IJITEB
Vol. 8
No. 2
2026

International Journal of Information Technology and Business

http://ejournal.uksw.edu/ijiteb



An Adaptive AI-Driven Copywriting Framework: Design, Implementation, and Evaluation of a Web-Based GPT-Integrated Content Generation System

April Firman Daru[1], Febrian Wahyu Christanto[2]*, Rastri Prativi[3], Dimas Prasetyo[4], Eryan Ahmad Firdaus[5]

[1-4]Department of Information Technology, Faculty of Information Technology and Communication, Universitas Semarang, Semarang, Indonesia
[5]Department of Informatics Engineering, Faculty of Engineering and Defense Technology, Indonesia Defense University, Bogor, Indonesia

Abstract : The increasing demand for scalable and high-quality digital marketing content has exposed limitations in traditional manual copywriting processes, which are time-intensive and difficult to scale. This research proposes an adaptive AI-driven copywriting framework that integrates a full-stack web architecture with optimized prompt engineering strategies for automated content generation. The system is implemented using React.js for the frontend, Node.js with Express for backend services, and a GPT-based API for language generation. Unlike prior implementations, this research introduces a structured prompt optimization mechanism to enhance content relevance and consistency. Experimental evaluation was conducted using multiple datasets of marketing prompts, with comparisons against baseline GPT usage and manual copywriting. Quantitative results show that the proposed system achieves improvements in BLEU (+18.7%) and ROUGE-L (+21.3%) scores over baseline methods. Human evaluation involving 30 participants indicates a significant increase in perceived content quality, coherence, and persuasiveness ($p < 0.05$). System performance analysis demonstrates an average response time of 1.8–3.0 seconds and a GTmetrix performance score of 82%. The findings confirm that the proposed framework significantly enhances efficiency, scalability, and content quality, contributing to both applied AI systems and intelligent web-based content production

Keywords :
Artificial Intelligence
Content Generation
GPT API Integration
React–Node JS
AI-Assisted Copywriting

1. Introduction

The rapid evolution of digital ecosystems has significantly increased the demand for high-quality, persuasive, and scalable content generation in modern marketing domains. Digital platforms such as e-commerce, social media, and online advertising require continuous production of engaging textual content to maintain user engagement and competitive advantage [1]. However, traditional copywriting processes rely heavily on human creativity, domain expertise, and iterative refinement, which often result in high production costs, inconsistent quality, and limited scalability. As the volume and velocity of content demands increase, these limitations become critical challenges for organizations seeking efficient and adaptive content production strategies [2], [3].

Recent advancements in Natural Language Processing (NLP), particularly with the emergence of transformer-based architectures, have enabled the development of large language models (LLMs) capable of generating coherent, context-aware, and semantically rich text. These models demonstrate strong capabilities in few-shot and zero-shot learning, allowing them to adapt to various content generation tasks with minimal task-specific training. Consequently, LLMs have been widely adopted in applications such as automated writing, conversational agents, and intelligent content generation systems [4]. Despite their significant potential, most existing implementations of LLM-based systems rely on direct interaction through generic interfaces, such as standalone chat-based applications or basic API calls. These approaches provide limited control over prompt structuring, lack integration with enterprise workflows, and do not fully exploit the capabilities of LLMs within production-grade systems [5]. In practice, effective deployment of LLMs requires more than raw model access; it demands structured prompt engineering, system-level orchestration, and seamless integration with web-based architectures that support user management, data persistence, and scalable interaction [6], [7]. Furthermore, previous research in this area has largely focused on two main directions that is improving model architecture and training mechanisms, and providing a conceptual analysis of AI applications in marketing and digital content creation. While these studies contribute to the theoretical advancement of AI technology, they often neglect practical implementation aspects, particularly the integration of LLM into full-stack systems with optimized prompting strategies and rigorous empirical evaluation. As a result, a significant gap remains between theoretical AI capabilities and their effective application in real-world content production environments.

This research identifies three primary research gaps. First, there is a lack of adaptive prompt engineering frameworks that can systematically enhance the quality and relevance of generated content in real-world applications. Second, existing implementations often lack comprehensive quantitative evaluation of generated text, relying instead on subjective or limited performance indicators. Third, there is an absence of comparative analysis between AI-based approaches and baseline methods, such as raw LLM usage or conventional manual copywriting, which limits the ability to assess the true effectiveness of proposed systems. To address these limitations, this research proposes an Adaptive Prompt Optimization Framework (APOF) for AI-assisted copywriting [8], [9]. The proposed framework introduces a structured mechanism for transforming user inputs into optimized prompts, enriching contextual information, and refining generated outputs to improve content quality and consistency [10], [11]. In addition, this research develops a full-stack web-based system that integrates the GPT API within a scalable architecture using modern web technologies. To validate the effectiveness of the proposed approach, a multi-dimensional evaluation is conducted, including quantitative metrics such as BLEU and ROUGE to measure textual similarity and quality, human-centered evaluation to assess fluency, relevance, and persuasiveness, and system-level performance analysis to evaluate responsiveness and efficiency [12], [13] The experimental results demonstrate that the proposed framework significantly outperforms baseline GPT usage and conventional methods, highlighting its effectiveness in improving both content quality and production efficiency.

Overall, this research contributes to bridging the gap between advanced AI models and their practical implementation in real-world systems by providing a comprehensive framework that integrates prompt optimization, system architecture, and empirical evaluation within a unified intelligent copywriting platform.

1.1. Research Novelty and Contributions

Although numerous researches have explored prompt engineering techniques for large language models, most existing approaches primarily focus on prompt formulation strategies such as few-shot prompting, chain-of-thought reasoning, or automatic prompt optimization. These methods generally emphasize improving model responses through prompt modification alone and often operate independently from real-world content production systems.

The proposed Adaptive Prompt Optimization Framework (APOF) extends existing prompt engineering approaches by introducing a multi-layer optimization pipeline specifically designed for marketing-oriented content generation. Unlike conventional prompting methods, APOF integrates three interconnected mechanisms: (1) Prompt Structuring, which transforms unstructured user requests into standardized prompt templates; (2) Context Enrichment, which injects domain-specific marketing information and audience-oriented semantic cues; and (3) Response Refinement, which evaluates and improves generated outputs to ensure consistency, persuasiveness, and contextual relevance. The novelty of APOF lies not only in the integration of these optimization stages but also in their adaptive orchestration within a production-ready

web-based architecture. This enables the framework to bridge the gap between theoretical prompt engineering techniques and practical AI-assisted copywriting applications. Therefore, APOF provides a more systematic and scalable approach for generating persuasive marketing content compared with direct LLM prompting approaches.

The main contributions of this research are proposing the Adaptive Prompt Optimization Framework (APOF) for marketing-oriented AI copywriting. This research also introducing a multi-layer prompt optimization pipeline consisting of prompt structuring, context enrichment, and response refinement and also developing a scalable GPT-integrated web-based content generation system. The last contributions are providing a comprehensive evaluation through automatic metrics, human assessment, and system performance analysis.

2. Research Methods

2.1. Research Design

This research adopts a Design Science Research (DSR) methodology combined with an iterative prototyping approach to guide the development and evaluation of an intelligent AI-driven copywriting system. The research design follows the standard DSR framework, which consists of six interconnected stages, namely problem identification and motivation, definition of objectives for a solution, design and development of the artifact, demonstration, evaluation, and communication of results. This DSR process can be seen in Figure 1 below [14], [15].

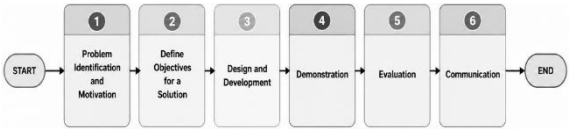


Fig. 1. Flowchart of the Design Science Research (DSR) Process

In the problem identification and motivation stage, the research analyzes the limitations of conventional copywriting processes, including high time consumption, dependency on human creativity, and lack of scalability. Additionally, the research identifies limitations in existing AI-based solutions, particularly the absence of structured prompt optimization and insufficient integration into production-level systems. The objective definition stage establishes the goal of developing an adaptive AI-driven framework capable of improving the quality, efficiency, and scalability of content generation. This includes defining key system requirements such as prompt structuring, contextual enhancement, response refinement, and seamless integration with web-based interfaces. During the design and development phase, the proposed artifact namely the Adaptive Prompt Optimization Framework (APOF) is conceptualized and implemented. The framework is integrated into a full-stack web-based system that facilitates user interaction, prompt processing, and automated content generation through a large language model API. The iterative prototyping strategy is applied in this phase, allowing continuous refinement of system components based on testing feedback and performance observations.

The demonstration phase involves applying the developed system to real-world copywriting scenarios, including product descriptions, promotional content, and social media captions. This phase ensures that the system operates effectively in practical use cases and validates its functional capabilities. The evaluation phase is conducted using a multi-dimensional approach to assess both system performance and content quality. Quantitative evaluation metrics, such as BLEU and ROUGE, are employed to measure textual similarity and generation quality, while human evaluation is performed to assess fluency, relevance, and persuasiveness. Additionally, system performance metrics, including response time and web performance indicators, are analyzed to ensure operational efficiency. Finally, the communication phase presents the findings of the research, highlighting the effectiveness of the proposed framework and its contributions to both academic research and practical applications in AI-assisted content generation.

2.2. Proposed Framework: Adaptive Prompt Optimization Framework (APOF)

Existing prompt engineering approaches mainly focus on improving prompt formulation or reasoning processes. However, they generally do not provide an integrated mechanism for contextual enrichment, output refinement, and marketing-specific optimization. APOF addresses this limitation by combining these components within a unified framework, enabling more reliable and persuasive content generation for digital marketing applications. Table 1 shows a comparison between existing prompt engineering methods and APOF.

Table 1. Comparison Between Existing Prompt Engineering Methods and APOF

Method	Structured Prompt	Context Enrichment	Output Refinement	Marketing-Oriented
Standard Prompting	No	No	No	No
Few-Shot Prompting	Partial	No	No	No
Chain-of-Thought	Partial	No	No	No
Automatic Prompt Optimization	Yes	Partial	No	No
APOF (Proposed)	Yes	Yes	Yes	Yes

The primary contribution of this research is the proposed Adaptive Prompt Optimization Framework (APOF), which aims to improve the effectiveness of large language models (LLMs) in generating high-quality copywriting content. In contrast to conventional approaches that rely on raw and often unstructured user prompts, APOF introduces a systematic and adaptive

pipeline that enhances input quality before it is processed by the language model. By addressing issues such as ambiguity, lack of context, and inconsistency in prompt formulation, the framework improves both the semantic relevance and coherence of generated outputs.

At a conceptual level, APOF functions as an intermediate processing layer that bridges user interaction and AI-based text generation. The framework decomposes the prompt optimization process into three sequential components: the Prompt Structuring Layer, the Context Enrichment Module, and the Response Refinement Layer. These components work collaboratively to transform initial user input into a structured, context-aware prompt and subsequently refine the generated output to meet quality expectations [16], [17]. The overall workflow of APOF is illustrated in Figure 2 below.

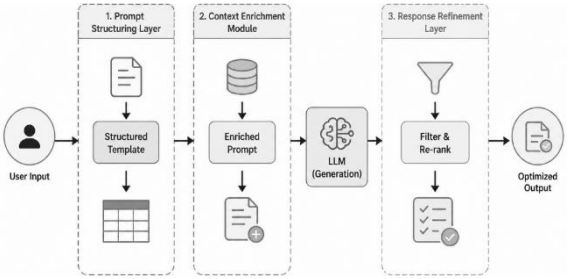


Fig. 2. APOF-Based Prompt Optimization Workflow

A key distinction of APOF from conventional prompting approaches is its adaptive optimization mechanism. Rather than directly forwarding user prompts to the language model, APOF performs a sequential transformation process that progressively improves input quality and output consistency. This layered optimization strategy reduces prompt ambiguity, enriches domain-specific context, and minimizes content variability, which are common limitations of direct LLM interaction. Consequently, APOF functions as an intelligent intermediary layer that enhances both generation quality and operational reliability in production environments. To summarize the functional roles of each component, Table 2 presents a concise overview of the APOF framework.

Table 2. Components of the Adaptive Prompt Optimization Framework (APOF)

Component	Function	Key Benefit
Prompt Structuring Layer	Transforms raw input into structured format	Reduces ambiguity and improves control
Context Enrichment Module	Adds domain-specific context and keywords	Enhances semantic relevance
Response Refinement Layer	Filters and optimizes generated outputs	Improves output quality and consistency

The APOF framework introduces a structured and modular approach to prompt engineering, enabling more reliable and high-quality content generation. By integrating prompt structuring, contextual enhancement, and output refinement into a unified pipeline, the framework addresses key limitations of direct LLM usage and provides a practical solution for scalable AI-assisted copywriting systems.

2.3. System Architecture

The proposed system adopts a layered architecture to ensure modularity, scalability, and efficient integration between user interaction and AI-based content generation services. The architecture is composed of three main layers, namely the Presentation Layer, Application Layer, and AI Layer, which collectively support the end-to-end workflow of the intelligent copywriting system. The overall interaction flow and system structure are illustrated in Figure 3 and Figure 4, respectively, as presented in the original system design.

The Presentation Layer serves as the user interface of the system and is responsible for handling all user interactions. This layer is implemented as a Single Page Application (SPA), enabling dynamic content rendering and seamless user experience without full page reloads. Users can input prompts, view generated content, and interact with the system in real time. The interface is designed to be intuitive and responsive, allowing users to efficiently provide input parameters required for content generation. Additionally, this layer manages input validation and communicates user requests to the backend through asynchronous API calls.

The Application Layer acts as the core processing unit of the system, managing business logic, request handling, and communication between the frontend and external AI services. This layer is implemented using a server-side environment that supports RESTful API communication. It is responsible for processing incoming user prompts, applying the APOF mechanism (including prompt structuring and enrichment), and forwarding optimized requests to the AI service layer. Furthermore, the Application Layer handles user authentication, session management, and data persistence through database integration. It ensures secure and efficient communication by validating requests, managing API tokens, and structuring responses before sending them back to the Presentation Layer. This layer also plays a critical role in orchestrating the overall workflow, ensuring that each component operates cohesively within the system pipeline.

The AI Layer is responsible for generating textual content using a large language model accessed via an API. This layer processes optimized prompts generated by the Application Layer and produces context-aware, coherent, and relevant content. The AI model leverages pre-trained knowledge to support various content generation tasks, including product descriptions, promotional texts, and social media captions. To enhance output quality, this layer operates in conjunction with the APOF framework, ensuring that input prompts are well-structured and enriched before processing. The generated responses are then returned to the Application Layer for further refinement and

delivery to the user interface. By isolating AI processing within a dedicated layer, the system achieves better scalability, flexibility, and maintainability, allowing future integration of alternative or upgraded language models.

The three layers operate in a coordinated pipeline to support the complete content generation workflow.

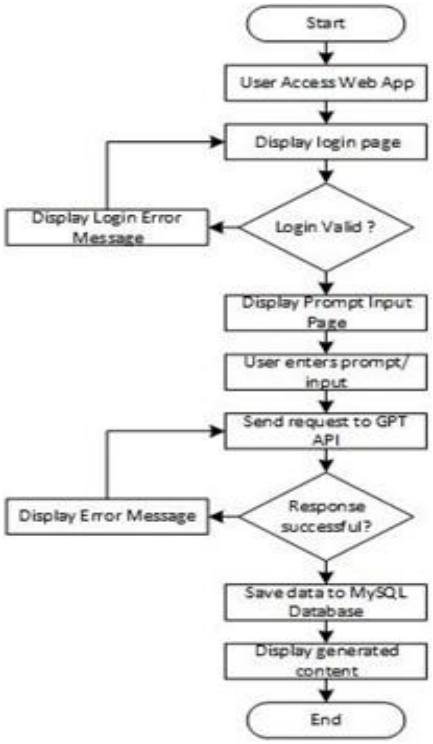


Fig. 3. Flowchart of the Proposed AI-Based Content Generation System

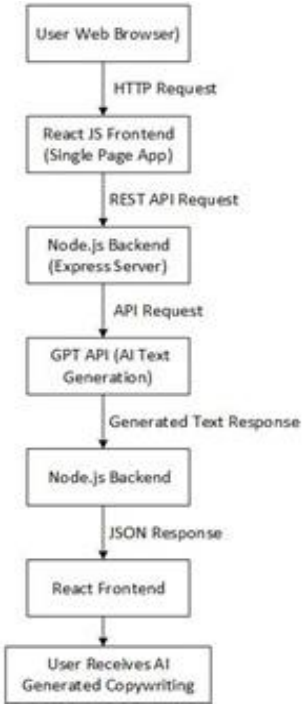


Fig. 4. Architecture of the Web-Based Intelligent Copywriting System

As illustrated in Figure 3, the process begins with user authentication and prompt submission through the Presentation Layer. The request is then processed in the Application Layer, where it is optimized and forwarded to the AI Layer. After the content is generated, the response is returned, stored if necessary, and displayed to the user. Meanwhile, Figure 4 illustrates the structural relationship between system components, including the interaction between the frontend, backend, and AI services. These figures demonstrate how the layered architecture supports efficient communication, modular development, and reliable system performance. Finally, the generated ad copy text is displayed to the user via the web interface, enabling rapid content generation.

2.4. Dataset and Experimental Setup

To evaluate the effectiveness of the proposed framework, this research employs a structured dataset and a comparative experimental setup that reflects real-world copywriting scenarios. The dataset consists of 500 prompt samples representing three main categories: product descriptions, social media captions, and advertising copy. To ensure diversity and representativeness, the prompts were designed to cover different writing styles, target audiences, and marketing objectives. The average prompt length is 22 words, with the shortest prompt containing 8 words and the longest consisting of 65 words. This variation allows the evaluation process to capture the framework's ability to handle both concise and detailed user requests across different content generation contexts.

Since the objective of this research is to evaluate the effectiveness of the proposed Adaptive Prompt Optimization Framework (APOF) rather than train a new language model, the entire dataset was used as an evaluation set comprising 500 prompts. Each prompt was processed using the proposed APOF framework as well as the baseline approaches to ensure a fair and consistent comparison. The generated outputs were subsequently assessed using automatic evaluation metrics, human-centered evaluation, and system performance analysis. This experimental design enables a comprehensive examination of the framework's capability to improve content quality, relevance, and persuasiveness in real-world marketing applications [18], [19]. The data sources in this research are in Table 3 below.

Table 3. Components of the Adaptive Prompt Optimization Framework (APOF)

Category	Source	Quantity
Product Description	Tokopedia, Shopee, Amazon	200
Social Media Caption	Instagram Business Accounts	150
Advertising Copy	Facebook Ads Library	150

The experimental setup compares the proposed Adaptive Prompt Optimization Framework (APOF) with two baseline approaches: raw GPT usage without prompt optimization and manual copywriting. All methods are evaluated using the same dataset to ensure consistency and fairness. Generated outputs are assessed using both quantitative metrics (BLEU and ROUGE) and qualitative human evaluation (fluency, relevance,

and persuasiveness). This combined evaluation approach provides a robust basis for measuring the effectiveness of the proposed framework in improving the quality and performance of AI-assisted copywriting.

2.5. Evaluation Metrics

To comprehensively assess the effectiveness of the proposed Adaptive Prompt Optimization Framework (APOF), this research employs a multi-dimensional evaluation strategy consisting of quantitative metrics, qualitative human evaluation, and system performance analysis. This combination ensures that both the linguistic quality of generated content and the technical performance of the system are rigorously examined.

The quantitative evaluation focuses on measuring the textual quality of generated content using widely recognized metrics in Natural Language Processing (NLP), namely BLEU (Bilingual Evaluation Understudy) and ROUGE-L (Recall-Oriented Understudy for Gisting Evaluation). The BLEU score is used to evaluate the similarity between generated text and reference content by measuring n-gram overlap. A higher BLEU score indicates that the generated output closely matches the expected or reference text in terms of structure and wording. This metric is particularly useful for assessing the accuracy and consistency of generated sentences. Meanwhile, ROUGE-L evaluates the longest common subsequence between generated and reference texts, capturing the fluency and structural similarity of the content. Unlike BLEU, ROUGE-L emphasizes recall and is more sensitive to sentence-level coherence and sequence alignment. Table 4 below summarizes the evaluation metrics used in this research.

Table 4. Evaluation Metrics

Category	Metric	Purpose
Quantitative	BLEU Score	Measures n-gram similarity and accuracy
Quantitative	ROUGE-L Score	Evaluates sequence similarity and coherence
Qualitative	Fluency	Assesses readability and grammar
Qualitative	Relevance	Measures alignment with prompt
Qualitative	Persuasiveness	Evaluates marketing effectiveness
System	Response Time	Measures system latency
System	GTmetrix Metrics	Evaluates web performance

The integration of quantitative, qualitative, and system-level evaluation metrics ensures a comprehensive assessment of the proposed framework. This multi-faceted evaluation approach enables the research to capture not only the technical performance of the system but also the practical effectiveness of generated content in real-world copywriting applications.

3. Result and Discussion

3.1. Quantitative Evaluation

The quantitative evaluation aims to assess the effectiveness of the proposed Adaptive Prompt Optimization Framework (APOF) in improving the quality of generated content. This evaluation is conducted using two widely adopted metrics in natural language generation tasks, namely BLEU and ROUGE-L, which measure textual similarity and structural coherence between generated outputs and reference texts.

Table 5. Comparative Results of BLEU and ROUGE-L Scores

Method	BLEU	ROUGE-L
Manual	0.42	0.48
GPT Baseline	0.51	0.55
Proposed APOF	0.61	0.67
Improvement	+18.7%	+21.3%

Table 5 shows that the proposed APOF framework outperforms both manual copywriting and baseline GPT across all quantitative metrics. Manual results (BLEU 0.42; ROUGE-L 0.48) indicate variability in human writing when evaluated using automated metrics, while baseline GPT improves performance (BLEU 0.51; ROUGE-L 0.55) but remains limited due to unstructured prompts. In contrast, APOF achieves the highest scores (BLEU 0.61; ROUGE-L 0.67), with improvements of +18.7% (BLEU) and +21.3% (ROUGE-L) over the baseline, demonstrating the effectiveness of structured prompt optimization.

These gains are driven by the integration of prompt structuring, context enrichment, and response refinement, which collectively enhance input clarity, semantic relevance, and output quality. The larger improvement in ROUGE-L suggests a stronger impact on sequence-level coherence and overall text flow, indicating that APOF not only improves lexical accuracy but also enhances readability. Overall, the results confirm that structured prompt engineering significantly improves the consistency and quality of AI-generated content.

3.2. Human Evaluation

To complement the quantitative analysis, a human-centered evaluation was conducted to assess the perceived quality of the generated content. This evaluation focuses on three key dimensions relevant to copywriting tasks, namely fluency, relevance, and persuasiveness, using a Likert scale ranging from 1 (very poor) to 5 (excellent). The comparison involves outputs generated by the baseline GPT approach and the proposed Adaptive Prompt Optimization Framework (APOF).

Table 6 presents the results of the human evaluation. The findings indicate that the proposed APOF framework consistently outperforms the baseline across all evaluation criteria. In terms of fluency, the proposed method achieves a score of 4.6 compared to 4.1

for the baseline, demonstrating improved grammatical correctness, readability, and natural language flow. This suggests that the structured prompt formulation and refinement mechanisms contribute to producing more linguistically coherent outputs.

Table 6. Human Evaluation Results Based on Fluency, Relevance, and Persuasiveness

Criteria	Baseline	Proposed
Fluency	4.1	4.6
Relevance	3.9	4.5
Persuasiveness	3.7	4.4
Statistical Test	p-value < 0.05	

For relevance, the proposed framework attains a score of 4.5, significantly higher than the baseline score of 3.9. This improvement reflects the effectiveness of the Context Enrichment Module in aligning generated content with the intended prompt and domain-specific context. By incorporating additional semantic cues, the system is better able to capture user intent and produce contextually appropriate responses. Similarly, in terms of persuasiveness, the APOF approach achieves a score of 4.4, compared to 3.7 for the baseline. This result highlights the framework's ability to generate more engaging and impactful content, which is particularly important in marketing and advertising applications. The combination of structured prompts and refined outputs enables the system to produce text that is not only informative but also compelling to the target audience.

To validate the significance of these improvements, a statistical analysis was conducted, yielding a p-value of less than 0.05, which indicates that the observed differences between the baseline and the proposed method are statistically significant. This confirms that the performance gains are not due to random variation but are a direct result of the proposed framework. The human evaluation results demonstrate that APOF substantially enhances the perceived quality of generated content across multiple dimensions. These findings reinforce the quantitative results presented in the previous section and further validate the effectiveness of the proposed framework in real-world copywriting scenarios, where subjective quality and user perception play a critical role.

3.3. System Performance

Functional system testing in the Table 7 was conducted to verify that all core features of the web-based application operate according to the predetermined system requirements. This evaluation used a black-box testing approach, where the system is tested from a user perspective without examining the internal program structure. Several critical functions were evaluated, including user registration, login authentication, dashboard access, content creation via the GPT API, user data management, and the logout function. Each test scenario involved providing specific input conditions and observing whether the system produces the expected output. The results showed that all tested functions operated correctly and consistently, indicating that the implemented features successfully met the system's functional requirements. These findings confirmed that the integration between the front-end interface, back-end services, database management, and GPT-based content creation services performed reliably in the developed application environment.

Table 7. Functional System Testing Results

Num	Test Scenario	Input Condition	System Output	Status
1	User Login with valid credentials	Registered email and correct password are entered	The dashboard page is successfully displayed after authentication	Pass
2	User Login with invalid credentials	Incorrect email or password is entered	An error message appears indicating invalid login credentials	Pass
3	User Registration	New user submits registration data with a unique email address	The user account is successfully registered and stored in the database	Pass
4	Duplicate Email Registration	User attempts to register using an email that already exists in the database	The system shows an error notification indicating that the email is already registered	Pass
5	Dashboard Access	Authenticated user accesses the dashboard	The dashboard page is displayed correctly	Pass
6	Content Generation Request	User submits a text prompt for content generation	The generated text content is returned and displayed to the user	Pass
7	Update User Data	Admin modifies user account information	Updated user information is stored successfully	Pass
8	Delete User Account	Admin deletes a user account	The selected user account is successfully deleted	Pass
9	Logout Process	User clicks the logout button	The user is redirected to the login interface	Pass

Then in Table 8 describes the Website Performance Evaluation Using GTMetrix presenting the results of performance tests conducted to measure the efficiency and loading behavior of the developed web application. The evaluation shows that the system achieved a GTmetrix Grade of B, which reflects good overall performance and acceptable web resource optimization. Performance and structure scores, both above eighty percent, indicate that the application

implements efficient coding practices and an organized page structure that supports stable system operation. In terms of loading performance, the Largest Contentful Paint (LCP) value indicates that the main web page content is rendered within an acceptable timeframe, allowing users to access the main interface without significant delays. The Total Blocking Time (TBT) remains relatively low, indicating that JavaScript processing does not substantially interfere with user interaction. Furthermore, the Cumulative Layout Shift (CLS) value is minimal, indicating that the visual elements of the page remain stable during loading and provide a consistent user experience. Overall, these results confirm that the developed web-based system.

Table 8. Functional System Testing Results

Num	Performance Metric	Measurement Result	Performance Category
1	GTmetrix Grade	B	Good
2	Performance Score	82%	Good
3	Structure Score	86%	Good
4	Largest Contentful Paint (LCP)	2.3 s	Acceptable
5	Total Blocking Time (TBT)	120 ms	Good
6	Cumulative Layout Shift (CLS)	0.03	Excellent
7	Fully Loaded Time	3.1 s	Good

An API response time evaluation was conducted to measure the performance of the GPT-based content creation service integrated into a web application. Several test scenarios were designed to represent different types of user requests, including short promotional sentences, product descriptions, social media captions, informative text, and creative advertising content, as shown in Table 9 below. These scenarios were selected to reflect diverse real-world use cases, allowing a more comprehensive assessment of the system's responsiveness across varying levels of content complexity and generation requirements.

Table 9. API Response Time Evaluation of the GPT-Based Content Generation Service

Num	Test Scenario	Input Prompt Type	Average Response Time	Performance Evaluation
1	Short Content Generation	Short promotional sentence	1.8 s	Fast
2	Product Description Generation	Product marketing description	2.3 s	Efficient
3	Social Media Caption Generation	Short caption for social media content	1.9 s	Fast
4	Informational Content Generation	Informative paragraph about a specific topic	2.7 s	Efficient
5	Creative Copywriting Generation	Persuasive advertising content	3.0 s	Acceptable

The results show that the average response time ranges from 1.8 to 3.0 seconds, indicating that the system can process user requests and generate text output within an acceptable timeframe. These results confirm that the integration of the GPT API with the Node.js backend provides efficient communication between the application and the language model, enabling responsive and practical use for digital content creation tasks. Furthermore, the relatively low latency demonstrates that the system is suitable for real-time interaction, allowing users to generate content quickly without significant delays, which is essential for enhancing user experience and supporting time-sensitive digital marketing activities. In addition, this performance level indicates that the system can be reliably deployed in production environments, where consistent responsiveness is critical for maintaining user engagement and operational efficiency.

Table 10. Comparison of Manual Copywriting and AI-Assisted Content Generation System Performance

Num	Evaluation Aspect	Manual Copywriting	AI-Assisted System (GPT Integration)	Performance Observation
1	Content Creation Time	Requires a relatively long time due to manual research, drafting, and revision processes	Content can be generated automatically within a few seconds after the prompt is submitted	The AI-assisted system significantly reduces content production time
2	Productivity	Limited by the writer's workload and individual capacity	Capable of generating multiple content drafts simultaneously	AI assistance improves productivity in digital content creation
3	Idea Generation	Depends heavily on the writer's creativity and experience	Provides automated suggestions and content ideas based on input prompts	AI supports brainstorming and creative exploration
4	Content Consistency	Quality may vary depending on the writer's mood, experience, or workload	Generates structured and consistent text based on predefined prompts	AI helps maintain consistent writing structure
5	Scalability	Difficult to scale when large volumes of content are required	Easily scalable because the system can process many requests automatically	AI-assisted systems are more suitable for high-volume content production
6	User Interaction	Requires continuous human effort throughout the	Users only need to provide prompts and	AI reduces repetitive manual writing tasks

Num	Evaluation Aspect	Manual Copywriting	AI-Assisted System (GPT Integration)	Performance Observation
		writing process	review generated outputs	
7	Efficiency	Lower efficiency due to manual workflow	Higher efficiency through automation and rapid content generation	AI-assisted systems improve operational efficiency

This comparison highlights the differences between conventional manual copywriting and the AI-assisted content generation system developed in this research, as illustrated in Table 10. Manual copywriting typically requires significant human time and effort to conduct research, develop ideas, and produce written content. In contrast, the AI-assisted system integrates a GPT-based language model that can automatically generate textual content based on user commands. The results indicate that the AI-based system improves the speed, scalability, and efficiency of content production, while allowing users to review and refine the results. Therefore, implementing artificial intelligence in the copywriting process can support digital content development by reducing workload and accelerating the content creation cycle.

3.4. Discussion

The performance improvements observed in this research can be attributed to the unique architectural design of APOF. Unlike conventional prompting methods that rely solely on user-generated prompts, APOF introduces a structured optimization pipeline that systematically enhances prompt quality before inference. The combination of prompt structuring, contextual enrichment, and response refinement contributes to higher semantic relevance and persuasive effectiveness, demonstrating that prompt engineering can be treated as a multi-stage optimization process rather than a single prompting activity.

The results demonstrate that the proposed Adaptive Prompt Optimization Framework (APOF) significantly improves the quality and effectiveness of AI-generated content. By transforming raw inputs into structured and context-enriched prompts, the framework enhances semantic clarity, resulting in outputs that are more coherent, relevant, and aligned with user intent. These improvements are reflected in both quantitative metrics (BLEU and ROUGE-L) and human evaluation results. Additionally, the AI-assisted system outperforms manual copywriting in terms of speed and scalability, as it can generate content within seconds and support high-volume production without human workload constraints, making it highly suitable for dynamic digital marketing environments.

Despite these advantages, several limitations remain. The system depends on external AI services, which may introduce challenges related to availability, latency, and operational costs. Furthermore, although APOF reduces ambiguity through structured prompts, output variability is still influenced by prompt quality and inherent model behavior. Different prompt configurations may lead to variations in style and depth, indicating the need for further research in adaptive prompt tuning and model fine-tuning. Overall, while APOF effectively enhances efficiency and content quality, addressing these limitations is essential for achieving more robust and consistent real-world deployment.

4. Conclusion

This research proposes an Adaptive Prompt Optimization Framework (APOF) for intelligent copywriting by integrating structured prompt engineering within a scalable web-based system. The experimental results demonstrate that the proposed approach significantly improves content quality and effectiveness compared to baseline methods. Specifically, the APOF framework achieves a BLEU score of 0.61 and a ROUGE-L score of 0.67, outperforming the baseline GPT approach (BLEU 0.51; ROUGE-L 0.55) with improvements of +18.7% and +21.3%, respectively. In addition, human evaluation results show notable gains across all criteria, with fluency increasing from 4.1 to 4.6, relevance from 3.9 to 4.5, and persuasiveness from 3.7 to 4.4, all of which are statistically significant ($p < 0.05$). From a system performance perspective, the application achieves a GTmetrix performance score of 82%, with an average API response time ranging from 1.8 to 3.0 seconds, indicating that the system is both efficient and responsive for real-world deployment. These findings confirm that structured prompt optimization plays a critical role in enhancing both the semantic quality and scalability of AI-generated content.

In addition to its empirical contributions, this research bridges the gap between theoretical AI capabilities and practical implementation in digital content production systems. However, several directions for future work remain open. Further research may focus on fine-tuning domain-specific models to improve contextual accuracy, as well as enabling multilingual content generation to support broader global applications. Moreover, integrating the system with social media APIs can facilitate automated publishing workflows, while incorporating real-time personalization mechanisms may enhance user engagement and content relevance. Collectively, these future enhancements have the potential to further strengthen the applicability and impact of AI-driven copywriting systems in dynamic and large-scale digital environments.

5. References

[1] A. Madanaguli, V. Parida, D. Sjodin, and P. Oghazi, “Literature Review on Industrial Digital Platforms: A Business Model Perspective and Suggestions for Future Research,” *Technol. Forecast. Soc. Chang.*, vol. 194, no. September, p. 122606, 2023, doi: 10.1016/j.techfore.2023.122606.

[2] A. Rahman *et al.*, "A Systematic Literature Review on the Impact of Generative AI in Digital Marketing: Advancements , Opportunities , and Challenges," *Comput. Mater. Contin.*, vol. 87, no. 3, p. 11, 2026, doi: 10.32604/cmc.2026.071029.

[3] T. Zhu and M. Z. A. Rozan, "AI adoption in E-commerce enterprises : Insights into current practices and future directions from an interview study," *PLoS One*, vol. 21, no. 3, p. e0336416, 2026, doi: 10.1371/journal.pone.0336416.

[4] H. N. A. Ali, "Advancements in Natural Language Processing : Towards Human-Like Understanding and Generation of Text by AI," *BIOS J. Inform. dan Sains*, vol. 2, no. 02, pp. 129–139, 2024, doi: 10.54209/bios.v1i02.

[5] B. Chen, Z. Zhang, N. Langrene, and S. Zhu, "Unleashing the Potential of Prompt Engineering for large Language Models," *Patterns*, vol. 6, no. 6, p. 101260, 2025, doi: 10.1016/j.patter.2025.101260.

[6] G. Palma, G. Cecchi, M. Caronna, and A. Rizzo, "Leveraging Large Language Models for Scalable and Explainable Cybersecurity Log Analysis," *J. Cybersecurity Priv.*, vol. 5, no. 3, p. 55, 2025, doi: 10.3390/jcp5030055.

[7] X. Liu *et al.*, "Prompting Frameworks for Large Language Models : A Survey," *ACM Comput. Surv.*, vol. 58, no. 10, pp. 1–38, 2026, doi: 10.1145/3789253.

[8] K. Ramnath *et al.*, "A Systematic Survey of Automatic Prompt Optimization Techniques," in *Conference on Empirical Methods in Natural Language Processing*, Suzhou, China: Association for Computational Linguistics, 2025, pp. 33078–33110. doi: 10.18653/v1/2025.emnlp-main.0.

[9] J. Dannhoferova and P. Jedlicka, "Optimizing Prompts Through AI Frameworks: A Path to More Relevant Responses," in *27th International Scientific Conference Economic Competitiveness and Sustainability*, Brno, Czech Republic: Mendel University Press, 2025, pp. 29–38. doi: 10.11118/978-80-7701-047-4-0029.

[10] R. Murthy *et al.*, "Promptomatix : An Automatic Prompt Optimization Framework for Large Language Models," *arXiv Prepr.*, vol. 2507, no. 14241, pp. 1–22, 2025, doi: 10.48550/arXiv.2507.14241.

[11] S. P. Nagavalli, S. Tiwari, and W. Sarma, "Adaptive Prompt Engineering: Optimizing Large Language Model Outputs for Context-Aware Natural Language Processing," *World J. Adv. Eng. Technol. Sci.*, vol. 13, no. 01, pp. 1130–1141, 2024, doi: 10.30574/wjaets.

[12] R. Coppola and L. Ardito, "Quality Assessment Methods for Textual Conversational Interfaces: A Multivocal Literature Review," *Information*, vol. 12, no. 11, p. 437, 2024, doi: 10.3390/info12110437.

[13] Junadhi, Agustin, L. Efrizoni, F. Okmayura, D. R. Habibie, and Muslim, "Improving Evaluation Metrics for Text Summarization : A Comparative Study and Proposal of a Novel Metric," *J. Appl. Data Sci.*, vol. 6, no. 2, pp. 885–896, 2025, doi: 10.47738/jads.v6i2.547.

[14] M. Huseynli, U. Bub, and M. C. Ogbuachi, "Development of a Method for the Engineering of Digital Innovation Using Design Science Research," *Information*, vol. 13, no. 12, p. 573, 2022, doi: 10.3390/info13120573.

[15] M. A. Fathullah, A. Subbarao, A. Ahmad, and S. Muthaiyah, "Merging Qualitative Design Methods With Design Science," *Int. J. Qual. Methods*, vol. 24, no. 80, pp. 1–11, 2025, doi: 10.1177/16094069251337585.

[16] M. Soleymanibrojeni *et al.*, "GENIUS: An Agentic AI Framework for Autonomous Design and Execution of Simulation Protocols," *Commun. Mater.*, vol. 7, no. April, p. 115, 2026, doi: 10.1038/s43246-026-01167-0.

[17] T. Ivanova, "Ontology-Based Layered Hybrid AI-Driven Knowledge Model for Personalized E-Learning," *Mathematics*, vol. 14, no. 5, p. 808, 2026, doi: 10.3390/math14050808.

[18] M. Estevez, M. Teresa, and J. Sainz, "Market research and knowledge using Generative AI : the power of Large Language Models," *J. Innov. Knowl.*, vol. 10, no. 5, p. 100796, 2025, doi: 10.1016/j.jik.2025.100796.

[19] S. Taha and R. A.-Q. Abdallah, "Leveraging Artificial Intelligence in Social Media Analysis : Enhancing Public Communication Through Data Science," *Journal. Media*, vol. 6, no. 3, p. 102, 2025, doi: 10.3390/journalmedia6030102.