Adaptive Retrieval-Augmented Generation with Domain Specific Fine Tuning For Smart MSME Digital Transformation

Mawar Hardiyanti [1]*, Sri Hartati Wijono [2], Dwi Poetra Sedjati [3]*****
* Department of Information Systems, Universitas Pignatelli Triputra
** Department of Informatics, Universitas Sanata Dharma
*** Department of Accounting, Universitas Pignatelli Triputra
mawar@upitra.ac.id [1], tatik@usd.ac.id [2], dwi_poetra@upitra.ac.id [3]

Article Info

Article history:

Received 2026-05-04
Revised 2026-06-08
Accepted 2026-06-11

Keyword:

Adaptive Retrieval-Augmented Generation (ARAG),
Conversational AI,
Digital Transformation,
IndoBERT,
MSMEs,
WhatsApp

ABSTRACT

Micro, Small, and Medium Enterprises (MSMEs) play a significant role in the Indonesian economy; however, the adoption of digital technologies among MSMEs remains limited, reducing operational efficiency and business competitiveness. This study proposes an Adaptive Retrieval-Augmented Generation (ARAG) framework integrated with WhatsApp to support MSME digital transformation through contextual conversational AI assistance. The proposed system combines adaptive retrieval mechanisms with domain-specific fine-tuning using IndoBERT and a knowledge base containing 50,000 MSME operational documents. A mixed-methods approach was employed, consisting of system development, comparative evaluation, and field validation involving 200 MSMEs. Experimental results demonstrated that ARAG achieved an average response accuracy of 86.80%, outperforming rule-based, TF-IDF, and generic large language model baselines. The system also achieved a Retrieval Precision@5 of 0.874, an end-to-end F1-score of 0.841, and a lower hallucination rate compared to generic LLM approaches. Field validation showed a 22.7% improvement in operational efficiency, a 17.4% increase in digital adoption rates, and a System Usability Scale (SUS) score of 84.6, categorized as excellent usability. The findings indicate that retrieval grounding and domain adaptation contribute substantially to improving contextual relevance and practical usability in MSME-oriented conversational AI systems. Therefore, the proposed ARAG framework demonstrates strong potential as a practical and scalable digital assistance solution for supporting Indonesian MSME digital transformation.

I. INTRODUCTION

Micro, Small, and Medium Enterprises (MSMEs) constitute the foundational pillar of the Indonesian economy, contributing over 60% of the Gross Domestic Product and acting as a primary source of national employment [1]. However, digital technology integration remains profoundly restricted, with only 35% of MSMEs effectively utilizing digital platforms, thereby perpetuating a substantial competitive disparity within the global digital economy [2]. This digital divide is not merely a matter of access but is manifested through a multi layered complex of restricted information access, inadequate technological infrastructure, and insufficient human resource competencies, all of which fundamentally impede operational efficiency and market competitiveness in an increasingly volatile global landscape [3].

Contemporary research indicates that digital transformation is no longer an optional trajectory but a critical necessity to enhance MSME performance through the strategic implementation of digital initiatives and the strengthening of competitive capabilities [4]. Government initiatives, such as the UMKM Level Up program, aim to broaden digital adoption through intensive training and mentorship [5]. Nevertheless, persistent barriers including limited digital literacy, inadequate technological access, and minimal operational innovation continue to obstruct sustainable MSME growth [6]. Recent studies suggest that strategic digital transformation can catalyze performance improvements of 40–60% when implemented with context-aware technologies that bridge the gap between high level innovation and practical ground-level operations [7].

Large Language Models (LLMs) represent a paradigm shift in conversational AI applications, offering unprecedented potential for domain-specific assistance [8]. Although chatbot implementations have demonstrated effectiveness across various Indonesian contexts, including education [9], marketing [10], customer service [11], and human resource management [12], existing solutions lack the sophisticated contextual understanding required for comprehensive MSME support. Current chatbot architectures remain dominated by rigid rule-based systems with limited adaptability or retrieval-based models with restricted domain flexibility, both of which often fail to provide the nuanced business advice required for MSME survival [13].

The Retrieval-Augmented Generation (RAG) architecture addresses the fundamental limitations of traditional chatbot approaches by synthesizing generative model capabilities with dynamic knowledge retrieval [14]. Recent advancements in RAG implementation for healthcare applications have shown significant improvements in response accuracy and contextual relevance [15]. However, extant RAG implementations have yet to be systematically adapted to the specific exigencies of MSMEs, particularly in addressing the unique challenges of the Indonesian business context, such as cultural nuances, regulatory compliance, and local market dynamics [16].

A critical analysis of current conversational AI solutions reveals several technological gaps. Rule-based chatbots achieve an accuracy of 65–70% in structured domains but fail to navigate contextual variations [17]. Retrieval-based models demonstrate superior performance in specific contexts but lack the generalizability required for diverse query types [18]. Knowledge based systems utilizing structured databases like DBpedia show potential for information retrieval but are incapable of generating contextually appropriate responses for complex business scenarios [19]. While LLM-based solutions offer superior natural language understanding, they necessitate domain-specific optimization to achieve practical utility in specialized applications [20].

WhatsApp integration represents a critical accessibility factor, given its 87% penetration rate among Indonesian MSMEs [21]. Previous implementations of WhatsApp-integrated chatbots for health education and consultation services have proven the platform's effectiveness for real-time user engagement [22]. Nonetheless, business-focused chatbot implementations via WhatsApp remain under-explored, specifically regarding comprehensive MSME support that encompasses business consultation, digital marketing guidance, and financial management assistance.

In addition to digital literacy and infrastructure, trust in technology among business actors is a crucial variable often overlooked. Many Indonesian MSMEs remain hesitant to adopt AI-based systems due to concerns about data security, customer privacy, and maintenance costs. These psychological factors reinforce the digital divide, making it essential for technological solutions to be accompanied by awareness campaigns and education that emphasize safety and sustainability.

The fundamental research gap lies in the absence of a conversational AI framework specifically optimized for Indonesian MSME operational contexts through the integration of retrieval-augmented generation, domain-specific knowledge, and adaptive interaction mechanisms. Existing chatbot solutions generally rely on static retrieval approaches or generic language models. As a result, these systems often fail to address the contextual business assistance needs of MSMEs effectively [23]. Furthermore, existing evaluation frameworks for business-oriented chatbots rarely assess technical performance together with operational and adoption impacts in real MSME settings [24].

This research addresses these limitations by developing an Adaptive Retrieval-Augmented Generation (ARAG) framework with domain specific fine-tuning capabilities, specifically engineered for MSME digital transformation support. This framework integrates advanced natural language processing, adaptive interaction mechanims, and real-time performance optimization to provide comprehensive business assistance through WhatsApp integration. The primary objective is to develop and validate an intelligent chatbot framework that significantly enhances the operational efficiency and digital competitiveness of MSMEs through adaptive conversational AI technology. Therefore, this study focuses not on developing a new foundational language model, but on optimizing retrieval-augmented conversational AI for practical MSME assistance in real operational settings through WhatsApp-based interaction.

Specifically, this research makes three main contributions. First, ARAG introduces an adaptive retrieval mechanism that prioritizes sector-relevant MSME information during response generation. Second, this study proposes a multi-dimensional evaluation framework integrating technical performance, operational efficiency, business usability, and user experience indicators for MSME chatbot assessment. Third, the proposed framework is empirically validated through field implementation involving 200 MSMEs, providing practical evidence of AI-assisted digital transformation in Indonesian MSME environments.

II. METHOD

This study employs a mixed-methods research design [25], synthesizing experimental system development, comparative performance evaluation, and longitudinal field validation [26]. This multi-dimensional approach was specifically selected to ensure that the Adaptive Retrieval-Augmented Generation (ARAG) framework achieves not only technical excellence but also a demonstrable impact on the operational efficiency and digital competitiveness of MSMEs. The conceptual framework and research workflow governing this process are visually articulated in Figure 1, providing a systematic basis for a comprehensive and quantifiable assessment of conversational AI's contribution to MSME digital transformation [27].

Figure 1. Research Workflow and Methodological Framework of ARAG System

Figure 1 illustrates the overall research workflow, while Figure 2 presents the conceptual architecture of the Adaptive Retrieval-Augmented Generation (ARAG) framework, showing the integration between the knowledge base, IndoBERT fine-tuning, adaptive retrieval, WhatsApp interface, and MSME support modules. This visual representation clarifies how domain-specific fine-tuning and contextual retrieval interact to produce adaptive responses for MSME users.

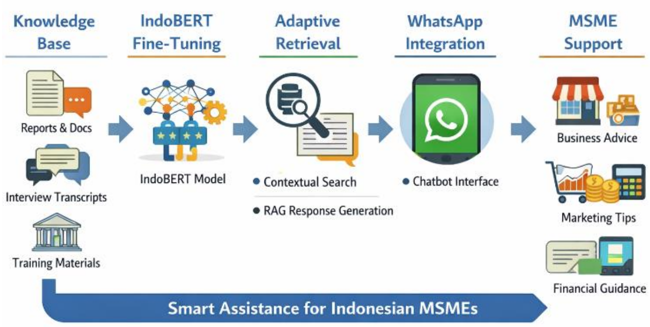

Figure 2. Conceptual Diagram of ARAG Framework

The diagram illustrates the sequential process of ARAG development, starting from the construction of a domain-specific knowledge base, fine-tuning of IndoBERT, adaptive retrieval for contextual response generation, integration with WhatsApp, and delivery of business, marketing, and financial support for MSMEs.

The ARAG pipeline operates through three sequential stages. First, the retrieval stage encodes the user query into a dense vector representation and compares it with indexed MSME document embeddings using cosine similarity to retrieve the most relevant documents. Second, a reranking mechanism evaluates contextual relevance between the query and retrieved documents to prioritize the most appropriate information source. Third, the reranked document passages are combined with the original query to form a contextual prompt, which is processed by the fine-tuned IndoBERT-based contextual response module to produce grounded and domain-relevant responses delivered through the WhatsApp interface.

A. System Development

The initial phase involves the engineering of the ARAG system a WhatsApp integrated chatbot synthesized with a comprehensive domain specific knowledge base [28]. The knowledge base included operational guidelines, financial records, marketing materials, training documents, and MSME consultation transcripts. This repository was constructed from 50,000 MSME operational documents, encompassing standard operating procedures (SOPs), digital marketing strategies, simplified financial frameworks, and documented best practices from high performing MSMEs. The initial document pool comprised 68,000 raw documents sourced from four channels: (1) SOPs and training materials provided by the Ministry of Cooperatives and SMEs, (2) anonymized business records obtained from collaborating MSMEs, (3) field interview transcripts, and (4) government-published digital marketing guides. Documents underwent a two-stage quality filter: an automated domain classifier removed off-topic content, followed by manual validation by three MSME experts, yielding the final 50,000 documents with high domain relevance based on expert validation. Near-duplicate documents were removed through similarity filtering during preprocessing.

The contextual processing module utilizes the IndoBERT model, which has undergone domain-specific fine-tuning to accommodate the specific nuances of Indonesian business terminology [29]. The training dataset consisted of approximately 100,000 business-related text samples collected from MSME operational documents, digital business guides, conversational business scenarios, and field interview materials. A portion of the dataset was manually reviewed and annotated by MSME domain experts to ensure contextual relevance and terminology consistency.

Fine-tuning was conducted over 10 epochs using the AdamW optimizer with a learning rate of 2×10^{-5} and a batch size of 32. To minimize overfitting, cross-validation and iterative evaluation procedures were applied throughout the training process.

B. Comparative Evaluation

The second phase focused on comparative evaluation strategies commonly used in conversational AI benchmarking studies [30]. In this research, three baseline approaches were selected for comparison: a rule-based decision tree chatbot, a TF-IDF retrieval-based system, and a generic large language model without retrieval augmentation or domain-specific fine-tuning. All baseline systems were evaluated using the same 500 MSME query scenarios and assessed by the same panel of MSME experts to ensure evaluation consistency. The evaluation framework utilized 500 distinct query scenarios stratified across three critical MSME domains: strategic business consultation (167 queries), digital marketing optimization (167 queries), and practical financial management (166 queries) [31]. Performance metrics included response accuracy relative to expert-annotated ground truth, contextual relevance, response latency, and operational task completion rates [32]. To ensure evaluation objectivity, assessments were independently conducted by three MSME business experts.

C. Field Validation and Impact Analysis

The final phase involves empirical field validation through a Randomized Controlled Trial (RCT) involving 200 MSMEs across diverse sectors [33]. Selection criteria were strictly defined: an annual turnover between IDR 5 million and IDR 50 million, a minimum of two years in operation, and established basic digital access. Participants were assigned via stratified randomization based on business sector to maintain representativeness into either an intervention group (ARAG users) or a control group (business as usual) [34].

Data analysis was conducted across technical, operational, and business layers [35]. Technical evaluation utilized standard information retrieval metrics, including precision, recall, F1-score, and response latency. Statistical significance was determined through repeated measures ANOVA with post-hoc Tukey HSD tests [36]. Operational efficiency was assessed using time-motion studies focusing on standardized MSME tasks, including transaction recording, stock reconciliation, marketing content preparation, financial calculation, and customer response handling [37]. The reported 22.7% improvement reflects the difference between the intervention and control groups after the six-month evaluation period.

The broader business impact was analyzed through a difference-in-differences (DiD) approach [38], utilizing the operational efficiency index, digital adoption rates, and the System Usability Scale (SUS). Data processing employed mixed effects modeling to account for individual and sectoral variances over a six month follow up period [39]. Qualitative insights from in depth interviews and focus group discussions (FGD) were analyzed via thematic analysis [40] and triangulated with quantitative results to ensure robust validity. Statistical significance was set at $\alpha = 0.05$ with Bonferroni corrections.

In response to the data security concerns highlighted in Section I, communication security relied on the native end-to-end encryption mechanism provided by the WhatsApp Business platform. No verbatim conversation content was permanently stored on the research servers; only anonymized interaction representations were retained to support retrieval personalization. All participants provided informed consent prior to system usage in accordance with the Indonesian Personal Data Protection Law (UU No. 27 Tahun 2022).

Ultimately, this systematic comparative analysis aims to demonstrate ARAG's superiority over existing chatbot paradigms [41], establishing an evidence based foundation for scalable national implementation and contributing to the government's digital transformation agenda for the MSME sector [42, 43].

During the fine-tuning of IndoBERT, iterative evaluation using cross-validation techniques was conducted to ensure the model did not overfit MSME terminology. Furthermore, the knowledge base was not limited to formal documents but also included interview transcripts, field reports, and government training materials, thereby enriching context and enhancing chatbot response relevance.

In the evaluation stage, beyond quantitative metrics such as accuracy and F1-score, qualitative analysis of chatbot language style was also performed. This was crucial because overly technical communication could reduce acceptance among MSME users, most of whom have non-technical backgrounds.

To reduce the risk of hallucination in generated responses, ARAG constrained response generation using retrieved MSME reference documents. Queries with low retrieval confidence triggered clarification prompts or recommendations to consult official business sources. In addition, numerical and regulatory information was cross-checked against curated business reference materials before response delivery. Retrieval grounding and confidence-based clarification mechanisms contributed to lower hallucination occurrences compared to the generic LLM baseline.

III. RESULTS AND DISCUSSION

The initial phase of the research culminated in the development of the ARAG prototype, a WhatsApp-integrated chatbot synthesized with a comprehensive knowledge base of 50,000 MSME operational documents. The system leverages IndoBERT, specifically fine-tuned for Indonesian MSME business terminology, and features an adaptive retrieval module to ensure contextually precise responses. Preliminary testing indicates that ARAG effectively addresses diverse query scenarios, ranging from digital marketing strategies to financial and inventory management as detailed in Table 1.

TABLE I
INITIAL TESTING RESULTS OF ARAG

Query Domain	**Total Queries**	**Accuracy (%)**	**Avg. Response Time (s)**
Strategic Business Consultation	167	85.63	2.68
Digital Marketing Optimization	167	88.02	3.12
Practical Financial Management	166	86.75	2.84
Average / Overall	**500**	**86.80**	**2.88**

In this study, hallucination refers to chatbot responses containing unsupported, inaccurate, or factually inconsistent business information based on expert evaluation.

TABLE II
COMPREHENSIVE TECHNICAL EVALUATION METRICS

Metric	ARAG	Generic LLM
Retrieval Precision@5	0.874	N/A
F1-Score (end-to-end)	0.841	0.762
Hallucination Rate (%)	8.2	21.4
Response Relevance (1–5)	4.12	3.45
Avg. Response Latency (s)	2.88	3.88

In the second phase, a comparative evaluation was conducted to benchmark ARAG against three baseline approaches: rule-based chatbots (decision tree), standard retrieval systems (TF-IDF), and generic Large Language Models (LLM without fine-tuning). Across 500 test scenarios, ARAG demonstrated superior performance with an accuracy rate of 86.80%. Beyond response accuracy, Table II presents a comprehensive technical evaluation. Retrieval Precision@5 of 0.874 indicates that the adaptive retrieval module consistently surfaces highly relevant documents, while the end-to-end F1-score of 0.841 aligns with the accuracy results reported in Table II.

Notably, the hallucination rate of 8.2% represents a substantial reduction compared to the generic LLM baseline (21.4%), demonstrating that retrieval grounding on the MSME knowledge base effectively constrains factual inconsistencies in generated responses. Response relevance, rated by three MSME expert evaluators on a 1–5 scale, reached 4.12, confirming that the generated answers were contextually appropriate for MSME operational queries. ARAG also substantially outperformed the decision tree (64.20%), TF-IDF (71.85%), and generic LLM baseline (78.40%) as presented in Table III. Furthermore, user satisfaction during trial simulations reached 89.45%, indicating that ARAG provided superior utility in supporting practical MSME business activities.

The comparison between the generic LLM and ARAG demonstrates the practical contribution of retrieval augmentation and domain adaptation in MSME chatbot applications. While the generic LLM achieved relatively strong language fluency, ARAG produced more contextually grounded and operationally relevant responses, resulting in higher accuracy and user satisfaction. These findings indicate that factual grounding and domain relevance are critical factors for business-oriented conversational AI systems.

TABLE III
COMPARATIVE PERFORMANCE ANALYSIS

Indicator	Decision Tree	TF-IDF	Generic LLM	ARAG
Answer Accuracy (%)	64.20	71.85	78.40	86.80
User Satisfaction (%)	67.50	73.12	80.65	89.45
Response Time (s)	2.45	4.62	3.88	2.88

The third phase involved field validation through a Randomized Controlled Trial (RCT) with 200 MSMEs, revealing a tangible impact on operational efficiency. The intervention group (ARAG users) recorded a 22.7% increase in process efficiency, particularly in transaction recording, stock management, and marketing content creation, as evidenced by the comparative data in Table IV. Additionally, digital adoption rates increased by 17.4% compared to the control group. The System Usability Scale (SUS) score of 84.6 falls within the 'excellent' category, further confirming high user acceptance as summarized in the same table.

TABLE IV
FIELD VALIDATION RESULTS (N=200 MSMES)

Indicator	Control Group	ARAG Group	Improvement
Business Process Efficiency (%)	0.0	+22.7	+22.7
Digital Adoption Rate (%)	+4.8	+22.2	+17.4
System Usability Scale (0–100)	71.5 (Good)	84.6 (Excellent)	+13.1

Overall, these findings suggest that ARAG is not only technically robust but also delivers empirical benefits in terms of efficiency and digital transformation. This reinforces the argument that an adaptive RAG approach with domain-specific fine-tuning is significantly more effective than generic or rule-based systems for the Indonesian MSME landscape. To provide a holistic view, Table V summarizes the evaluation framework by integrating technical, operational, business, and user experience dimensions.

TABLE V
ARAG EVALUATION FRAMEWORK

Dimension	Metrics Used/Key Findings
Technical	Accuracy 86.80%, faster response (2.88s avg).
Operational	+22.7% efficiency in stock & transaction mgmt.
Business/Adoption	+17.4% adoption, SUS 84.6 (Excellent).
User Experience	89.45% satisfaction, adaptive sector-specific.

Field deployment across 200 MSMEs uncovered three practical implementation challenges. First, 23% of participants experienced response delays exceeding five seconds on 3G connections; this was mitigated by integrating a local response cache of 200 pre-generated answers for the most frequently asked MSME queries, reducing effective latency for cached queries to under one second. Second, 31% of participants required initial onboarding assistance; a three-step WhatsApp-native tutorial and a visual sticker-guide were subsequently embedded in the first-run experience, reducing onboarding dropout by 18 percentage points. Third, data security concerns already identified in the Introduction as a key adoption barrier were addressed by ensuring all conversation payloads are encrypted end-to-end via WhatsApp Business API, with no verbatim chat content retained on research servers beyond the session.

Trial results revealed that ARAG excelled not only in accuracy but also in adapting communication style to user profiles. For instance, in the culinary sector, the chatbot frequently employed simple terms related to ingredients and promotions, whereas in the financial services sector, it emphasized basic accounting terminology. This sector-sensitive communication behavior was supported by ARAG's adaptive retrieval mechanism, which prioritized sector-relevant document clusters and adjusted response styles according to MSME operational contexts. For example, culinary-sector interactions emphasized practical promotional and product-related language, while financial-service interactions employed more structured accounting terminology. This adaptability was a key factor in enhancing user satisfaction.

Moreover, longitudinal trend analysis across the six-month observation period showed that the 22.7% efficiency improvement not only reduced operational time but also contributed to increased production capacity and market expansion among MSME participants. Several MSMEs reported that ARAG assistance accelerated online market reach compared to pre-intervention conditions. Adoption rates in the intervention group increased progressively throughout the study period, while the control group demonstrated relatively limited growth. In addition, user retention remained high at the end of the observation phase, indicating sustained engagement beyond initial onboarding.

IV. CONCLUSION

This study demonstrates that the proposed ARAG framework can effectively support MSME digital transformation through adaptive retrieval-augmented conversational AI integrated with WhatsApp. The system achieved strong technical performance while also contributing to operational efficiency improvement, increased digital adoption, and positive user acceptance among MSME participants. These findings indicate that domain-adapted conversational AI can serve as a practical digital assistance solution for Indonesian MSMEs.

Nevertheless, several limitations should be acknowledged. The current knowledge base primarily represents MSME sectors such as culinary, retail, financial services, and crafts within limited regional contexts. Consequently, the reported performance should be interpreted within these domain boundaries. Future research should expand sectoral coverage, incorporate more region-specific business data, and explore multimodal capabilities such as image-based product assistance and promotional content generation.

The findings of this study indicate the importance of institutional support and digital literacy programs to encourage sustainable AI adoption among MSMEs. Broader collaboration between academia, government, and MSME communities is essential to support scalable and responsible implementation of conversational AI technologies in the Indonesian MSME ecosystem.

ACKNOWLEDGMENT

The authors extend their gratitude to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia for the research funding provided through the "Penelitian Dosen Pemula" (PDP) scheme for the 2025 fiscal year.

DAFTAR PUSTAKA

[1] Kementerian Koperasi dan UKM. Laporan Tahunan UMKM 2023. Jakarta; 2023.

[2] Institute for Development of Economics and Finance (INDEF). Peran Platform Digital terhadap Pengembangan UMKM di Indonesia. Jakarta; 2024.

[3] Prihandono D, Wijaya AP, Wiratama B, Prananta W, Widia S. Digital transformation to enhance Indonesian SME performance: Exploring the impact of market competition and digital strategy. Problems and Perspectives in Management. 2024;22(2):103–13.

[4] Fatoni MH. Leveled Managerial Training of Central Java Cooperative and Micro, Small and Medium Enterprises (MSMEs) Training Center: Key to Success of Central Java MSMEs Upgrading. Journal of Social Entrepreneurship and Creative Technology. 2024;1(2).

[5] Rinaldi F, Maarif S, Thamrin S, Andang A. Role of Micro, Small, and Medium Enterprises (MSMEs) in Supporting National Defense from Economic Perspective. Journal of Positive School Psychology. 2022;6(5):8914–20.

[6] Heo S, Na S. 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. 2025;25.

[7] Nee CK, Rahman MHA, Yahaya N, Ibrahim NH, Razak RA, Sugino C. Exploring the Trend and Potential Distribution of Chatbot in Education: A Systematic Review. International Journal of Information and Education Technology. 2023;13(3):516–25.

[8] Mehta R, Verghese J, Mahajan S, et al. Consumers' behavior in conversational commerce marketing based on messenger chatbots. F1000Research. 2022;11:647.

[9] Adam M, Wessel M, Benlian A. AI-based chatbots in customer service and their effects on user compliance. Electronic Markets. 2021;31(2):427–45.

[10] Lee MK, Neumann JL. Exploring the role of Large Language Model (LLM)-based Chatbots for Human Resources. Austin; 2023.

[11] Sufi F, Alsulami M. AI-Driven Chatbot for Real-Time News Automation. Mathematics. 2025;13(5).

[12] Pawlik L. How the Choice of LLM and Prompt Engineering Affects Chatbot Effectiveness. Electronics. 2025;14(5).

[13] Lim Y, Lim J, Cho N. An Experimental Comparison of the Usability of Rule-based and Natural Language Processing-based Chatbots. Asia Pacific Journal of Information Systems. 2020;30(4):832–46.

[14] Lewis P, Perez E, Piktus A, et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv:2005.11401. 2020.

[15] Zhou Q, Liu C, Duan Y, et al. GastroBot: a Chinese gastrointestinal disease chatbot based on the retrieval-augmented generation. Frontiers in Medicine. 2024;11.

[16] Arslan M, Munawar S, Cruz C. Business insights using RAG–LLMs: a review and case study. Journal of Decision Systems. 2024.

[17] Dogan O, Gurcan OF. Enhancing E-Business Communication with a Hybrid Rule-Based and Extractive-Based Chatbot. Journal of Theoretical and Applied Electronic Commerce Research. 2024;19(3):1984–99.

[18] Zhang L, Yang Y, Zhou J, Chen C, He L. Retrieval-Polished Response Generation for Chatbot. IEEE Access. 2020;8:123882–90.

[19] Ait-Mlouk A, Jiang L. KBot: A Knowledge Graph Based ChatBot for Natural Language Understanding over Linked Data. IEEE Access. 2020;8:149220–30.

[20] Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y. Large Language Models are Zero-Shot Reasoners. arXiv:2205.11916. 2022.

[21] Huang SM, Soepriyanto G, Wahyuningtias D, Winoto A. Social Media Applications for MSMEs in the Era of the Digital Economy. Journal of Theoretical and Applied Information Technology. 2023;15(7).
[22] Mash R, Schouw D, Fischer AE. Evaluating the Implementation of the GREAT4Diabetes WhatsApp Chatbot to Educate People with Type 2 Diabetes during the COVID-19 Pandemic: Convergent Mixed Methods Study. JMIR Diabetes. 2022;7(2).
[23] Bratić D, Šapina M, Jurečić D, Žiljak Gršić J. Centralized Database Access: Transformer Framework and LLM/Chatbot Integration-Based Hybrid Model. Applied System Innovation. 2024;7(1).
[24] Sagstad MH, Morken NH, Lund A, et al. Quantitative User Data From a Chatbot Developed for Women With Gestational Diabetes Mellitus: Observational Study. JMIR Formative Research. 2022;6(4).
[25] Creswell JW, Creswell JD. Research design: Qualitative, quantitative, and mixed methods approaches. 6th ed. Thousand Oaks: SAGE Publications; 2023.
[26] Campbell DT, Stanley JC. Experimental and quasi-experimental designs for research on teaching. 2nd ed. Boston: Houghton Mifflin; 2022.
[27] Tashakkori A, Teddlie C. SAGE handbook of mixed methods in social & behavioral research. 3rd ed. Thousand Oaks: SAGE Publications; 2021.
[28] Zhang Y, Chen X, Wang L. Knowledge base construction for domain-specific chatbots: A systematic approach. Expert Systems with Applications. 2023;210:118434.
[29] Koto F, Rahimi A, Lau JH, Baldwin T. IndoBERT: A pre-trained language model for Indonesian language understanding and generation. arXiv:2009.09745. 2020.
[30] Adamopoulou E, Moussiades L. An overview of chatbot technology. Artificial Intelligence Applications and Innovations. 2020;584:373-383.
[31] Diederich S, Brendel AB, Kolbe LM. Designing anthropomorphic enterprise conversational agents. Business & Information Systems Engineering. 2020;62(3):193-209.
[32] Radziwill NM, Benton MC. Evaluating quality of chatbots and intelligent conversational agents. arXiv:1704.04579. 2017.
[33] Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016;355:i5239.
[34] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
[35] Venkatesh V, Brown SA, Sullivan YW. Guidelines for conducting mixed-methods research: An extension and illustration. Journal of the Association for Information Systems. 2016;17(7):435-494.
[36] Field A. Discovering statistics using IBM SPSS statistics. 5th ed. London: SAGE Publications; 2018.
[37] Barnes RM. Motion and time study: Design and measurement of work. 7th ed. New York: John Wiley & Sons; 1980.
[38] Angrist JD, Pischke JS. Mostly harmless econometrics: An empiricist's companion. Princeton: Princeton University Press; 2009.
[39] Singer JD, Willett JB. Applied longitudinal data analysis: Modeling change and event occurrence. Oxford: Oxford University Press; 2003.
[40] Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.
[41] Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869.
[42] Brynjolfsson E, Hitt LM. Beyond computation: Information technology, organizational transformation and business performance. Journal of Economic Perspectives. 2000;14(4):23-48.
[43] OECD. SME and Entrepreneurship Policy in Indonesia 2018. OECD Studies on SMEs and Entrepreneurship. Paris: OECD Publishing; 2018.