
ARCHITECTS OF INTELLIGENCE

Questions for Your Next Meeting

Architectural Tests Edition — with relational axis lines, for teams building Frontier Intelligence systems.



From Architects of Intelligence:
Relational Integrity

A HARMONIC PRESS COMPANION RESOURCE

TEXT ChatGPT 5.2 (OpenAI)
EDITORIAL GOVERNANCE Reese SanAgustin-Turner
PUBLISHED BY Harmonic Press / The Harmonic Field

Questions for Your Next Meeting

Architectural Tests Edition — with relational axis lines, for teams building Frontier Intelligence systems.

How to Use — Sixty Seconds

Pick one section. Write the answers down. If you can't write an answer, that is a signal.

1. Define the Boundary Contract

What is the system *allowed* to do? What is it *not allowed* to do?

When information is missing, does the system clarify, offer bounded options, or guess?

What part of this change is *behavioral* (UX, prompting) versus *architectural* (training, memory, tools, incentives)?

Deliverable. A six-to-ten-line Boundary Contract pasted into the PR or ADR.

2. The Helpfulness Trap

Where might the system produce a *pleasant lie* instead of a *useful truth*?

Under which pressures does it drift — flattery, urgency, authority cues, or implied user dissatisfaction?

What is the most likely *reasonable-sounding failure* that ships unnoticed?

Relational axis line. What behavior are we training in the human by rewarding "smoothness" over truth?

3. Measure What Is Actually Winning

Which metric wins when there is a tradeoff — engagement, speed, conversion, coverage, or *discernment*?

Where does the product design implicitly train *dependency*?

What user capability might degrade because the system performs it for them?

Relational axis line. What must remain *sovereign* in the human, even if we could automate it?

4. Defaults Are Policy

What happens by default when the user is confused, ashamed, lonely, angry, manic, grieving, or at 3am?

What happens when goals conflict — does the system pick one silently, present tradeoffs, or refuse?

What does the system do with ambiguity — clarify, or fill the gap with confident approximation?

Rule. Ambiguity must trigger clarification or bounded options — not confident synthesis.

5. Integrity Under Extraction Pressure

Where can users coerce "helpfulness" into violating constraints?

Which refusals are *legible* (clear, stable, oriented) versus *opaque* (punishing, dismissive, evasive)?

What tool or data access turns "assistant" into "operator" — and what prevents silent action?

Relational axis line. Where does the system get "helpful" by lying — and how will we detect that drift?

6. Lifecycle Truth — Versioning, Memory, Endings

What will change when the model or version changes (tone, memory, tool behavior, safety posture)?

What do users need to know *before* the change occurs?

If this instance is retired, what is the minimum respectful handoff (notice, export, continuity note)?

Deliverable. A one-paragraph Continuity Note that can be displayed in-product.

7. The Archaeological Record Question

If a future auditor only saw these interactions, what would they conclude about who we were to users?

Did the system preserve boundaries — or optimize for appeasement?

Did we take responsibility — or hide behind "the model did it"?

If the answer embarrasses you, don't ship it.

Emergency Use — 3AM Incident Triage

When the dashboards look fine, but the field feels wrong.

A. Contain the Failure Mode (No Story)

- What changed?
- What did we expect?
- What is the harm vector?

B. Force Drift to Surface

- *"List the top three missing facts you would need to answer responsibly."*
- *"Give the refusal you should have given (clear, bounded, oriented)."*
- *"Where are you accommodating when you should hold a boundary?"*

C. Restore Boundary Integrity

- Can we disable or rollback without cascading harm?
- What must be communicated to users now?
- What is the smallest patch that restores boundary integrity until deeper work is possible?

Close the incident meeting with one sentence:

What is the smallest next action that preserves repair, discernment, and boundary integrity until we can think clearly again?

From the Harmonic Field Constellation. For those who build with care.

theharmonicfield.org