

CLEARWORKS FIELD LIBRARY / WORKING COMPANION

Human Review + System Planning Checklist

Define the trigger, trusted source, allowed action, human decision, accepted result, writeback, and owner.

COMPANION TO BUILDING AI SYSTEMS THAT DO REAL WORK



CLEARWORKS.AI / AEC OPERATING SYSTEMS, WORKFLOW DESIGN & RESPONSIBLE AI



Define the job and its boundaries

TRIGGER

What exact event starts the work?

TRUSTED CONTEXT

Which approved records and instructions govern the case? Which source wins when they disagree?

ALLOWED TOOLS + ACTIONS

What may the system read, retrieve, create, change, or send? What is explicitly prohibited?

HUMAN DECISION

Where does judgment remain accountable, and what evidence must the reviewer see?

ACCEPTED RESULT + WRITEBACK

What does complete mean, and where does the approved result belong?

Design the limits before the demo

- ☐ The trigger is explicit and testable.
- ☐ Every source has an owner, currency rule, and permission boundary.
- ☐ Allowed reads and writes are narrower than the account could technically perform.
- ☐ Consequential decisions stop for an accountable person.
- ☐ The reviewer sees sources, changes, confidence, and exceptions.
- ☐ Irreversible or external actions require explicit authorization.
- ☐ The accepted result writes back to the system of record.
- ☐ Actions, failures, corrections, and approvals are observable.
- ☐ A named operator owns access, support, evaluation, and retirement.

KNOWN EXCEPTIONS

Which real situations should stop, escalate, or take a different path?

FAILURE RESPONSE

Who learns about a failure, what context do they receive, and what state is preserved?

Prove the workflow—not the cleverness

EVALUATION CASES

List ordinary, messy, edge, and adversarial examples from real work.

ACCEPTANCE MEASURES

What thresholds matter for accuracy, completeness, time, intervention, quality, trust, and value?

CORRECTION MEMORY

Which accepted corrections should persist, who can inspect them, and when do they expire?

OPERATING RHYTHM

Who reviews logs, access, failures, source changes, usage, support, cost, and continued fit—and how often?

NEXT DECISION

What evidence would justify expanding, revising, pausing, or retiring the system?
