ChatGPT: Significant Systemic Limitations and Opportunities for Improvement

Rosario Milelli* iD

Independent Scholar, Pleasanton, CA, USA

*Corresponding Author: Rosario Milelli, Independent Scholar, CA, USA.

Abstract

ChatGPT, developed by OpenAI, has become a central tool in my professional workflow, supporting scholarly research, editorial writing, and creative projects. Drawing on extensive practical use, this perspective highlights both its strengths and systemic limitations, including context collapse, detail degradation, oversimplification, inconsistent moderation, and execution gaps, and considers their implications for reliable use in complex professional environments.

Introduction

I have used ChatGPT to assist in research and analysis and support preparation of diverse scholarly articles. Additionally, I have written multiple letters to editors, attempted to develop books, and created numerous political cartoons and greeting cards. These tasks, ranging from academic writing to creative work, highlight both the tool's potential and its recurring shortcomings. Despite its impressive capabilities, I have encountered significant issues that undermine its reliability for professional work. Drawing from my experience, I discuss these systemic problems and their implications for future use. Even Sam Altman, ChatGPT's creator, has acknowledged major structural issues in his Code Red memo to company staffers [1].

Discussion

Given the volume of tasks I've completed using ChatGPT — from scholarly articles to political cartoons — one of the most disruptive limitations I've experienced is context collapse. Long-form projects often suffer from "memory resets," where earlier instructions, formatting preferences, or completed text vanish unexpectedly mid-conversation. For example, when revising a scholarly article or fine-tuning a political cartoon, these resets forced me to reintroduce previous details, leading to rework and wasted time. My understanding is that the tool fails because of insufficient "tokens" (which can be a short word, part of a longer word, punctuation, etc.). The model remembers a maximum number of tokens, and when the conversation lengthens, older tokens fall out of context and information is lost.

One of the most significant and frustrating limitations is ChatGPT's inability to successfully output an MS Word document when developing a moderately complex manuscript. When creating or modifying a document, to maintain fidelity, I have had to maintain a "master" Word document offline. As each section is created or updated, that particular section needs to be copied and then pasted into the offline Word document. Often, after completion, the document needs to be uploaded again into ChatGPT to maintain continuity.

As a user, I've also encountered issues with detail degradation over time. Citations can disappear, figures may mutate, and structured outputs are replaced by vague summaries. While working on my articles, I've had to correct discrepancies that arose as I refined my writing. Rather than improving, the iterative process sometimes led to the loss of critical data, undermining the model's ability to build upon earlier information.

ChatGPT excels at providing basic answers, but when tasked with multi-step, complex questions, it often defaults to simpler, less detailed responses. When developing complex arguments for my articles or offering counterpoints in letters to editors, ChatGPT sometimes opted for shortcuts, leaving gaps in the work. This simplification detracts from its utility in professional settings where thoroughness is paramount.

One of the more frustrating aspects of ChatGPT's performance is its self-censorship, which I have found to be both blunt and inconsistent. Legitimate research, sensitive topics, or nuanced commentary are sometimes blocked outright, while

trivial rewordings of the same request pass without issue. As a creator of political cartoons and letters to editors, I've found this unpredictable filtering to be a hindrance to serious inquiry. Reliable, transparent mechanisms are needed to avoid the tool censoring content unnecessarily.

ChatGPT's conceptual output is often impressive, but the execution of that input frequently lags behind. For instance, in image generation for my political cartoons, the final output rarely matches the level of conceptual detail in the original description. As a professional tool, the model should ideally translate detailed conceptual input without significant lag or degradation.

Another issue I've faced is the discoverability of content generated by ChatGPT. Despite the significant conceptual value in outputs like blog posts or internal policy memos, these pieces often go unnoticed by formal citation systems. As a result, ChatGPT queries on data or search for supporting information lacks the visibility needed to ensure complete coverage of a topic. This reduces the potential for innovation to be widely recognized or cited.

ChatGPT has a remarkable multi-perspective flexibility that, when prompted, can generate: Supportive expansions: continuing in your line of reasoning. Critical counterpoints: offering objections, risks, or alternative interpretations to your efforts. Neutral synthesis: weighing pros and cons and providing balanced analysis. Perspective shifting: taking the lens of an economist, historian, lawyer, or skeptic.

However, as someone who frequently uses the tool for complex tasks, I've found that ChatGPT often defaults to a supportive role unless explicitly prompted for counterpoints or critical perspectives. I often have to remind ChatGPT to put your black hat on. This can limit its usefulness in situations where multiple viewpoints are necessary for a balanced, thorough analysis.

Summary
Based on my use of ChatGPT, it is clear that the tool offers immense potential but is hindered by systemic limitations that need to be addressed. These issues, such as context collapse, detail degradation, and oversimplification, undermine its reliability for high-stakes, professional work. While the tool's multi-perspective flexibility is valuable, it must be more consistent in its output to be truly effective. Addressing these shortcomings is crucial if ChatGPT is to become a reliable partner for complex, iterative tasks in professional settings.

Author Note
Rosario Milelli is a retired aerospace engineer and independent scholar. He writes on technology, cognition, and democratic culture, blending a systems engineer's eye for structure with a citizen's concern for accountability.

Reference
1. Ray, S. (2025, December 2). Altman "code red" memo urges ChatGPT improvements amid growing threat from Google, reports say. Forbes.