

Forum:[Hack Forums](#)

Thread: [The internet may be poisoning the next generation of AI](#)

Maxpro - September 27, 2026, 9:34 am

? loading encrypted image...

Researchers from Oxford and Cambridge have highlighted a major problem for future AI models: model collapse.

As AI-generated text floods the internet, future models could increasingly train on content produced by previous models. The result? AI may gradually lose the rare, creative and unusual patterns found in genuine human data.

The researchers found that repeatedly training models on generated data can cause them to lose the “tails” of the original data distribution, essentially forgetting the weird, diverse examples that make human knowledge so rich.

And this isn't limited to ChatGPT. The researchers showed the phenomenon can affect multiple types of generative models, suggesting it's a broader problem with learning from recursively generated data.

AI companies depend heavily on massive amounts of web data. If the web becomes increasingly synthetic, authentic human-generated data could become one of the most valuable resources in AI.

[Source](#)

Infinity. - September 27, 2026, 9:35 am

How about AI poisoning and dumbing the Kids?

octanitrocubane - October 2, 2026, 12:06 am

Most of the content on the internet is mediocre, additionally models isn't particularly good at logical, temporal or spatial thinking. So mediocrity * mediocrity = mediocrity square

Maxpro - October 5, 2026, 4:05 am

(Sep 27, 2026 09:35 AM)Infinity. Wrote: [\[CITE\]](#) How about AI poisoning and dumbing the Kids?

That's a real concern too—if kids rely on AI for everything instead of learning to think, question, and create for themselves, it could weaken critical thinking.

(Oct 2, 2026 12:06 AM)octanitrocubane Wrote: [\[CITE\]](#) Most of the content on the internet is mediocre, additionally models isn't particularly good at logical, temporal or spatial thinking. So mediocrity * mediocrity = mediocrity square Exactly. If models keep learning from mediocre content and then generating more mediocre content, the feedback loop could make the overall quality worse over time—especially for reasoning-heavy tasks.

Lowkey - October 5, 2026, 4:20 am

Uhhhhh "AI IS BAD!!! BECAUSE UMMM"

The models will be fine, half the parameters for data intake are pentaogn tier propoganda filters anyways.

I'd be more worried about not keeping up with how the models work and what they can do for you.

Maxpro - October 5, 2026, 6:31 am

(Oct 5, 2026 04:20 AM)Lowkey Wrote: [\[CITE\]](#) Uhhhhh "AI IS BAD!!! BECAUSE UMMM"

The models will be fine, half the parameters for data intake are pentaogn tier propoganda filters anyways.

I'd be more worried about not keeping up with how the models work and what they can do for you.

Yeah, I don't think the point is "AI is bad" either. The concern is more that if models keep training on AI-generated data, they could slowly lose the weird, rare stuff that makes human data valuable. Better filters definitely help, but they don't completely remove that problem.