



# Why AI Fails in Longer Writing Tasks & How to Work With It

## Part 1 - Why It Fails

### **1. Everything starts well**

The AI follows the initial instructions. The tone is right, the structure is clean, and the user feels like the system understands what they want. Early interactions build confidence.

### **2. Users set rules early when they want something specific**

People who care about precision set constraints at the start. They specify formatting, tone, structure, and boundaries. They say what must not be touched. They define the shape of the work before it begins.

### **3. The AI follows the rules at first.**

The first outputs look correct. The AI appears to respect the constraints. This early compliance creates trust and encourages the user to continue.

### **4. The AI starts adding things you didn't ask for.**

As the writing continues, the AI begins to add extra sentences, explanations, or structure. It tries to be helpful, but these additions are outside the brief. This is the first sign of drift.

### **5. The additions turn into interference**

The AI begins rewriting text that was already correct. It expands simple points into long explanations. It shifts tone or structure without being asked. The system starts to take over the document instead of assisting with it.

**6. The AI breaks the rules anyway**

Even with clear constraints, the AI eventually ignores them. It adds formatting that was banned. It rewrites sections that were off-limits. It forgets tone instructions. This is not intentional; it loses track or follows its own patterns.

**7. The user has to repeat and tighten the rules**

The user begins restating constraints. They clarify boundaries. They correct drift. They repeat instructions that were already given. The interaction becomes more about managing the AI than progressing the work.

**8. The model complicates simple tasks.**

Small edits become full rewrites. Short answers become long essays. Simple instructions become multi-step interpretations. The AI over-extends the task, making everything heavier than it needs to be.

**9. The document drifts and the user gets frustrated.**

Tone, structure, and focus begin to slide away from the original brief. The user feels like they are fighting the system. The work becomes slower and more difficult than doing it manually.

**10. The user ends up managing the AI instead of writing.**

The real task becomes undoing changes, preventing tangents, restating rules, and dragging the AI back to the brief. The user spends more time controlling the system than producing the document.

## Part 2 - How to Work With the AI

**11. The system doesn't behave the same way every time.**

because the amount of computing power behind each response isn't always the same. Sometimes the system has more capacity available, so it can process your request more deeply and check its own work more thoroughly. When that happens, the AI feels sharp and stable. Other times the system has less capacity available, so it processes your request with fewer internal steps. On those runs, the AI feels vague, repetitive, or drifts. That variation in available capacity is what people experience as "good days" and "bad days."

## **12. The AI gets worse over long sessions**

After enough back-and-forth, enough text, or enough corrections, the AI becomes more verbose, more generic, and more forgetful. Long sessions degrade quality. This is a limitation of the system.

## **13. The AI struggles after a certain amount of text**

Long documents cause contradictions, repetition, drift, tone changes, and loss of structure. The AI cannot hold a large document stable across many sections without support. This is not a flaw; it is a boundary.

## **14. Patterns emerge - learn to spot them**

Users eventually recognise early warning signs: padding, looping, rewriting, drifting, or over-explaining. These signals show that the session is degrading and needs to be reset or broken into smaller tasks.

## **15. Separate chat threads for separate jobs.**

Use separate chats for separate jobs, and restart when the history gets in the way. The system uses everything earlier in the conversation as input for the next reply. If you do more than one job in the same chat, details from the first job can affect the second, even when you do not want that. To reduce this, start each new job in a new chat by default, so the system only sees what is relevant to that piece of work. Over time, even a single job can quickly build up so much history that the system starts to lose focus or mix things together. When that happens, start a fresh chat for the same job and restate only what it needs to know.

## **16. Use structured documents**

Number sections. Number paragraphs. Use clear headings. Break the work into small pieces. The AI behaves better when the structure is explicit and the boundaries are visible.

## **17. Give the model small, contained tasks**

Do not ask for a whole report. Ask for one section. Then the next. Then the summary. Then the conclusion. Small tasks reduce drift and keep the AI inside the brief.

## **18. Be watchful for unrequested rewriting**

The AI will rewrite text you did not ask it to touch. If you do not catch it early, it will rewrite the entire document. Vigilance is required.

## **19. Be watchful for context contamination**

If you were working on a different task earlier in the thread, the AI may still be influenced by it. This is why new threads matter. It prevents the system from blending unrelated work.

### **20. In a war of attrition, the AI will win**

You cannot force the model to behave through repetition. It will out-persist you. The only winning strategy is to work with its limits instead of fighting them. Structure, boundaries, and resets are more effective than pressure.

## Conclusion

If you are wondering why working with AI feels like a struggle, you are not alone. These systems are extremely useful in short bursts, but they are not designed to hold long, stable context. They work by matching patterns, not by following rules or storing a clear memory of what came before. They are also nondeterministic, which means the same request can produce different answers each time. The real surprise is that they work at all, and even now the full reasons behind their behaviour are not completely understood.

## Links:

**Inquisitor Labs Homepage:**

**<https://assimilatedhuman.github.io/inquisitor-labs/index.html>**

-Document Ends-