

# ProxyLLM : LLM-Driven Framework for Customer Support Through Text-Style Transfer

Sehyeong Jo<sup>1</sup> and Jungwon Seo<sup>2</sup> ✉

<sup>1</sup> University of Colorado Boulder, Boulder, CO, USA [se.jo@colorado.edu](mailto:se.jo@colorado.edu)

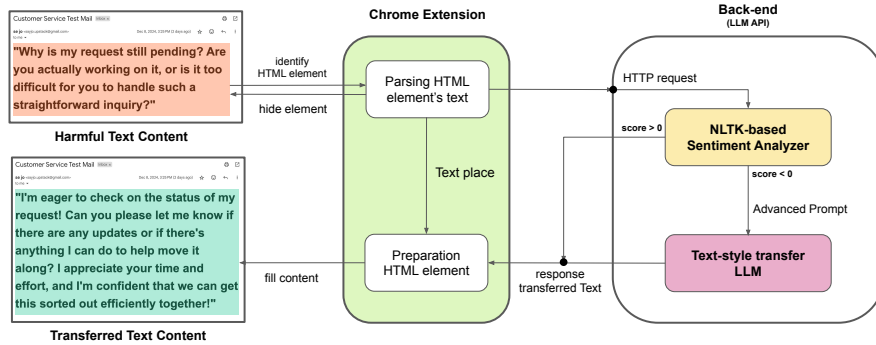
<sup>2</sup> University of Stavanger, Stavanger, Norway [jungwon.seo@uis.no](mailto:jungwon.seo@uis.no)

**Abstract.** LLM-based chatbots have improved response quality and reduced costs in customer support, while the experiences of human agents, essential to the service ecosystem, have remained largely overlooked. Stress from harmful texts poses a major challenge for agents, undermining their efficiency, customer satisfaction, and business outcomes. In this work, we propose an LLM-powered system designed to enhance the working conditions of customer service agents by addressing emotionally intensive communications. Our proposed system leverages LLMs to transform the tone of customer messages, preserving actionable content while mitigating the emotional impact on human agents. Furthermore, the application is implemented as a Chrome extension, making it highly adaptable and easy to integrate into existing systems. Our method aims to enhance the overall service experience for businesses, customers, and agents. The code related to this paper is available at: <https://github.com/sehyeongjo/Proxy-LLM>

**Keywords:** Large Language Models · Human-centered Interfaces · Sentiment Analysis.

## 1 Introduction

Advancements in large language models (LLMs) have driven the widespread adoption of chatbot-based customer support services across industries, enabling companies to reduce operational costs and enhance response accuracy and contextual relevance. While these improvements have contributed to higher customer satisfaction, human agents remain indispensable for managing complex interactions. Agents are required to address the actionable aspects of customer interactions while also withstanding emotionally charged or negative responses, all within the constraints of strict organizational protocols. This ongoing burden underscores the need to reevaluate the current customer interaction framework. Therefore, we propose ProxyLLM, a system that functions similarly to a proxy server, positioned between human agents and existing systems to prevent human agents from being directly exposed to raw harmful messages. Specifically, we propose using LLMs to adjust the tone of incoming messages using text-style transfer [3] while preserving their core content. This adjustment aims to shield agents from unnecessary emotional distress, enabling them to work more



**Fig. 1.** Overview of the ProxyLLM Application System.

efficiently and with greater composure. Furthermore, to ensure ease of use and seamless integration, ProxyLLM is implemented as a Chrome extension. This implementation enables effortless integration into diverse web-based customer service systems without necessitating modifications to existing system pipelines or the deployment of additional components, such as databases.

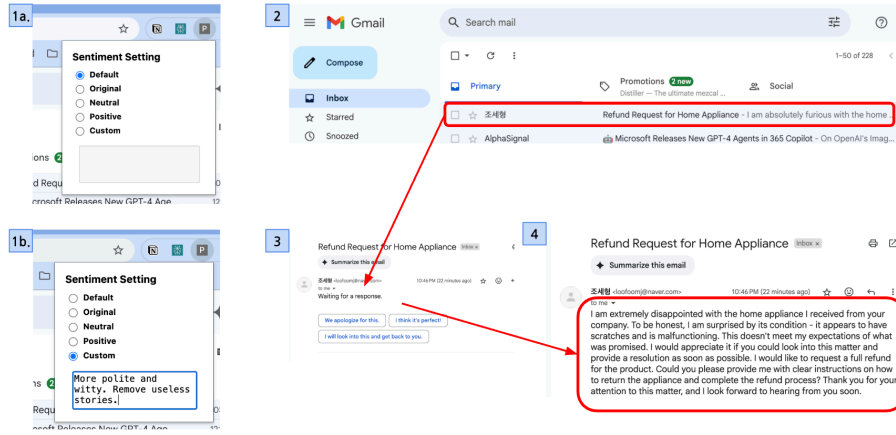
**Contributions** We propose a novel application that leverages large language models (LLMs) to enhance mental health support for human agents in high-stress environments. By utilizing the style-transfer capabilities of LLMs, the system preserves the critical content of communications while filtering harmful or distressing material before it reaches agents. Implemented as a lightweight Chrome extension, it integrates seamlessly into existing workflows without causing disruptions.

## 2 ProxyLLM

ProxyLLM is a proposed system that consists of two primary components: (1) a Chrome extension and (2) a back-end machine learning model server, as depicted in Figure 1. An introductory video can be viewed at <https://youtu.be/MTCMmlmOTpQ>

### 2.1 Chrome Extension

A Chrome extension is a third-party application that enhances browser functionality through packaged HTML, JavaScript, and CSS files, operating under user-granted permissions. The ProxyLLM extension automatically identifies and obscures HTML elements containing harmful text, shielding human agents from



**Fig. 2.** ProxyLLM UI components and Chrome Extension workflow example. (1a) Users can select a predefined sentiment preset, and (1b) optionally apply a custom prompt. (2) After completing the setup, selecting email content triggers (3) the original text to be hidden and a waiting message to be displayed. Subsequently, (4) the content is replaced with the prompt-applied output.

exposure, and transmits the content to a back-end API for style transformation. It also provides a graphical user interface (GUI) for customizing prompts, enabling context-sensitive style-transfer responses.

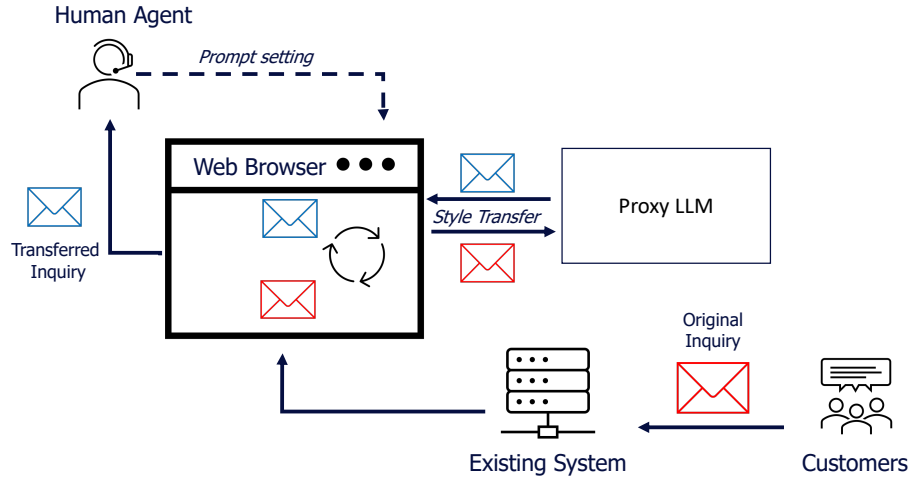
## 2.2 Back-End Processing

The back-end component processes incoming requests by first applying sentiment analysis using the Natural Language Toolkit (NLTK) [1] and then refining the sentiment scores through the LLM’s self-feedback mechanism, ultimately producing scores ranging from -1 to 1. The lightweight sentiment analyzer reduces computational overhead by bypassing the LLM for scores outside predefined thresholds. Requests requiring style transfer are processed through a text-style transfer module using simple prompting, and the transformed text is seamlessly injected into the agent’s interface via the Chrome extension. The overall workflow is illustrated in Figure 3.

## 3 Text-style transfer and Implementation

ProxyLLM achieves text-style transfer by leveraging fine-tuned LLMs through prompt engineering, appending instructions to rephrase original text with greater politeness while minimizing training costs and supporting customization. The basic prompt structure is :

*This is original text. Change this text style. [Original Text]. Change this text content to be more polite.*



**Fig. 3.** ProxyLLM Workflow Diagram

While the LLM determines sentiment based on its inherent standards, basic prompting leaves the criteria fully dependent on the model without personalization. ProxyLLM addresses this by embedding personalized configuration parameters into the prompt to tailor outputs for human agents. The advanced prompt structure is:

*This is original text. Change this text style. [Original Text]. Change this text content to be more [Custom Parameter].*

Text modification is guided by personalized configuration settings. Predefined presets trigger the use of corresponding predesigned prompts, whereas custom configurations employ dynamically constructed prompts to realize the specified transformation. An example of a prompt corresponding to a custom parameter is as follows:

- **Neutral :** *a neutral tone to remove any emotional, biased, or subjective elements while preserving the original meaning.*
- **Positive :** *a positive tone, enhancing the optimism and uplifting language while preserving the original meaning.*

**Implementation Details** The application’s back-end, built with Flask, enables efficient communication with the Chrome extension through a lightweight API. Text-style transfer and sentiment-aware transformations are handled by a Llama 3.1 8B model [2], fine-tuned on sentiment data to optimize performance under limited GPU resource(Test Server : RTX 2060 GPU).

### 3.1 Evaluation

To evaluate ProxyLLM’s style transfer, we employed commercial LLMs (GPT-4o-mini, Claude Sonnet, Gemini) as proxy evaluators. Each was given 10 pairs of negative inquiries and their corresponding positive versions. The structure of the prompt for requesting the score is as follows:

*Please mark the sentimental score I present from -1.0 (very negative) to 1.0 (very positive). [Original Text].*

**Table 1.** The average sentiment score variations original and transferred applying LLMs to sample texts.

LLMs	Original	Transferred
GPT-4o-mini	-0.48	0.28
Claude Sonnet	-0.66	0.24
Gemini	-0.6	0.18
NLTK	-0.48	0.83

Table 1 shows that our fine-tuned LLM consistently achieved positive tone conversion across all LLMs using a simple prompt. The NLTK-based sentiment analysis also aligned closely with LLM scores, confirming its efficiency and reliability for detecting negative tone.

## 4 Conclusion

In this study, we introduced ProxyLLM, a Chrome extension designed to enhance customer support by leveraging LLM capabilities to improve service quality while safeguarding the mental well-being of human agents. By transforming harmful or emotionally charged customer messages into positive tones, ProxyLLM fosters a more constructive interaction environment without requiring additional infrastructure changes. Its modular and adaptable design also positions it for broader applications across various web services and platforms. Future work could enhance ProxyLLM by advancing personalization, model efficiency, and multilingual capabilities to better support human-centric service delivery.

## References

1. Bird, S.: Nltk: the natural language toolkit. In: Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions. pp. 69–72 (2006)
2. Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., et al.: The llama 3 herd of models. arXiv preprint arXiv:2407.21783 (2024)
3. Toshevsk, M., Gievska, S.: A review of text style transfer using deep learning. IEEE Transactions on Artificial Intelligence **3**(5), 669–684 (2021)