INQUIRY · ETHICS AND LIMITS

The Future of Inquiry: Human Questions in an Age of Machine Answers

Machines can now produce fluent answers to almost any question. This article considers what that means for inquiry: what research shows about relying on automated answers, where machine answers are strong and weak, and why the human work of asking good questions and checking answers matters more, not less.

ModQu research article · 1,010 words · 11 references (APA 7th edition)

For most of history, the hard part of inquiry was finding an answer. Today, an answer to almost any question can be produced in seconds by search engines and by systems that generate fluent text. This changes the balance of inquiry. When answers are cheap, the scarce skills become asking the right question, judging whether an answer is trustworthy and knowing what to do with it. Research on how people use automated tools helps show both the opportunity and the risk.

A partnership foreseen

In 1960, the psychologist and computer scientist J. C. R. Licklider described a hoped-for *symbiosis* between people and computers, in which machines would handle routine work of calculation and retrieval while people set goals, formulated hypotheses and evaluated results (Licklider, 1960). In his vision, the human contribution was precisely the part of inquiry that involves asking and judging. That division of labour remains a useful guide.

What machine answers can do

Experiments suggest that automated assistance can improve productivity substantially on suitable tasks. Shakked Noy and Whitney Zhang (2023) gave professionals writing tasks and found that those given access to a text-generating system completed them considerably faster, with higher rated quality. Dell'Acqua et al. (2023) studied consultants working on realistic tasks with and without such a system. For tasks within the system's capabilities, consultants with access completed more tasks, more quickly and with higher quality. For a task deliberately chosen to fall outside its capabilities, consultants using the system were less likely to reach the correct answer than those working without it. The authors described the boundary of the system's competence as a *jagged frontier*: uneven, and not obvious to users.

Where machine answers fail

Systems that generate text can produce statements that are fluent and confident but false, a problem researchers call *hallucination*. Ji et al. (2023) surveyed the problem across many kinds of text-generating systems and found it widespread. Bender et al. (2021) warned that such systems produce language by reproducing statistical patterns in their training data, and that fluent text can create an impression of understanding and reliability that may not be justified.

How people rely on automation

People's trust in automated answers is not well calibrated. Raja Parasuraman and Dietrich Manzey (2010) reviewed research on *automation complacency* and *automation bias*: people who rely on automated aids monitor them less carefully and tend to accept their outputs even when other information suggests an error. On the other hand, Dietvorst et al. (2015) found that people who saw an algorithm make a mistake became reluctant to use it, even when it outperformed human judgment. Logg et al. (2019) found that in many tasks people actually preferred advice they believed came from an algorithm over advice from a person. The pattern is not simple distrust or trust, but trust that responds to the wrong signals.

Relying on machines may also change what people remember and learn. Sparrow et al. (2011) found that people who expected to be able to look information up later remembered the information less well, though they remembered where to find it. Evan Risko and Sam Gilbert (2016) reviewed research on *cognitive offloading*, the use of external tools to reduce mental effort, and concluded that it can improve performance while reducing what people retain, so the balance depends on what knowledge a person needs to hold internally.

The human part of inquiry

Figure 1 sets out the parts of inquiry that machines can assist and the parts that remain with people.



Figure 1. Dividing the work of inquiry between people and machines. Based on Licklider (1960) and the research cited in this article. Diagram by ModQu.

Neil Postman and Charles Weingartner (1969) argued that the most important intellectual skill is the ability to ask questions: once a person has learned to ask relevant and substantial questions, they have learned how to learn. In an age of machine answers, this skill becomes more valuable. A machine can answer the question it is given, but it cannot tell whether that was the right question to ask, whether the answer fits the specific situation, or what the answer means for a decision.

Several habits follow:

- **Frame the question carefully.** Be specific about the situation, the purpose and what kind of answer would be useful.
- **Ask for sources and check them.** An answer that cannot be traced to evidence should be treated as a hypothesis.
- **Know where the frontier lies.** Test tools on questions where the answer is already known before relying on them.
- **Keep thinking.** Use tools to extend inquiry, not to replace the judgment that inquiry requires.

Consider a hypothetical example. A market analyst asks an automated system for the size of a niche market and receives a precise figure with a plausible explanation. Checking the cited sources, she finds that one does not exist and another refers to a different country. She uses the system instead to list possible data sources and methods, then verifies each herself. The machine sped up her inquiry; it did not replace it.

What changes for organisations

For organisations, the arrival of cheap answers shifts where value lies in research. Generic information, available to anyone who asks a machine, gives no organisation an advantage. What remains valuable is evidence that machines cannot generate on their own: fresh data collected from real people about current questions, careful measurement of specific markets and situations, and the judgment to interpret findings in context. Machine-generated summaries of what has already been published can describe the past, but they cannot report what customers think today about a product launched last month. Organisations that combine fast automated tools with original, well-designed inquiry will be better placed than those that rely on either alone.

References

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality* (Working Paper No. 24-013). Harvard Business School.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), Article 248. https://doi.org/10.1145/3571730

Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, *HFE-1*(1), 4–11. https://doi.org/10.1109/THFE2.1960.4503259

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Postman, N., & Weingartner, C. (1969). *Teaching as a subversive activity*. Delacorte Press.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*(6043), 776–778. https://doi.org/10.1126/science.1207745