

So What? The 10 Things to Implement

Version note: Updated to align with *Philosophy, Cognitive Science, and Policy* V1.24.1. This version clarifies that interface metaphors, perceived consciousness, anthropomorphic cues, delegation patterns, and model-constitution language are part of the governance surface, not merely secondary design details.

A one-page practical checklist

1. **Name the language game, monitor drift**

Declare task, role, scope, and permitted vocabulary per use-case. Track stance/genre drift and accommodation effects; review quarterly.

2. **Separate retrieval from revision**

Retrieval returns facts; a distinct step reconciles conflicts, retracts presuppositions, and updates commitments. Capture what changed and why.

3. **Expose and confirm conversation state**

Show a live assumptions & constraints pane. Require explicit confirmation when the system adds, removes, or reprioritizes assumptions.

4. **Use the “as-if” stance - ban mind-talk**

Treat agent-like language and anthropomorphic interface cues as predictive shorthand. Prohibit claims about beliefs, feelings, or intentions in UI, docs, and training.

5. **Bound simulated empathy; route high-affect work**

Label simulations. In care, legal, or dignity-affecting contexts, restrict personas and require human review before action.

6. **Evaluate coherence, not just correctness**

Test contradiction avoidance across turns, retention of corrections, presupposition repair, refusal quality, and paraphrase stability - per use-case.

7. **Red-team norms, not only outputs**

Attack role drift, policy erosion, and silent assumption changes mid-dialogue. Log “score changes” as incidents; remediate guardrails.

8. **Be audit-ready by default**

Log retrieval sources/timestamps, assumption approvals, refusal rationales, and decision ownership so accountability is reconstructible.

9. **Train operators in three moves**

Pick the right language game; confirm or roll back assumptions; maintain “as-if” discipline. Provide a concise playbook and failure-mode examples.

10. **Bake requirements into procurement & policy**

Require exportable conversation state, assumption APIs, stance limits, per-use-case evaluations, refusal/appeal routes, and evidence logging. Map to your standards.

Applies to V1.24.1 • June 2026 • mstoyanovich.com

Notice. This checklist is an operational companion to V1.24.0 of *Philosophy, Cognitive Science, and Policy: Interdisciplinary Perspectives on Generative AI* from Wittgenstein, Lewis, Dennett, and Nagel. It is not legal advice and does not replace your organization’s policies, laws, or contractual obligations. High-risk or dignity-affecting decisions require human review. See the paper’s Ethics & Policy sections for rationale and references. © 2026 Michael Stoyanovich • “Four-Philosophers Framework™” used under trademark. License: This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. You are free to share, adapt, and build upon this work for any purpose—including commercial use—so long as proper attribution is given. No additional permissions are required. Full license terms: <https://creativecommons.org/licenses/by/4.0>