

Autonomous AI Control

A concise internal review brief for candidate judgment, delegated authority, result linkage, and versioned update activation.

Problem

Autonomous systems can possess tool access, produce locally valid but incompatible actions, repeat Human Gates, or over-generalize one outcome into a live rule.

Position



Core separation

Outcome evidence may inform a candidate update. It does not directly become a live operating rule.

Activation boundary

Validate applicability, authority, version, expiry, Human Gate, and rollback before a candidate update becomes an active version.

First evaluation

Choose one candidate action, delegated authority path, or result-to-update-candidate sequence.

Buyer questions

- Which action can change an external state?
- Who currently owns authority?
- How is the result linked to the original candidate and judgment?
- Can an activated version expire, be invalidated, or roll back?