

Research Insight: Will AI lead to the collapse of collective knowledge?

July 2026

The promise of agentic AI models is to provide context-specific information and recommendations to human decision-makers. While this may benefit individuals in the short run, it discourages the formation of collective knowledge by disincentivizing learning effort.

In a recent [paper](#), Daron Acemoglu, Dingwen Kong, and Asuman Ozdaglar develop a theoretical model for how agentic AI interacts with human learning incentives and long-run collective knowledge. **While powerful AI models can improve the efficiency and quality of decision-making today, they replace human effort in a way that threatens to erode long-run community-level knowledge.**

Individual human learning contributes to general knowledge

Humans require two types of knowledge to make well-informed decisions: general knowledge and individual, context-specific information. For example, to make investment decisions, people need general knowledge about different financial instruments (e.g., stocks versus treasury bonds), how stock markets and economies are performing, and macroeconomic risks. But they also need information specific to their individual context, such as their personal risk tolerance, planning horizon, and other existing sources of income.

Human decision-makers typically acquire general and context-specific knowledge at the same time when learning about a topic. **As humans learn, they contribute to growing the stock of general knowledge.**

Key Findings

Individual human learning contributes to general knowledge. **Agentic AI disincentivizes individual learning efforts, indirectly eroding public knowledge.**

AI could even trigger a state of **knowledge collapse**, in which society's stock of useful general knowledge gradually disappears over time.

To combat these effects, policy could attempt to reduce the extent to which individuals rely heavily on AI recommendations.

Larger and better-connected communities also reduce the risk of knowledge collapse and make people better off.

Source: Daron Acemoglu, Dingwen Kong, and Asuman Ozdaglar, "AI, Human Cognition and Knowledge Collapse," NBER Working Paper 34910 (2026), nber.org/papers/w34910



And when the quality and availability of general knowledge is higher, individuals find it more valuable to invest in their context-specific learning. Consider, again, the example of investment decisions: the more you understand markets in general, the more useful it is to figure out your own risk tolerance.

While individuals benefit from the collective knowledge that already exists on a topic, because this collective knowledge is the result of the contributions of hundreds of thousands of people, each individual's direct benefit from their own contribution is small. For example, when a software engineer diagnoses a rare bug in their code, they benefit mainly from fixing their own system (a context-specific fix). If they write up a summary of the problem and its solution on a public forum like Stack Overflow, this will be mostly for the benefit of others in the future.

Importantly, general knowledge is not static. It is dynamically updated over time as the frontier of knowledge changes – for example, as environmental changes cause new diseases and medical advances lead to new treatments.

AI disincentivizes learning effort

The promise of agentic AI is that it provides context-specific, individualized information to human decision-makers. Returning to the example of investing, existing textbooks and online resources allow people to teach themselves about finance and gain general knowledge. But future agentic AI might recommend a specific investment choice based on an individual's context (or even trade autonomously on their behalf). **Since humans can now receive this information without investing in their own learning on a topic, the AI is a substitute for their effort.** The better the recommendations from the AI, the less value humans will gain from putting effort into their own learning.

Agentic AI thereby disincentivizes individual learning efforts. And since those individual learning efforts build up the stock of collective knowledge, **AI indirectly erodes public knowledge.**

At first glance, it is unclear whether these combined effects will make people better or worse off. On the one hand, individuals benefit personally from high-quality information provided by AI. On the other hand, the erosion of general knowledge makes people (and society) worse off.

The researchers' model suggests that there is a "sweet spot" of AI accuracy from a welfare perspective. Modestly accurate AI can have a net benefit. But when agentic AI is very accurate, the negative effect dominates.

These dynamics can create a tipping point. Because general knowledge is not static, ongoing human learning effort is required to sustain it. If a community's stock of general knowledge stays above a critical threshold, continued learning will maintain it. But if the stock dips below that threshold, learning incentives and knowledge can fall together, creating a downward spiral.

AI could trigger the collapse of general knowledge

All together, these forces could reduce collective knowledge in society or even lead to a state of knowledge collapse, in which society's stock of useful general knowledge gradually disappears over time. AI may still provide context-specific information to individuals. But because good decisions require *combining* general and context-specific knowledge, personalized AI advice becomes useless once the underlying general knowledge has eroded. Individuals can't apply tailored information well without understanding the broader context.



How likely is this knowledge-collapse scenario? A key factor is how well the knowledge individuals generate gets pooled and made widely available. Well-aggregated general knowledge – for example, well-functioning public forums, professional networks, open repositories, and Internet search – reduces the risk of knowledge collapse and unambiguously makes people better off. **Larger and better-connected communities not only spread existing knowledge, but also motivate people to invest in their learning by making learning more worthwhile.**

On the flip side, as agentic AI improves, the erosion of general knowledge could paradoxically become more likely. The risk increases when people respond to high-quality AI by sharply reducing their own learning. Where learning is more ‘sticky’ (e.g., in professions with mandatory training), AI erodes but doesn’t eliminate collective knowledge.

If AI itself adds to the stock of available knowledge or produces novel insights, then knowledge is less prone to collapse – but unless these channels are very strong, public knowledge will still erode.

Policy should consider the optimal level of AI accuracy

There is already suggestive evidence that generative AI has hindered collective knowledge-building, including declining activity and human engagement on knowledge-sharing platforms like Stack Overflow and Wikipedia.

To combat these effects, policy could attempt to reduce the extent to which individuals rely heavily on AI recommendations.

Because *highly accurate* agentic AI is what discourages human learning, one policy lever could be to manage its accuracy directly. This approach would limit the precision of AI recommendations in order to target the ‘sweet spot’ of AI accuracy for preserving human learning and knowledge formation. In general, the analysis shows that a two-phase policy is necessary. Such a policy would start with a temporary moratorium on agentic AI to rebuild the stock of public knowledge above the ‘tipping point’ for knowledge collapse. The second phase would set a permanent cap on the AI’s accuracy at an optimal level.

Other types of policy may also be effective to reduce individual reliance on AI recommendations. As agentic AI models become more powerful, policymakers should carefully consider how AI could degrade collective knowledge in the long run – and not be blinded by its short-term benefits.

 [Read the paper](#)

 [Learn more about the Stone Center on Inequality and Shaping the Future of Work](#)

We gratefully acknowledge support from the James M. and Cathleen D. Stone Foundation and the William and Flora Hewlett Foundation.

