Cheat Sheet: Temporal for AI

Temporal has an answer for every one of the key elements of an AI application or agent. Use this cheat sheet to map AI capabilities to Temporal features, so you can build production-ready AI applications.

Key element in AI	How Temporal addresses it
We stitch together a bunch of steps into **chains, graphs or agentic loops**.	This is a Temporal Workflow. Workflows are equally well-suited to designs that structure steps at design time as those that are dynamic.
We use **Large Language Models** (LLMs).	You invoke these through Temporal Activities, delivering resilience out of the box.
LLMs determine the **execution flow** for agents.	Temporal Workflows implement the agentic loop, while the LLM drives what happens in each turn and the number of turns executed. This LLM-driven flow is recorded in the Workflow Event History and is used to recover execution in the event of failure.
We invoke **tools**, **craft prompts**, and access **resources**. We also invoke MCP servers via **MCP clients**.	You invoke these through Temporal Activities, delivering resilience out of the box. When tools are exposed via MCP servers, the MCP client is implemented within an Activity.
We implement **tools via MCP Servers.**	These are implemented as Temporal Primitives (Workflow, Activity, Signal, Query, Update, etc).
AI applications, especially AI agents, are responsible for providing **memory**.	You just manage your application state in variables in your Workflow. As an added bonus, with Temporal, that state is durable.
We use **checkpointing** to keep from having to rerun steps if a process crashes.	Temporal delivers this implicitly through its event sourcing and state management architecture. You never have to think about checkpointing.
We must allow for **humans in the loop**.	This is achieved through Temporal Signals & Updates, and Temporal Queries.
AI applications are often **long-running**.	Temporal Workflows are state-of-the-art for long-running operations, with a Worker architecture that allows Workflows to wait for arbitrary lengths of time, without consuming resources.

Temporal