

Your own audit

The worksheet from Chapter 9 of *AI That Earns Its Keep* · Book 1 of The Implementation Files

You do not need a consultant to start. You need two honest weeks of looking at your own shop. Print this. Write on it. If you only highlight it on your phone, you will skip the hard lines.

No tool recommendations here. Map first. Buy later, or never.

How to use this in seven days

Day 1. Fill Sections A, B and C. Forty-five to sixty minutes.

Days 2-15. Capture Section D metrics only. Do not fix anything yet.

Day 16. Score Sections E and F. Make a go or no-go call in writing.

If you change three tools during the measurement window, you poisoned the baseline. Stop. Restart the clock.

Section A · Map your repetitive text work

List jobs where the words repeat more than the judgment.

Task 1

What

How often

Who does it

Same template most times? Y / N Safe to draft with AI help? Y / N / Not sure

Task 2

What

How often

Who does it

Same template most times? Y / N Safe to draft with AI help? Y / N / Not sure

Task 3

What

How often

Who does it

Same template most times? Y / N Safe to draft with AI help? Y / N / Not sure

Task 4

What

How often

Who does it

Same template most times? Y / N Safe to draft with AI help? Y / N / Not sure

Task 5

What

How often

Who does it

Same template most times? Y / N Safe to draft with AI help? Y / N / Not sure

Prompts to jog memory

- What emails do you send every week that start the same way?
- What reports are 70% last month's report with new numbers?
- What social captions, menu notes, donor updates or recall reminders follow a pattern?
- What intake forms get re-typed into nicer language?

Rule of thumb. If a trained high-school hire could draft it with a good checklist, it belongs in Section A. If a mistake could cost a license, a grant or a lawsuit, park it until Sections C and F are clear.

Section B • Map judgment-only work

List jobs where the value is the decision, not the paragraph.

Judgment 1

Decision

Owner

What breaks if a model “helps”

Keep 100% human? Y / N

Judgment 2

Decision

Owner

What breaks if a model “helps”

Keep 100% human? Y / N

Judgment 3

Decision

Owner

What breaks if a model “helps”

Keep 100% human? Y / N

Judgment 4

Decision

Owner

What breaks if a model “helps”

Keep 100% human? Y / N

Examples from the five shops, for calibration, not copying

- Ray: which route gets the last emergency slot on a 110° day.
- Dr. Ahn: clinical recommendations and anything that smells like diagnosis.
- Teresa: who gets scarce housing help when the list is longer than the rooms.
- Ruiz family: cutting a shift when labor is high but a regular event is tonight.
- Hartnett & Cole: fiduciary advice and final document language.

If you catch yourself writing “AI could probably do this” next to a Section B item, rewrite the sentence as “AI could hide who is accountable.” Then leave it human.

Section C · Map sensitive data

Check every type of data that appears in your weekly workflow.

- | | |
|---|---|
| ☐ Customer, patient or client names and contact info | ☐ Employee records, wages, disciplinary notes |
| ☐ Payment details, account numbers, SSNs, Tax IDs | ☐ Donor or grant beneficiary case notes |
| ☐ Health information, even “just scheduling” | ☐ Attorney work product, confidential client facts |
| ☐ Children's information | ☐ Passwords, badge codes, vendor logins |
| | ☐ Unpublished financials |
| | ☐ Other _____ |

For each checked item, answer

1. May this ever be pasted into a third-party AI tool? Yes / No / Only if de-identified

2. Who has authority to say yes

3. Where the written rule is posted

If you cannot name the person in question 2, your answer to question 1 is No until you can.

De-identify checklist, use before any draft help

- | | |
|---|---|
| ☐ Names removed or replaced with role labels (Patient A, Donor 12) | ☐ Addresses and school or employer details removed |
| ☐ Dates shifted or generalized if they identify someone | ☐ Unique story details that only one person would know removed |
| ☐ Account, chart or case numbers removed | ☐ Second human scanned the prompt before submit |

Section D • Baseline metrics, two weeks

Pick three to five metrics. More than five and you will quit. Example — metric: minutes on after-hours text triage. Count: timer plus tally.

Metric 1

What you track

How you count it

Week 1 total

Week 2 total

Notes

Metric 2

What you track

How you count it

Week 1 total

Week 2 total

Notes

Metric 3

What you track

How you count it

Week 1 total

Week 2 total

Notes

Metric 4, optional

What you track

How you count it

Week 1 total

Week 2 total

Notes

Metric 5, optional

What you track

How you count it

Week 1 total

Week 2 total

Notes

Suggested metric menu, steal only what fits

- Minutes spent rewriting the same email type
- Number of no-shows or late cancels
- Hours to produce the weekly schedule
- Hours to produce the monthly donor or board update
- Number of review replies older than 48 hours
- Count of wrong-list or wrong-person contact errors
- Hours spent fixing incomplete field logs
- Number of drafts that needed factual correction before send

Do not track feelings as your only metric. “Team seems happier” is allowed as a note, not as proof.

Section E • Red-flag list

If any of these show up, pause the experiment the same day.

- ☐ Real sensitive data pasted into a tool that is not approved in writing
- ☐ A citation, statistic, case cite or policy claim you did not verify from a primary source
- ☐ Customer-facing copy published without a human tone check
- ☐ Staff saying they feel watched, replaced or set up to fail, and leadership shrugging

- ☐ A near miss in legal, medical, financial or safety language
- ☐ Owner radio-silence after announcing a new process
- ☐ Lists used for mail, email or text not reconciled to the current system of record
- ☐ Free-tier output treated as grant-ready or court-ready
- ☐ Family or partner conflict turning into secret pilots
- ☐ Anyone saying “the AI said so” as a reason to skip review

Add your industry-specific flags

Section F · Go / no-go thresholds

Fill these before week 1 ends. Do not invent them after you see results you like.

Go, proceed to a tiny pilot, if all are true

- ☐ Sections A, B and C complete
- ☐ At least 10 business days of Section D data captured
- ☐ One repetitive text task identified with a named human reviewer
- ☐ Sensitive-data rule posted and initialed by the owner

Success metric, one sentence

Stop metric, one sentence

Time box, example 14 days

Cash box, example \$0 / \$50 / \$100

No-go, do not pilot yet, if any are true

- ☐ Owner will not put their name on the data rules
- ☐ Staff who must review drafts are already past capacity with no relief
- ☐ The only proposed task is judgment-only (Section B)
- ☐ You cannot explain the pilot in one paragraph to a new hire
- ☐ Week 1 and 2 baseline is missing or fake
- ☐ A red flag from Section E is already active

Pilot scope sentence, write one

For _____ days, we will use AI only to draft _____. A human named _____ will review every output before it is used. We will not touch _____ data. We will stop if _____.

If that sentence is hard to write, you are not ready. That is a successful audit result. Readiness is the point.

One-page Monday starter

If you are overwhelmed, do only this tomorrow morning.

1. List five repetitive text tasks (Section A).
2. Circle one that does not involve sensitive data.
3. Start a two-week tally of how long that task takes.
4. Write the sensitive-data rule in six lines and tape it near the keyboard.
5. Put a 30-minute meeting on the calendar for Day 16 titled Go / No-Go. Invite the people who actually do the work.

That is the whole job for week one of your own audit. Your version only needs to be honest.

From AI That Earns Its Keep: Where AI Actually Fits in a Small Business, and Where It Does Not, Book 1 of The Implementation Files, by Patrick Albert.

© 2026 Patrick Albert / Albert Publishing House. Free to print and use inside your own business.

patrickalbert.com/implementationfiles