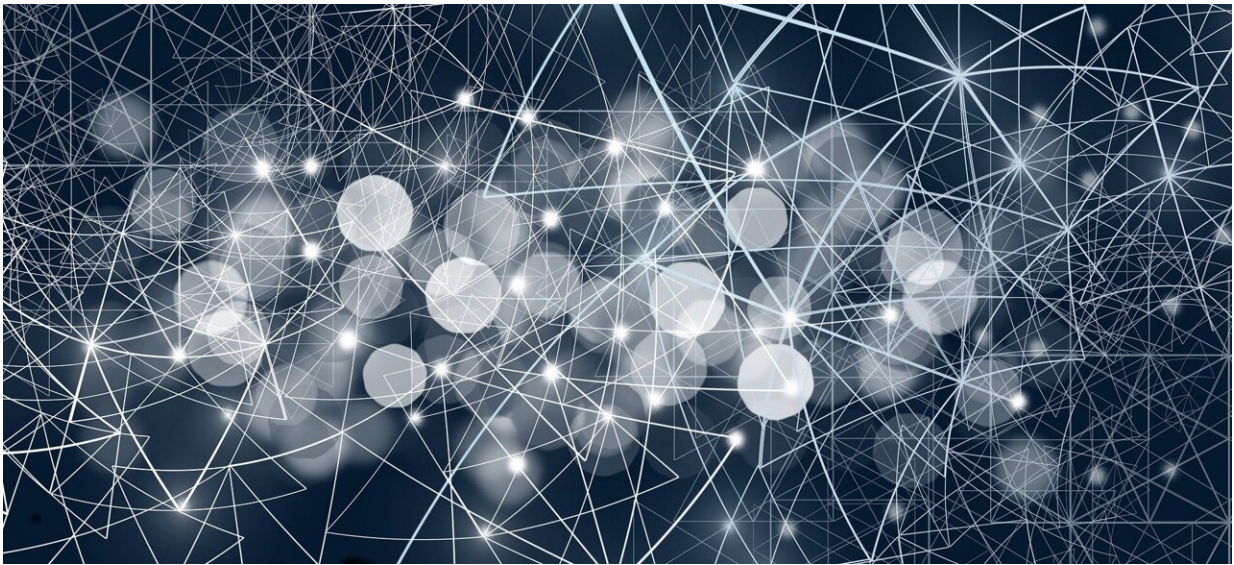


# What is 'model collapse?' An expert explains the rumors about an impending AI doom

August 19 2024, by Aaron J. Snoswell

---



Credit: Pixabay/CC0 Public Domain

Artificial intelligence (AI) [prophets](#) and [newsmongers](#) are forecasting the end of the generative AI hype, with talk of an impending catastrophic "model collapse."

But how realistic are these predictions? And what is model collapse anyway?

Discussed in [2023](#), but popularized [more recently](#), "model collapse"

refers to a hypothetical scenario where future AI systems get progressively dumber due to the increase of AI-generated data on the internet.

## The need for data

Modern AI systems are built using machine learning. Programmers set up the underlying mathematical structure, but the actual "intelligence" comes from training the system to mimic patterns in data.

But not just any data. The current crop of generative AI systems needs [high quality data](#), and lots of it.

To source this data, big [tech companies](#) such as OpenAI, Google, Meta and Nvidia continually scour the internet, scooping up [terabytes of content](#) to feed the machines. But since the advent of [widely available](#) and [useful](#) generative AI systems in 2022, people are increasingly uploading and sharing content that is made, in part or whole, by AI.

In 2023, researchers started wondering if they could get away with only relying on AI-created data for training, instead of human-generated data.

There are huge incentives to make this work. In addition to proliferating on the internet, AI-made content is [much cheaper](#) than human data to source. It also isn't [ethically and legally questionable](#) to collect en masse.

However, researchers found that without high-quality human data, AI systems trained on AI-made data [get dumber and dumber](#) as each model learns from the previous one. It's like a digital version of the problem of inbreeding.

This "[regurgitive training](#)" seems to lead to a reduction in the quality and diversity of model behavior. Quality here roughly means some

combination of being helpful, harmless and honest. Diversity refers to the variation in responses, and which people's cultural and social perspectives are represented in the AI outputs.

In short: by using AI systems so much, we could be polluting the very data source we need to make them useful in the first place.

## **Avoiding collapse**

Can't big tech just filter out AI-generated content? Not really. Tech companies already spend a lot of time and money cleaning and filtering the data they scrape, with one industry insider recently sharing they sometimes discard [as much as 90%](#) of the data they initially collect for training models.

These efforts might get more demanding as the need to specifically remove AI-generated content increases. But more importantly, in the long term, it will actually get harder and harder to distinguish AI content. This will make the filtering and removal of synthetic data a game of diminishing (financial) returns.

Ultimately, the research so far shows we just can't completely do away with human data. After all, it's where the "I" in AI is coming from.

## **Are we headed for a catastrophe?**

There are hints developers are already having to work harder to source high-quality data. For instance, [the documentation](#) accompanying the GPT-4 release credited an unprecedented number of staff involved in the data-related parts of the project.

We may also be running out of new human data. [Some estimates](#) say the

pool of human-generated text data might be tapped out as soon as 2026.

It's likely why OpenAI and others are [racing to shore up exclusive partnerships](#) with industry behemoths such as [Shutterstock](#), [Associated Press](#) and [NewsCorp](#). They own large proprietary collections of human data that aren't readily available on the public internet.

However, the prospects of catastrophic model collapse might be overstated. Most research so far looks at cases where synthetic data replaces human data. In practice, human and AI data are likely to accumulate in parallel, which [reduces the likelihood of collapse](#).

The most likely future scenario will also see an ecosystem of somewhat diverse generative AI platforms being used to create and publish content, rather than one monolithic model. This also increases robustness against collapse.

It's a good reason for regulators to promote healthy competition by [limiting monopolies](#) in the AI sector, and to fund [public interest technology development](#).

## **The real concerns**

There are also more subtle risks from too much AI-made content.

A flood of synthetic content might not pose an [existential threat](#) to the progress of AI development, but it does threaten the digital public good of the (human) internet.

For instance, researchers [found a 16% drop](#) in activity on the coding website StackOverflow one year after the release of ChatGPT. This suggests AI assistance may already be reducing person-to-person interactions in some online communities.

[Hyperproduction](#) from AI-powered content farms is also making it harder to find content that isn't [clickbait stuffed with advertisements](#).

It's becoming impossible to reliably distinguish between human-generated and AI-generated content. One method to remedy this would be watermarking or labeling AI-generated content, as I and many others have [recently highlighted](#), and as reflected in recent Australian government [interim legislation](#).

There's another risk, too. As AI-generated content becomes systematically homogeneous, we risk losing [socio-cultural diversity](#) and some groups of people could even experience [cultural erasure](#). We urgently need [cross-disciplinary research](#) on the [social and cultural challenges](#) posed by AI systems.

Human interactions and human data are important, and we should protect them. For our own sakes, and maybe also for the sake of the possible risk of a future [model](#) collapse.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: What is 'model collapse?' An expert explains the rumors about an impending AI doom (2024, August 19) retrieved 3 September 2024 from <https://techxplore.com/news/2024-08-collapse-expert-rumors-impending-ai.html>

|  |
|--|
| <p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p> |
|--|