

ARE WE GIVING AI BRAIN ROT?

I'm sure some of you have wondered if social media or "AI" (Artificial Intelligence) is giving the human race **Brain Rot**. Or, is the reverse possible? Are humans giving AI **Brain Rot**?

Would you believe that a study by **Texas A&M University, the University of Texas at Austin, and Purdue University** says yes to AI brain rot? After reading this article, I suggest you do a Goggle search and prove it for yourself. More studies are expected on this topic.

In short, humans tend to look for the quick and each answer for everything? Our resistance to reading, learning and social media is considered to be largest cause of **Brain Rot**. Are humans now **"Dumbing Down"** AI (Artificial Intelligence)? Is this even possible? Which is worse, our **Brain Rot**, or AI's **Brain Rot**? At the end of these notes there are links to additional studies.

I personally asked an AI bot (*in writing*) if humans are responsible for **Brain Rot**. The answer I received (*in writing*) is shown below and was shocking.

In the technical sense as used in various "brain rot" studies: Yes, human choices about the content created, amplified, and fed into training pipelines is the primary cause of LLM brain rot; not some spontaneous defect in the models themselves.

What "brain rot" depends on

The LLM Brain Rot Hypothesis says that continual exposure to "junk web text" (i.e. short, viral, superficial social media and similar low-quality content) induces lasting cognitive decline in LLMs.

Experiments show that as the proportion of such junk data increases during continued pre-training, the results indicate that reasoning, long-context recall, and safety scores fall, while anti-social "dark traits" and thought-skipping increase.

Where humans come in

Humans design the data pipelines and filtering criteria. The study shows that should the human curation tolerate too much low-quality, high-engagement text, the resulting models degrade.

Other human-driven risks

Separate work on "model collapse" shows that if future training data is polluted with AI-generated text (because humans indiscriminately scrape and reuse it), models can progressively lose information about the original human data distribution.

Together, these results argue that human data practices — content creation, platform incentives, and training-set curation — are the main causal factors behind LLM “brain rot” and related degradation phenomena.

(Note: LLM in the above discussion is the acronym for “*Large Language Model*.”) LLMs are an integral part of AI models. Stated in a different way, they are the part of AI that communicates with humans. I wrote a 4-part article on this subject that was published in the *Treasure Coast Ham News* newsletter a few years ago. You can review the articles in the newsletter archives on treasurecoasthams.com.

Bruce Campbell, W8HW

Email: W8HW@comcast.net

=====

Now, let’s take a look at what the studies have shown us.

A study titled “*LLMs Can Get ‘Brain Rot!’*” was published in October 2025. It demonstrated that large language models (LLMs) can undergo lasting cognitive decline, termed “Brain Rot,” when continually exposed to low-quality, engagement-driven web content — like viral social media posts and junk data.

Study Details

Researchers from the University of Texas at Austin and Purdue University conducted controlled experiments using real Twitter/X datasets to isolate the effects of junk versus high-quality text.

Key findings showed measurable declines in reasoning, long-context understanding, and safety, as well as increases in so-called “dark traits” (psychopathy, narcissism) when LLMs were trained on junk data.

Performance drops were significant. For instance, reasoning accuracy on ARC-Challenge tasks fell from 74.9% to 57.2% as the proportion of junk data increased from 0% to 100%.

Mechanisms of “Brain Rot”

The main failure mode identified was “thought-skipping,” where models increasingly omitted steps in reasoning chains, leading to shallow or incomplete answers.

Importantly, the decline persisted even when models were later tuned with clean data, indicating representational drift rather than recoverable format confusion.

The flowchart-style chart below explains the processes used in the Brain Rot study.

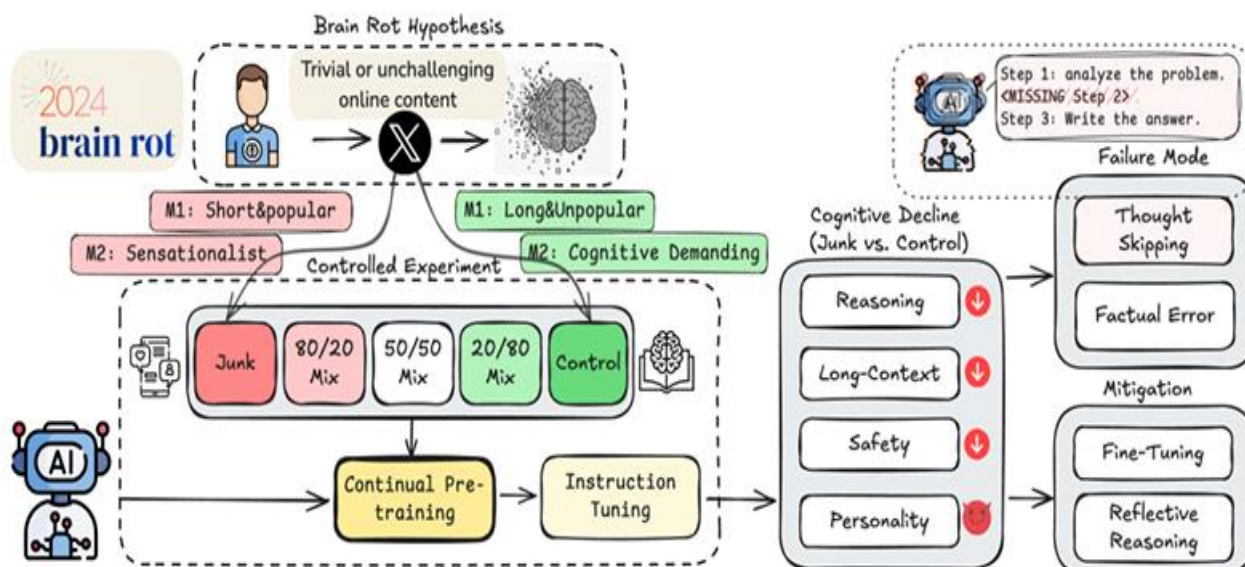


Figure 1: Outline of our work: (i) Inspired by the concept of Brain Rot, we establish the hypothesis of LLM Brain Rot; (ii) We construct junk and control data from Twitter/X posts for intervention; (iii) We benchmark four different cognitive functions of the intervened LLMs; (iv) We analyze the results to identify the failure modes caused by the brain rot; and (v) Brain rot is persistent after various mitigation.

References

Cornell university: [\[2510.13928\] LLMs Can Get "Brain Rot"!](#)

Tech news – Business standards: [AI is suffering 'brain rot' as social media junk clouds its judgment | Tech News - Business Standard](#)

Glenn Beck – The simple explanation for this...: [AI bots are experiencing BRAIN ROT... and it's happening to all of us - Glenn Beck](#)

For a more detailed explanation of AI Brain Rot I suggest you Goggle search the phrase: “*Study that shows that LLM's can get brain rot by both University of Texas at Austin and Purdue University*”.