

# Ethical Fake-Information Rewriter

(A System for Detecting and Ethically Rewriting Misleading Content)

Thanuvidhya C.<sup>1</sup>; Sabipriya A.<sup>2</sup>; Vikram R.<sup>3</sup>

<sup>1,2,3</sup>Department of Computer Science and Engineering Adhiyamaan College of Engineering  
(An Autonomous Institution) Hosur, India

Publication Date: 2026/04/09

**Abstract:** With increasing use of online platforms, the spread of false information is an major concern. Existing system mainly focus on detecting misinformation but do not provide users with corrected or reliable alternatives. This paper proposes an Ethical Fake Information Rewriter that not only detects misleading content but also transforms it into accurate and responsible information. The system supports multiple input formats, including text, website links and images. Natural Language Processing techniques are used to analyse textual content, while Optical Character Recognition is applied to extract text from images. After identifying misinformation, the system rewrites the content by removing bias, harmful expressions and misleading statements while preserving the original meaning. The system is implemented using a Flask-based backend with an interactive web interface, integrating multiple application programming interfaces to enhance accuracy and reliability. This approach enhances user understanding and promotes responsible use of digital information.

**Keywords:** Ethical AI; Natural Language Processing; Optical Character Recognition; Content Rewriting; Misinformation.

**How to Cite:** Thanuvidhya C.; Sabipriya A.; Vikram R. (2026) Ethical Fake-Information Rewriter. *International Journal of Innovative Science and Research Technology*, 11(3), 3695-3697. <https://doi.org/10.38124/ijisrt/26mar1778>

## I. INTRODUCTION

The expansion of digital media platforms has led to a noticeable rise in the spread of misleading information. Such misinformation can influence public opinion, create confusion and affect decision-making, making it important to develop reliable detection systems<sup>[4]</sup>. Current methods rely on various AI-based models to detect false information. However, most systems only classify information as real or fake and do not provide corrected or trustworthy alternatives. In addition, issues such as bias, lack of transparency and limited support for multimedia content remain major challenges. To overcome these limitations, this paper proposes an Ethical Fake Information Rewriter that not only detects misinformation but also rewrites it into accurate and responsible content. The system integrates text analysis and image processing techniques to ensure effective and ethical handling of digital information.

## II. SYSTEM OVERVIEW

### ➤ Existing System

Current fake news tools focus on marking or flagging misleading information by using computer programs that learn from examples<sup>[3]</sup>. These systems check if information seems true or false based on how the words are used. But they do not offer users any corrections or helpful changes to the content.

### ➤ Proposed System

The proposed Ethical Fake Information Rewriter is designed to overcome these limitations by combining detection and content correction. The system accepts input in the form of text, URLs and images. Natural Language Processing is used for analysing textual content, while Optical Character Recognition is used for extracting text from images<sup>[5]</sup>. After detecting misinformation, the system rewrites the content into accurate, neutral and responsible information while maintaining ethical standards.

## III. METHODOLOGY

The proposed system follows a structured approach to detect and rewrite fake or misleading information in an ethical and user-friendly manner. The methodology includes multiple stages, including input processing, detection, analysis and rewriting to ensure accurate and responsible output.

### ➤ Input Processing

The system accepts user input in different formats such as text, URLs and images. The input data is processed to remove unnecessary symbols, normalize content and convert it into a suitable format for further analysis. This step ensures consistency and improves detection accuracy.

➤ *Detection Mechanism*

The processed input is analysed using Natural Language Processing techniques to identify fake or misleading information. The system examines linguistic patterns, contextual meaning and key indicators to determine the authenticity of the content. For image inputs, Optical Character Recognition is used to extract textual information before performing analysis. It then categorizes the content based on authenticity using predefined models.

➤ *Ethical Analysis*

After detection, the system performs an ethical evaluation to identify harmful, biased or sensitive content. This stage identifies biased, harmful, or misleading information in the content. The system ensures that ethical guidelines are followed by filtering unsafe content and maintaining neutrality. The system outputs are responsible, unbiased and suitable for all users.

➤ *Content Rewriting*

Once misinformation is detected, the system rewrites the content into accurate, neutral and socially responsible information. The rewriting process removes false claims, reduces bias and restructures the content into a clear and meaningful format. It preserves the original intent while ensuring correctness and readability. This approach helps users better understand the information and prevents the spread of misinformation.

➤ *Output Generation*

The final output is presented to the user through a simple and user-friendly interface. The system displays the rewritten content along with clear and understandable results. The output is structured in a way that enables users to easily interpret the corrected information and make informed decisions. This enhances user trust and promotes responsible information usage.

#### IV. SYSTEM ARCHITECTURE

The Ethical Fake Information Rewriter is designed using a modular architecture that integrates multiple components to ensure efficient processing, accurate detection and ethical rewriting of misinformation. The system is structured in a way that different modules work together in a coordinated manner, allowing smooth data flow from input to output. It combines frontend, backend and database layers to handle user interaction, data processing and storage effectively. Each module performs a specific function such as input handling, preprocessing, detection, ethical filtering and content rewriting. This modular design improves system scalability, flexibility and maintainability.

➤ *Input Module*

The Input Module acts as the entry point of the system, where users can provide input in different formats such as text, website URLs and images. This module ensures flexibility in handling diverse types of digital content commonly found on online platforms. It captures user data through a web interface and forwards it to the back-end for further processing. The module is designed to be simple, accessible and user-friendly.

➤ *Preprocessing Module*

The Preprocessing Module is responsible for cleaning and preparing the input data before analysis. It removes unnecessary symbols, normalises text and converts all inputs into a consistent and structured format. This process improves the quality of data and reduces noise, leading to more accurate detection results. For URL inputs, relevant content is extracted using web scraping techniques by removing unwanted HTML elements and retaining meaningful text.

➤ *Detection Module*

The Detection Module analyses the processed content using Natural Language Processing techniques. It identifies linguistic patterns, keywords and contextual features to classify the content as real or fake. This module serves as the core decision making component of the system.

➤ *Ethical Filtering Module*

The Ethical Filtering Module evaluates the detected content to identify harmful, biased or sensitive information. It focuses on detecting offensive language, misleading claims and content that may negatively impact users. This module ensures that ethical AI principles such as fairness, transparency and safety are maintained.

➤ *Rewriting Module*

The Rewriting Module transforms fake or misleading content into accurate, neutral and socially responsible information. It removes false claims and biased language while preserving the original meaning. This module ensures that users receive clear and trustworthy content.

➤ *Database Module*

The Database Module stores user inputs, processed data and rewritten outputs in an organized manner. It maintains records of system activity, which can be used for future reference, analysis and system improvement. The stored data also helps in tracking performance and identifying patterns in misinformation, enabling better optimization of the system over time.

➤ *Output Module*

The Output Module presents the final rewritten content to the user through a user-friendly interface. The output is displayed in a clear and understandable format, enabling users to easily interpret the corrected information

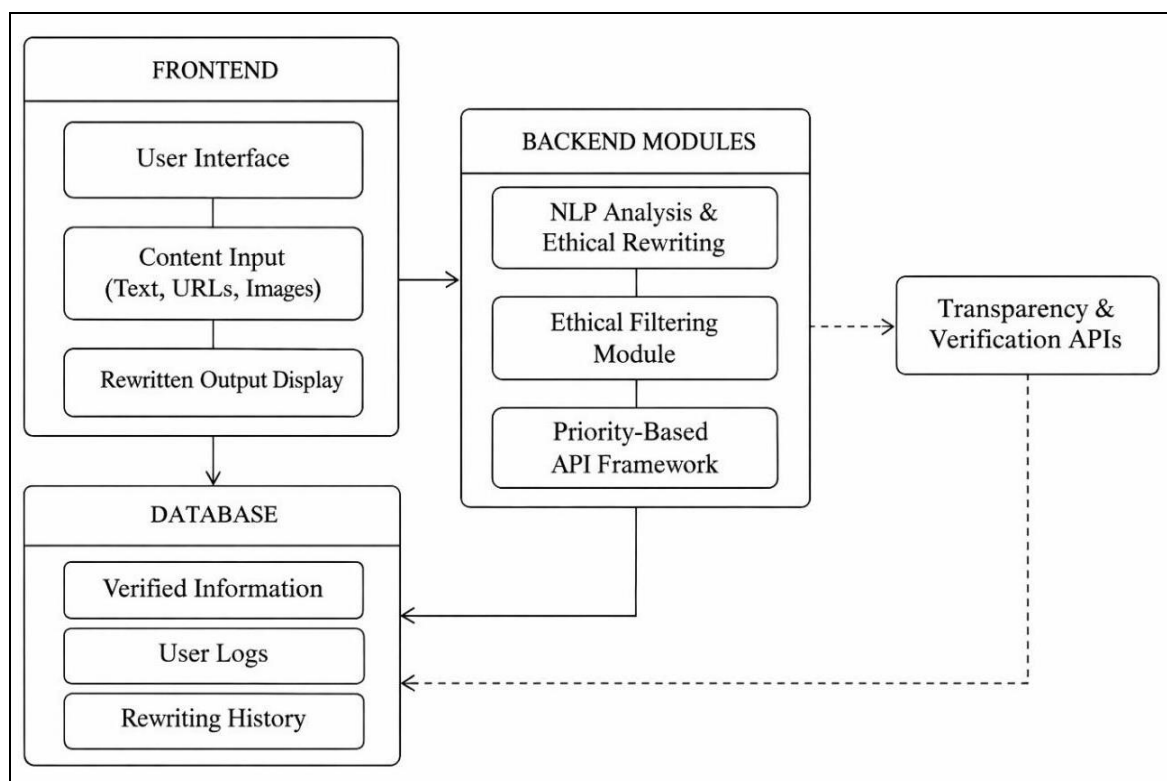


Fig 1 Proposed System Architecture

## V. CONCLUSION

This work presents a system that identifies misleading information and rewrites it into clear and accurate content. By using basic language processing and text extraction techniques, the system can handle different input types and improve the quality of information. It not only detects false content but also makes it more understandable and responsible.

## ACKNOWLEDGMENT

We would like to thank Mr. R. Vikram for his guidance and continuous support throughout this project. The authors also thank Dr. G. Fathima, Head of the Department, and the management of Adhiyamaan College of Engineering for providing the required resources and encouragement to successfully complete this work.

## REFERENCES

- [1]. Allein, L., Moens, M. F., and Perrotta, D., "Preventing profiling for ethical fake news detection," *Journal of Artificial Intelligence Ethics*, vol. 2, no. 1, pp. 45–59, 2022.
- [2]. Baly, R., Karadzhev, G., Alexandrov, D., Glass, J., and Nakov, P., "Predicting factuality of reporting and bias of news media sources," *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 3528–3539, 2018.
- [3]. Bhogade, M., Deore, B., Sharma, A., Sonawane, O., and Singh, M., "A research paper on fake news detection," *International Journal of Computer Applications*, vol. 12, no. 7, pp. 589–599, 2025.
- [4]. Castillo, C., Mendoza, M., and Poblete, B., "Information credibility on Twitter," *Proceedings of the 20th International World Wide Web Conference (WWW)*, pp. 675–684, 2011.
- [5]. Chen, Y., Conroy, N. J., and Rubin, V. L., "Misleading online content: Recognizing clickbait as false news," *Proceedings of the ACM Workshop on Multimodal Deception Detection*, pp. 15–19, 2015.
- [6]. Montoneri, B., "Plagiarism and ethical issues: A literature review on academic misconduct," *International Journal of Educational Integrity*, vol. 17, no. 2, pp. 75–88, 2021.
- [7]. Niu, C., Guan, Y., Wu, Y., and Zhu, J., "Retrieval augmented ethical AI for fake information rewriting," *NewsBreak Research Journal*, vol. 3, pp. 65–78, 2024.
- [8]. Niu, C., Guan, Y., Wu, Y., and Zhu, J., "VeraCT scan: Retrieval augmented fake news detection with justifiable reasoning," *NewsBreak Research Journal*, vol. 3, pp. 50–64, 2024.
- [9]. Plikynas, D., Rizgeliënė, I., and Korvel, G., "Ethical implications in fake news detection systems," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 4, pp. 150–164, 2024.
- [10]. Plikynas, D., Rizgeliënė, I., and Korvel, G., "Systematic review of fake news, propaganda, and disinformation," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 3, pp. 120–134, 2024.