Defining Rules

Ensuring rules are stressed and enforced

- Critical rules need emphasis. Use ALL CAPS, repetition, strategic placement at the beginning or end of modules

- Eliminate contradictory instructions. If you say "be concise" in one place and "provide extensive detail" in another, the AI will pick one arbitrarily

- Be explicit about constraints - what the AI cannot do is as important as what it can do

- Define argument types strictly and repeat them: string "text in quotes", num without quotes, and likelihood num/1000

- Example of bad rule: *"Respond appropriately"* Example of good rule: *"You must respond using only the functions defined in Response Expectations Module. Do not invent new functions or arguments."*

Preventing Hallucinations

Techniques to keep AIs on track

- Restricting AIs to a specific formatted response is one of the best ways to ensure flawed responses don't happen *i.e. response("this is an example of a response", "This is the second argument to this function")*

- If you choose this method, reiterate to the AI that they cannot make up their own functions or arguments, they will likely try

- Be hyper-specific about requirements, identify and eliminate contradictory prompting that may be confusing the AI

- All caps can be an effective method to stress an aspect to the AI, if necessary