The Easton Way. How we build AI agents that get installed.

A short walkthrough of the methodology we run on every build. Eight phases, two hard gates, and the 93 steps in between. What you'll see, what we need from you, and why it pays back inside the first month.

PREPARED BY
Jordan Bailey · Easton Consulting House

VERSION
1.0 · May 2026

FOR
Operators evaluating an AI install partner

The shortcut.
Why projects stall, and what we do about it.

Most AI projects don't fail on the tech. **They fail on the install.** Models are ready, integrations are ready, the missing piece is the operating discipline that turns a prototype into a system the team actually uses.

The Easton Way is the methodology we run on every build. Eight phases, two hard gates that block the next step, a 93-step checklist that runs end to end. The asset you're reading is the field guide.

BUILD LABOUR PER AGENT

~48 hours

From signed PRD to closed retro. Includes the two-week stabilization window scoped into every build.

OPERATIONAL CHECKLIST

93 steps

Cloned from the master template at the start of every project. Reproduced in full in the appendix of this document.

What changes when there's process behind the prompts

- **One signed PRD per project.** Produced in week one. Signs off goal, inputs, outputs, success criteria, and the out-of-scope list before any code gets written.
- **Two hard gates that block forward motion.** Discovery blocks Build, UAT blocks Deploy. No verbal sign-offs, no silent slips into production.
- **Two-week stabilization scoped into every build.** Daily log review for the first three days, weekly after. Cost reconciliation on day seven. Bug-only window.
- **Mandatory retro before the project closes.** Locks hours, profitability, and learnings so the next build benefits.

The methodology.
Eight phases, two hard gates.

Roughly forty-eight hours of consulting labour per agent, plus two weeks of stabilization scoped into every build. The two hard gates are non-negotiable; the rest is visual order.

#	PHASE	KEY OUTPUT	TIME
01	Discovery & Spec HARD GATE	Signed PRD, golden examples, stakeholder map	~6 h
02	Architecture & Credentials	Flow diagram, API list, cost estimate	~4 h
03	Build	Working agent, integrations, structured logs	~20 h
04	Internal QA	Edge cases tested, zero P0/P1 bugs	~8 h
05	Client UAT HARD GATE	Pass/fail per criterion, written sign-off	~6 h
06	Deploy & Handoff	Production live, Loom walkthrough, runbook	~4 h
07	Stabilization	Bug-only window, cost reconciled, clear log	2 weeks
08	Closeout & Retro	Hours locked, internal retro, follow-on proposal	~3 h

PHASE 01 · HARD GATE

Discovery & spec.
One job: a signed PRD.

~6 hours · blocks Build until client sign-off

The first phase has one job. Produce a written PRD the client signs off before any code gets written. No PRD, no Build. It's the cheapest place to catch a misunderstanding, and the most expensive place to discover one is week three.

We interview the SME who owns the workflow, map it step by step, and collect five to ten golden examples (real input, real expected output) so QA later has a fixed bar to test against. The phase closes when the client has signed off on the PRD in writing.

KEY OUTPUTS

- Signed PRD attached to the project
- Process map of today's manual workflow
- 5–10 golden examples (real input → expected output)
- Stakeholder + system list
- Hours allocated and expected monthly cost

WHAT YOU BRING

One hour with the workflow owner. A handful of real examples. A list of systems the agent will touch. Sign-off authority on the spec.

> **WHY THIS GATE EXISTS**
>
> Skipping a signed PRD is the single biggest source of unprofitable projects in this industry. A 90-minute conversation in week one is worth a 90-hour rewrite in week six.

PHASE 02

Architecture & credentials.
The half-day that saves three weeks.

~4 hours · prep work before the build starts

With the PRD signed, the next four hours go into the small but boring decisions that wreck projects when they're skipped. We pick specific tools inside the chosen category, draw the agent flow as a single diagram, and request every credential and API key formally so nothing dies in a Slack DM.

Token cost gets estimated against the golden set so the expected monthly run cost is grounded in maths, not vibes. Cost alerts get wired before a single line of agent code ships.

KEY OUTPUTS

- Flow diagram for the agent
- Credential list requested and stored in the vault
- Dev environment, sub-account, GitHub repo provisioned
- Token cost estimate vs expected monthly cost
- Cost monitoring and alerts configured

WHAT YOU BRING

Credentials for the systems the agent connects to. A nominated IT contact for access requests. Confirmation of which platforms are off-limits.

PHASE 03

Build.
The part most agencies sell as the whole project.

~20 hours · only starts once Phase 01 is signed off

At Easton this is roughly forty percent of the work, and it never starts without a signed PRD upstream. Happy path first, tested against golden examples. Integrations get wired. Error handling, retries, and structured logs are first-class citizens, not afterthoughts.

Prompts get iterated until every golden example passes. Cost gets optimised where the wins are obvious. Commits go to GitHub at the end of every session. Hours go into the project the same day, not at the end of the month. Both habits exist so a budget overrun shows up while there's still time to do something about it.

KEY OUTPUTS

- Agent passing every golden example from Discovery
- All integrations wired and authenticated
- Error handling, retries, structured logs in place
- Trigger live (cron, webhook, manual, email, Slack)
- Code committed to a private GitHub repo

WHAT YOU BRING

A nominated SME on call for prompt-iteration sessions. Sign-off authority for small in-scope tweaks that surface during the build.

PHASE 04

Internal QA.
Everything we can break before the client sees it.

~8 hours · runs before UAT opens

Before the client opens the agent, we break it. Every golden example, every edge case identified in Discovery, every failure mode the team can think of. Network pulled, bad input shoved in, auth tokens expired, rate limits pushed.

A 100-run load test confirms cost per run lines up with the estimate. Logs get reviewed for signal vs noise so the runbook isn't reading tea leaves at 2 a.m. Client UAT doesn't open until zero P0 and P1 bugs are open.

KEY OUTPUTS

- Every golden example passes
- Every Discovery edge case tested
- Forced failures handled gracefully
- Cost-per-run measured against estimate at 100× volume
- Zero P0/P1 bugs at UAT opening

WHAT YOU BRING

Nothing on this phase. We run internal QA without taking client time.

PHASE 05 · HARD GATE

Client UAT.
Pass each criterion, in writing, on the spot.

~6 hours · blocks Deploy until written sign-off

The agent stands up in staging. The client gets a live walkthrough, then runs their own tests against real data. Feedback gets captured in a shared doc in real time, and each success criterion from the PRD gets marked pass or fail on the spot.

Production deploy doesn't happen without written sign-off. If sign-off is blocked, the blocker gets named in writing and the project sits in UAT until it's cleared. No silent slips into production.

KEY OUTPUTS

- Live walkthrough recording saved to the project
- Shared feedback doc with pass/fail per success criterion
- Written client sign-off
- P0/P1 issues fixed; P2/P3 logged or deferred

WHAT YOU BRING

The people who'll actually use the agent. Real data to throw at it. A decision-maker who can sign off in writing the same day.

> **WHY THIS GATE EXISTS**
>
> A demo that lands in staging isn't an installed system. If we deploy without confirming success criteria against real data, the post-deploy noise costs us both more than a UAT session would have.

PHASE 06

Deploy & handoff.
Fast, because the work behind it is done.

~4 hours · only starts once Phase 05 is signed off

The agent goes to production, failure alerts get wired to the client and Easton, and a forced test failure confirms the alert actually fires before the team relies on it.

A 10–15 minute Loom walkthrough explains what the agent does, how to trigger it, and what to do when it breaks. A one-page runbook ships as both a ClickUp doc and a PDF. Optional 30-minute Q&A if the build warrants one.

KEY OUTPUTS

- Production-deployed agent
- Failure alerts wired and tested
- 10–15 min Loom walkthrough
- One-page runbook (doc + PDF)
- Final cost sanity-check on production

WHAT YOU BRING

A nominated owner on your side who'll keep the runbook handy and forward the alert thread. Thirty minutes for the optional handoff Q&A.

PHASE 07

Stabilization.
Two weeks, bug-only, included in the price.

2 weeks · scoped into every build

The two weeks after deploy are where most projects quietly go sideways elsewhere. At Easton they're scoped into every build. Logs reviewed daily for the first three days, weekly after. Every alert triaged the same day.

Bug-fixes only. Feature requests don't get folded in for free during this window. They get captured as new project candidates and quoted properly so the budget on the live agent stays predictable.

On day seven, actual monthly cost gets compared against the estimate from Phase 02 and the difference gets explained. On day fourteen, after seven consecutive days clear of P0/P1 issues, monitoring drops to weekly.

KEY OUTPUTS

- Daily log review days 1–3, weekly thereafter
- Actual monthly cost vs estimate (day 7)
- Seven-day clear log of P0/P1 issues
- Quoted follow-on list for deferred features

WHAT YOU BRING

Honest feedback when something looks off. A nominated person to forward alerts to if our monitoring catches anything first.

PHASE 08

Closeout & retro.
Lock the project, feed the next one.

~3 hours · the final phase

The final phase locks hours, profitability, and learnings. The internal retro captures what was estimated vs actual per phase, what broke unexpectedly, and what's worth lifting into the pattern library for the next build.

Final invoice goes out. Client satisfaction gets a final score. If the engagement is retainer-eligible, a follow-on proposal gets drafted while the relationship is warm and the work is fresh.

Skipping this phase is how an agency forgets what blew up and repeats it on the next project. We run it on every build, every time.

KEY OUTPUTS

- Locked hours and profitability number
- Internal retro doc
- Final client satisfaction score
- Final invoice
- Retainer proposal where eligible

WHAT YOU BRING

Honest answers on the final satisfaction check and whether the agent delivered against the success criteria from Phase 01.

The two hard gates.
Where most agencies cut, we won't.

Most of the eight phases are visual order. Two of them are non-negotiable. These are the gates that protect every build from the common patterns that kill AI projects elsewhere.

GATE 1 · DISCOVERY BLOCKS BUILD

No code gets written before the PRD is signed.

A 90-minute conversation in week one is worth a 90-hour rewrite in week six. We won't move into Build until the PRD has named the goal, the inputs, the outputs, the success criteria, and the out-of-scope list, and the client has signed off in writing.

GATE 2 · UAT BLOCKS DEPLOY

Nothing goes to production without written client sign-off.

An agent that demos well in staging isn't an installed system. We hold production deploy until the client has run the agent against real data, marked every success criterion pass or fail, and confirmed in writing they're ready to go live. If sign-off is blocked, the blocker gets named explicitly and the project sits in UAT until it's cleared.

The full checklist.
Ninety-three steps, every project.

Every project Easton runs works against this list end to end. Cloned from the master template in ClickUp, customised only where a client need requires it, and reviewed in retro at the end of every build.

01 Discovery & Spec — 15 steps

01 Schedule 60-min stakeholder interview with the SME
02 Record the interview and save URL to the project
03 Map the manual process being automated, step by step
04 Identify every stakeholder this agent touches
05 List every system or tool the agent will read from or write to
06 Draft lean PRD: goal · trigger · inputs · outputs · success criteria · out of scope
07 Collect 5–10 golden examples (real input → expected output)
08 Enumerate edge cases and failure modes explicitly
09 Estimate hours per phase → fill Hours Allocated
10 Estimate monthly cost → fill Expected Monthly Cost
11 Set Tech Stack field (Workflow / LLM / Voice / RPA / Hybrid)
12 Assign Project ID using [CLIENT_CODE]-[NNN] format
13 Send PRD to client for written sign-off
14 Get approval and save email/Slack screenshot to project
15 Attach PRD as a ClickUp Doc nested under the project

02 Architecture & Credentials — 13 steps

01 Choose specific tools within Tech Stack category
02 Sketch agent flow diagram and attach as image
03 List every credential and API key needed from client
04 Send formal credential request to client
05 Receive credentials and store in 1Password under [Client]/[Agent]
06 Provision dev environment, workspace, or sub-account
07 Create separate API workspaces if needed (OpenAI org, Anthropic project)
08 Provision infra (Modal / Railway / GCP) if applicable
09 Set up GitHub repo if code, private, named easton-[client]-[agent]
10 Confirm test data and golden set are captured in the PRD
11 Estimate token cost on golden set, confirm vs Expected Monthly Cost
12 Set up cost monitoring and alerts on provider dashboard
13 Document architecture decision as PRD addendum

03 Build

12 steps

01 Implement happy path against golden examples
02 Wire all integrations
03 Implement error handling (try/catch, exponential backoff, dead-letter queue if applicable)
04 Implement retry logic where appropriate
05 Add structured logging
06 Implement the trigger (cron / webhook / manual / email / Slack)
07 Manually run all golden examples and verify each passes
08 Iterate on prompts if LLM-based
09 Optimise for cost where there are obvious wins
10 Write inline code comments only where the why is non-obvious
11 Commit to GitHub at end of each work session
12 Update Hours Spent at end of each work session

04 Internal QA

11 steps

01 Re-run every golden example, verify each passes
02 Run every edge case identified in Discovery
03 Force failures (network down, bad input, auth expired) and verify graceful handling
04 Push to rate limits and verify retry/backoff behaviour
05 Run 100× load test, measure cost per run vs estimate
06 Verify trigger reliability (cron fires, webhook receives, etc.)
07 Review logs for signal vs noise
08 Document any known limitations
09 Self-review the runbook draft against actual behaviour
10 Fix issues found and re-run full golden set
11 Confirm zero P0/P1 bugs before opening UAT

05 Client UAT

10 steps

01 Stand up agent in staging environment
02 Schedule 30–60 min UAT session with client
03 Walk client through agent live
04 Have client run their own tests against real data
05 Capture feedback in real-time (shared doc)
06 Mark each success criterion pass/fail
07 Fix P0/P1 issues immediately
08 Defer P2/P3 to backlog or out-of-scope notes
09 Get written client sign-off (email or Slack screenshot saved)
10 If sign-off blocked, document blocker and update Dependencies field

06 Deploy & Handoff

12 steps

01 Deploy to production
02 Wire failure alerts to client and Easton (email / Slack / SMS path)
03 Force one test failure and verify alert fires
04 Update DNS or access for client if needed
05 Final cost sanity check on production
06 Record 10–15 min Loom walkthrough
07 Write one-page runbook (ClickUp Doc + PDF export)
08 Send handoff package to client (Loom URL + runbook PDF)
09 Optional 30-min Q&A call if complexity warrants
10 Add to client's master system inventory
11 Update status to in-progress (now in stabilization sub-state)
12 Update Next Milestone field to "Stabilization end (Day 14)"

07 Stabilization

12 steps

01 Day 1–3: monitor logs daily, watch failure alerts actively

02 Day 7: check actual cost vs Expected Monthly Cost

03 Day 7: fill Actual Monthly Cost field

04 Triage every incoming issue from the client

05 Fix bugs only; defer feature requests

06 Capture each feature request as a new-project candidate

07 Day 10: send weekly update flagging stabilization end approaching

08 Day 14: confirm zero P0/P1 for 7 consecutive days

09 Reduce monitoring to weekly

10 Notify client that stabilization is ending

11 Send deferred feature-request list with quoted prices

12 Final cost reconciliation

08 Closeout & Retro

8 steps

01 Lock Hours Spent (final number)

02 Verify Profitability formula reflects reality

03 Write internal retro doc as ClickUp Doc nested under project

04 Update Client Satisfaction final score

05 Send final invoice if not already sent

06 Move project task to Project Archive list

07 Update status to completed

08 If retainer-eligible, draft retainer proposal as follow-on project

Run your next build through the Easton Way. A field methodology from Easton Consulting House.

Built and operated for clients turning manual workflows into installed AI systems.

WEB

eastonconsultinghouse.com

EMAIL

hello@eastonconsulting.house

DOCUMENT

The Easton Way · v1.0 · May 2026