Implementation Of The LLM Model For Inventory Management In `UD Murni Bitung Company

Pattinaja Andreas Timothy[1], Debby E. Sondakh[2]
[1,2]Informatics, Universitas Klabat
[1]s22210352@student.unklab.ac.id, [2]debby.sondakh@unklab.ac.id

Abstract *in the digital era, efficiency in inventory management is a key factor for the success of commodity distribution companies. ud murni bitung, a salt distribution company, still applies a manual stock recording process, which is prone to errors and inefficiencies. this research proposes the development of a mobile-based inventory management system integrated with a large language model (llm) using retrieval-augmented generation (rag). the system allows real-time stock monitoring, automated reporting, and interactive chatbot features to assist administrators in making strategic decisions. the prototyping method is applied to iteratively develop the system based on user feedback. the results indicate that the integration of llm with rag improves operational accuracy, reduces recording errors, and optimizes decision-making processes in inventory management.*

Keywords: *Large Language Model (LLM), Inventory Management, Gemini, Mobile Application*

I. INTRODUCTION

In the digital era, efficient supply management is key to successful commodity distribution. One critical factor affecting operational efficiency is warehouse stock management, which ensures product availability, balances market demand, and stabilizes prices [1]. Warehouses serve as central storage hubs for various products, enabling smooth distribution processes.

This also applies to salt distribution companies, where stock recording accuracy greatly impacts operations. For example, UD Murni Bitung—a medium-sized salt distributor in North Sulawesi established on May 26, 2016—still records stock manually using notebooks. This system is prone to errors, delays, and difficulties in making strategic decisions.

According to [2], manual inventory recording carries a high risk of errors, misplaced data, and lost information, complicating future data retrieval. Therefore, conventional recording should be replaced with more user-friendly, structured systems that allow easier, faster, and more efficient processing, updating, searching, and deletion of inventory data. As shown in [3], companies like PT Mitra Tekno that still use manual systems face challenges in

tracking stock and processing data. Implementing an inventory system at UD Murni is expected to improve operational performance and data accuracy. Furthermore, [4] highlights that such systems can optimize inventory management and goods distribution, thereby increasing productivity and operational precision.

According to [5], application development is increasingly moving toward dynamic and interactive designs, applied across fields such as information, education, telecommunications, and trade. This helps
companies adapt to technological advancements and stay relevant in the digital era. However, [6] notes that Pharmacy Neofarma Sanggau still relies on manual methods for daily transactions and customer service, which are inefficient and prone to errors in data recording, sales reports, and stock management. The absence of online transaction support further limits customer access.

In the digital era, computer-based technology plays a crucial role in ensuring accuracy and efficiency. Digital inventory systems not only provide consistent performance but also reduce workload significantly. UD Murni seeks to enhance its system by integrating intelligent features powered by Large Language Models (LLMs) such as GPT, Claude, or Gemini, which are capable of processing natural language, generating summaries, and providing recommendations automatically. As shown in [7], real-time inventory systems enhanced with LLMs can greatly improve both user experience and operational efficiency.

This study adopts the Retrieval-Augmented Generation (RAG) approach, allowing the chatbot to respond to admin queries based on actual sales data. Through RAG, internal sales records are retrieved, processed with an LLM, and transformed into relevant, accurate, and comprehensible outputs—such as monthly sales recaps or predictive stock recommendations [8].

The significance of this research lies in demonstrating the application of LLMs within inventory management, enabling conversational interaction that supports admins in monitoring and managing stock with data-driven insights. Beyond improving operational efficiency at UD Murni, this study also contributes academically by showcasing the integration of LLM technology into inventory management within the salt distribution sector. The objective is to develop an LLM-based inventory management application integrated with RAG that can record stock, monitor sales in real-time, and provide predictive recommendations on stock requirements to support decision-making at UD Murni..

II.   LITERATURE REVIEW

o *Artificial Intelligence*

Artificial Intelligence (AI) allows computers to mimic human intelligence in processing data, learning patterns, and making decisions. Across different industries, AI has contributed to improving efficiency, minimizing errors, and speeding up workflows. The emergence of advanced models, particularly Large Language Models (LLMs) such as GPT, Gemini, and Claude, has enhanced AI's ability to comprehend and generate natural language, leading to more intuitive interactions between users and systems. Nonetheless, because AI-driven decision-making often relies on complex mechanisms, maintaining transparency and explainability is crucial to foster user trust and acceptance [10]. Leveraging these strengths, LLM-based AI holds significant potential to make inventory management systems more responsive, efficient, and better aligned with operational demands [11].

o *Natural Language Processing*

Natural Language Processing (NLP) focuses on enabling computers to analyze and understand human language automatically. NLP is widely applied in sectors like finance, business, and customer service due to its ability to process unstructured data such as reports or conversations [12]. In inventory management systems, NLP enables more communicative interfaces, such as chatbots that automatically respond to queries about stock, transaction history, or product information [13]. Ethical concerns such as bias and data privacy remain important considerations [14].

o *Natural Language Understanding*

Natural Language Understanding (NLU) is one of the branch from NLP that allows system computer understand meaning from the text input provided by the user. NLU works For analyze intent, context, and structure Language natural so that the system can extract relevant information in a way automatic [15]. This enables the application to classify inputs automatically and respond with relevant data, such as stock availability or restock recommendations.[16].

o *Natural Language Generation*

Natural Language Generation (NLG) is the process by which computers produce human-like text outputs [17]. While NLU interprets input, NLG constructs coherent and meaningful answers. In inventory management, NLG enables the system to provide natural responses to user queries about stock levels, transaction histories, or sales trends. This creates smoother and more intuitive interactions between the system and users [18].

o *Gemini*

Gemini, previously known as Bard [19], is a multimodal artificial intelligence model developed by Google. It is capable of understanding, integrating, and processing various types of data simultaneously, such as text, programming code, audio, images, and even video. The

first Gemini, Gemini 1.0, was released in three variants: Ultra, Pro, and Nano. In version 1.5, a new model called Gemini Flash was introduced. Gemini Flash is a general-purpose model, free for everyone to use, and optimized for speed [20]. It supports 1,048,576 input tokens and has an output limit of 8,192 tokens, where each token is roughly equivalent to 4 characters, and 100 tokens correspond to about 60–80 English words [21].Gemini Pro (also called Gemini Advanced) is an advanced model, Gemini Nano is the most efficient and works offline, and Gemini Ultra is the largest, scoring 90% on the MMLU and 59.4% on the MMMU benchmarks, which test expert-level performance across disciplines [20], [22] .

o *Large Language Models (LLM)*

Large Language Models (LLMs) are a major breakthrough in artificial intelligence, especially in Natural Language Processing (NLP). An LLM is a large-scale AI model capable of understanding, generating, and responding to text like a human. [23]. In inventory management, LLM can be used to support text-based communication interfaces, such as intelligent chatbots for warehouse staff or admins to query stock, view transaction history, or request reports. Using Natural Language Understanding (NLU) and Natural Language Generation (NLG)[24].

Implementation technology This make system inventory more interactive, flexible, and user-friendly, especially for previously non - technical users Possible difficulty adapt with interface system conventional. Example implementation of LLM which has been used in a way wide including ChatGPT (OpenAI), Gemini (Google), and Claude (Anthropic), which have ability answer questions, summarizing report , even give recommendation action based on available data [25].

o *Chatbot*

A chatbot is a computer program designed to simulate conversations with humans using natural language, either through text or voice, automatically. It applies Natural Language Processing (NLP) algorithms to understand user messages, process information, and provide relevant responses quickly [26]. In inventory management systems, the use of chatbots offers significant advantages because they can deliver real-time answers, reduce the manual workload of administrators when searching for data, and support faster data-based decision-making [27]. With the Retrieval-Augmented Generation (RAG) technique, chatbots can extract relevant information directly from the database, ensuring that responses are accurate and contextually relevant.

o *Inventory Management Concept*

Inventory management is the process of monitoring and controlling stock to meet company needs while avoiding shortages or excess [28]. Modern systems have replaced manual recording with digital applications that provide real-time stock data, including transaction

history and demand forecasts [29]. These systems reduce errors, simplify reporting, and enable faster decision-making, resulting in improved efficiency and alignment with market demand [30].

o *The Role of Mobile Technology (Android)*

Mobile technology enables real-time, broad data access in modern information systems [31]. Android, dominating over 70% of the smartphone market, facilitates app development and supports digital transformation [31]. Its mobility allows users to scan and update inventory anytime, enhancing speed, efficiency, and responsiveness [32].Industry digitalization and the growth of smart devices demand high mobility in inventory management, including faster stock processing across the supply chain [33]. Integrating Android into inventory systems improves efficiency and ensures real-time information availability for quicker decision-making.
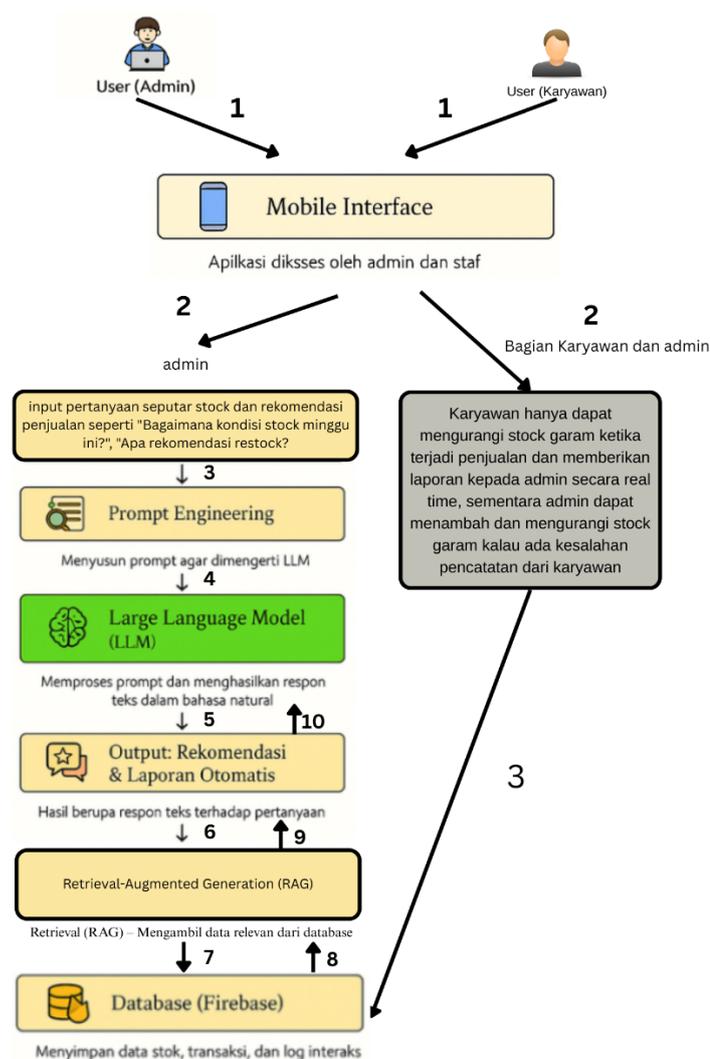

Figure 2.1

In Figure 2.1 it is shown framework conceptual where channel Work from Inventory Management Application which is held by 2 users , namely Admin and Employee . structure framework conceptual explained under This :

1. Application Development

Development process done using Visual Studio Code with the React Native framework to build appearance mobile interface . Android Studio is used as tool help emulator and debugging during the development process . Application designed For can accessed by two types Users : Admin and Employees .

2. Deploy to Mobile Devices

After the development process finished , the application is built and deployed to Android devices . Applications This Then used by Admin and Employees For operate activity inventory mobile .

3. Employee Use of Applications
- Employee log in to in application .
- When it happens sales , employees reduce stock in a way direct through application .
- Employees can also send report condition warehouse or stock to Admin in real-time.
- All activity This will stored in Firebase as the main database .

4. Use by Admin
- Admin login to application through the same device (mobile).
- Admin can :
    o Showing all stock data in real-time.
    o Add or delete items in the inventory list .
    o See chart or report stock .
    o Managing employee data, including right access of each user.

5. Data Synchronization with Firebase
- Firebase plays a role as a real-time database that stores all related data stock, transactions, and reports .
- Every data changes (either from Admin and Employees) directly saved and synchronized between device users in a way automatic .

6. Admin Interaction with LLM Features
- Admin can using a Large Language Model (LLM) based chatbot to access information with Language experience .
- Example interaction : " Show report stock Sunday this " or "What is the restock recommendation based on May data ?"
- Request the will processed through engineering prompts and LLM will produce response in form text experience .

7. Retrieval-Augmented Generation (RAG) Process
- For answer question from admin, system will do search data into the database first before (using RAG technique).
- Relevant data taken from Firebase and compiled become answer automatically by LLM.

8. Recommendation & Report Output Automatic

- LLM provides answer based on the data taken via RAG.
- The results are in the form of recommendations , analysis simple , or summary stock will displayed in form natural conversation in the app .

III. MATERIALS AND METHODS

o *3.1 Research Design*

This research uses a structured design to ensure each stage of application development aligns with the study's objectives. The process consists of five main stages, from problem identification to system evaluation, with each stage interconnected to deliver solutions suited to the company's inventory management needs.
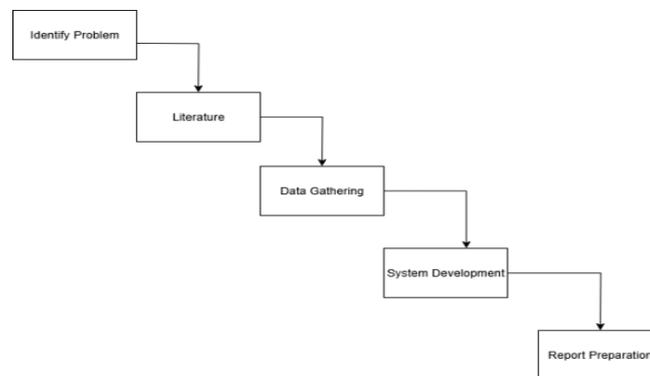

Figure 3.1 Research Design

o *3.2 System Development Methods*

This study uses the Prototyping method, creating an initial application model based on UD Murni's functional needs. The prototype is tested by admins and employees, and feedback guides gradual improvements.
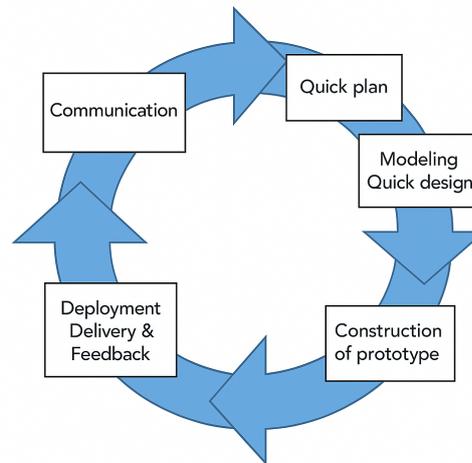
Figure 3 2 Development Method

The stages that exist in The prototype method is :

1. Communication

The first stage in the prototyping method is identifying and formulating user needs. In this study, interviews with the owner and admin of UD Murni revealed issues in manual stock recording, such as errors, delays in obtaining stock information, and the absence of a centralized reporting system. From these interviews, the researcher identified key needs, including easier recording of goods in/out, real-time stock reports, and an LLM-based chatbot to assist in stock data searches. This information forms the basis for the next development stages.

2. Quick Planning

After formulating system requirements, the researcher conducted initial planning for the application, identifying core components such as stock in/out recording, inventory reports, and LLM-based chatbot integration. The system uses mobile technology with Android Studio for testing, React Native for the user interface, and Firebase as a real-time database. The LLM is integrated into the chatbot to answer user queries, such as daily stock levels or weekly stock outflow.

3. Modeling

At the stage Modeling , researcher start compile design system in form visual representation for describe How system will functioning in a way conceptual and technical .

Modeling This important For bridge communication between researchers and users (UD Murni admin) so that understanding to system become more clear before the construction process application started .

4. Construction

This stage develops the initial prototype with core features such as stock input/output, stock reports, and an LLM-based chatbot. Firebase handles real-time storage, React Native builds the interface, and the functional prototype gathers user feedback before finalization.

5. Deployment & Feedback

The initial prototype was tested directly by the UD Murni admin to evaluate its functionality and usability. Users tried various features and provided feedback on the interface, chatbot response speed, and ease of stock recording. The feedback was collected for system improvements.

- ***FUTURE WORK***

At the current stage, the mobile-based inventory management application is already functional in terms of stock management and transaction history, which are currently supported by a dummy database. However, several components of the system have not yet been developed, namely:

1. **Firebase integration for transaction history** – At present, the transaction history still uses a dummy database and has not been connected to Firebase, resulting in the absence of real-time synchronization.

2. **AI chatbot integration** – Although the application design already accommodates the use of an AI-powered chatbot, the chatbot has not yet been integrated into the system.

3. **Connection between Firebase and the AI chatbot** – The chatbot has not been connected with the Firebase-based transaction history, preventing it from accessing real data to provide sales analysis and insights.

Future development will focus on completing these three components. Once Firebase integration and the AI chatbot are fully implemented and connected, the system will be able to provide real-time synchronization, AI-driven analysis, and more accurate decision-making support. These improvements are expected to enhance operational efficiency, minimize errors in stock recording, and allow the company to respond more effectively to market demands. Furthermore, this foundation will enable scalability for additional features such as safety stock analysis and demand forecasting.

REFERENCE

[1] D. Nur Azizah, " Emitter : Journal of Electrical Engineering Development Trading Company Goods Inventory System Website Based (Case Study: CV. Agung Nugraha)".

[2] NMZ Hashim and NAMM Arifin, "Laboratory Inventory System," 2013.

[3] DN Setianingsih , " Development System Inventory Web -Based at PT Mitra Tekno ," Thamrin University Repository, 2024.

[4] Subkhan Fadilah, Muhtajuddin Danny, Nurhadi Surojudin , " Optimization Management Inventory through System Information Web -Based at PT Herso ," Jurnal Technology and Systems Information Explore, Volume 14 No 2, pp 99-107, 2024

[5] Hidayat R, Asmadi I, " Design Get up Application Study Teaching at an English Tutoring Center Web- based ", Research and E- Journal Management Informatics Computers , Volume 5 No 2, pp 271-283, April 2021

[6] Puteri Noviandhiny , " Design Get up Application Sales and Purchases Web -Based in Pharmacy Neofarma Sanggau ", JUSTIN (Journal Systems and Technology Information), Volume 6 No 3 , pp 133-138, April 2018

[7] O. Patil, H. Kulkarni, R. Pottavathini , I. Dhole , and K. Salgaonkar, "Real Time Inventory Management System powered by Generative User Interface," International Journal of Scientific Research in Engineering and Management (IJSREM), vol. 8, no. 4, pp. 1–6, Apr. 2024. [Online]. Available: https://www.researchgate.net/publication/380268678

[8] N. Rachmat and DP Kesuma, " Implementation of Gemini Large Language Models in Software Development Android- Based Chatbot Application ," Journal Knowledge Computer (JUIK), vol. 4, no. 1, pp. 40–52, Feb. 2024. [Online].

[9] Annas , . . AN, Wijayanto , G. . , Cahyono, D. . , Safar, M. . , & Ilham, I. (2024). Technical Training on the Use of Artificial Intelligence (AI) Chat Gpt and Bard AI Applications as Tools for Students in Working on Task Lectures . Journal of Human and Education (JAHE), 4(1), 332–340. https://doi.org/10.31004/jh.v4i1.617 [11] PT Garam (Persero), " Report PT Garam Annual Report 2021. [Online]. Available: https://www.ptgaram.com/ . [Accessed: 5-Jun-2025].

[10] S. Heo and S. Na, "Ready for departure: Factors to adopt large language model (LLM)-based artificial intelligence (AI) technology in the architecture, engineering, and construction (AEC) industry," Results in Engineering, vol. 25, p. 101012, 2025. [Online]. Available: https://www.researchgate.net/publication/391368354

[11] R. Aghaei, AA Kiaei, M. Boush, J. Vahidi, Z. Barzegar, and M. Rofooseh , "The potential of Large Language Models in supply chain management: Advancing decision-making, efficiency, and innovation," arXiv preprint arXiv:2501.15411, Jan. 2025. [Online]. Available: https://arxiv.org/abs/2501.15411

[12] M. Bolanle and E. Frank, "AI Chatbots for Demand Forecasting and Inventory Management in Sustainable Supply Chains," ResearchGate, Jul. 2023. [Online]. Available: https://www.researchgate.net/publication/383603293_AI_Chatbots_for_Demand_Forecasting_and_Inventory_Management_in_Sustainable_Supply_Chains

[13] A. Zaremba and E. Demir, "ChatGPT: Unlocking the Future of NLP in Finance," Modern Finance, vol. 1, no. 1, pp. 93–98, Nov. 2023. [Online]. Available: https://doi.org/10.61351/mf.v1i1.43

[14] Pacific Data Integrators, "Transforming Retail Operations with LLM Technology: Chatbots, Inventory Management, and Security," Pacific Data Integrators, Nov. 8, 2024. [Online]. Available: https://www.pacificdataintegrators.com/blogs/transforming-retail-operations-with-llm-technology

[15] J. Parker, "Natural Language Understanding for Conversational AI," International Journal of Artificial Intelligence and Machine Learning in Engineering, vol. 20, no. 08, pp. 715–724, 2020. [Online].Available: https://www.researchgate.net/publication/391704435_Natural_Language_Understanding_for_Conversational_AI

[16] M. Bolanle, E. Frank, and K. Sheriffdeen , "AI Chatbots for Demand Forecasting and Inventory Management in Sustainable Supply Chains," ResearchGate, July 24, 2023. [Online]. Available : https://www.researchgate.net/publication/383603293

[17] E. Reiter and R. Dale, "Natural Language Generation," in Natural Language Generation, Summarization, & Translation. University of Wisconsin–Madison, 2021. [Online]. Available: https://wisconsin.pressbooks.pub/naturallanguage/chapter/natural-language-generation/

[18] Small Business Inventory Management, "How Natural Language Processing (NLP) is Transforming Inventory Management," Small Business Inventory Management Blog, 2024. [Online]. Available: https://small-business-inventory-management.com/blog/ai-inventory-automation/nlp-inventory-management.html

[19] S. Hsiao , "Google Bard is now Gemini: How to try Ultra 1.0 and new mobile app." [On line]. Available: https://blog.google/products/gemini/bard-gemini-advanced-app/

[20] D. Hassabis and S. Pichai, "Introducing Gemini: Google's most capable AI model yet." [On line]. Available: https://blog.google/technology/ai/google-gemini-ai/#sundar-note

[21] Google, "Gemini Models |Gemini API | Google AI for Developers." Accessed: Nov. 15, 2024. [Online]. Available: https://ai.google.dev/gemini-api/docs/models/gemini

[22] Gemini - Google DeepMind," Google Deepmind . [Online]. Available: https://deepmind.google/technologies/gemini/

] W. X. Zhao, K. Zhou, J. Li, T. Tang , 2023. [Online]. Available: https://arxiv.org/abs/2303.18223 .

[24] D. Narwade , A. Kanhere, S. Mulla, A. Sanap, and A. Patil, "LLM for Retail Business (Optimizing Clothing Sales with AI)," International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), vol. 11, no. 5, pp. 176–179, Sep.–Oct. 2024. [Online]. Available : https://doi.org/10.32628/IJSRSET24115108

[25] Mubashar Raza, Z. Jahangir, MB Riaz, MJ Saeed, and MA Sattar, "Industrial applications of large language models," Scientific Reports, vol. 15, no. 13755, pp. 1–23, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-98483-1

[26] E. Adamopoulou and L. Moussiades , "An Overview of Chatbot Technology," in *IFIP International Conference on Artificial Intelligence Applications and Innovations* , vol. 584, pp. 373–383, 2020, doi : 10.1007/978-3-030-49186-4_31 .

[27] BAT de Freitas and R. de A. Lotufo , "Retail-GPT: Leveraging Retrieval Augmented Generation (RAG) for Building E-commerce Chat Assistants," * arXiv preprint* arXiv:2408.08925, Aug. 2024. [Online]. Available: https://arxiv.org/pdf/2408.08925

[28] EBJ Janson and IN Nurcaya , " Just In Time Application for Efficiency Cost Inventory ," E- Journal Management Unud , vol. 8, no. 3, pp. 1755-1783, 2019.

[29] OA Madamidola , OA Daramola, KG Akintola, and OT Adeboje , "A Review of Existing Inventory Management Systems," International Journal of Research in Engineering and Science (IJRES), vol. 12, no. 9, pp. 40–50, Sept. 2024.

[30] ZN Ahmad, " Design "PT. KFC Goods Inventory Application ," JRAMI (Journal of Research and Application) Student Informatics), vol. 01, no. 01, pp. 142–149, 2020. [Online]. Available: https://jurnal.unindra.ac.id/index.php/jrami/article/view/945

[31] Mobile operating system," Wikipedia, Mar. 2025. [Online]. Available: https://en.wikipedia.org/wiki/Mobile_operating_system

[32] RFgen Software, "Mobile Inventory Management," RFgen Blog, [Online]. Available: https://www.rfgen.com/blog/mobile-inventory-management/.

[33] MT Ahad, "Mobile phone enabled Supply Chain Management in the RMG sector: A conceptual framework," arXiv preprint arXiv:2206.03560, Jun. 2022. [Online]. Available: https://arxiv.org/abs/2206.03560