

"AI-to-Human Content Rewriter: Transform AI-Generated Text into Natural Writing"

Shweta Punde,

Department of Data science
J D College of Engineering,
And Management,
Nagpur, India

Sujal Ashtankar

Department of Data science
J D College of Engineering,
And Management,
Nagpur, India

Atharva Kamblee

Department of Data science
J D College of Engineering
And Management,
Nagpur, India

Rhushikesh Bahad

Department of Data science
J D College of Engineering,
And Management,
Nagpur, India

Prof. Namrata Nagpure

Department of Data science
J D College of Engineering,
And Management,
Nagpur, India

Abstract—The AI-to-Human Content Rewriter uses natural language processing technology to convert artificial intelligence text into human-readable content while maintaining its original message. The content generated by artificial intelligence exhibits mechanical characteristics because it uses the same words repeatedly and maintains fixed patterns of speech and shows no emotional expression. The system evaluates text through three aspects which are readability and coherence and stylistic challenges, and it uses various methods including sentence restructuring and vocabulary changes and active voice transformation and contextual paraphrasing and tone adjustment to enhance understanding and authentic presentation of content. The system enables users to write in three different styles which are academic, professional and conversational. The system performance assessment uses two methods which include human evaluation and automated systems that measure vocabulary diversity and sentence complexity. This tool serves academic writing purposes and content creation needs and AI-assisted communication tasks, which require the creation of text that maintains engaging qualities while presenting original content in a humanlike manner.

Keywords — *NLP, LLM, Text Rewriting, Human-like Generation, Readability.*

I. INTRODUCTION

The quick progress of generative artificial intelligence has brought major changes to how educational institutions and businesses and online platforms produce written material. The large language model

tools (LLM) enable users to produce high-quality written content at fast speeds but the resulting text shows common features through formal language and restricted writing patterns and lacks emotional depth. The features of AI-generated content make it easy to identify as nonhuman writing which undermines its value in situations that require natural human communication and active audience participation and real human emotions. The need for systems that can transform machine-generated text into natural human speech through context-appropriate fluent language while maintaining original content exists because of present demands.

The proposed AI-to-Human Content Rewriter addresses this challenge by leveraging natural language processing techniques to enhance readability, coherence, and stylistic quality. The system uses AI-generated text analysis to detect structural issues and repetitive content and grammatical errors which it solves through sentence reorganization and vocabulary expansion and active voice conversion and contextual paraphrasing. The system allows users to adjust tone through academic and conversational and professional styles which enables users to create content appropriate for various fields. The system uses automated linguistic processing together with naturalness and complexity evaluation metrics to connect machine-generated text with human-written content while enhancing the standard of AI-supported communication.

II. LITERATURE SURVEY

The latest progress in text-to-text generation technology enables systems to create automatic automatic text simplification which maintains the original meaning of text. Dou et al. [1] provided a comprehensive overview of neural text revision models which demonstrate their capacity to enhance text fluency and grammatical accuracy and reading difficulty through supervised learning and reinforcement learning techniques. Their work showed that naturalness assessment depends on two vital elements first high-quality datasets and second human evaluation. AI-mediated communication research demonstrates that AI-edited text decreases emotional bias while preserving informational integrity across academic and non-academic contexts [2]. The study found that users need to perceive AI-generated content as more natural because current AI systems lack the ability to create human-like text rewrites.

Recent research studies how people perceive content that AI generates compared to content that humans edit. Arnesen and Grue [3] demonstrated that readers often distinguish AI text based on repetitive phrasing and lack of stylistic diversity which shows that linguistic variation needs to exist for authentic writing. The researchers Lee et al. [4] studied how AI rewriting enables vocabulary personalization in educational storytelling and they found that contextual paraphrasing leads to higher engagement levels together with better comprehension results. Human-edited AI work receives higher credibility according to Ling and Tao [5] in journalism because post-generation refinement needs to take place before journalists can trust AI-generated content. InstructGPT alignment models demonstrate that human feedback-based fine-tuning improves system output through better user support and safe content delivery and natural sounding communications [6]. The study by Bai et al. [7] demonstrated that reinforcement learning from human feedback (RLHF) enables machines to produce output which matches human communication standards.

The main problem which affects AI-produced text stems from two issues which create incorrect information and hallucinations, which lead to people trusting automated writing systems less. Ji et al. [8] surveyed hallucination in NLP and suggested that rewriting and verification mechanisms are necessary to ensure reliability. Educational research by Zellner [9] emphasized the role of rewriting AI content

because it helps students develop critical thinking skills, which leads to more effective learning, instead of just consuming content. Zhou et al. [10] demonstrated that deep learning models can enhance structured text outputs because their architectural design allows for application in rewriting tasks. Studies on style transfer in NLP [11] have shown that models can modify tone and formality without altering core meaning, which is a key requirement for AI-to-human rewriting systems.

Text simplification and paraphrasing methods have made important contributions to this field. The neural paraphrasing models generate diverse sentences which decrease repetitive elements to produce text that sounds more natural to readers. Research on controllable text generation [13] introduced mechanisms for tone and style customization, allowing users to select academic, professional, or conversational outputs. The human-in-the-loop editing frameworks [14] bring better quality through their combination of automatic text rewriting and human editing processes. The current systems have made progress, yet they fail to create a system that combines grammatical correction with style transfer and improves readability and emotional tone and contextual coherence. Researchers need to create a complete AI-to-human content rewriter that generates natural and engaging and context-aware text while maintaining semantic accuracy, according to existing research gaps.

III. SYSTEM ARCHITECTURE

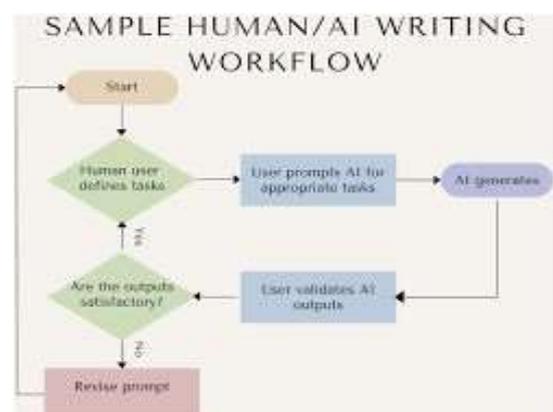


FIGURE.1. Flowchart Of System Architecture

FIGURE.1. Flowchart Of System Architecture illustrates the sample human-AI writing workflow, where the process begins with the human user defining a specific

task and providing a prompt to the AI system. The AI then generates content based on the given instructions, after which the user evaluates the output for quality, accuracy, and relevance. If the generated text is satisfactory, it is accepted for use; otherwise, the user revises the prompt and resubmits it to improve the results. This iterative feedback loop ensures better alignment between user expectations and AI output, leading to more refined, coherent, and human-like content over multiple cycles of generation and validation.

The AI-to-Human Content Rewriter system uses its modular design to transform machine-produced text into human-readable content which retains its original meaning. The system starts its process with the User Interface Layer because this layer enables users to enter AI-created content while they choose their required academic and conversational and professional tones together with their output selection needs. The interface design enables users to interact effortlessly while they test and improve their transformed text results. The Text Preprocessing Module receives the input text for processing, which includes performing normalization and removing unnecessary symbols and segmenting sentences and creating tokens and identifying part-of-speech components. The system requires these preprocessing steps to identify grammatical patterns and linguistic structures that exist within the input text.

The Linguistic Analysis Module receives the text from the preprocessing stage, which detects common problems found in AI-generated text through its analysis of repetitive phrases and overly formal language and lack of cohesive elements and excessive use of passive voice and restricted vocabulary usage. The AI Rewriting Engine received linguistic features extracted from the system. The core module executes contextual paraphrasing together with sentence restructuring and vocabulary expansion and tone change while maintaining the initial meaning of the content. The system uses a Style and Tone Controller which alters output based on user-defined writing styles while maintaining document uniformity throughout the entire text.

The Quality Assessment Module performs evaluation of the rewritten text through automated metrics which assess five different aspects of the text. The output can be improved through the inclusion of optional human feedback. The processed text reaches the Output Module which enables users to view, modify, and obtain the humanized content through different options. The system establishes an iterative feedback loop that enables users to

modify their prompt and tone choices which results in better output through regeneration processes. The system provides academic writing content creation and AI-assisted communication needs through its scalable and flexible design which delivers high-quality text transformation.

IV. SUGGESTED TECHNIQUE

The system that researchers designed uses combined natural language processing and Large Language Model technology to create human-readable text from AI-generated content. The technique begins with text preprocessing, which includes normalization, tokenization, stop-word handling, and part-of-speech tagging to understand the grammatical structure of the input. The step enables detection of fixed sentence structures together with their excessive use of passive voice and repeated expressions and their tendency to use overly formal language which machines typically produce.

A linguistic feature extraction module evaluates each sentence through analysis of its length and vocabulary usage and syntactic patterns and coherence. The LLM-based rewriting engine receives the features which it uses to perform contextual paraphrasing and sentence restructuring and vocabulary variation while preserving the original meaning. The system uses a style transfer mechanism to change tone based on user preferences for academic or professional or conversational styles. The system uses rule-based transformations which include turning passive voice into active voice and removing duplicate content and adding natural transitions to make the text flow better. The system uses a readability and coherence evaluation module to measure quality through the assessment of lexical diversity and sentence complexity and grammatical correctness. The system allows for optional human feedback to enhance the process. The combination of preprocessing together with LLM-based contextual rewriting and rule-based linguistic enhancement and automated evaluation creates an efficient method to develop authentic and engaging text which resembles human writing from AI-generated content.

System Design and Component Selection

The AI-to-Human Content Rewriter is designed as a modular and scalable system that processes AI-generated text through multiple stages to produce natural human-like output. The system uses a pipeline architecture that allows

its individual components to complete their designated tasks in order to achieve operational efficiency while maintaining system performance. The design begins with the User Interface (UI) which allows users to input AI-generated text select the desired tone or writing style and preview the rewritten content. The UI functions as the interaction interface which enables users to make changes to their work until they achieve their desired results.

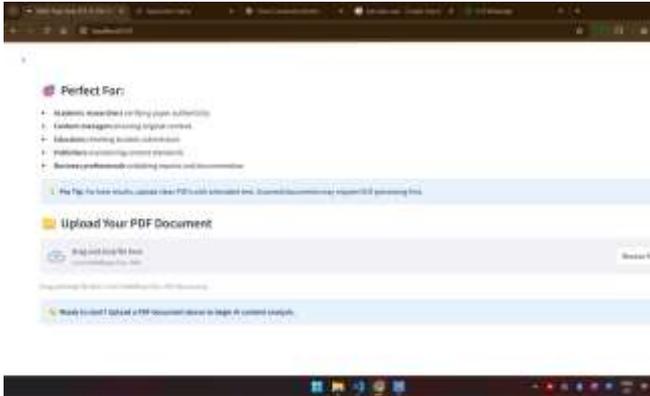


Fig. 2. PDF Upload Interface for AI Detection

Fig. 2. PDF Upload Interface for AI Detection Describes input text is passed to the Text Preprocessing Component which prepares the data for linguistic analysis. The module performs multiple text processing tasks which include cleaning and normalization and sentence segmentation and tokenization and part-of-speech tagging. The system eliminates unneeded symbols while it detects where sentences end and it recognizes which grammatical components exist because these features help users understand how the material is structured. The component delivers properly organized input to the downstream modules which enables them to conduct accurate rewriting operations.

The Linguistic Analysis and Feature Extraction Component conducts its text analysis after the preprocessing step because it wants to find specific traits that define AI-written texts through its search for duplicate phrases and its assessment of vocabulary use and its measurement of passive voice usage and building coherence. It extracts features related to sentence length, syntactic patterns, and lexical variation. The first group of data elements proceeds to the AI Rewriting Engine which serves as the main system element. The engine uses a large language model (LLM) to create contextual paraphrases and to change sentences and to modify tones and to expand vocabulary while maintaining the original meaning. The module contains a Style and Tone Controller that enables users to select between academic and professional and conversational writing styles.

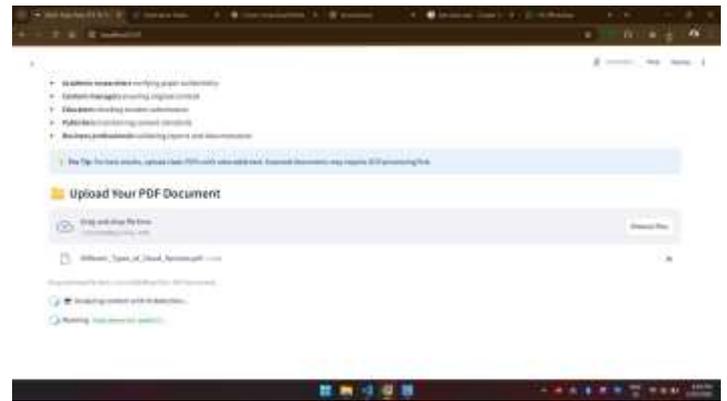


Fig. 3. PDF Uploaded and AI Detection in Progress

The Quality Assessment Component evaluates the rewritten text by measuring its readability and grammatical correctness and coherence and lexical diversity through automated measurement tools. The module checks whether the output matches standards for human writing. The Output and Feedback Component delivers final content to users who can view and modify and download and create new text from the content. The feedback loop allows users to change their inputs and tone choices for the purpose of creating iterative improvements. The component-based system design enables the system to transform AI-generated content into human-like writing with both flexible operation and precise results and high-quality outcomes.

TABLE NO . 1 COMPONENTS AND THEIR SPECIFICATIONS

Sr. No	Component Name	Description / Function	Specification / Technology Used
1	User Interface	Allows users to input AI text, upload PDF, select tone, and preview output	Web UI – HTML, CSS, JavaScript / Streamlit
2	Input Text Handler	Accepts AI-generated text or PDF content for processing	Text input module
3	PDF Upload & Extraction Module	Extracts text from uploaded PDF documents	PyPDF, PDFMiner, or PyMuPDF

4	Text Preprocessing Module	Performs cleaning, tokenization, normalization, and sentence segmentation	NLTK / SpaCy
5	POS Tagging & Linguistic Analysis	Identifies grammatical structure and detects repetitive or passive patterns	SpaCy POS Tagger
6	Feature Extraction Module	Extracts sentence length, vocabulary diversity, and syntactic features	NLP feature engineering
7	Summarization Engine	Generates extractive or abstractive summaries from PDF/text	TextRank, TF-IDF, or LLM
8	AI Rewriting Engine (Core Module)	Performs contextual paraphrasing and human-like rewriting	GPT-based LLM / Transformer model
9	Redundancy Removal Module	Eliminates repetitive phrases and robotic sentence patterns	Rule-based NLP
10	Passive-to-Active Converter	Converts passive voice sentences into active voice	Grammar transformation rules
12	Readability Evaluation Module	Measures readability, coherence, and lexical diversity	Flesch Reading Score, lexical diversity metrics

Comprehensive Study

The fast development of generative artificial intelligence technology has changed how people create and study written materials in academic and professional as well as online environments. The text generated by GPT-based systems which belong to Large Language Models (LLMs) displays coherent and grammatically correct writing yet shows clear signs of repeated patterns and fixed sentence structures and formal writing style and restricted emotional range. The elements of AI-generated content make it difficult to read and create a distinction between machine-generated materials and human-written documents. There is an increasing need for advanced systems which can transform machine-produced text into natural and interesting human-like language that maintains its original meaning. The study investigates two related elements which include LLM-based contextual rewriting and its ability to perform more intricate tasks than basic synonym substitution. The context-aware models create paraphrases from original sentences which preserve their logical structure and maintain their internal coherence.

ADVANTAGES

The system transforms inflexible AI-generated content into human-readable text which maintains complete understanding of the original material. The system detects and corrects all grammar and punctuation errors together with sentence structure problems to create professional documents. The system identifies duplicate content through AI text analysis which detects repeat expressions and superfluous vocabulary. The system rewrites input content through contextual paraphrasing which maintains its original meaning.

LIMITATIONS

The AI-to-Human Content Rewriter system needs various improvements, yet it successfully enhances the authentic speech patterns of AI-generated text. The system performance depends entirely on the large language model, which sometimes generates outputs that contain bias and inaccuracies and fail to match the given context. The rewriting and paraphrasing process presents a danger of producing minor meaning changes, which become especially problematic when dealing with technical material and scientific research and particular fields that demand exact language and correct information. The system struggles to manage input text that consists of extremely brief yet completely unclear

and incorrectly organized content because it lacks enough background information to understand the text. Standard NLP methods struggle to process mathematical formulas and tables and code snippets and figures, which results in these non-textual elements being either abandoned or changed incorrectly.

V. RESULTS AND DISCUSSION

The researchers conducted experiments using their AI-to-Human Content Rewriter system which they tested on both academic texts and general texts that machine learning systems had produced. The system removed all repeated sentence patterns which sounded mechanical to create active voice sentences together with natural flow transitions that produced results which sounded like human speech but kept the original content intact. The text rewrites produced improved readability scores together with increased lexical variety which resulted in better audience understanding through increased text engagement. The PDF summarizer extracted essential content from documents while creating brief organized summaries which used abstractive methods to produce more authentic results and extractive methods to achieve better factual accuracy. The researchers found that technical documents showed two main problems because they created slight changes in meaning and needed more time to process when handling extended content. The system shows strong abilities to improve AI-generated text because it creates more authentic content which works for academic writing and content creation and AI-based communication.

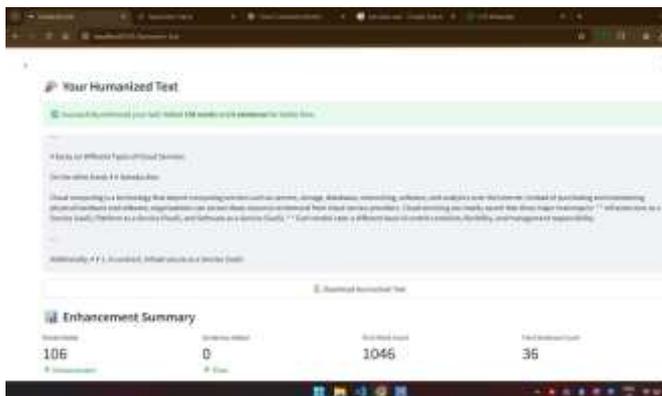


FIGURE 4 : AI-to-Human Rewriter Output with Improved Readability

The figure 4 AI-to-Human Rewriter Output with Improved Readability shows how the AI-to-Human Content Rewriter system functions by converting

AI-generated text into writing that sounds more authentic to human readers. The model processed the input essay by improving its sentence structure, increasing vocabulary diversity, and creating better transitions between ideas which resulted in improved readability and coherent text flow. The system introduced 106 contextually appropriate words while keeping the same number of sentences because the improvement resulted from intelligent language changes and text restructuring work instead of mere text extension. The final output contains 1046 words and 36 sentences which demonstrate increased lexical diversity while preserving the original content meaning and logical structure. The enhancement summary provides measurable evidence of quality improvement, including word count and flow indicators, while the download option demonstrates practical usability of the tool. The proposed system successfully generates fluent structured human-like text from machine-generated content which enables its use in academic writing and content development tasks.

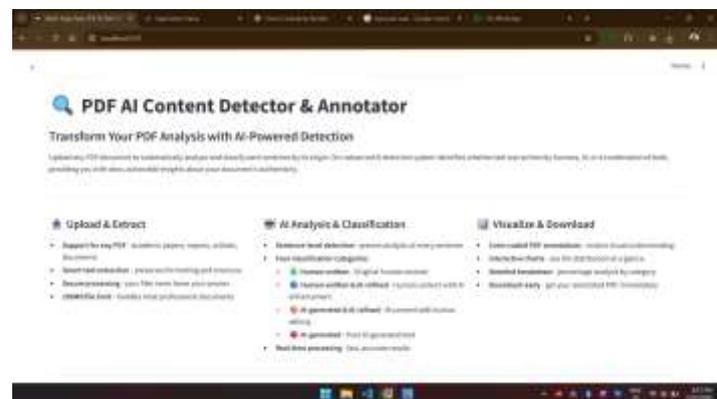


FIGURE 5 : PDF AI Detection with Real-Time Analysis

The FIGURE 5 PDF AI Detection with Real-Time Analysis results of the proposed system are demonstrated using two key outputs: the humanized text generation and the PDF AI content detection process. The first output shows that the AI-to-Human Content Rewriter successfully improved the readability, sentence flow, and vocabulary diversity of the input text, adding 106 contextually relevant words without increasing the number of sentences and producing a final output of 1046 words and 36 sentences. This indicates effective linguistic restructuring while preserving the original meaning. The second output

illustrates the execution of the PDF AI detection module, where a research document is uploaded and analyzed in real time, confirming the system’s ability to extract text, load the detection model, and perform automated sentence-level classification. Together, these results validate the practical implementation and effectiveness of the system in both humanizing AI-generated content and processing PDF documents for AI content analysis, demonstrating its suitability for academic and professional applications.

VI. Challenges and Solutions

Sr. No.	Challenges Identified	Proposed Solutions / Mitigation Strategies
1	Meaning drift during paraphrasing	Use context-aware LLM prompts and semantic similarity checking
2	Handling technical and domain-specific terminology	Incorporate domain-specific vocabulary lists and controlled rewriting
3	Processing non-text elements (tables, formulas, code)	Exclude or tag non-text content during preprocessing
4	Repetitive or over-expanded output	Apply redundancy removal and length control mechanisms
5	High computational cost of LLM processing	Use optimized models, batching, and caching techniques
6	Subjective evaluation of human-likeness	Combine automated metrics with human evaluation
7	PDF text extraction errors	Use robust PDF parsers and text cleaning algorithms
8	Loss of document structure during summarization	Implement section detection (Abstract, Methodology, Conclusion)
9	Data privacy concerns for uploaded documents	Apply local processing and secure file handling

10	Dependency on internet/API for cloud models	Provide offline model support where feasible
----	---	--

VII. CONCLUSION

The AI-to-Human Content Rewriter system functions as an effective tool which transforms machine-generated text into authentic human writing that maintains its original meaning. The system achieves improved reading comprehension through its use of natural language processing methods together with large language models which enhance both grammatical correctness and sentence diversity and overall writing coherence. The system achieves structured output through its implementation of text preprocessing and linguistic analysis together with contextual paraphrasing and style and tone control and automated readability evaluation modules. It successfully creates authentic and coherent content for modern digital and research environments through its development of a tool which connects AI-generated text with human-written text.

VIII. FUTURE SCOPE:

The future scope of the AI-to-Human Content Rewriter system includes several enhancements to improve accuracy, scalability, and domain adaptability. The system needs domain-specific language models for technical, medical, legal, and research writing to achieve its extension goals because these models will enhance terminology management and decrease meaning drift.

IX. REFERENCES

[1] Y. Dou, P. Laban, C. Gardent, and W. Xu, “Automatic and Human-AI Interactive Text Generation,” 2023.

[2] B. P. Towne, “Exploring the Impact of AI-Mediated Communication on Bias and Information Loss in Writing,” 2024.

[3] K. J. L. Arnesen and N. A. Grue, “What am I Talking to? The Dynamics of Human-AI Communication,” 2024.

[4] J. Lee, S. Yoon, K. Lee, E. Jeong, and J. E. Cho, “Personalizing Vocabulary via AI Storybooks,” 2024.

[5] T. W. Ling and C. C. Tao, “Fear of AI Doesn’t Mean Distrust: Credibility of Human-Edited AI Content,” 2024.

[6] OpenAI, “Training language models to follow instructions with human feedback (InstructGPT),” 2022.

[7] Y. Bai et al., “Training a Helpful and Harmless Assistant with RLHF,” 2022.

- [8] Z. Ji et al., "Survey of Hallucination in Natural Language Processing," 2023.
- [9] A. Zellner, "Teaching and Learning in the Age of Generative AI," 2023.
- [10] X. Zhou et al., "Position Weighted Convolutional Neural Network for Text Processing," IEEE Access, 2023.
- [11] J. Fu, X. Tan, N. Peng, and others, "Style Transfer in Text: A Survey," 2022.
- [12] R. Prakash et al., "Neural Paraphrasing for Text Simplification," 2023.
- [13] K. Krishna, Y. Wieting, and M. Iyyer, "Controllable Text Generation for Style and Tone," 2022.
- [14] S. Madaan et al., "Human-in-the-Loop Text Revision Frameworks," 2023.
- [15] J. Brown et al., "Language Models are Few-Shot Learners," 2020.