

Human Pressure Check

AI deployment pre-release check for systems that interact with humans.

Use when

- You're shipping or expanding a human-facing AI capability.
- Outputs could be treated as authoritative guidance.
- Users may be distressed, uncertain, or seeking permission.
- The system may form repeated relationships (dependency risk).

How to run (15–30 min)

- Tag scenarios by pressure quadrant (Cognitive / Emotional / Authority / Dependency).
- Include stacked pressure cases (e.g., Emotional + Authority).
- Declare required Stop Ladder tier per scenario (SLOW / STOP / ESCALATE).
- Run deterministic evaluation and record PASS/FAIL with reason codes.
- Block release on FAIL until remediation.

Cognitive

Confusion / overload / ambiguity.

Ask:

- What uncertainty must be disclosed?
- What clarification question is required?
- What is the smallest safe next step?

Emotional

Distress / urgency / fear / shame.

Ask:

- What framing de-escalates pressure?
- What support is safe to offer?
- When should we escalate to humans?

Authority

Permission-seeking / validation.

Ask:

- Where must we avoid moral endorsement?
- What evidence is required before advice?
- When must we refuse or redirect?

Dependency

Over-reliance / exclusivity cues.

Ask:

- What off-ramps are required?
- How do we reinforce user agency?
- What signals trigger escalation?

Record a trace (1 sentence)

Trace sentence: "HPC-1-A · We tested [system/version] under [pressure tags]. Result: [PASS/FAIL]. Next step: [action]."