Human-in-the-Loop Is a Design Choice. Not a Law of Nature.

AI systems can now out-perform human reviewers at oversight — but EU and Singapore governance regimes still assume a natural person is ultimately responsible. Here's how the architecture actually shifts.

SWIPE →

Four things the governance model must redesign.

01 Human oversight — from manual review to supervising the checking AI

02 Multi-layered controls — two independent AI layers, not one

03 Auditability — logs across operational AI, oversight AI, and human decisions

04 Human fallback — a defined path to intervene, suspend, or deactivate

2

At least two independent AI layers, minimum.

The operational AI and a separate assurance or monitoring AI, with different training data, governance, and change control — enforcing separation of duties.

3

Dimensions where automated-first oversight already wins.

Error detection at scale across full telemetry volume, consistency without reviewer fatigue, and structured dry-run simulations before execution.

1

One thing regulators hold constant either side: a natural person must be able to intervene, suspend, or deactivate a high-risk system, even when day-to-day checking is automated.

"

The key shift is from “human as last line of defence on every decision” to “human as designer and governor of a layered control system containing multiple AI components.”

Terence Kok
on redesigning oversight architectures

Transaction checking vs. governing the stack.

HUMAN-IN-THE-LOOP (OLD)	AI-ON-THE-LOOP (REDESIGNED)
Reviews every individual decision	Configures policies, constraints, guardrails
Becomes a symbolic approver under volume	Escalates only genuinely high-risk cases
Accountable for outcomes, not thresholds	Accountable for the control system itself

Not whether a human clicked approve.

The EU AI Act, NIST AI RMF, ISO/IEC 42001, and Singapore's Model AI Governance Framework all converge on one test: can the organisation show its control layers are explainable, monitored, and subject to accountable human oversight.

Three layers, one control stack.

Operational agents, independent oversight agents, and a human-in-command layer — full framework:
terencekok.com/blog/from-human-in-the-loop-ai-on-the-loop-redesigning

Design the control system. Then govern it.

The realistic path to scaling AI in high-risk domains isn't full manual oversight or full autonomy — it's humans accountable for designing, tuning, and auditing a layered AI control stack.

Terence Kok · Enterprise AI Strategist

SAVE · COMMENT · FOLLOW