Choosing your first policy

The defaults are chosen so that switching Sentilai on changes nothing about how your developers work. That is deliberate. The mistake to avoid is turning everything to **block** on day one: the tools break in ways your team cannot diagnose, and they learn that the way to get work done is to stop using the governed path.

Here is a sequence that works.

Week 1 — see everything, block almost nothing

- **MCP default action: Report.** You will get a full inventory of the MCP servers your developers actually use, without interrupting anyone.
- **Risk classifier sensitivity: Medium.**
- **Risk detectors:** leave at their defaults. Secret scanning warns, credentials and personal data report, lethal-trifecta warns.
- **Prompt capture: off.** Turn it on later, deliberately, and tell your team first.

At the end of the week, **MCP Inventory** shows you what your team is really connected to. Most organizations find at least one thing they did not know about.

Week 2 — decide about MCP

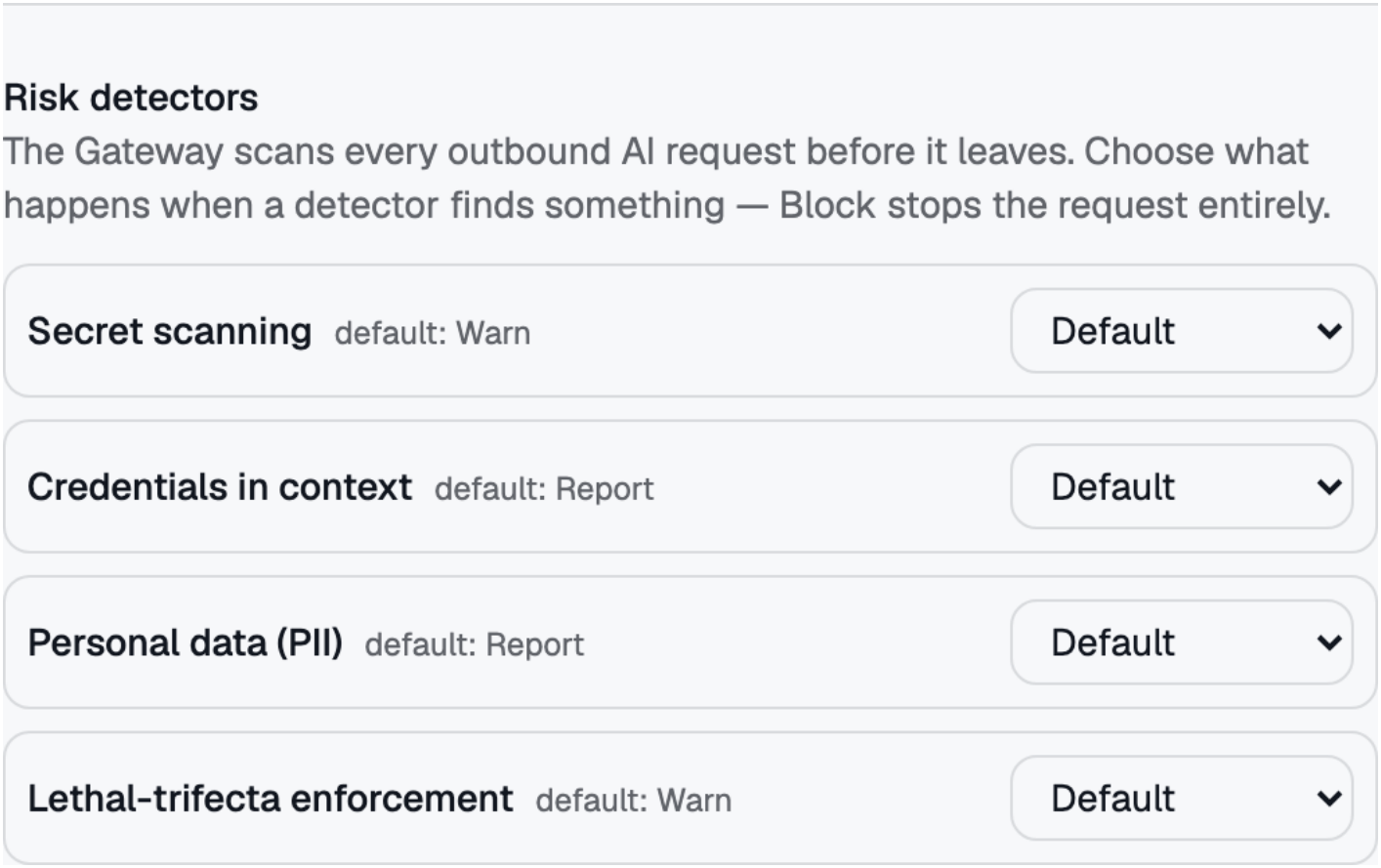
Go through the inventory and set an explicit rule per server: allow the ones you recognize, block the ones you do not. Then turn on **Require approval for new MCP servers** with a pending action of **Report** or **Warn**, so anything new shows up before it becomes normal.

Week 3 — start enforcing

- Move **Secret scanning** to **Block**. This is the one nearly everybody agrees on: an API key pasted into a prompt should not leave your network. Note that detected secrets are blocked regardless of the sensitivity dial.
- Move **Lethal-trifecta enforcement** to **Block** if your team uses MCP heavily.
- Consider **MCP default action: Block** now that the servers you approve have explicit allow rules.

What to tell your team

Tell them before you start, not after the first block. Two facts do most of the work: their subscription still pays for their usage in subscription mode, and nobody is reading their prompts unless prompt capture is on — in which case say so explicitly.



The Risk detectors section. Starting on Report or Warn, and moving to Block once you have seen a week of real traffic, is the sequence that avoids teaching your team to route around you.

Revision #11
Created 2026-08-02 10:19:50 UTC by Sentilai Docs
Updated 2026-08-06 04:31:31 UTC by Sentilai Docs