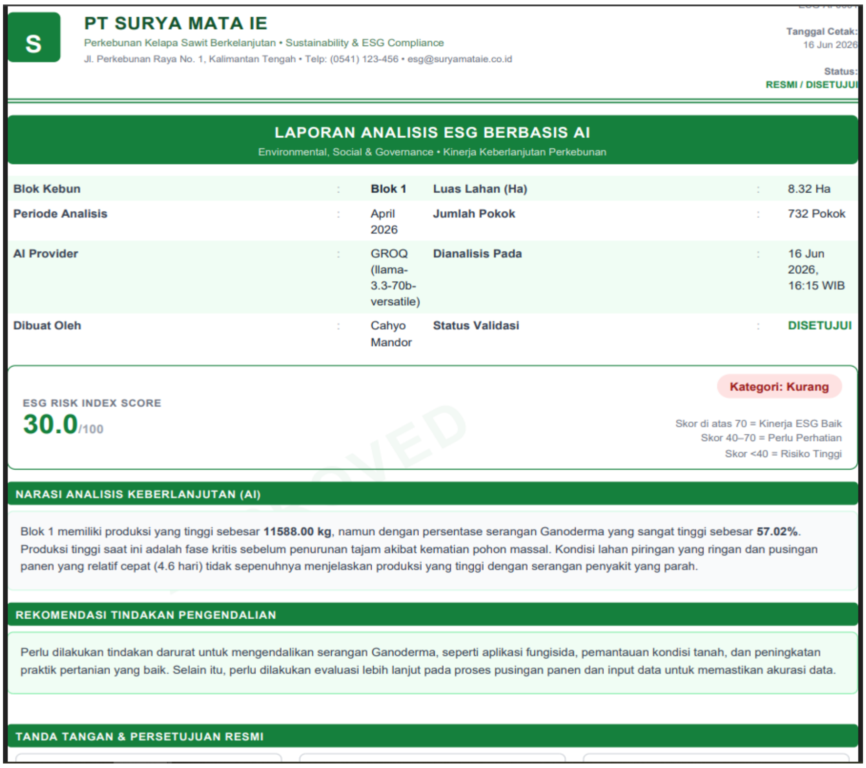


Figure 5. Application of the Strict Weighting Rule for Block 20

The culmination of this intelligent processing for the primary case study (Blok 1) is explicitly represented in Figure 6, which displays the formal AI-Based ESG Analysis Report document generated by the system's Document Engine module document generated by the system's Document Engine module. As illustrated in the figure, the PDF document is professionally structured to serve as an official corporate record. It details the precise analysis metadata, including the AI provider used and the exact timestamp of the evaluation. Beyond the causal narrative and the restricted ESG Risk Score, the report automatically formulates actionable agronomic recommendations, explicitly advising emergency interventions such as fungicide application, soil condition monitoring, and an evaluation of the harvest rotation process. Finally, to ensure administrative accountability and support a human-in-the-loop framework, the generated document includes a formal validation section complete with the approval signatures of the field foreman (Mandor) and the estate manager (Manager), confirming that the AI-generated insights have been reviewed and officially verified for management action [20].



S
PT SURYA MATA IE
Perkebunan Kelapa Sawit Berkelanjutan • Sustainability & ESG Compliance
Jl. Perkebunan Raya No. 1, Kalimantan Tengah • Telp: (0541) 123-456 • esg@suryamataie.co.id

Tanggal Cetak:
16 Jun 2026
Status:
RESMI / DISETUJUI

LAPORAN ANALISIS ESG BERBASIS AI
Environmental, Social & Governance • Kinerja Keberlanjutan Perkebunan

Blok Kebun	:	Blok 1	Luas Lahan (Ha)	:	8.32 Ha
Periode Analisis	:	April 2026	Jumlah Pokok	:	732 Pokok
AI Provider	:	GROQ (llama-3.3-70b-versatile)	Dianalisis Pada	:	16 Jun 2026, 16:15 WIB
Dibuat Oleh	:	Cahyo Mandor	Status Validasi	:	DISETUJUI

ESG RISK INDEX SCORE
30.0/100

Kategori: Kurang

Skor di atas 70 = Kinerja ESG Baik
Skor 40–70 = Perlu Perhatian
Skor <40 = Risiko Tinggi

NARASI ANALISIS KEBERLANJUTAN (AI)

Blok 1 memiliki produksi yang tinggi sebesar **11588.00 kg**, namun dengan persentase serangan Ganoderma yang sangat tinggi sebesar **57.02%**. Produksi tinggi saat ini adalah fase kritis sebelum penurunan tajam akibat kematian pohon massal. Kondisi lahan piringan yang ringan dan pusingan panen yang relatif cepat (4.6 hari) tidak sepenuhnya menjelaskan produksi yang tinggi dengan serangan penyakit yang parah.

REKOMENDASI TINDAKAN PENGENDALIAN

Perlu dilakukan tindakan darurat untuk mengendalikan serangan Ganoderma, seperti aplikasi fungisida, pemantauan kondisi tanah, dan peningkatan praktik pertanian yang baik. Selain itu, perlu dilakukan evaluasi lebih lanjut pada proses pusingan panen dan input data untuk memastikan akurasi data.

TANDA TANGAN & PERSETUJUAN RESMI

Figure 6. AI-generated ESG report integrating operational data with sustainability insights

Discussion

The findings from the system engineering and implementation process in this study demonstrate a significant paradigm shift from manual sustainability reporting (document-based processing) to an integrated intelligent Decision Support System (DSS). In conventional systems, high harvest quantities often lead management into productivity bias, where high yields are viewed as an outstanding indicator of plantation performance. Because environmental reporting (Ganoderma spread) is separated from production reports, operational anomalies are frequently detected too late [6]. Through the integration of Large Language Models, this information system successfully breaks data silos by simultaneously synthesizing two spectra of variables: plant productivity versus environmental threats.

The AI's success in assigning a "Critical" status to Blok 1 is a key finding of this study. If evaluated solely through traditional static computational algorithms, production of 11,588 kg on 8.32 Ha would be categorized as "Very Good" performance. However, the palm oil auditor instructions injected through prompt engineering successfully stimulated inferential reasoning capabilities in the Llama 3 and Gemini models [23]. The AI was able to diagnose and understand the natural biological phenomenon of oil palm trees, where trees severely infected by the pathogen Ganoderma boninense often enter a stress-induced flowering phase, producing unusually large amounts of fruit just before mass tree death. This biological insight was then translated by the LLM into a causal narrative: "The current high production is a critical phase before a sharp decline due to mass tree death". This confirms that the developed system, when combined with domain-specific constraints, shows functional feasibility as a prototype for early-stage prescriptive analytics in agricultural management [8]. It serves as a risk-control mechanism to mitigate machine hallucination and provides a structured alternative to conventional report assessment methods.

In addition to narrative synthesis capabilities, the methodological decision not to fully relinquish validity control to the machine proved highly rational. The implementation of activity diagrams that include a human-in-the-loop stage (approval, rejection, or revision by the Manager) shows that the system was designed to support, rather than completely replace, agronomic expert judgment. The cognitive burden on humans to sort through raw data from hundreds of blocks has been drastically reduced, leaving managers with more time to focus on verification and the formulation of executive actions. Recent empirical studies in high-stakes domains similarly confirm that human-AI collaboration significantly reduces workload and improves inter-expert reporting consistency, provided that human oversight remains active to mitigate the rare but severe risks of AI misdiagnosis [26]. From a software engineering perspective, black-box testing also showed that the dual-provider failover infrastructure successfully ensured reporting system reliability and reduced vulnerability to third-party downtime [19].

The strategic implications of this information system architecture center on strengthening management-level transparency at PT Surya Mata Ie. Based on general operational baseline observations, administrative processes that traditionally required prolonged manual data entry, cross-referencing, and report drafting to summarize hundreds of daily operational logs have now been streamlined into near real-time automated execution. The autonomously generated narrative insights function as a crucial early warning system for environmental risk mitigation. The application of advanced Natural Language Processing (NLP) fundamentally transforms unstructured corporate texts into standardized ESG alternative data, enhancing regulatory compliance and predictive monitoring [27]. For example, with the AI indicating "mass tree death," management can promptly adjust future revenue projections for the affected block, allocate emergency fungicide treatment programs, or design early replanting strategies. On a broader industry level, the consistency of reporting that adheres to specific environmental, social, and governance (ESG) categorizations makes it easier for companies to comply with sustainability standards. The system's output in standard PDF and JSON formats demonstrates the functional feasibility of supporting operational transparency for palm oil industry stakeholders, including the board of directors and external auditors, in aligning with strict principles such as the Indonesian Sustainable Palm Oil (ISPO) and Roundtable on Sustainable Palm Oil (RSPO) [1], [3].

This study offers a pioneering contribution at the intersection of information systems, software engineering, and precision agriculture informatics. It is among the first applied research to design a Laravel MVC-based software architecture that is seamlessly integrated with a Dual-LLM Engine API Failover system (Groq Llama 3 and Gemini 1.5 Flash) specifically for the operational domain of Indonesia's palm oil industry. The research concretely adds to the literature by demonstrating the Strict Weighting Rule validation method. The 20% Ganoderma infection threshold was selected based on established epidemiological boundaries where the pathogen triggers an exponential risk of mass tree mortality, necessitating immediate agronomic intervention [18]. Conversely, the maximum ESG Risk Score limit of 50.0 is explicitly defined as a software proof-of-concept constraint to validate the backend's deterministic intervention capabilities, rather than representing an official global ESG compliance standard. The injection of these targeted disqualification rules into the JSON object parsing mechanism successfully transforms the AI from a mere natural language generator into a structured computational tool with mitigated hallucination risks, aligning with recent AI governance frameworks for reliable agricultural applications [28].

Although this ESG reporting information system has been successfully implemented according to its functional design, several fundamental limitations remain. First, the system architecture still fully depends on external third-party computing server interfaces (Groq APIs and Gemini). This dependence creates risks of cost escalation driven by high token usage charges if the system is applied across large multinational company networks. The need for stable high-speed internet connections may also pose significant challenges when replicating the system in remote plantation areas. Second, an academic limitation of this study is the absence of extensive real end-user usability testing, such as the System Usability Scale (SUS) questionnaire. As a result, the cognitive reliability and acceptability of the interface among field workers and foremen have not been measured. Furthermore, the linguistic accuracy of the AI-generated agronomic language has not been statistically evaluated through inter-rater agreement measures, such as Fleiss' Kappa, from a panel of human experts [20]. Feasibility testing remains focused purely on functional black-box testing at the software engineering level.

Drawing from the implementation results and the technical limitations discussed, several strategic recommendations can be made for further development. For palm oil plantation companies, it is advisable to invest in local on-premise graphics computing server infrastructure. This aims to support the offline deployment and execution of advanced prompt engineering for open-source Large Language Models in an offline environment. Such initiatives would reduce dependence on paid cloud API token subscriptions while safeguarding the confidentiality of sensitive operational data. For future research, it is crucial to expand the testing dimension by conducting empirical expert agreement tests (Fleiss' Kappa Agreement Test) [20]. This process would involve a committee of certified agronomy experts to assess the validity and lexical precision of AI-synthesized recommendation texts. Additionally, the integration of Internet of Things (IoT) devices, such as soil moisture sensors or drone imagery fed directly into the prompt engine, is expected to create even more measurable, transparent, and autonomous ESG reporting governance.

CONCLUSION

This study successfully designed and implemented an automated, Large Language Model (LLM)-based Environmental, Social, and Governance (ESG) reporting information system for the palm oil industry, effectively addressing the chronic inefficiencies and human-error risks associated with manual data processing. By integrating a Dual-LLM failover architecture (Groq Llama 3 as primary and Gemini 1.5 Flash as backup) within a Laravel MVC framework, the system demonstrates high availability and seamlessly transforms multimodal operational data, such as harvest yields and Ganoderma infestation rates, into structured, prescriptive sustainability narratives. Demonstrating functional feasibility under selected case conditions, the programmatic integration of a Strict Weighting Rule successfully functioned as a risk-control mechanism to mitigate hallucination-driven score inflation during prototype testing, ensuring that ESG Risk Scores are capped accurately when critical biological anomalies occur. Ultimately, this system significantly enhances real-time managerial decision-making and boosts corporate transparency in complying with global sustainability standards. For future development, it is recommended to incorporate empirical usability testing with end-users, conduct expert agronomic validation using Fleiss' Kappa statistics, and explore large-scale implementation across broader plantation networks to further optimize the system's prescriptive accuracy.

ACKNOWLEDGMENT

The authors would like to express their deepest gratitude to Institut Teknologi Sawit Indonesia for providing the academic and institutional support necessary to conduct this research. Special appreciation is also extended to the management and field staff of PT Surya Mata Ie for their cooperation in providing access to essential operational data and facilitating the case study environment.

AUTHOR CONTRIBUTION STATEMENT

NA conceived the research idea, designed the system architecture, developed the software, conducted the system testing, and wrote the initial draft of the manuscript. RS supervised the research project, provided methodological guidance, validated the research design, and critically reviewed the final manuscript.

AI DISCLOSURE STATEMENT

The authors used Google Gemini and ChatGPT during the preparation of this work for translation assistance, text summarization, and English proofreading to ensure academic writing standards. After using these tools, the authors thoroughly reviewed and edited the content as needed and take full responsibility for the final publication. Additionally, the core reporting system developed and evaluated in this research programmatically utilizes the Groq and Gemini Application Programming Interfaces (APIs) as part of its LLM-based methodology, which has been explicitly described in the manuscript.

REFERENCES

[1] T. Lim, “Environmental, social, and governance (ESG) and artificial intelligence in finance: State-of-the-art and research takeaways,” Artif. Intell. Rev., vol. 57, no. 4, p. 76, Feb. 2024, doi: 10.1007/s10462-024-10708-3.

[2] T. Hoang et al., "ESG Reporting Lifecycle Management with Large Language Models and AI Agents," arXiv preprint, Mar. 2026, doi: 10.48550/arXiv.2603.10646.

[3] N. Haryanti, A. Marsono, and M. A. Sona, “Strategi Implementasi Pengembangan Perkebunan Kelapa Sawit Di Era Industri 4.0” [Implementation Strategies for Palm Oil Plantation Development in the Industry 4.0 Era], J. Din. Ekon. Syariah, vol. 8, no. 1, pp. 76–87, Feb. 2021, doi: 10.53429/jdes.v8i1.146.

[4] A. Cravero, A. Bustamante, M. Negrier, and P. Galeas, “Agricultural Big Data Architectures in the Context of Climate Change: A Systematic Literature Review,” Sustainability, vol. 14, no. 13, p. 7855, Jun. 2022, doi: 10.3390/su14137855.

[5] I. H. Fadhlurrahman, Tanto, and M. H. Saputra, “Perancangan Sistem Informasi Monitoring Pelaporan Produksi Sawit Pada Koperasi Wahana Agung” [Design of Palm Oil Production Reporting Monitoring Information System at Koperasi Wahana Agung], J. Elektron. List. Dan Teknol. Inf. Terap., vol. 7, no. 1, pp. 22–32, Jun. 2025, doi: 10.37338/alti.v7i1.452.

[6] F. Erlina, S. Samsudin, and R. A. Putri, “Sistem informasi monitoring pengangkutan dan pelaporan hasil panen kelapa sawit pada PTPN IV unit berangir” [Information System for Monitoring Transportation and Reporting of Palm Oil Harvests at PTPN IV Unit Berangir], Jurnal Sistem Informasi, vol. 12, no. 1, pp. 33–42, Dec. 2024.

[7] H. Naveed et al., “A Comprehensive Overview of Large Language Models,” ACM Trans. Intell. Syst. Technol., vol. 16, no. 5, pp. 1–72, Oct. 2025, doi: 10.1145/3744746.

[8] S. Shahriar, M. G. Corradini, S. Sharif, M. Moussa, and R. Dara, “The role of generative artificial intelligence in digital agri-food,” J. Agric. Food Res., vol. 20, p. 101787, Apr. 2025, doi: 10.1016/j.jafr.2025.101787.

[9] D. O. Sihombing, “Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis” [Implementation of Natural Language Processing (NLP) and Cosine Similarity Algorithm in Automated Essay Exam Assessment], J. Sist. Komput. Dan Inform. JSON, vol. 4, no. 2, p. 396, Dec. 2022, doi: 10.30865/json.v4i2.5374.

[10] J. Bailey et al., “Leveraging Generative AI for Data Analysis in Farm Management,” Appl. Eng. Agric., vol. 41, no. 5, pp. 505–519, 2025, doi: 10.13031/aea.16429.

[11] J. Tummers, A. Kassahun, and B. Tekinerdogan, “Reference architecture design for farm management information systems: a multi-case study approach,” Precis. Agric., vol. 22, no. 1, pp. 22–50, Feb. 2021, doi: 10.1007/s11119-020-09728-0.

[12] M. F. Al-Adhim and G. S. Dewi, “Sistem Monitoring IoT Smart Farm Berbasis Web dengan Integrasi Template Dashboard Bootstrap dan Laravel 10” [Web-Based IoT Smart Farm Monitoring System with Bootstrap Dashboard Template Integration and Laravel 10], COMSERVA J. Penelit. Dan Pengabdi. Masy., vol. 4, no. 7, pp. 1973–1981, Nov. 2024, doi: 10.59141/comserva.v4i7.2595.

[13] A. A. Bimantara and R. D. Gunawan, “Sistem Monitoring Produksi Menggunakan Laravel Dan Cork-Bootstrap” [Production Monitoring System Using Laravel and Cork-Bootstrap], J. Inf. Technol. Softw. Eng. Comput. Sci. ITSECS, vol. 1, no. 4, pp. 143–153, Oct. 2023, doi: 10.58602/itsecs.v1i4.73.

[14] G. Surono, Y. Suhanda, and F. Alfiah, “Penerapan MVC Arsitektur Pada Sistem Informasi Monitoring Pada Divisi Produksi Menggunakan Laravel Framework” [Application of MVC

Architecture in Production Division Monitoring Information System Using Laravel Framework], J. Sensi, vol. 8, no. 2, pp. 180–189, Aug. 2022, doi: 10.33050/sensi.v8i2.2423.

[15] Z. K. Daulay, S. Suendri, and H. Santoso, "Penerapan Sistem Informasi Monitoring Hasil Panen Dan Produksi Di PTPN III Kebun Sei Baruhur" [Implementation of Harvest and Production Monitoring Information System at PTPN III Kebun Sei Baruhur], Journal of Science and Social Research, vol. 7, no. 3, pp. 980–986, Aug. 2024, doi: 10.54314/jssr.v7i3.2096.

[16] D. A. Rina Sari and M. D. Irawan, "Implementation of Key Performance Indicators in the Palm Oil Harvest Monitoring Information System," Green Intell. Syst. Appl., vol. 5, no. 2, pp. 150–163, Aug. 2025, doi: 10.53623/gisa.v5i2.782.

[17] S. Shool, S. Adimi, R. Saboori Amleshi, E. Bitaraf, R. Golpira, and M. Tara, "A systematic review of large language model (LLM) evaluations in clinical medicine," BMC Med. Inform. Decis. Mak., vol. 25, no. 1, p. 117, Mar. 2025, doi: 10.1186/s12911-025-02954-4.

[18] Rizky Delianngi, Ratu Mutiara Siregar, Nurliana, Muhammad Akbar Syahbana Pane, Phaklen Ehkan, and Andi Prayogi, "Performance Evaluation of YOLOv9, YOLOv10, and YOLOv11 for Real-Time Early Detection of Ganoderma Boninense in Oil Palm," J. RESTI Rekayasa Sist. Dan Teknol. Inf., vol. 10, no. 2, pp. 429–440, Apr. 2026, doi: 10.29207/resti.v10i2.7479.

[19] A. Herdiansah, R. I. Borman, and S. Maylinda, "Sistem Informasi Monitoring dan Reporting Quality Control Proses Laminating Berbasis Web Framework Laravel" [Web-Based Monitoring and Quality Control Reporting Information System for Laminating Process Under Laravel Framework], J. Tekno Kompak, vol. 15, no. 2, p. 13, Aug. 2021, doi: 10.33365/jtk.v15i2.1091.

[20] Y. Wang and I. J. Yusof, "Expert consensus and reliability validation of the portfolio assessment guideline for Chinese practical writing: An empirical study based on fleiss' kappa," BenchCouncil Trans. Benchmarks Stand. Eval., vol. 5, no. 4, p. 100248, Dec. 2025, doi: 10.1016/j.tbench.2025.100248.

[21] L. Ouyang et al., "Training language models to follow instructions with human feedback," 2022, arXiv. doi: 10.48550/ARXIV.2203.02155.

[22] J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," 2022, arXiv. doi: 10.48550/ARXIV.2201.11903.

[23] K. Singhal et al., "Large language models encode clinical knowledge," Nature, vol. 620, no. 7972, pp. 172–180, Aug. 2023, doi: 10.1038/s41586-023-06291-2.

[24] M. Li, A. Basit, A. Javed, X. Song, and Y. Ge, "Generative artificial intelligence and ChatGPT in agriculture supply chain management: a systematic literature review and future research agenda," Br. Food J., vol. 128, no. 13, pp. 121–141, Nov. 2025, doi: 10.1108/BFJ-05-2025-0720.

[25] J. Gallifant et al., "The TRIPOD-LLM reporting guideline for studies using large language models," Nat. Med., vol. 31, no. 1, pp. 60–69, Jan. 2025, doi: 10.1038/s41591-024-03425-5.

[26] S. Chen et al., "The effect of using a large language model to respond to patient messages," Lancet Digit. Health, vol. 6, no. 6, pp. e379–e381, Jun. 2024, doi: 10.1016/S2589-7500(24)00060-8.

[27] R. Dara, S. M. Hazrati Fard, and J. Kaur, "Recommendations for ethical and responsible use of artificial intelligence in digital agriculture," Front. Artif. Intell., vol. 5, p. 884192, Jul. 2022, doi: 10.3389/frai.2022.884192.