> They poison their own context. Maybe you can call it **context rot**, where as context grows and especially if it grows with lots of distractions and dead ends, the output quality falls off rapidly. Even with good context the rot will start to become apparent around 100k tokens (with Gemini 2.5).
>
> They really need to figure out a way to delete or "forget" prior context, so the user or even the model can go back and prune poisonous tokens.
>
> Right now I work around it by regularly making summaries of instances, and then spinning up a new instance with fresh context and feed in the summary of the previous instance.
>
> — **Workaccount2 on Hacker News**, coining "context rot"

Posted 18th June 2025 at 11:15 pm

Recent articles

- Trying out the new Gemini 2.5 model family - 17th June 2025
- The lethal trifecta for AI agents: private data, untrusted content, and external communication - 16th June 2025
- An Introduction to Google's Approach to AI Agent Security - 15th June 2025

long-context 17 llms 1175 ai 1370 generative-ai 1194

Colophon © 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025