

TECHNOLOGY

## Computers as Creative Collaborators for Businesses?

by Rohit Nishant, Deepti Singh, Vivek Kumar Singh, and Robert D. Austin



Image Credit | Simon Abrams

*Large Language Models (LLMs) can generate useful content, but creative collaboration requires more than that.*

✔ INSIGHT | FRONTIER 11 Sep 2023

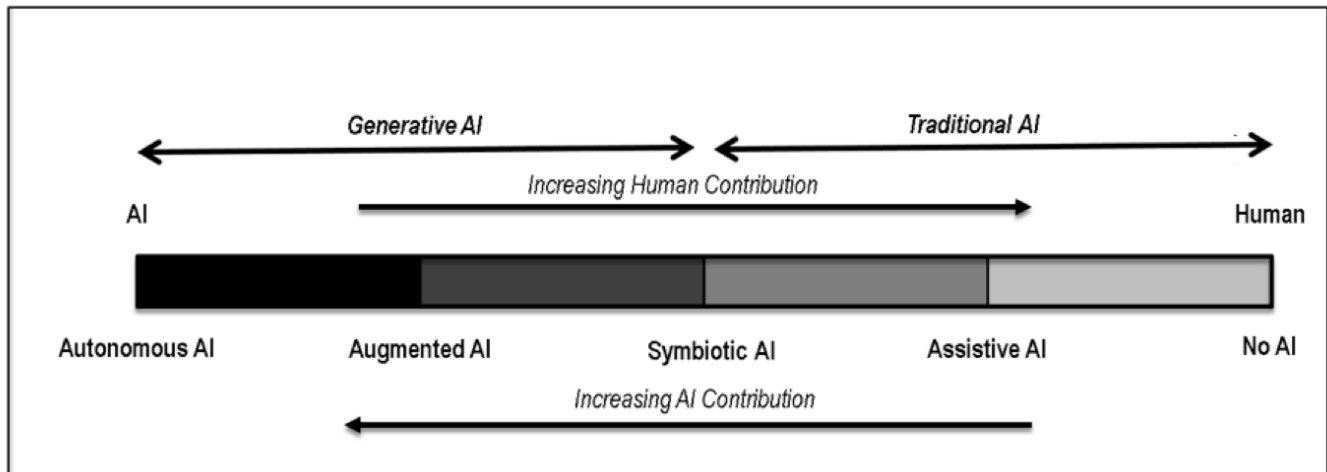
---

If you've been following the news about generative AI, you won't be surprised that people are starting to work with large language models (LLMs) in ways that seem like human-AI partnerships. It's become common, for example, for people to start writing by prompting ChatGPT for a rough first draft. But this is just one of a range of ways a person might collaborate with an AI partner, as shown in **Figure 1**.

RELATED CMR ARTICLES

**“Orchestrating Human-Machine Designer Ensembles During Product Innovation”**

by Jan Recker, Frederik von Briel, Youngjin Yoo, Varun Nagaraj, & Mickey McManus.  
(Vol. 65/3) 2023.



**Figure 1:** Continuum of Human AI Collaboration

In late 2022, CNET started using generative AI to produce news articles that humans edited into final copy.<sup>2</sup> We were curious about how articles produced this way differed from human-written CNET articles<sup>3</sup> in how they might appeal to human readers. It wasn't so much content that we were curious about, however.

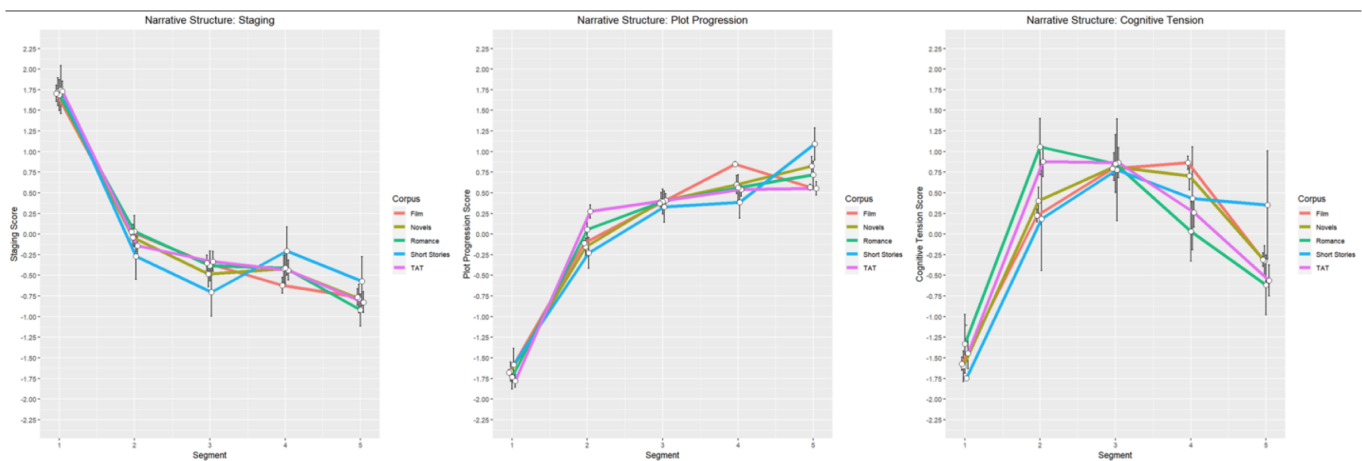
Many have noted that ChatGPT and other similar models can tend toward making up facts, giving weird responses, even “hallucinating,” which at least partially explains the need for human editors. What we wanted to explore is whether a piece primarily written by an AI would resemble human-written articles in the more stylistic ways that generate subjective impressions in human readers. This, we reasoned, has also to do with the *shape* of a piece of writing, not just its content.

The idea that writing has a shape goes back to Aristotle. In *Poetics*, he argued that a good story has a beginning, a middle, and an end. By “beginning” he meant a written expression by the author that the reader would come to prepared already to recognize and understand, without any preparation by the author, and that would move the story toward subsequent issues. A beginning, then, has no antecedents, but does generate implications: new issues that need to be dealt with. Those next new issues to be dealt with are middles. A middle takes up issues that flow from beginnings and progresses them further. Middles don’t resolve the issues; they might even raise more of them. A middle, then, is both implied and has implications. Middles flow toward ends. And ends have the characteristic that they follow naturally from middles, but they *do* resolve issues, achieving some kind of closure. An end has antecedents (middles), but no succedents.

Aristotle was pointing out something about how a good piece of writing hangs together, how its pieces interrelate to create a sense of unity. Or, as he put it: “the structural union of the parts being such that, if any one of them is displaced or removed, the whole will be disjointed and disturbed.”<sup>4</sup> *Plot*, he said, is the *soul* of a good story. Many writers have described “shapes” in stories, such as Joseph Campbell’s “hero’s journey”<sup>5</sup> or Kurt Vonnegut’s “8 shapes” of narratives (e.g., “Man in a Hole” and “Girl Meets Boy”).<sup>6</sup>

But news articles aren’t exactly stories and don’t, usually, have shapes like them. There is a journalistic plot that starts with “The Lead” (spelled “lede” by many journalists—a provocative hook with essential facts), moves to “The Body” (background details), and then to “The Tail” (extra info that adds richness). There might be similarities between the shapes of stories and news articles written by humans and AI, but there might also be differences. We examined this question empirically.

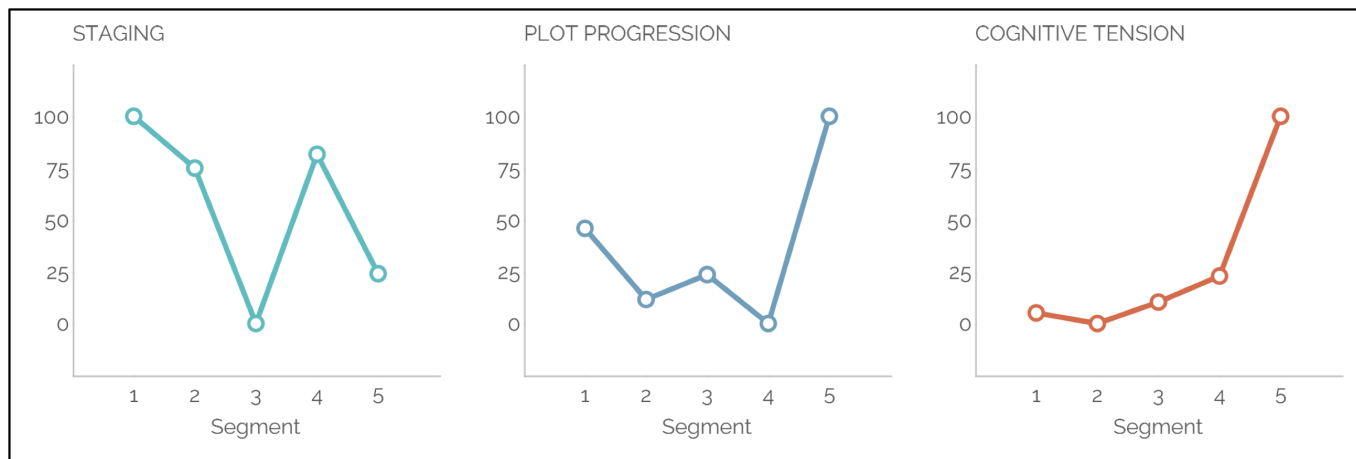
Psychologists Ryan Boyd, Kate Blackburn, and James Pennebaker have developed a way of accessing the shape of a piece of writing based on the frequency of certain kinds of words at different points throughout the piece. They use “narrative arc analysis” accomplished by software called “Linguistic Inquiry and Word Count” (LIWC)<sup>7</sup> to show that the early parts of most stories contain a lot of “words that pertain to nouns and how they [relate] to one another,” which Boyd et al. label “Staging language.” As they move to the middle of a story, authors “use more words that signal action,” which Boyd et al. call “Plot Progression language.” Moving further through a story, an author employs words that signal tension and conflict, and uncertainty about whether a character’s goals will be reached, from which Boyd et al. discern the level of “Cognitive Tension.” These measures provide a quantitative, graphical description of what the authors call a “narrative arc,” a reasonable approximation of what we have called plot. They find consistency in narrative arc when they analyze many stories from films, novels, short stories, etc. (see **Figure 2**).



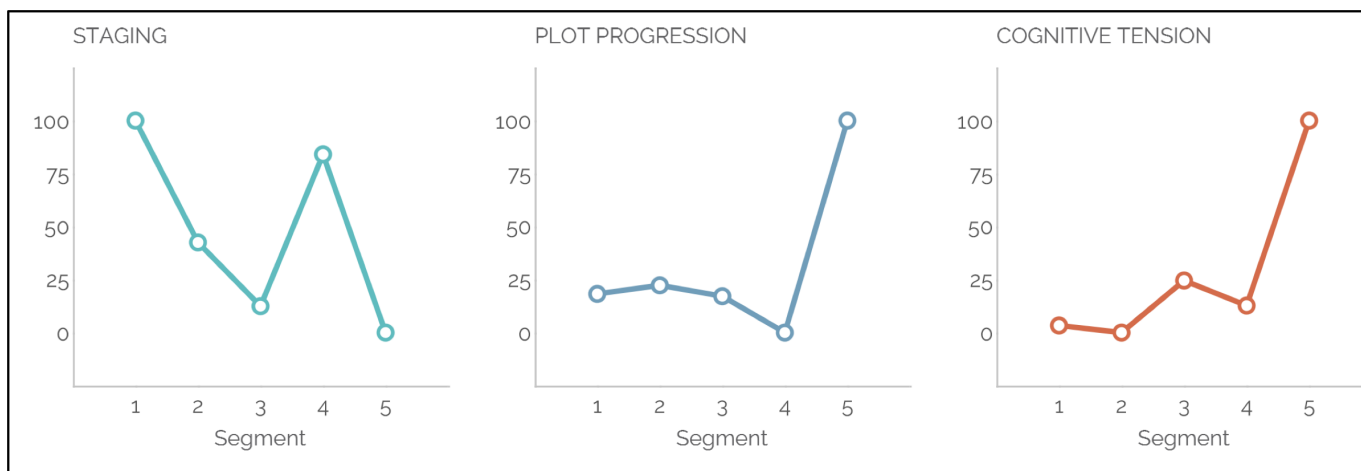
**Figure 2:** Narrative Arc Analysis of stories from Boyd et al. 2020

Narrative arc analysis does not “read” a story in the way that humans do. The LIWC software measures the frequency of words of different types. Staging language is determined by the frequency of prepositions and articles. Plot progression is measured by counting pronouns, auxiliary verbs, and other function words, and so on. Also, ChatGPT and its relatives are not “writing” the way humans do. LLMs generate structures in natural language that are like, but not a replica of, training data. The creators of ChatGPT (OpenAI) tell us that its LLM has trained primarily on a large corpus of available data through 2021 and extracts from that a sense of what words more appropriately follow the words it has already committed to.

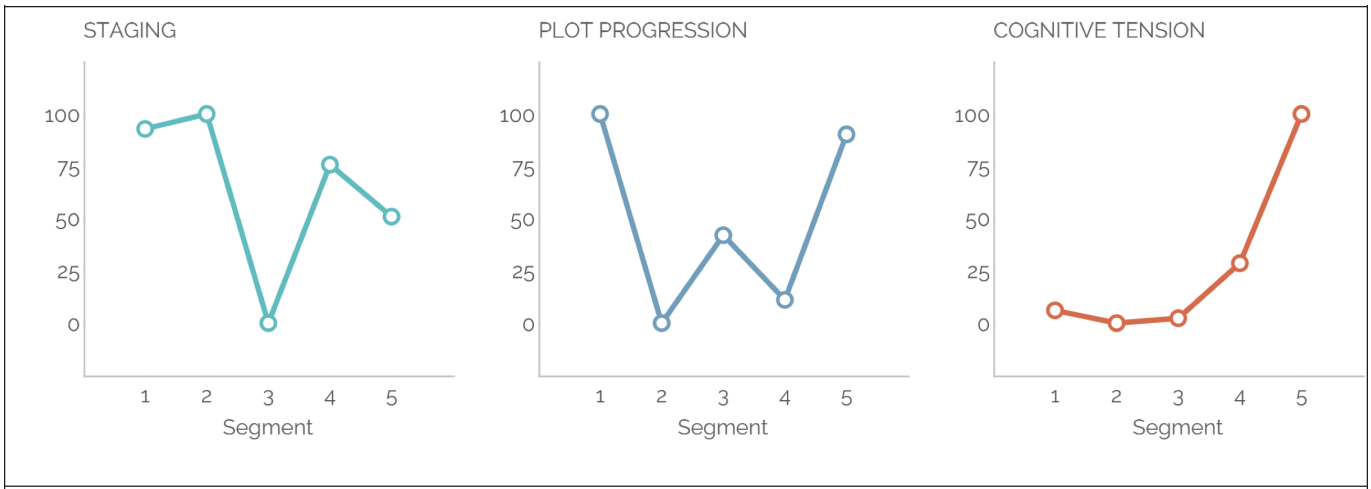
**Figure 3** shows the narrative arc analysis for 77 AI-written articles from the CNET platform that were edited and fact-checked by two human editors.<sup>8</sup> We also analyzed the narrative arc of AI-written articles reviewed by each of the human editors separately (see **Figures 4a** and **4b**), and of a representative sample of CNET posts written by additional human authors (not just the two who edited the AI posts) during a period parallel to the composition of the AI-written posts (**Figure 5**).



**Figure 3** Narrative Arc Analysis for CNET posts written by AI, edited by humans



**Figure 4a** Narrative Arc Analysis for CNET posts written by AI, edited by human editor #1 (n=27)



**Figure 4b** Narrative Arc Analysis for CNET posts written by AI, edited by human editor #2 (n=50)



**Figure 5** Narrative Arc Analysis for all CNET posts written by all human authors during an interval parallel to the composition of AI-written articles

There are not huge differences across these plots. Staging roughly decreases as the article progresses, as we might expect, as The Lead (the provocative hook) progresses to The Body (background details) and The Tail (additional info). Plot Progression shows some variation but generally increases through the article. Cognitive Tension increases too, though there is an interesting dip at the end in the graph for all human writers, perhaps suggesting that human writers are more inclined toward achieving closure at the end, as in a story narrative.

CNET articles need not have a similar structure to that of a film, novel, or short story. They are different genres of writing. But it is interesting that there seems to be in news articles an echo of the shape of stories, perhaps indicating a shape within writing that transcends genre. It makes sense that you'd need to do more Staging early in any kind of written work, and that writers might use Plot Progression as an attention-keeping mechanism, regardless of genre. It makes sense, too, that Cognitive Tension is not the same for a CNET's article as in a classic narrative story. Maybe a news article doesn't achieve closure at the end because there's an aim to make you want to come back or read the next story. Whereas, in a classic film or play, the audience will be unhappy if the story doesn't reach closure.

From the similarities in these graphs, we might conclude that AI is up to the task of generating acceptable plots for news articles, with a human editor at least. In writing articles that are then edited by humans, then, an AI-partner seems to perform competently. But can it be an equal partner in a creative collaboration?

## LLMs as Equal Partners in Symbiotic, Creative Collaboration

Researchers agree that to be considered creative an outcome must be original (unlike outcomes that have preceded it) and useful (or, more generally, valuable – usefulness is not the only kind of value).<sup>9</sup> People and organizations, however, struggle to generate original ideas and recognize value in unfamiliar outcomes.<sup>10</sup> We tend to cling to past ideas and old ways of assessing what is valuable. How might an LLM contribute to breaking these patterns as part of an innovative partnership?

One might presume that since an LLM derives its responses entirely from data about the past, its outputs would be derivative rather than original. However, Professor Ethan Mollick of Wharton Business School has found that LLMs can make surprising connections that yield unusual, seemingly non-intuitive, juxtapositions.<sup>11</sup> A glance at the images DALL-E can generate will disabuse you of the idea that nothing original can come of inferences from training data. Mollick demonstrates how we can use LLMs to get unstuck from a too-

narrow set of ideas. Continuous experimentation with such tools, in which you start with a prompt, receive a result, and then build on that by asking subsequent prompts, exemplifies Symbiotic AI to at least some degree.

In great creative collaboration, however, you might well want more than this from a partner. We draw on here the idea of “ensemble,” a notion from the world of the arts. In an ensemble, each collaborator makes fresh choices that, when shared, provoke additional fresh choices, which, in turn, provoke additional fresh choices. This chain reaction continues among collaborators until they are “successful in combining their voices into a coherent whole...[and] individual members relinquish sovereignty over their work and thus create something none could have made alone.”<sup>12</sup> In a string quartet, for example:

*“[T]he relentless sequence of new actions and reactions...[creates] an intense and joyful experience for [collaborators who are] confident in their abilities to solve the problems that [come] at them... Leadership passes from member to member, from moment to moment... All of this depends on a high standard of listening to each other... The quality of communication in the ensemble [grows] so high that members [can] sense when another member [is] about to make a mistake, and even when he recover[s] and avoid[s] it.”<sup>13</sup>*

We can see elements of this kind of symbiosis in the iterative interactions that Mollick describes, but not at the level apparent in this description. While LLMs can be helpful collaborators, they currently fall short of achieving the highest levels of symbiosis.

## Implications for Business

Our analysis suggests several actions that managers should take to prepare their organizations to maximize the potential of creative collaboration with generative AI.

1. **Pay attention to the “shape” of work generated by AI, not just the content** – How receptive human audiences are to AI-generated content is not just a matter of content, but also factors like narrative arc or plot. The mainstream conversation



about generative AI has been very content focused. Get ahead of this by promoting an awareness within your organization that it matters how messages are conveyed, not just what they say.

2. **Use automated tools to help monitor the shape of AI-generated content** –

Evaluating creative works, such as writing, has often largely relied on human judgment. But there are now analytics-based techniques, such as narrative arc analysis, that can help check if AI-generated content conforms to human writing conventions for specific genres.

3. **But maintain human editorial oversight** – Just because AI appears to be pretty good

at generating the narrative structures of human writing genres does *not* mean that you should let AI generate content unsupervised. It is well documented, of course, that AI-generated content can contain factual errors<sup>14</sup> or exhibit bias.<sup>15</sup> But you might also want to monitor for differences in shape, in the inclination to achieve some closure at the end of an article, for example, to make sure your AI-generated content is palatable for human audiences.

4. **Develop methods for more symbiotic AI collaboration, that move toward ensemble** –

Generative AI tools are useful in creative collaboration, but they do not yet achieve the highest levels of symbiosis, such as ensemble. Ensemble occurs only when individual collaborators make fresh choices that provoke further fresh choices from others in a highly dynamic chain reaction. Creative collaboration with AI partners will require, then, development of increasingly dynamic processes of interaction between human and AI models. The iterative, prompt-based nature of LLMs is already conducive to this, but there will likely be a learning curve involved in moving toward true ensemble. Make moving down this learning curve a formal project within your organization.

5. **Monitor stylistic trends in your AI-generated content** – Today, LLMs train

predominantly on human-generated content. But as more LLMs are released “into the wild,” they will increasingly train also on AI-generated content. Research suggests that this might affect the quality and shape of the content.<sup>16</sup> Differences between AI and human-generated content may become amplified, resulting in narrative shapes that start to diverge from what is natural for humans. This could generate new levels

of creativity. Or it could begin to generate content that humans experience as strange or unappealing. Organizations engaged in creative collaboration with AI will need to keep an eye on this.

AI is becoming capable of performing tasks that were once exclusive to humans. This does not diminish the importance of human activity but rather brings new meaning to it. It is important for today's managers to be thinking about this now, to prepare their people for new roles and their organizations for an increasingly exciting future with new kinds of creative partnerships.

## References

1. The idea of augmenting human intellectual capabilities by using computers has a long and distinguished history. See Douglas Engelbart's seminal 1962 paper, "Augmenting Human Intellect: A Conceptual Framework," [Englebart](#), accessed July 4, 2023.
2. Harrington, C. (2023). CNET Published AI-Generated Stories. Then Its Staff Pushed Back, May 16, [CNET](#), accessed July 11, 2023.
3. When they were first written, CNET's AI-written articles were clearly labeled on their website as AI-generated (see, for example, [CNET](#)), which allowed us to easily distinguish between human- and AI-written pieces. We collected data in February 2023, when articles were labelled as written by CNET Money and edited by human editor. However, in June, CNET overhauled its AI policy and updated past stories ([The Verge](#)). These changes make it harder, now, to distinguish between articles that are AI- and human-generated. The same URL that once took you to an article clearly identified as AI-generated might today take you to a modified version of that same article that shows a human byline. For those interested, early versions labeled "AI-generated" can be independently recovered and verified from the Wayback Machine web archive.
4. Aristotle. (1895). *Poetics*. trans. S. H. Butcher, no longer under copyright, p. 16 in original edition.

5. Campbell, J. (1990). *The Hero's Journey*. New World Library.
6. Johnson, S. (2022). Kurt Vonnegut on the 8 “shapes” of stories. [Big Think] (<https://bigthink.com/high-culture/vonnegut-shapes/>), accessed July 8, 2023; see also Frye, N. (1957). *Anatomy of Criticism*. Princeton University Press.
7. Boyd, R.L., Blackburn, K.G. & Pennebaker, J.W. (2020). The narrative arc: Revealing core narrative structures through text analysis. *Science Advances*, 6(32), p.eaba2196.
8. The narrative arc analysis requires that we specify the number of “segments” (or parts) we want to divide an article into for the purposes of the analysis. The prior literature recommends five segments for this analysis. Although we present our analysis for five segments (represented on x-axes of Figures 3, 4, and 5), we have also performed this analysis for different number of segments and found similar results.
9. Amabile, T.M. (2018). *Creativity in Context: Update to the social psychology of creativity*. Routledge.
10. Simonton, D.K. (1999). Creativity as blind variation and selective retention: Is the creative process Darwinian? *Psychological Inquiry*, pp.309-328.
11. Mollick, E. (2023). *One Useful Thing*, **oneusefulthing**, accessed June 23, 2023.
12. Austin, R. D. & Devin, L. (2003). *Artful Making*. Financial Times Prentice Hall, p. 16.
13. Austin, R. D. & O'Donnell, S. (2007). Paul Robertson and the Medici String Quartet. Harvard Business School case 607-083, p. 8; see also, Peat, F. D. (2000) *The Blackwinged Night: Creativity in Nature and the Mind*. Basic Books, p. 191.
14. Sato, M. & Roth, E. (2023). CNET found errors in more than half of its AI-written stories. *The Verge*, January 25, **The Verge**, accessed May 28, 2023.
15. Shanklin, W. (2023) AI Seinfeld was surreal fun until it called being trans an illness. *engadget*, February 6, [Engadget] (<https://www.engadget.com/ai-seinfeld-bigoted-transphobic-hateful-moderation-193449772.html>), accessed May 28, 2023.

16. See Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023). The Curse of Recursion: Training on Generated Data Makes Models Forget. *ArXiv*, *abs/2305.17493*.
- 



Rohit Nishant [Follow](#)

Rohit Nishant's research interests include AI. He has published in top academic journals, including MIS Quarterly, Research Policy, Journal of the Association for Information Systems, Journal of Strategic Information Systems, Journal of Information Technology, and Information Systems Journal, and practitioner journals, including MIS Quarterly Executive.



Deepti Singh [Follow](#)

Deepti Singh holds graduate degrees in Economics and Informatics and has nine years of corporate experience. Her research is focused on artificial intelligence, analytics, and information security, targeted for premier IS journals and conferences. She currently serves as Senior Treasurer in the Forum for Women in OR/MS (WORMS).



Vivek Kumar Singh

Vivek K. Singh's research focuses on applications of data mining and Artificial Intelligence to enhance cloud computing security, healthcare analytics, and online learning. He has published articles in leading journals, such as Decision Support Systems, Computers in Human Behavior, Journal of the American Medical Informatics Association, Psychology and Marketing, and Psychological Studies.



Robert D. Austin [Follow](#)

Robert D. Austin is a professor of Information Systems at Ivey Business School and an affiliated faculty member at Harvard Medical School. He has authored many articles in academic and professional venue, eight books, many cases and notes, three Harvard online products, and two popular Massively Open Online Courses.