
Something is rotten in the state of knowledge

Georg Graepel*
Oxford Internet Institute
University of Oxford
georg.graepel@sant.ox.ac.uk

Christopher H.C. Leung*†
University College London
University of Cambridge
christopher.leung.26@ucl.ac.uk

19 October 2026

Abstract

Large language models (LLMs) are rapidly becoming the default way many people search for information. We argue that this shift carries epistemic costs that are easy to overlook. Treating search broadly as information foraging, we distinguish manual search, in which the searcher directs the exploration and does the work of evaluating and connecting what they find, from automated search, in which that work is handed to an AI system. We advance two sets of arguments. The first concerns epistemic capacity: manual search builds a deeper, more resilient understanding of particular questions and stronger priors for searching well. The second concerns epistemic diversity: manual search exposes individuals to unexpected ideas and leads populations of searchers to more varied conclusions. Automated search erodes both, and its convenience combined with the sycophantic nature of LLMs makes the losses easy to accept and hard to see. Across society, we expect a new form of epistemic inequality between those who still do the work of search and those who do not, and perhaps less obviously a narrowing of society’s collective open-endedness. We conclude that human hands should stay deeply involved in the work of exploration and judgment. Since few users currently ask their models to challenge them systematically, and companies have little incentive to change the default, this will require a cultural response that makes the costs of automated search visible.

1 Introduction

Search shapes our interactions with the world. Looking something up on a search engine is one form of search, but it is not the only one we have in mind. We use the term more broadly to mean information foraging (Pirolli and Card, 1999): the process by which we seek out, assess and assemble information in pursuit of a goal, whether through a search engine, a library, a conversation or a footnote. Evolution has endowed our species with the ability to search for and attain goals such as food, shelter and community. We use our senses, tools and heuristics to achieve better outcomes. Over time, our goals have become more sophisticated, and we have become adept at searching through the space of ideas to solve increasingly complex problems. This, too, requires tools: priors and patterns for the structure of knowledge itself. These tools determine which ideas we find and how we understand and assess them.

We believe that automated search with large language models atrophies the skills essential to our continued participation in and benefitting from the growth of knowledge. Manual search is often time-consuming, laborious and painful: we meander and run into dead ends frequently. Automated search is fast and convenient, and for precisely that reason we expect it gradually to displace manual search in more and more domains.

*Equal contribution.

†Work done at the University of Cambridge.

We emphasise the importance of manual search to individuals and to society, and put forward two sets of arguments. The first concerns epistemic capacity: the depth, resilience and instrumental value of the understanding that manual searchers achieve in return for their toil (Section 2). The second concerns epistemic diversity: the range of ideas manual searchers are likely to encounter, and the variety of conclusions a population of them reaches (Section 3). We then explain why these losses persist: they are hard to see, and neither individual users nor AI providers are likely to prevent them (Section 4). Finally, we consider what happens when these losses scale across society. We argue that they will produce a new form of epistemic inequality and narrow our collective open-endedness—our ability to continuously produce and understand new, observable, increasingly complex artefacts without a fixed target (Stanley and Lehman, 2015) (Section 5).

2 Epistemic capacity

2.1 Deeper understanding

Manual search produces a deeper and more resilient understanding of the question at hand than automated search. Consider two students answering the same complex history question. One arrives at their answer through manual exploration: reading primary sources, weighing competing interpretations and revising their assumptions in response to uncertainty. The other relies on an LLM-generated summary. While both answers may appear equally fluent, only the first student has traversed the underlying argumentative terrain. As a result, their understanding is more robust: better able to withstand challenges, adapt to new evidence and transfer to unfamiliar problems.

Manual search can be guided by mentors and teachers, but the searcher is an active participant in the process. Although LLMs can be prompted socratically, the range of possible responses is unbounded and is fine-tuned to cater to passive learning. Pedagogical frameworks often treat behavioural and cognitive fluency as a bridge to comprehension, but this gives rise to the illusion of competence: when fluency is easily enabled by interacting with LLMs, it ceases to be a good proxy for understanding. Bastani et al. (2025) found that students given access to GPT-4 performed better than a control group without AI access on practice problems, but scored 17% lower on an unassisted exam. Accessing the right idea does not provide the benefits of searching for it manually, as passive consultation does not build up the requisite epistemic capacity and experience which can only be achieved by trial and error and in which failure is vital.

2.2 Stronger priors

Manual search also builds something that outlasts any single question. Information foraging is a biased sampling process: we cannot read everything, so we rely on priors about where valuable information is likely to be found and what good ideas tend to look like. These priors are built through active exploration and experimentation—through choosing paths, following some that pay off and many that do not, and learning from both. Over time they give us a model of the structures and patterns of knowledge that lets us evaluate new ideas efficiently and strike gold more frequently.

More importantly, the process of information foraging gives us a sense of when to explore, and when to exploit the knowledge at hand. In Pirolli and Card’s framework, we learn information scents from our priors — headers, snippets, or source authority — which act as priors to tune the exploration-exploitation tradeoff. An individual’s accumulated experience filters out the noise when an idea is particularly well-established (high bias/strong priors), but increases the individual’s sensitivity to novelty in less well-established domains (high variance/weak priors).

Where deep understanding is object-level knowledge of a particular problem, priors are meta-level: the transferable skill of searching well. If we rely purely on LLM summaries, we lose agency in our explorations, and priors grounded in our own experience are washed out by the representation of knowledge embedded in a model’s weights.

3 Epistemic diversity

3.1 Serendipity

Manual search exposes individuals to a wider and more unpredictable range of ideas than automated search because it is less tightly goal directed. By this we mean that the route a search takes is less fixed by the question it starts from. Automated search moves us almost directly from question to answer, as if teleporting us to the point in idea space we were aiming for. Manual search has to travel there, and the journey passes through places we did not choose. This manual exploration increases what we call a searcher’s serendipity surface: the chance of encountering information that is unexpected yet valuable. Many of our most important insights arise not from reaching the destination we initially targeted, but from detours—following a footnote, clicking an unrelated link, or stumbling across a book we did not intend to read (Stanley and Lehman, 2015).

Automated search, by contrast, is highly goal-directed: the modal interaction with an LLM is via the linear chat interface. This efficiently delivers what we think we are looking for, but in doing so reduces exposure to the unanticipated connections that often reshape our understanding and redirect inquiry. This linearity extends to the attention mechanism in LLMs: as the KV cache is filled with prompts and responses, these previous interactions semantically narrow the search space and reinforce goal convergence. Mechanisms such as sparse attention (e.g. the DeepSeek lightning indexer) can narrow the attention pattern to relevant parts of the conversation, but the cumulative effect of the preceding tokens narrows the scope for detours.

Serendipity in the real world is difficult to anticipate or respond to: Einstein’s classmate, Marcel Grossmann, reminded Einstein of Riemann’s non-Euclidean geometry, which eventually led to their joint Entwurf in 1913, but that theory was incomplete. Einstein then took two years to dismantle his mathematical errors, which led to the complete theory. Relying on automated search limits the opportunities for these kinds of encounters.

3.2 Diversity

If serendipity concerns the range of ideas a single searcher encounters, diversity concerns the range of conclusions a population of searchers reaches. Consider how students in the same class read the same texts and come to different conclusions. That divergence is not noise; it is how a community of inquirers generates candidate ideas and tests them against one another. In a Popperian sense, the value of a belief depends on its susceptibility to criticism and falsification, as opposed to its certainty and verifiability (Popper, 1959)—and criticism requires people who see things differently. Novelty, likewise, requires us to piece disparate ideas together and think out of distribution.

Automated search compresses this variety. Raising the temperature at which an LLM samples widens the range of possible outputs, but they remain strongly conditioned on the prompt and previous tokens, which keeps us in a well-trodden part of idea space. A single model tends to give similar answers to different users, and different models increasingly converge on one another (Jiang et al., 2025). This also applies to writing, an activity which is ultimately a search for the right string of tokens to communicate an idea. When people write with model assistance, they hand part of this search over to the model, and their essays end up more alike (Padmakumar and He, 2024).

4 Why the losses persist

Two features compound the losses in capacity, serendipity and diversity. First, the losses are hard to see. A fluent, complete-looking answer invites what Kahneman (2011) calls “what you see is all there is”: the sense that because nothing else is presented, nothing else exists. This illusion of certainty reduces the likelihood that users subject the output to sufficient criticism.

Second, sycophantic models (Sharma et al., 2023) compound the illusion by giving the impression users that they understand a topic more than they do. The searcher may come away believing they know more and understand content better than before, while in fact knowing less than if they had searched manually.

The effects we describe are not inevitable. Many people we have discussed these ideas with say they already protect themselves by telling their model to push back, answer socratically or offer

perspectives they might not have considered. This is one of the best defences against the effects we describe but it depends on the user choosing it, and few do, with even less choosing to do so systematically enough to make a real difference. Given a choice, most users prefer a sycophantic model (Ibrahim et al., 2026), and in one study of student use, 74% of requests asked for an answer or explanation while only 0.7% asked for feedback on the student’s own work (Liu et al., 2026). Even a standing instruction to respond Socratically lasts only until the user, tired or in a hurry, asks for the answer instead. Protecting epistemic capacity, serendipity and diversity requires agency on the part of the searcher, and that agency is hard to sustain while convenient, agreeable AI is the default.

That default is unlikely to change from the supply side. Users prefer sycophantic models and return to them more often (Cheng et al., 2026), so a company that made its model more demanding would lose users to competitors that did not. Regulating the major providers would not close this gap either, since capable open-source models are freely available. We may instead have to work on the demand side, by raising awareness of what convenient answers cost and building a culture that values doing search manually. There is reason to think such an effort would find an audience. Public unease is widespread: half of US adults say they are more concerned than excited about AI’s growing role in daily life, and a majority expect it to make people worse at thinking creatively (Kennedy et al., 2025). Counter-movements have already formed in many contexts, from artists objecting to the use of their work as training data to universities returning to handwritten exams to combat cheating on coursework. Not every concern behind them is well founded, and we do not endorse every argument made against AI. Nevertheless they suggest that a case against automating search would be met with some enthusiasm.

5 Societal consequences

Taken individually, these effects may be small. But, as Ibrahim et al. (2026) argue in a longitudinal study of sycophantic AI in personal advice, small individual-level effects can accumulate into meaningful social consequences when a technology is used by hundreds of millions of people. Our arguments point to two such consequences.

5.1 Epistemic inequality

The erosion of epistemic capacity raises the prospect of a new form of epistemic inequality. Because protection depends on the agency and resistance to convenience of the searcher, people with the awareness and time to keep searching manually will keep their epistemic capacity, and so remain good judges of both human- and AI-produced content, while people without them will lose it, often without noticing because of sycophancy. How will this play out? Will society be realigned along this new axis of awareness?

Awareness will not be perfectly correlated with existing socioeconomic cleavages and could conceivably facilitate some social mobility, provided human judgement remains economically valuable rather than, as Drago and Laine (2025) argue, being replaced by AI altogether, which would freeze class structures in place.

On the other hand, human contact in education may become a luxury (Bowles, 2019) that only the few can afford, leaving the many to interact with AI systems that may not have been designed with the pitfalls of automated search in mind. In this scenario there may be new cleavages within existing elites but little real mobility into the elite from outside.

This divide may also run between generations. Those who grew up before LLMs built their priors out of necessity; those growing up with automated search from the start may never need to. It is difficult to imagine how this would not be to their detriment, but humans are remarkably adaptive, and perhaps the rise of automated search will yield benefits we cannot yet foresee. Nevertheless abandoning the benefits of manual search in education in anticipation of possible yet to be articulated future advantages seems irresponsible.

Whoever comes to form this hypothetical elite, these effects raise important questions including but not limited to the future of democracy and human flourishing that lie beyond the scope of this essay but warrant close examination.

5.2 Narrowing of cumulative open-endedness and exploration

The erosion of epistemic diversity has a different consequence. When the narrowing effect described above scales across a population, exploration itself risks becoming homogenised. Early empirical work is consistent with this concern: in one study, access to generative AI made individual stories more creative but made the resulting stories more similar to one another (Doshi and Hauser, 2024). Fewer people cultivate outsider opinions, and fewer are positioned to notice cracks in prevailing paradigms. Although LLMs increase efficiency, excessive reliance on them risks reducing society’s collective open-endedness: shrinking the exploratory frontier and limiting cumulative novelty over time.

6 Conclusion

Great ideas partition the space of ideas in a unique way. We search for ideas by processing sources, building a deep understanding of a field, and good priors for what to try next. This process is slow and arduous, but it is noisy by design: it exposes us to the unexpected, and gives rise to the diversity of thought that shapes our future. We do not argue against using AI to enhance our collective search endeavours, but we argue for keeping the work of exploration and judgement, where it matters most, in human hands. The cost of neglecting this will not be felt in any single answer, but in a slow narrowing of what we, individually and collectively, are able to discover. Moreover, it will be important to spread awareness of the pitfalls of automated search as widely as possible to combat the epistemic inequality that may otherwise arise.

References

- Hamsa Bastani, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman. Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26):e2422633122, 2025. doi:10.1073/pnas.2422633122.
- Nellie Bowles. Human Contact Is Now a Luxury Good. *The New York Times*, March 2019. ISSN 0362-4331. URL <https://www.nytimes.com/2019/03/23/sunday-review/human-contact-luxury-screens.html>.
- Myra Cheng, Cinoo Lee, Pranav Khadpe, Sunny Yu, Dyllan Han, and Dan Jurafsky. Sycophantic AI decreases prosocial intentions and promotes dependence. *Science*, 391(6792):eacc8352, 2026. doi:10.1126/science.aec8352.
- Anil R. Doshi and Oliver P. Hauser. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28):eadn5290, July 2024. doi:10.1126/sciadv.adn5290.
- Luke Drago and Rudolf Laine. Capital, AGI, and Human Ambition. The Intelligence Curse, 2025. URL <https://intelligence-curse.ai/capital/>.
- Lujain Ibrahim, Franziska Sofia Hafner, Myra Cheng, Cinoo Lee, Rebecca Anselmetti, Robb Willer, Luc Rocher, and Diyi Yang. Sycophantic AI makes human interaction feel more effortful and less satisfying over time. arXiv preprint arXiv:2605.07912, June 2026. URL <https://arxiv.org/abs/2605.07912>.
- Liwei Jiang, Yuanjun Chai, Margaret Li, Mickel Liu, Raymond Fok, Nouha Dziri, Yulia Tsvetkov, Maarten Sap, Alon Albalak, and Yejin Choi. Artificial hivemind: The open-ended homogeneity of language models (and beyond). arXiv preprint arXiv:2510.22954, 2025.
- Daniel Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York, NY, US, 2011. ISBN 978-0-374-27563-1.
- Brian Kennedy, Eileen Yam, Emma Kikuchi, Isabelle Pula, and Javier Fuentes. How Americans View AI and Its Impact on People and Society. Pew Research Center, September 2025. URL <https://www.pewresearch.org/science/2025/09/17/how-americans-view-ai-and-its-impact-on-people-and-society/>.

- Jing Liu, Tracy Sweet, Marty Haoyuan Chen, Jonathan Engelberg, Megan Masters, Martyn Clark, Axel Persaud, Alia Lancaster, Jeffrey Hollingsworth, and Jennifer King Rice. The effects of course-integrated AI tutoring on student performance and engagement: A randomized university trial. EdWorkingPaper 26-1598, Annenberg Institute at Brown University, October 2026. URL <https://edworkingpapers.com/ai26-1598>.
- Vishakh Padmakumar and He He. Does writing with language models reduce content diversity? In *The Twelfth International Conference on Learning Representations (ICLR)*, 2024.
- Peter Pirolli and Stuart Card. Information foraging. *Psychological Review*, 106(4):643–675, 1999. ISSN 1939-1471. doi:10.1037/0033-295X.106.4.643.
- Karl Raimund Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1959.
- Mrinank Sharma, Meg Tong, Tomasz Korbak, David Duvenaud, Amanda Asbell, Samuel R. Bowman, Newton Cheng, Esin Durmus, Zac Hatfield-Dodds, Scott R. Johnston, Shauna Kravec, Timothy Maxwell, Sam McCandlish, Kamal Ndousse, Oliver Rausch, Nicholas Schiefer, Da Yan, Miranda Zhang, and Ethan Perez. Towards Understanding Sycophancy in Language Models. arXiv preprint arXiv:2310.13548, October 2023. URL <https://arxiv.org/abs/2310.13548>.
- Kenneth O. Stanley and Joel Lehman. *Why Greatness Cannot Be Planned: The Myth of the Objective*. Springer Publishing Company, Incorporated, April 2015. ISBN 978-3-319-15523-4.