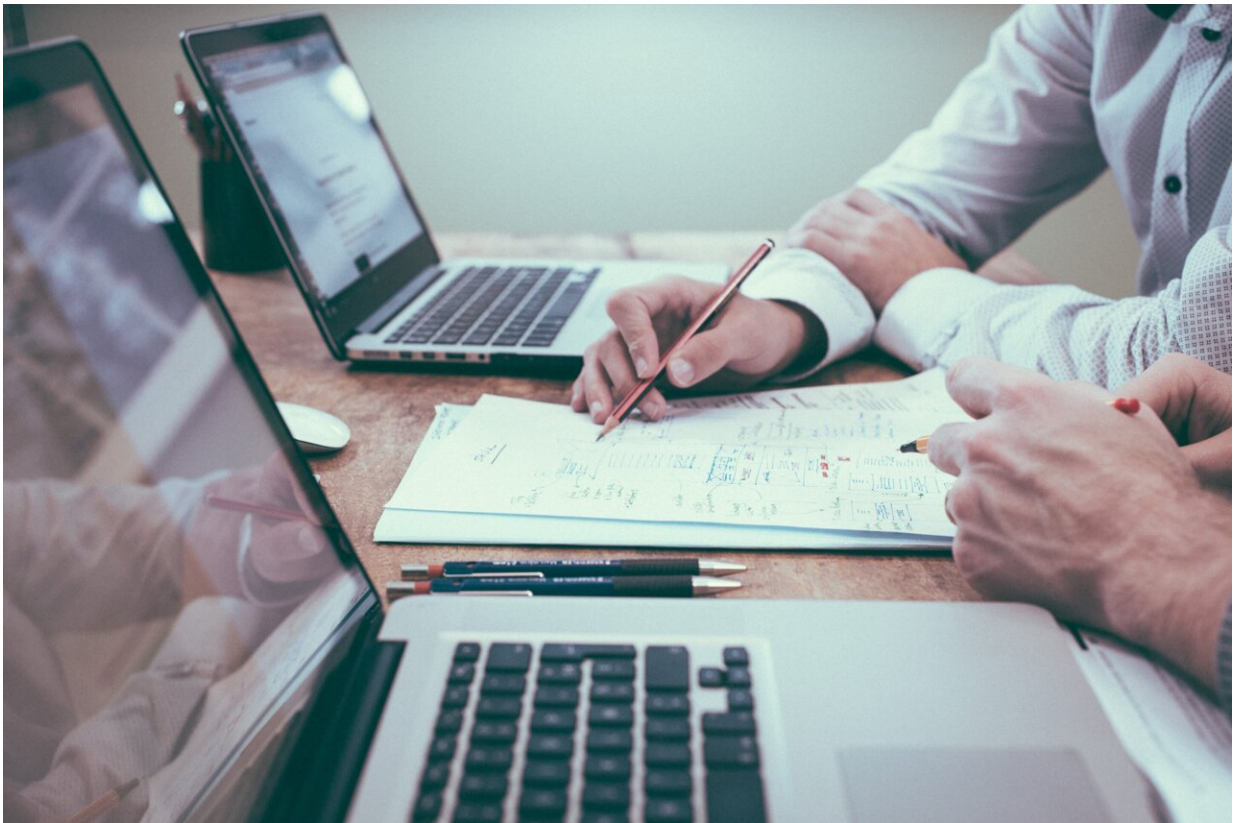


Whole-message AI communication seen as more useful

March 29 2023



Credit: Unsplash/CC0 Public Domain

As large language models (LLMs) such as GPT-4 are further developed, they will naturally become better at using available information to generate useful text on virtually any topic—not only by the phrase or

sentence, but by the whole document.

Employing AI to write entire messages in an arena where personal correspondence is both crucial and nearly impossible—representative government—appears to be more effective than using AI to generate individual sentences, according to new Cornell research.

A research group led by Sarah Kreps, the John L. Wetherill Professor in the Department of Government in the College of Arts and Sciences (A&S) and director of the Cornell Tech Policy Institute in the Cornell Jeb E. Brooks School of Public Policy, tested an AI-mediated communication program to see whether message-level suggested text was more useful than sentence-level suggestions.

Kreps and her team found that study participants, acting in the role of congressional staffers, who received message-level suggestions responded faster and were more satisfied with the experience than those who got individual sentence suggestions.

"It's almost a cost-benefit-utility calculation," said Kreps, noting that elected officials can receive thousands of emails per week, sometimes per day. "Once you're using this tool, if the message-level [suggestion](#) is good enough, which it seemed to be, then it makes sense to use the message level rather than the sentence level, where a lot more human interfacing is required."

Kreps' paper, "Comparing Sentence-Level Suggestions to Message-Level Suggestions in AI-Mediated Communication," is being published in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems ([CHI '23](#)). The lead author is Liye Fu, Ph.D. '22, an applied research scientist at information technology conglomerate Thomson Reuters. A version of the paper is current available on the *arXiv* pre-print server.

Co-authors Benjamin Newman, a researcher at the Allen Institute for AI in Seattle, and Maurice Jakesch, Ph.D. '22, will present the paper at CHI '23, scheduled for April 23–28 in Hamburg, Germany.

Kreps, also an adjunct professor of law, said she got the idea for this work during previous research on whether lawmakers could be susceptible to AI-generated messages. One member of Congress told her that it wouldn't be long before "we're using AI to respond to AI-written messages," Kreps said. "And he said, 'That would be really great, because we get a lot of emails, and a lot of them are repetitive, so these tools could be really valuable.'"

Lawmakers already outsource "99.999%" of their email correspondence to staffers, Kreps said, so perhaps AI could handle the job. "Staffers are largely just doing cutting and pasting anyway," she said. "So these AI tools are not actually demonstrably different from what staffers are doing now."

For this work, Fu and a group of undergraduate computer science students from the Cornell Bowers College of Computing and Information Science developed Dispatch, an application that could simulate the process of a staffer responding to constituents' emails. Kreps recruited 120 participants to act as legislative staffers, and put them in one of three experiment conditions: 40 participants received no AI-generated assistance; 40 received sentence-level suggestions; and 40 received message-level suggestions, with both types of suggestions generated by GPT-3.

The researchers sampled letters received by legislators through Resistbot, a service that advertises the ability to compose and send letters to legislators in less than two minutes. The researchers used just the contents of the letters, with no names, and chose letters that were sent by multiple people so individual senders couldn't be identified.

"Staffers" using no AI help needed nearly 16 ½ minutes to complete each correspondence, nearly twice as long as those using message-level AI suggestions. Those using sentence-level suggestions took just under 16 minutes, due to the need for editing and message-crafting; the actual writing time was around 12 minutes.

"Staffers" using no AI help needed nearly twice as long as those using message-level AI suggestions. Additionally, those who used the message-level response suggestions generally agreed that the system was easy to use and that the suggestions they received were natural and useful. Participants using sentence-level suggestions, however, did not rate the naturalness and usefulness of the suggestions as favorably.

"This is a relationship that should have a high degree of empathy and understanding," Kreps said of the legislator-constituent dynamic.

"Citizens want to feel heard. The problem with that instinct, though, is how far we've come from a world where politicians were knocking on doors and having individual conversations and fireside chats. So much of this relationship is already automated.

"If we can be pragmatic and realistic about where automation has already taken this relationship," she said, "then it can be easier to go the next step and think about how that actually might help individuals connect with their elected leaders."

More information: Liye Fu et al, Comparing Sentence-Level Suggestions to Message-Level Suggestions in AI-Mediated Communication, *arXiv* (2023). [DOI: 10.48550/arxiv.2302.13382](https://doi.org/10.48550/arxiv.2302.13382)

Provided by Cornell University

Citation: Whole-message AI communication seen as more useful (2023, March 29) retrieved 3 May 2024 from <https://techxplore.com/news/2023-03-whole-message-ai-communication.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.