

PRACTICAL SYSTEMS GUIDE

Building AI Systems That Do Real Work

How bounded AI systems use context, tools, review, memory, and writeback without losing human accountability.

DESIGNED GUIDE + HUMAN-REVIEW AND SYSTEM-PLANNING CHECKLIST



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



More than a chat window—less than a digital employee

A useful system finds the right information, does only the steps it is allowed to do, asks a person when judgment matters, and saves the approved result where the team can find it.

EXAMPLE

A proposal-preparation system might collect current project proof and resumes, draft an assembly, flag missing evidence, and route it to a pursuit lead. It should not invent credentials, choose the final fee, or send the proposal.

Five production patterns

01 Document processing

Receive, identify, extract defined fields, validate against rules or source records, and route exceptions—for invoices, intake forms, contracts, submittals, and reports.

02 Client communication

Assemble meeting briefs, summarize long threads, draft routine replies, and capture commitments using relationship, project, and decision context.

03 Operations monitoring

Watch overdue approvals, missing records, budget drift, failed automations, or stalled queues—and escalate with the evidence needed to act.

04 Research + intelligence

Collect from an approved source set, remove duplicates, compare claims, and produce a cited brief that preserves links, dates, and uncertainty.

05 Workflow orchestration

Move work safely between systems after known events, preserving review gates and an accepted record.

Seven pieces every useful system needs

The limits matter as much as the capability.

LAYER	QUESTION	FAILURE WHEN MISSING
Trigger	What starts the work?	The system runs at the wrong time.
Context	Which records and instructions govern this case?	Fluent output with the wrong facts.
Tools	What may the system read or change?	No useful action—or excessive access.
Review	Where does human judgment remain accountable?	Unreviewed commitments or professional decisions.
Memory	What approved state should persist?	The system repeats mistakes and loses decisions.
Writeback	Where does the accepted result belong?	Another draft stranded in a chat window.
Owner	Who operates and improves it afterward?	Access, errors, and workflow drift go unmanaged.

ACCEPTANCE

The work is not done because AI produced an answer. It is done when the right person accepts it and the team's record shows what happened.

Start with the work—not the model

Follow one real workflow from trigger to accepted outcome. If the team disagrees about what should happen, who owns it, or which record is correct, clarify the work before automating it.

- 01 Inventory recurring coordination, assembly, checking, searching, and follow-up.
- 02 Score frequency, delay, error cost, data readiness, judgment, and reversibility.
- 03 Choose one trigger, one team, one accepted output, one review owner, and real evaluation cases.
- 04 Name the authoritative record, normal path, exception path, and destination for the reviewed result.
- 05 Test with real, messy cases before expanding tools, permissions, or audience.

Common mistakes

Automating instability

Agree on owner, normal path, exceptions, and accepted result first.

Buying before choosing

Choose the workflow and success measure before choosing a platform.

Testing only the happy path

Build an evaluation set from real cases and known failure modes.

Removing judgment

Keep people responsible for commitments, sensitive communication, professional decisions, and irreversible actions.

No destination

A draft in a chat window is not finished work. Decide where the reviewed result belongs.

No operating owner

Keep an action record, limit permissions, and name who maintains the system when the work changes.

Budget for the system—not only the model

\$5k–\$25k

COMMON PLANNING RANGE FOR A BOUNDED PILOT

\$25k–\$100k+

COMMON PLANNING RANGE FOR A PRODUCTION SYSTEM

These are planning ranges, not quotes. Existing systems, integration quality, security requirements, and exception volume change the work substantially. Production systems connect real data, permissions, interfaces, evaluation, monitoring, and writeback.

Budget for ongoing operation: model and API usage, monitoring, source and prompt updates, access reviews, evaluation, user support, and ownership when the underlying workflow changes.

USEFUL FIRST BUILD

Choose one bounded system that makes one piece of work visibly better without asking the whole company to change at once. Bring the job that is giving people trouble—not a shopping list of AI tools.

Canonical web edition: <https://clearworks.ai/resources/building-ai-agent-systems>

Source snapshot SHA-256: 2ffb22bc6c8ec8055243d162a57b7fb36797175ac4b6687e2579ca77f427e239