NEBIS.AI

A NEBIS FIELD GUIDE

The AI Operating Playbook

For Singapore accounting practices: AI in the firm without compromising quality management, confidentiality or the file.

EDITION B — PRACTITIONERS, PARTNERS & PRACTICE OWNERS

Where it pays, how to design it, what the file has to show, and which judgements can never be delegated — written from inside a practice built this way.

> THE OPERATING IDEA
>
> AI creates durable value when it is attached to a real engagement outcome, bounded by clear rules, and improved through feedback. Everything else is a demonstration.

VERSION 1.0 · AUGUST 2026 · SINGAPORE

Why this exists

There is no shortage of writing about AI in accounting. Almost all of it is written for practitioners rather than by them, and it stops precisely where the difficult questions begin: what goes in the file, who is responsible, what happens at inspection, and which judgements can never be delegated.

We rebuilt a practice around AI systems from the ground up, under the same standards you work to. That forced answers to questions a pilot project never has to face. If a machine prepared the schedule, what is the evidence that a person exercised professional scepticism over it? What does the documentation have to show? How does client confidentiality survive a third-party tool? Which decisions belong irreducibly to the engagement partner?

This playbook is what we would have wanted at the start. It is deliberately platform-agnostic — it names capabilities, not products — because the tool you pick this year will not be the tool you run in three years, and the operating decisions are what survive a change of vendor.

It is written collegially. The argument here is never with practitioners; it is with a set of workflows the profession inherited from an era of paper, junior staff and billable hours, and has not revisited since.

How to use this playbook

Start with an engagement problem, select one workflow, and make the smallest safe change that produces evidence. Do not start with an audit. Start where a mistake is cheap and visible.

Read it once end to end. Then use Chapters 2, 3 and 11 as working tools, and return to Chapters 7 and 8 whenever you are about to widen the scope of anything.

> **START HERE**
>
> If you only have one hour: complete the Opportunity Scan in Chapter 2 for one non-assurance workflow, write the Success Contract in Appendix C, and put a 30-minute review in the calendar with the person who owns that workflow.

Contents

The AI operating model

Find the work worth improving

Prioritise opportunities

Design reliable workflows

Instructions as specifications

Data, confidentiality and the professional layer

Governance, risk and the accountability line

What the file has to show

Adoption, measurement and improvement

Eight ways this goes wrong

Practical checklists

A 30-day implementation roadmap

Appendix A — The readiness scorecard

Appendix B — Starter task specifications

Appendix C — The Success Contract

1. The AI operating model

AI should be managed as an operating capability, not a software purchase: a set of choices about work, information, judgement, controls and learning. Firms that treat it as a purchase get a subscription, a champion, and a quiet reversion to old habits by the second busy season.

LAYER	CORE QUESTION	TYPICAL OUTPUT
Value	What measurable engagement or firm outcome matters?	Outcome statement and baseline
Work	Which recurring task or judgement is constrained?	Workflow map and pain points
Intelligence	Where can AI assist, generate, classify, extract or retrieve?	Use-case design
Control	What must remain the professional's judgement, and how is that evidenced?	Risk tier, guardrails, documentation standard
Learning	How will we know it is working?	Metrics, feedback and review cadence

The Control layer carries more weight in a practice than it does anywhere else, because your output is an opinion or a conclusion that someone else relies on, and because a third party will one day read the file and form a view about whether the work was done properly.

The five operating principles

Outcome before novelty. Begin with turnaround, capacity, quality, recoverability or risk — not with a feature. "We should use AI" is not an objective; "our audit files are assembled six weeks after signing and our review points repeat every year" is.

Preparation is delegable; judgement is not. This is the whole architecture in one line. AI can prepare almost anything in an engagement. It cannot form the conclusion, and it cannot be the accountable party.

Small, observable steps. Make the work legible so a reviewer — and later an inspector — can see the inputs, the outputs and the exceptions.

Data minimisation. Client information is held under a duty of confidentiality that predates every tool you are evaluating. Use only what the task needs.

Feedback is a product input. Review points are your richest training data. A review point that recurs across engagements is a defect in the process, not in the preparer.

> **THE PARTNER TEST**
>
> Can the engagement partner explain, in one minute, what the system prepared, what the professional judged, and how the file evidences the difference? If not, the workflow is not ready for a live engagement — however good the demonstration was.

2. Find the work worth improving

The best first opportunities are **frequent, bounded, measurable and annoying**. In a practice they are also the ones where a mistake surfaces immediately rather than eighteen months later at inspection.

Where it pays in a practice

WHERE THE WORK SITS	TYPICAL CANDIDATES	WHY IT WORKS
Client intake	Identification-document extraction, structure charts from filed records, screening result triage, acceptance file assembly	Structured inputs, checkable outputs, currently done under time pressure
Information gathering	Building the information request list, chasing outstanding items, tracking what has arrived, reconciling submissions against the list	Pure administration; the single biggest source of engagement delay
Preparation	Lead schedules from the trial balance, reconciliations, ageing analyses, movement schedules, sample selection, tie-back of figures to source	High volume, deterministic, verifiable
Drafting	Financial statement drafts from a prior period, disclosure checklists, tax computation working papers, management letters, minutes and registers	Template-shaped, rules-driven, tedious to do well
Review support	Casting and cross-casting, internal consistency checks between statements and notes, prior-year comparability, completeness of the disclosure set	Machines are simply better at this than tired people are
Practice administration	Deadline calendars, engagement letters, billing narratives, filing status tracking, correspondence logs	Recurring, low-judgement, quietly expensive

Two observations about that table.

First, **none of these is the audit.** They are the preparation that surrounds it. That is where the hours are, and it is where the risk of automating is lowest.

Second, the biggest single win in most firms is not in the technical work at all. It is in **information gathering** — the weeks lost to chasing documents, re-requesting things already sent, and discovering on day nine that the schedule provided does not agree to the ledger. That workflow is almost entirely administrative, and almost entirely automatable.

The Opportunity Scan

SIGNAL	QUESTIONS TO ASK	EVIDENCE TO CAPTURE
Frequency	Does this happen on every engagement?	Volume per period; cycle time
Friction	Where do people copy, reformat, search or wait?	Steps; handoffs; rework; review points
Value	What improves if this is faster or better?	Turnaround, capacity, recoverability, quality
Feasibility	Are inputs accessible and reasonably consistent?	Sources; formats; who controls them
Risk	What could go wrong, and who is affected?	Failure modes; client sensitivity; file impact
Adoption	Will the preparer and the reviewer actually use it?	Current tool; incentives; habits

Workflow observation worksheet

- Workflow name: _____
- Owner and users: _____
- Trigger and desired outcome: _____
- Inputs and systems touched: _____
- Current steps (aim for 5–9): _____
- Most expensive delay or error: _____
- What a good result looks like: _____
- What must never happen: _____
- What the file must show either way: _____

That last line is the one that distinguishes a practice from any other business, and the one most often added too late.

> **GOOD FIRST USE CASES**
>
> Information request lists and chasing · extraction from client-supplied documents · schedule preparation from a trial balance · internal consistency checks on a draft · disclosure checklist first pass · meeting notes to action registers · engagement administration.
>
> **BAD FIRST USE CASES** Anything that produces a conclusion rather than an input to one: materiality judgements, risk assessment conclusions, sufficiency of evidence, going concern, acceptance decisions, the opinion itself. Not because these are technically impossible to assist — but because they are exactly where you do not want to be learning how the technology fails.

3. Prioritise opportunities

The V-R-R score

Score each dimension from 1 (low) to 5 (high).

Priority = Value × Readiness × (6 − Risk)

DIMENSION	1	3	5
Value	Nice to have; no owner	Useful improvement	Directly tied to turnaround, capacity or quality
Readiness	Messy inputs; unclear process	Known owner; some structure	Stable inputs; repeatable across engagements
Risk	Touches a conclusion or sensitive data	Moderate; review catches errors	Administrative; easily reversed

Decision rules

- **Start now** — high value, high readiness, low risk. Almost always an administrative or preparation workflow.
- **Prepare first** — high value, low readiness. Usually means your templates, your file structure or your client data are inconsistent. Fix that first; it pays for itself even if you never automate anything.
- **Pilot with controls** — moderate value or moderate risk. Narrow scope, one engagement, mandatory review, documented.
- **Do not automate** — anything where a wrong answer is indistinguishable from a right one on the face of the file. This is the test that matters most in assurance work.

> **A NOTE ON SEQUENCING**
>
> Prove the operating model on non-assurance work first — bookkeeping, corporate secretarial administration, tax computation preparation, management reporting. The controls, the documentation habits and the review discipline transfer. The consequences of getting it wrong do not.

4. Design reliable workflows

A reliable workflow is a controlled loop, not a clever prompt.

The workflow loop

1. **Trigger** — what starts the work?
2. **Prepare** — what inputs are gathered, cleaned and authorised?
3. **Assist** — what should AI produce or recommend?
4. **Review** — who checks accuracy, completeness, reasoning and compliance?
5. **Act** — what is filed, sent, recorded or concluded?
6. **Learn** — what feedback is captured, and when is the workflow reviewed?

Control points to define

CONTROL POINT	MINIMUM DESIGN DECISION
Input boundary	Which client information may be used? What must be excluded or masked? What consent or engagement-letter basis covers it?
Output standard	What fields, format, referencing and source support are required?
Human review	Who reviews, at what level of seniority, against what checklist?
Exception path	What happens when confidence is low, sources conflict, or the item is unusual?
Audit trail	What is retained: source reference, output, preparer, reviewer, decision, date?
Stop condition	What signal pauses the workflow for investigation?

Worked example: the information request loop

The situation. The request list is built from last year's list. Items arrive by email in no particular order. Nobody is certain what is outstanding until someone rebuilds the list by hand. The client says they sent it; they did, to a different person, three weeks ago.

Trigger — engagement planning concludes.

Prepare — prior-period file, this period's trial balance, the entity's filed records, the engagement scope.

Assist — generate the request list mapped to the account balances and disclosures it supports, in the client's language rather than internal shorthand. As items arrive: identify what each document is, match it

to a request line, check it against the balance it supports, and update the outstanding list. Draft the chaser for what remains, escalating in tone by age.

Review — the in-charge confirms the list is complete for the engagement scope before it goes, and reviews every mismatch flagged. The client-facing chaser is sent by a person.

Act — the list is issued; the status is visible to both sides; receipts are acknowledged.

Learn — track which items are chronically late and which are never used. Both are defects: the first is a client-education problem, the second is a request list that has been copied forward without thought for years.

What must never happen — a request for information the engagement does not need; a document accepted as agreeing to a balance when it does not; confidential information sent to the wrong recipient; an outstanding item silently dropped.

Worked example: schedule preparation and review

The situation. A preparer builds lead schedules by hand from the trial balance, and the reviewer spends the first hour checking arithmetic rather than reasoning.

Trigger — the trial balance is received and agreed to the ledger.

Prepare — trial balance, prior-period file, supporting listings, the firm's template set.

Assist — build the schedules to the firm's format; agree each to source; compute movements and ratios; identify items requiring explanation against defined thresholds; flag anything that does not tie, is missing a source, or is inconsistent with a related schedule.

Review — the professional reviews the reasoning and the exceptions. Their attention goes to judgement, not to casting.

Act — the schedules are placed on file with their source references and the preparer and reviewer record.

Learn — review points are categorised. Anything recurring becomes a rule, a template change, or a training point.

> **THE POINT OF THE EXAMPLE**
>
> The value is not that the schedule appears faster. It is that the reviewer's hour is spent on the items that need a professional to look at them. Automation of preparation is what makes room for the judgement the standards actually require.
>
> **DESIGN FOR EXCEPTIONS** Before going live, write down three exceptions and how each is handled: a missing input, two sources that disagree, and an item outside anything the workflow has seen. A workflow that only describes the happy path is a demonstration.

5. Instructions as specifications

A good instruction makes the work explicit: role, objective, context, constraints, output shape and review standard. The test is not "did it produce something good?" but **"could a competent new joiner follow this and produce the same thing?"**

The CLEAR pattern

ELEMENT	PURPOSE	EXAMPLE
Context	Relevant background only	"You are preparing working papers for a Singapore private company."
Limit	Boundaries and exclusions	"Use only the supplied trial balance and listings. Do not use outside sources."
Expected output	The artefact	"Return a lead schedule in the firm's format with source references."
Audience	Depth and language	"For review by the engagement in-charge."
Review	Checks and uncertainty	"Flag every item you could not agree to source. Do not estimate."

The reusable template

> **PROMPT**
>
> You are [role]. Complete [task] for [audience] using only [approved inputs]. The desired outcome is [measurable result]. Follow these constraints: [rules, exclusions, format, referencing]. Return the result as [format]. Before finalising, check [quality criteria]. If information is missing or sources conflict, state the gap and ask for the smallest clarification needed — do not assume, estimate or infer.

The final clause is doing most of the work. In professional practice, a confidently fabricated figure is far more dangerous than a blank one, because a blank invites a question and a figure invites reliance.

Specification quality checklist

- [ ] One primary objective
- [ ] Source material and its authority are explicit
- [ ] Output format is easy to review and to file
- [ ] Facts, assumptions and recommendations are distinguished
- [ ] The point at which a professional must review is stated
- [ ] Test cases exist with known correct answers

Keep a small library of real engagements with known outcomes. When an instruction changes, re-run them. This is unglamorous and it is the only honest way to know whether a change improved anything.

From instruction to system

A well-written instruction is where this starts. It is not what you operate. The difference between the two is not the wording — it is everything wrapped around it.

- **The instruction is versioned.** It lives as a file with a history, not as text in someone's head or in a chat window they scroll back through. You can see what changed, when, and why. An instruction that is not versioned cannot be described in a quality-management response, because you cannot say which version produced the working paper on the file.
- **There is a regression test set.** A fixed set of real engagements with known correct outcomes, re-run every time the instruction changes. Without this you cannot tell an improvement from a coincidence.
- **Deterministic checks wrap non-deterministic output.** Anything mechanically checkable is checked mechanically, not by reading: does the schedule cast, does the balance agree to the trial balance, does the note cross-reference resolve, is the date inside the period. The system proposes; ordinary code verifies. This is the single largest difference between something that works in a demonstration and something that works in the third week of a busy season.
- **Steps are orchestrated, not pasted.** In multi-step work, each step's output is validated before it becomes the next step's input, and there is a defined path for what happens when a step fails.
- **Changes are owned.** A named person approves changes to the instruction, and the change is logged — the same discipline the firm already applies to its templates and its methodology.
- **There is an audit trail.** Source reference, output, preparer, reviewer, decision, date — retained and retrievable for as long as the file must survive.

A small firm can start with the first, second and last of these and capture most of the value. But it is worth knowing what the full shape looks like, because the gap between "we wrote a good prompt" and "we operate a workflow" is exactly this list — and it is where both the reliability and the defensibility live.

6. Data, confidentiality and the professional layer

Capability map

CAPABILITY	WHAT TO SPECIFY	FALLBACK IF UNAVAILABLE
Generate	Draft, transform, explain, create variations	Templates and checklists
Retrieve	Find relevant approved information	Curated knowledge base
Classify	Assign category, priority or route	Rules-based triage
Extract	Turn documents into structured fields	Standard schedules, manual capture
Analyse	Compare, reconcile, identify patterns and outliers	Spreadsheet review with sampling
Orchestrate	Move work between people and stages	Checklist and status meeting

> **VENDOR-NEUTRAL REQUIREMENT**
>
> Document the workflow in terms of inputs, outputs, permissions, review gates and metrics. The tool will change. These requirements should not — and a workflow described only as "we use [product]" cannot be inspected, transferred or defended.

The professional layer

Four obligations shape every AI decision in a Singapore practice. None of them mentions AI. All of them apply.

Confidentiality. Client information is held under a professional duty that is broader than data protection law and older than any of these tools. Putting client information into a third-party system is a disclosure decision. It requires a considered basis — engagement terms, client consent where appropriate, and contractual protection from the provider — and it requires knowing where the information goes and whether it is retained or used for anything beyond your instruction. Consumer-grade and business-grade services differ mainly on this point, and the difference is not visible in the interface.

Quality management. The firm-level quality management standard requires you to identify quality objectives and risks and to respond to them — and technological resources sit squarely within that. In practical terms: if a system is part of how engagements are performed, the firm must have decided that it is appropriate, that people are competent to use it, that its outputs are subject to appropriate review, and that the arrangement is monitored and remediated when it fails. This is not an obstacle. It is the same discipline you already apply to any other resource, written down.

Documentation. The file must enable an experienced practitioner with no prior connection to the engagement to understand the work performed, the evidence obtained, and the significant judgements reached. Nothing about that changes because a system prepared the schedule. What changes is that the

file must make clear what was prepared, what was reviewed, on what basis the professional accepted or challenged it, and where the exceptions went. A file that is silent on this describes a process nobody can reconstruct.

Personal data. Where engagement information includes personal data — and it nearly always does — data protection obligations apply on top of confidentiality: purpose limitation, protection, retention limits, responsibility for third parties processing on your behalf, a designated data protection officer, and conditions on transferring personal data out of Singapore. Breach notification obligations apply where a breach is likely to cause significant harm or affects a significant number of individuals.

> **THE CONSTRUCTIVE READING**
>
> Every one of these obligations is satisfiable. None of them prohibits using AI in practice. What they prohibit is using it *without having decided anything* — which is, in fairness, how most technology enters most firms.

7. Governance, risk and the accountability line

Four risk tiers

TIER	TYPICAL WORK	REQUIRED CONTROLS
1 — Assistive	Internal drafting, summarising, formatting, administration	Approved inputs; user reviews before use; nothing leaves the firm unreviewed
2 — Preparatory	Extraction, schedule preparation, reconciliation, checklist first pass, consistency checks	Test set; quality threshold; exception path; documented review; audit sample
3 — Engagement-facing	Anything reaching the client, entering the file as evidence, or informing a conclusion	Named accountable professional; documented review of reasoning, not just output; retention and referencing standard; recurring monitoring
4 — Reserved	Conclusions, opinions, acceptance and continuance, sufficiency of evidence	Not delegable. Redesign the workflow so the system prepares and the professional concludes.

The accountability line

These are not technology limitations. They are the things a professional is engaged to do, and they cannot be delegated to a system regardless of how capable it becomes:

- **Forming and signing the opinion or conclusion**
- **Deciding acceptance and continuance**, including independence and integrity considerations
- **Concluding on the sufficiency and appropriateness of evidence**
- **Setting materiality and assessing risk** — the computation may be prepared; the judgement is the partner's
- **Concluding on going concern**
- **The engagement quality review**
- **Filing with a regulator** — ACRA, IRAS and every other submission is made by the licensed person responsible for it
- **Any judgement where the file could not later show that a professional actually applied their mind**

Preparation can be delegated almost without limit. Accountability cannot be delegated at all. Every firm that gets into difficulty with this technology will have crossed this line somewhere, and will have done it gradually.

Minimum governance checklist

- [ ] Named responsible partner and workflow owner
- [ ] Purpose, permitted engagements and intended users documented
- [ ] Client information categories and permitted destinations recorded
- [ ] Confidentiality basis considered and documented
- [ ] Known failure modes and mitigations listed
- [ ] Review level and escalation criteria defined by risk tier
- [ ] Test examples include edge cases and deliberately awkward inputs
- [ ] Monitoring metric and review cadence agreed, feeding firm-level monitoring
- [ ] Change log maintained for instructions, rules and templates

> **RED FLAGS**
>
> Output enters a file without a documented review · nobody can say which client information went where · the risk tier changed without re-assessment · the only person who understands the workflow built it · a reviewer approves faster than they could have read · a system's output is cited as evidence without a source reference.

8. What the file has to show

This chapter has no equivalent in a general business playbook, and it is the one practitioners ask about first.

An inspector, a peer reviewer or a successor practitioner is not going to ask whether you used AI. They are going to ask the questions they have always asked, and the file has to answer them.

Six questions to test any AI-assisted working paper against:

1. **What was the source?** Every figure traces to a document or a system extract that is identified and retained.
2. **What was prepared, and by what?** The file distinguishes what was produced as an input from what was concluded.
3. **Who reviewed it, and what did they do?** Not a tick. Evidence of engagement with the reasoning — the challenge raised, the exception followed up, the item sent back.
4. **What did the reviewer do about the exceptions?** Exceptions flagged and then silently accepted are worse than exceptions never flagged, because the file now records that you were told.
5. **Could someone reconstruct the conclusion without you?** The experienced-practitioner test, unchanged.
6. **Is it still retrievable and legible?** Records must survive for years, in a form that can be produced. An output that lives only inside a tool you no longer subscribe to is not a record.

Documentation habits that make this straightforward

- Record the preparer and reviewer, whether the preparer was a person or a system, and the date of each.
- Retain the input set actually used, not just the output.
- Keep exception logs on file. They are evidence of scepticism, not evidence of sloppiness — a file with no exceptions raises more questions than one with exceptions that were properly cleared.
- Render final working papers into a stable, self-contained format for the file. A document that depends on a live connection to reproduce itself is not an archive.
- Document once, at firm level, how the technology is used, who approved it, and how it is monitored. Then reference it, rather than re-explaining it on every engagement.

> **THE SUMMARY POSITION**
>
> Nothing in the standards prevents a system from preparing a working paper. What the standards require is that a professional applied their mind, and that the file shows it. Design for that from the first day, and inspection is a description of what you already do rather than an event you prepare for.

9. Adoption, measurement and improvement

Adoption in a practice has a particular obstacle: the people best placed to redesign the work are the people whose value has historically been measured in hours spent doing it. Address that directly or it will be addressed quietly, through non-adoption.

Adoption design

- Give the workflow a visible owner and a small group of early users, including at least one sceptic.
- Train on **the task**, not on the technology.
- Provide a good input, a good output and a good review as worked examples.
- Make the compliant path the easy path — approved templates, pre-approved sources, defaults that satisfy the documentation standard without anyone thinking about it.
- **Treat review points as data, not as blame.** A recurring review point is a process defect.
- Be explicit about what this means for people's roles. If capacity released is going into more engagements, more advisory work, or shorter hours, say so. Ambiguity here is read as threat, and it should be.

Measurement stack

LEVEL	MEASURE	EXAMPLES
Outcome	Did the engagement result improve?	Turnaround, file assembly time, recoverability, capacity
Workflow	Did the process improve?	Preparation time, handoffs, rework, adoption
Quality	Was the output fit for purpose?	Review points per engagement, exception rate, accuracy sample
Risk	Did exposure stay within bounds?	Escalations, confidentiality events, documentation gaps
Learning	Is it getting better?	Recurring review-point themes, instruction changes, test-set score

The number worth defending: **review points per engagement, by category.** Falling in the mechanical categories and stable or rising in the judgement categories is exactly the shape you want — it means the reviewer's attention has moved to where it belongs.

REVIEW RHYTHM

Weekly during the pilot: ten representative cases. Monthly after adoption: metrics, exceptions and changes. Each cycle: confirm the workflow still deserves its risk and its maintenance. Feed the results into firm-level monitoring rather than keeping them in a separate place — one system of record, not two.

10. Eight ways this goes wrong

1. Shadow use. Staff paste client information into whatever consumer tool is on their phone, because the approved route is slower or does not exist. This is a confidentiality breach with a professional dimension, not merely an IT policy matter. Prohibition alone has never worked anywhere. Provide an approved option that is genuinely easier, then set the rule.

2. The confident wrong figure. Fluent, correctly formatted, wrong. In a file this is uniquely dangerous, because presentation quality is one of the heuristics reviewers unconsciously use. Require source referencing on every figure and keep test cases with known answers.

3. Review theatre. The file records a review that did not happen in any meaningful sense. This is worse than no review: it manufactures documentary evidence of oversight. If review is real, work goes back sometimes.

4. Documentation that outruns the work. The workflow produces beautifully structured papers describing procedures more thorough than what actually occurred. The file must describe what was done, not what the template says is usually done.

5. Automating an inherited process. Last year's request list, copied forward for a decade, now generated automatically. Map the work and question it before you speed it up.

6. Scope creep without re-assessment. A workflow approved for internal preparation starts producing client-facing deliverables, or its output starts being relied on as evidence. The tier changed; the assessment did not.

7. The one-person system. It runs through the enthusiast. They leave. Anything on the critical path needs documentation and a second competent operator — and firm-level quality management expects exactly that of any resource.

8. No off switch. Nobody defined the stop condition while everyone was calm. Define it before you need it.

11. Practical checklists

Before a pilot

- [ ] Outcome and baseline written down
- [ ] Workflow owner named; responsible partner identified
- [ ] Scope narrow enough to test — one workflow, ideally non-assurance
- [ ] Client information categories and permitted destinations decided
- [ ] Confidentiality basis considered
- [ ] Review level and documentation standard explicit
- [ ] Success and stop conditions defined
- [ ] Test cases include normal and awkward examples

Before anything enters a file

- [ ] Every figure traces to an identified, retained source
- [ ] What was prepared and what was concluded are distinguishable
- [ ] The reviewer's engagement with the reasoning is evidenced
- [ ] Exceptions are logged and their resolution recorded
- [ ] The paper is legible and reconstructable by someone with no prior connection
- [ ] The retained form is stable and self-contained

Before anything reaches a client

- [ ] The risk tier has been re-assessed for an external audience
- [ ] A named professional has approved this instance
- [ ] Confidential information belonging to another client is definitively absent
- [ ] There is a way to correct or retract
- [ ] The client can reach a person easily

After an incident or near miss

- [ ] Pause the affected workflow if recurrence could cause harm
- [ ] Preserve the input, output and decision record
- [ ] Describe what happened without assigning blame
- [ ] Identify the failed control, not only the bad answer
- [ ] Consider whether any engagement already signed is affected
- [ ] Consider whether any notification obligation is triggered
- [ ] Add a test case or a guardrail; feed it into firm monitoring
- [ ] Re-enable only after the responsible partner signs off

12. A 30-day implementation roadmap

For a small firm with no spare capacity. Each week ends with evidence and a decision. Keep the team small: one responsible partner, one workflow owner, two to five users, one person responsible for measurement.

DAYS	FOCUS	ACTIONS	EXIT EVIDENCE
1–3	Align	Name the responsible partner; choose the outcome; set guardrails; select one non-assurance workflow	One-page Success Contract
4–7	Observe	Map the work; collect real examples; baseline time and review points; list failure modes	Workflow map + test set
8–14	Design	Draft the process; define review and documentation standards; build templates; train users	Working pilot + review checklist
15–21	Run	Use on live low-risk work; sample outputs; log review points; fix friction	10–20 reviewed cases + findings
22–26	Evaluate	Compare to baseline; review risk and documentation quality; interview users; decide adopt / revise / stop	Pilot scorecard + decision
27–30	Institutionalise	Document the procedure; assign the owner; set monitoring cadence; choose the next workflow	Published procedure + 60-day backlog

Pilot scorecard

METRIC	BASELINE	PILOT	TARGET	NOTES
Preparation time				
Review points (mechanical)				
Review points (judgement)				
Exception rate				
Documentation completeness				
Adoption / usage				
Incidents / escalations				

THE DECISION AT DAY 30

Adopt when the outcome improved, quality held, the documentation stands on its own and the owner can maintain it. **Revise** when value is visible but the inputs or the controls are not ready. **Stop** when the risk or the maintenance cost outweighs the benefit — you bought that information cheaply, and stopping is a successful outcome.

Appendix A — The readiness scorecard

Score each statement 0 (not true), 1 (partly true) or 2 (true).

#	STATEMENT	SCORE
	Value	
1	We can name one workflow where delay or rework costs us measurable time or recoverability.	
2	We know roughly how long that workflow takes today.	
3	A partner would notice if it improved.	
	Readiness	
4	The inputs are in a consistent, findable form across engagements.	
5	One named person owns the workflow end to end.	
6	Our templates are consistent enough to describe to a new joiner.	
	Confidentiality and data	
7	We have decided which client information may go into which systems, and written it down.	
8	We know where that information is held, and whether it is retained.	
9	Nothing in the intended workflow produces a conclusion without a professional.	
	Quality management and the file	
10	There is a named reviewer with the time to genuinely review.	
11	We know what the file will show about what was prepared and what was judged.	
12	We have real examples with known correct answers to test against.	
	Total (out of 24)	

BAND	SCORE	WHAT IT MEANS	DO THIS NEXT
Not yet	0–8	The groundwork isn't there. Automation would propagate existing inconsistency across every engagement.	Standardise one template set and settle the confidentiality question. Chapters 2 and 6.
Prepare	9–15	A real opportunity exists, but the inputs or the documentation standard aren't ready.	Write the Success Contract; fix the input boundary and the file standard; then pilot.
Pilot ready	16–20	You can start safely on a bounded non-assurance workflow.	Run the 30-day roadmap in Chapter 12, unmodified.
Scale ready	21–24	You have the discipline to run a portfolio.	Move from single pilots to prioritised sequencing using Chapter 3, and fold monitoring into firm-level quality management.

Appendix B — Starter task specifications

The specifications below are written as prompts because that is the portable form — you can paste one into any capable system and get useful work back today. Treat them as starting points, not artefacts. In an operating workflow each of these carries a version number, a named owner, a set of test cases with known outcomes, and a change log; the text is the smallest part of it, and it is the part an inspector will care about least. Replace the placeholders with your own templates, approved sources and review standards.

1. Information request list

> Using the prior-period file and the current trial balance, draft an information request list for [engagement type]. Map each request to the balance or disclosure it supports. Write in plain language a client's finance staff can act on — no internal shorthand. Group by who in the client organisation is likely to hold it. Mark each item as required or conditional. Do not request anything the engagement scope does not support.

2. Document identification and matching

> Identify what this document is, which entity and period it relates to, and which request-list line it satisfies. Extract [fields]. Check the totals against [reference balance] and report the difference. If the document is illegible, incomplete, or relates to a different period or entity, say so and do not attempt to reconcile it.

3. Lead schedule preparation

> Build a lead schedule for [account grouping] from the attached trial balance in the firm's format. Agree each balance to source and reference it. Compute current-versus-prior movements and flag every movement above [threshold]. List separately: any balance you could not agree to source, any account not mapped to a grouping, and any inconsistency with a related schedule. Do not estimate or plug differences.

4. Internal consistency check

> Check this draft set of financial statements for internal consistency only. Verify that: every note cross-reference resolves; figures in the notes agree to the primary statements; the primary statements cast and cross-cast; comparatives agree to the prior period; and the disclosure set is complete against the attached checklist. Report findings as a table of location, issue and suggested correction. Do not assess accounting treatment or make judgements about appropriateness.

5. Variance and analytical narrative

> Compare this period to the prior period and to expectation. For every variance above [threshold], explain the movement citing the specific accounts or transactions responsible. Distinguish clearly between (a) what the data shows, (b) what you infer, and (c) what cannot be determined from the information provided. Do not speculate beyond the data or propose an explanation you cannot evidence.

6. Meeting to action register

> Turn these notes into an action register. Extract decisions, open questions, actions, owners, due dates, dependencies and matters requiring partner attention. Do not invent owners or dates. Mark missing fields "Not specified." Return a table followed by the three most significant unresolved matters.

7. Quality reviewer

> Review this working paper against the following rubric: source support and referencing, arithmetic accuracy, internal consistency, completeness against scope, clarity for an experienced practitioner with no prior connection, and whether the conclusion follows from the evidence presented. Return pass or revise for each, quote the smallest problematic excerpt, explain the issue and propose a correction. Do not rewrite the paper.

8. Procedure builder

> Convert the approved workflow notes into a firm procedure with purpose, scope, roles, prerequisites, steps, decision points, exception handling, review and documentation requirements, records retained, and review cadence. Keep every step observable and testable. Mark unresolved ambiguity as an open item rather than resolving it.

Appendix C — The Success Contract

One page. Written before anything is built. Signed by the responsible partner and the workflow owner.

- Outcome to improve: _____
- Workflow and owner: _____
- Responsible partner: _____
- Pilot scope, engagements and dates: _____
- Baseline measure: _____
- Target measure: _____
- Review level and documentation standard: _____
- Client information allowed / excluded, and destination: _____
- What must never happen: _____
- Stop condition: _____
- Review cadence and decision date: _____

About this playbook

Nebis is a Singapore accounting and professional services firm built AI-native from the start. AI systems perform the preparation — extraction, reconciliation, scheduling, drafting — while qualified professionals direct the work, review it, and take personal responsibility for every regulated outcome. Audit opinions are signed by the firm's registered public accountant; statutory filings are made by the licensed people responsible for them. AI prepares; people decide, sign and file.

This playbook contains general guidance, not professional advice, and does not take account of your firm's particular circumstances. It is not a substitute for the applicable standards, which should be read directly. Requirements change; verify current obligations with the relevant authority before acting.

Nebis.AI · Singapore

Edition B — for practitioners, partners and practice owners. A companion edition for business owners, founders and finance leads is also available.