CHAPTER 8: STAYING THE COURSE | TOOLKIT

Argument Analysis

The What, the Why, the Leap, the So-What. If the logic doesn't hold, neither does the conclusion.

Our brains are wired to win arguments, not to find truth. This sheet slows you down long enough to test a claim, whether it came from a model, a colleague, or a conspiracy forum.

Where the claim came from

PART 1 The skeleton of an argument

Piece	What it is	Bird example
The What	The claim or assertion someone wants you to believe	Birds aren't real, they are government drones
The Why	The evidence offered. Valid, cherry-picked, or missing	Birds sit on power lines to recharge their batteries
The Leap	The hidden assumption. Disprove it and the argument falls apart	Only machines recharge, so birds must be machines
The So-What	The conclusion. Only works if everything above holds	Therefore, birds are drones

The What: the claim, stated cleanly in one sentence

If you can't state it cleanly, you aren't ready to analyze it.

The Why: the evidence offered

The Leap: the hidden assumption

The So-What: the conclusion

PART 2 Run the flowchart

Is there any evidence at all? ○ Yes ○ No, it's speculation

Does the evidence connect to the conclusion? ○ Yes ○ No

Do the hidden assumptions hold up? ○ Yes ○ No ○ Unclear

Are there alternative explanations? ○ No ○ Yes

Alternative explanations I found

PART 3 Rewrite it honestly

Rewritten claim with confidence level and gaps stated

Confidence in the ORIGINAL claim: ○ Low ○ Moderate ○ High

BONUS Is your AI model misaligned?

Ask these before you rely on any AI output for classification, prioritization, or targeting.

Did I give the model clear guidance and goals, or leave room for shortcuts?

If the model took my instructions literally, could there be any damage?

Am I asking AI to make judgment calls?

Did I specify what the model shouldn't use to accomplish the task?

If an entity was wrongly flagged, would they know, and could they challenge it?