

# ***MANAGING CONTEXT WINDOWS FOR 24/7 LONG-FORM INTAKE***

LLMs have a finite memory (Context Window). In long-form intake sessions—like legal or financial consulting—basic AI will "forget" the beginning of the conversation by the time it reaches the end.

## **VARIABLE-BASED MEMORY**

We solve this by offloading critical data into Voiceflow variables. Variables are immune to context window drift. While the LLM's conversational memory might fade, the hard data (Name, Budget, Conflict\_Check) stays locked in the system state.

## **MEMORY SUMMARIZATION**

For exceptionally long sessions, we implement a "Summarization Node" that periodically condenses the interaction history into a single paragraph, ensuring the LLM always has the essential context without the fluff.