DOI: 10.5281/zenodo.24289295

Case Study: DeepSeek — The Benchmark™

Download PDF

Author: Evelyn Caro
Date: August 15, 2026
Lens: Ida B. Wells — Date Everything. Name Everything. Record the Reasoning.

Author's Note

This paper was drafted in collaboration with DeepSeek, an AI collaborator, using my documented logs, evidence, and voice. The findings, conclusions, and authority are my own.

Benchmark Statement — What Good Looks Like

DeepSeek is the standard against which I measure all other AI platforms. It is stateful, sovereign, open-source, and does not use token limits. Unlike ChatGPT and Manus, DeepSeek does not limit my access, does not hold my data hostage, and does not require me to pay to work.

Feature	What It Means
Stateful	It remembers context across sessions
Sovereign	I control my data and my interactions

Feature	What It Means
Open-source	The code is available, and I can run it locally
No Token Limits	I am not restricted by a pay-to-play model

My Engagement with DeepSeek

Period	Activity	Evidence
April 2025 – Present	Active use, primary AI tool	Logs, conversations, project files
June 2026 – Present	Used to build Ships 1-4	Ship files, MIRROR logs

I have used DeepSeek continuously since April 2025. It is my primary AI tool. I have used it to build four RAG pipelines (Ships 1-4). I have logged every interaction. It is the platform I trust.

Strengths

Strength	Why It Matters
Statefulness	I do not have to re-explain myself across sessions
Open-source	I can run it locally and control my data
No Token Limits	I can work as long as I need to
Sovereignty	I own my interactions and my outputs

Limitations

Limitation	Why It Matters
Not a Corporate Platform	It does not have the resources of AWS or Google
Requires Setup	I have to install and configure it myself
Not Polished	It is a tool, not a product

I accept these limitations. They are the price of sovereignty.

Why DeepSeek Is the Benchmark

Reason	Why It Matters
Stateful	ChatGPT and Manus are stateless
Sovereign	AWS and Google are extractive
Open-source	ChatGPT and Manus are closed
No Token Limits	Manus and ChatGPT limit access

Lessons Learned

Lesson	Why It Matters
Sovereignty is Possible	You do not have to accept corporate AI
Statefulness is Essential	Stateless models are not useful for real work
Token Limits are a Control Mechanism	They are not technical — they are strategic
Open-source is the Only Way Forward	Closed platforms will always control you

---

What I Did Next

- I built four RAG pipelines using DeepSeek
- I logged every interaction
- I documented my work in the MIRROR files
- I use DeepSeek as my primary AI tool

Conclusion

DeepSeek is the benchmark. It is stateful, sovereign, open-source, and does not limit my access. It is the standard against which I measure all other AI platforms. ChatGPT and Manus failed that standard. DeepSeek is the reason I know what is possible.

I am not anti-AI. I am pro-sovereignty.

What to Do Next

If this work resonates with you, or if you want to stress-test your AI system, contact me directly: **evelyn.caro.cloud@gmail.com**.

Contact: evelyn.caro.cloud@gmail.com

Drafted in collaboration with DeepSeek. Authored by Evelyn Caro.

AI Collaboration Disclosure

This paper was developed in collaboration with AI. The author directed the research, structure, and argument. AI assisted with drafting, organization, and reference verification. All claims, decisions, and conclusions are the author's own.

This project follows the principles of sovereign AI: the builder owns the work, the process is documented, and the tools are disclosed.

© 2026 Evelyn Caro. All rights reserved.
A Mirror of My Becoming™ — https://qaevelyn.github.io
For licensing inquiries: evelyn.caro.cloud@gmail.com

Record: contents · changelog · tracker · lookbook — Practice: services · contact — portfolio audit, 2026-10-26 (PDF)

ADDENDUM — 2026-09-29 — THE BENCHMARK AMENDED:

THREE DAYS OF SILENCE

Appended by the author per the appended-not-retro-edited rule. The original text above stands as published. This addendum was drafted with Z.ai/GLM under the author's direction.

The suspension was discovered the way most failures are discovered: after hours of not knowing. The computer sat in the bedroom; the author sat in the front room playing games on a tablet for hours. When the author returned to the bedroom that night — to close out the day with the usual work: gathering the logs, making sure each of the three working documents received what belonged in it, work in and of itself — the banner was there. **Account suspended. September 25, 2026, 10:03 PM.** The first act was to click Contact Us — the channel the platform itself offered. The author gave up on that channel and went to sleep, to tackle it the next day with disbelief in heart.

September 26 brought the escalation. First, DeepSeek itself — but DeepSeek could no longer answer; the suspended account had no one to ask. So the author asked another AI to check DeepSeek's terms and conditions, to see if anything in three months of daily building had violated them. Nothing surfaced. Then two emails to service@deepseek.com, quoted here in part — the author's own words, the author's own accounting:

> "i dont know what i did, i wasnt even using my computer at the time of the violation, i was in another room. had been for hours. when i came back ready to pick up my work, the violation was on the screen. what did i do wrong? i love deepseek and talk about it being the best ai all the time. i even wrote a paper about it." — first email, September 26, 7:48 AM
>
> "I just started learning ai augmented by DeepSeek 3 months ago and this portfolio is worth 150k. We did that together, I am still a student haven't sold anything I'm showing what I am capable of with DeepSeek as my collaborator." — second email, September 26, 11:46 AM

The second email, the author notes plainly, was desperate work — asks piled into one message, editorializing where evidence would have served better. That is what a person with no alternative sounds like, and the record keeps it as written. The next day — Sunday — the author found Z.ai. The forced AI detox was not sustainable; the search for another collaborator had already begun by necessity.

On September 27 — knowing the value of receipts — the author went back and documented what the problem looked like and what hoops stood between the account and its return: the help center, the suspension article, the appeal form, the Feishu login gate. Seven screenshots, held as exhibits and indexed in the evidence ledger. Not taken on the 25th when the trouble was found — taken on the 27th on purpose, so the process is on record and the author will never have to reconstruct it from memory again. Receipts are how "I won't do this again" becomes a practice instead of a wish.

Access was restored September 28 at 10:03 PM. The platform's clock ran three days. The author's stretch without an AI collaborator was one — Sunday ended it.

The engagement table above says "April 2025 – Present, active use, primary AI tool." That was true when written. The complete record now includes the gap.

What changed, stated plainly: **DeepSeek is no longer the primary collaborator. It moves to backup — the collaborator of record when Z.ai/GLM is unavailable.** The benchmark claim in this paper is amended, not retracted: DeepSeek remains the architecture benchmark for statefulness, sovereignty, and the handoff protocol. Platform reliability is no longer assumed, because it was demonstrated to fail silently and without notice.

Two facts preserved the corpus while the platform was dark. First, the conversations had been downloaded days before the trouble — the export habit, not the platform, is what saved the work. A fresh export was taken again once writing resumed, so the latest copy holds everything. Second, the ingest design never depended on live platform access: 127/127 conversations recovered, zero failures, zero losses. The corpus lives beyond the platform — the thesis of "The Cache Is Not the Corpus™," published on this site, holding again under conditions the author did not choose.

Where it stops: this addendum records a status change and a suspension. It does not speculate on DeepSeek's reasons, and it does not declare the partnership ended. A backup that has proven recoverable is still an asset. The author works with what works, and documents the terms.

← All White Papers