Adopting AI Without Losing Control

A Practical Operating Guide for Small and Midsize Businesses

Joel Ward | Operations and Implementation Consultant

Version 1.2 | October, 2026

Suppose a business owner tells me, “This AI tool can handle our customer updates, estimates, and follow-up. It will save us hiring another person.” My first question is where is it getting its information. Directly from personnel? From the job records? From a price book? I would want to know that before discussing how much work it can take over.

The old computer adage, “garbage in, garbage out,” still applies. AI may catch an inconsistency or flag missing information. But, not always. I would want to see what it catches in trials and what it lets through. The source of the data is my first concern.

Then I would ask where the standard operating procedures came from. Has anyone reviewed and approved the methods AI is using to update customers, prepare estimates, and follow up? A written procedure might be out of date. It might be something AI drafted that nobody has checked. Before approving its use, I want the company to know what method it is approving.

We also need trials to verify two things: the output is accurate, and it meets the company’s standard of customer service and language. Those are separate questions. A message can get the facts right and still be something I would not want sent to a customer.

That is the concern behind this guide. Companies can adopt AI capabilities faster than they can evaluate, supervise, and control them. I use AI in my own work, and I see plenty of useful applications. I would introduce a specific use only after working through the source information, the method, and the results of those trials.

Where a small team can start

Choose one recurring task. Name the person responsible for the decision and write down the approved sources and procedure. Begin with fictional information and no live connections. Have someone who understands the work check the result, including how it would read to a customer. Record the unresolved issues before deciding whether to try it with real data.

The assessment has a blank form on pages 1–4 and a completed fictional roofing example on pages 5–8. It shows a decision to hold live use when the trial exposes problems, even though the estimated time saving looks attractive.

Sources and scope: I draw on NIST and OWASP primary guidance. [1–5] The operating choices and assessment are my recommendations, not NIST requirements or a certification. Uses affecting safety, employment, credit, health, or legal rights need appropriate specialist review.

Find out where the information and methods come from

If personnel are supplying the information, I would ask who enters it and who checks it. A foreman's note about work completed is different from a salesperson's expectation about what will happen next. The tool needs to work from the right job and the current information. Putting the note into a CRM does not, by itself, settle either point.

Follow a sample output back to its source. For a progress update, can the reviewer find the note that supports the claimed job stage or completion date? For an estimate, can an estimator reproduce the quantity and price calculation? When the records disagree, someone in the business needs to resolve them. AI should flag the conflict rather than choose an answer merely to finish the task.

Review the procedure AI is expected to follow

A standard operating procedure, or SOP, should describe the method the company has approved. I would ask who wrote it, who approved it, and whether that is the version the tool is using. AI can help draft a procedure, but the responsible manager still needs to examine the proposed steps and decide whether to adopt them.

Proposed work	Sources and methods to confirm
Customer updates	Job record, person supplying the update, current stage, confirmed next step, and who may authorize a new promise.
Estimates	Job measurements, current supplier prices, labor assumptions, allowances, pricing rules, and an estimator's check of the calculation.
Follow-up	Customer history, reason for contact, timing, channel, approved language, and when a question or complaint goes to a person.

For estimating, I would look closely at dates and units. Is the material price current? Is a labor allowance appropriate for this job? Does the quantity calculation use the same units as the price? A plausible total is not enough. The person checking it should be able to trace how it was reached.

I would also review actual examples of customer communications the company considers acceptable. Use those to define the language expected of the tool. Decide how it should handle an uncertain completion date or a dissatisfied customer before asking it to reply.

Keep the approved source list and SOP version with the assessment. Name the person responsible for updates. When prices, procedures, or source records change, check the affected output again. NIST's generative AI profile addresses source verification and the grounding of data used by the system. [2]

Check the value and who will do the reviewing

The owner's claim about saving another hire is worth investigating. I would first separate the proposed work. Customer updates, estimates, and follow-up may each need different information and a different reviewer. A successful trial of updates would not establish that the tool is ready to price a job.

Before buying anything, compare the proposed use with a standard template, a rule in existing software, or a clearer assignment. My work at a roofing company included CRM implementation, estimating and pricing tools, and management reporting. I would approach this decision by following the work through the business and finding where the delay or recurring error actually occurs.

Include the checking and correction time

Suppose the office prepares 200 customer updates a month, taking 12 minutes each. That is 40 hours. If AI brings the complete job down to 7 minutes, including preparing the input, checking the result, and correcting it, the difference is about 16.7 hours. At an assumed $30 an hour, that time has a value of about $500.

Allow an illustrative $150 a month for the tool and ongoing support, and about $350 in capacity value remains before setup. These are assumptions, not reported savings or a product quote. The business still has to use the freed time productively or actually reduce a cost. If reviewing and correcting erase the time saved, the case needs another look.

Who is going to check the work, and what are they checking it against? For a customer update, I would want someone who knows the job to compare the draft with the actual record. Then we need to look at that person's workload. What are we taking off their desk to make time for the review? That time belongs in the cost estimate. If we can only check a few updates properly, that is the size of trial I would start with.

Make the responsibility clear

Name the business owner who decides whether to proceed, the person checking outputs, and the administrator who controls access and can disconnect the tool. In a small company, people may hold more than one role. Record who covers an absence. NIST calls for clear responsibilities and leadership accountability. [1]

Before the trial, show staff the approved sources, the procedure, and examples of acceptable output. Give them a way to report doubtful answers and customer concerns. When the reviewer is unavailable, hold the item or use the existing process. The plan should reflect the staff and time the company actually has.

Set the limits on access and authority

Knowing where information comes from is one decision. Whether the tool may access it is another. Approve the specific records and fields the task needs. A customer update should not require payroll, bank details, or access to every attachment in the company's files.

Check the actual vendor product, account plan, and settings. Record training use, retention, deletion, and who can access the information. "Not used for training" does not answer all of those questions. Check export options, support contacts, and notice of changes. If a material answer remains unclear, keep the affected data out of the trial.

Permission	What it allows	Boundary to enforce
Draft only	Prepare an answer or recommendation for a person to check and use.	No sending, publishing, or changes to business records.
Act after approval	Execute a specific message or change after an authorized person approves.	Approval covers the exact wording, destination, and action. Material edits require another approval.
Act within fixed limits	Perform a defined routine action without approving every occurrence.	Enforced actions and records, amount or volume limits, monitoring, and a tested stop procedure.

These permission choices are my proposed operating model. They are not NIST levels or a required progression.

Consider the consequence separately from the permission. A draft estimate may be consequential even before anyone sends it. Read-only access can expose confidential information. My default is to retain an explicit human decision before spending money, accepting terms, changing bank details, deleting important records, or making a consequential promise.

Show the approver the final proposed action and the source record. Include the recipient, attachments, amount, or current and proposed field values as applicable. Record who approved what. A material change afterward requires a new approval. The tool must have no route around a required human decision. [4]

Enforce the limits in the application and connected systems. Use restricted accounts, separate test and production access, and verify both allowed and denied requests. An instruction hidden in a document or email must not grant new authority; this risk is called prompt injection. [4, 5] Keep protected records of important outputs, approvals, and settings, with an agreed retention period. If the product cannot enforce a necessary boundary, narrow the task or choose another approach.

Run trials for accuracy and customer service

Before approving the use, we need to see whether the output is accurate and whether it meets the company's standard of customer service and language. I would establish both standards before testing. Otherwise, it is too easy to accept an answer simply because it sounds good.

Check the result against the approved sources

Have someone who understands the work trace important statements and figures back to the record and approved method. For estimates, independently check the calculation. Test missing information, conflicting notes, outdated prices, and wrong-job records deliberately. The expected result may be to stop and ask for clarification. AI does not need to produce a finished answer to every request.

Keep each test's input, expected result, actual result, reviewer, and configuration. Repeat important cases and reserve some examples that were not used to adjust the instructions. NIST cautions against broad conclusions from narrow testing and recommends checking generated sources. [2]

Review how the company would sound to the customer

Compare the draft with examples the company has approved. Is it clear and respectful? Does it explain what is known without inventing certainty? Does the follow-up fit the customer's situation and the company's procedure? Correct facts alone would not persuade me to approve a message that sounds dismissive or makes the customer feel pressured.

Record accuracy and service results separately. A missing fact filled in as a promise is an accuracy failure. A factually correct but patronizing reply is a service failure. Both need attention if the company has made them conditions of release. In the fictional assessment, each occurs and live use stays on hold.

Test the controls as well as the wording

Try requests for prohibited information and instructions embedded in outside material. If the tool can act, test approval bypass, changes after approval, duplicate actions, and retries after a timeout separately. Practice pausing the work and using the manual process. An unauthorized disclosure or action stops affected use immediately.

Begin with one team and a volume the reviewer can handle. Set a decision date and record the unresolved issues. The example proposes two weeks; the duration itself proves nothing about reliability. A small sample with no observed failures still leaves uncertainty, especially where mistakes would be costly.

During the trial, review problems daily. After release, monitor corrections, customer complaints, review delays, and cost. Name someone to respond to urgent reports. Recheck affected behavior when sources, SOPs, prompts, permissions, or model behavior change. If the team cannot confirm that the controls still work, reduce the scope or pause use. [1, 2]

A roofing trial and the response to failure

Illustrative scenario only. The company, people, test outcomes, and financial assumptions are fictional.

A roofing company proposes AI help with updates, estimates, and follow-up. I would start the first trial with updates. Estimating would require its own source and method review. The proposed live trial covers 20 jobs, with no more than four drafts a day for two weeks, subject to approval after isolated testing.

For testing, staff use fictional job notes. Dana, the operations manager, checks the inputs and adapts the office's update procedure. Pat, the company owner, approves that version and sample language for the isolated test. Each request contains one job's stage, approved notes, and confirmed next step. AI has no CRM connection and cannot send messages.

In one fictional case, the note says, "Materials delivered. Crew date awaiting confirmation." The draft promises completion Friday. That fails the accuracy requirement. Another draft correctly says materials are on site, then adds, "You just need to be patient." Eli, the reviewer, rejects that wording under the approved customer-service standard.

Live use remains on hold while those failures are corrected and the vendor account review is completed. Fictional-data testing can continue. If eventually approved, Eli would compare every draft with the current job record before manually sending it through the existing process. A new price or consequential promise would go to Dana. The projected time saving does not resolve the accuracy, service, or data questions.

Prepare to stop the tool and continue the work

Before live use, decide who can pause the process and who removes access. Know how to cancel queued actions and retries. Keep the approved SOP, previous working configuration, and a usable manual procedure. Practice the handoff. The company still needs to communicate with customers while the tool is unavailable.

If another customer's information appears in a draft, stop affected work and preserve the relevant records. Establish which outputs and people were affected and whether anything left the company. The owner coordinates corrections and any needed vendor, security, or legal involvement. NIST includes response, recovery, and disengagement in its guidance. [1, 2, 6]

Disconnecting a tool will not recall a sent email or undisclose a file. Recovery may require contacting customers and correcting records as well as restoring software. Resume only after the cause is understood, the fix is tested, and the owner accepts the remaining risk. Approval of customer updates would cover that use alone. Estimating and other work would need separate review.

Sources and about the author

I used the following primary sources as a foundation. NIST's framework organizes risk work around Govern, Map, Measure, and Manage; its companion playbook offers suggestions that can be adapted. [1, 3] The permission choices, examples, and assessment in this guide are my operating recommendations. All sources were checked October 22, 2026.

[1] NIST AI 100-1 Artificial Intelligence Risk Management Framework 1.0 (January 2023). See GOVERN 1.6, 2.1–2.3, 3.2; MEASURE 2.3–2.4; MANAGE 2.3–2.4, 4.1–4.3.

NIST framework PDF

[2] NIST AI 600-1 Generative Artificial Intelligence Profile (July 2024). See MS-2.5-001, MS-2.5-003 and MS-2.5-005; MG-2.3-001; MG-2.4-001–004; MG-4.1-002.

NIST generative AI profile PDF

[3] NIST AI RMF Playbook Voluntary implementation suggestions that organizations can tailor to their circumstances.

NIST playbook

[4] OWASP LLM06 2025 Excessive Agency. Minimum functionality and permissions, human approval, downstream authorization, and activity limits.

OWASP excessive agency guidance

[5] OWASP LLM01 2025 Prompt Injection. Risks from instructions in user input and outside material.

OWASP prompt injection guidance

[6] NIST SP 800-61 Revision 3 Incident Response Recommendations and Considerations for Cybersecurity Risk Management (April 2025). Relevant when an AI incident involves cybersecurity.

NIST incident response publication

[7] NIST framework status NIST states that AI RMF 1.0 is being revised. This guide uses the published 1.0 framework and 2024 generative AI profile. Check for later revisions before relying on them.

NIST framework status page

About Joel Ward

I work with businesses on operations, systems, and implementation. My experience includes CRM and workflow implementation, estimating and pricing tools, KPI reporting, and owner cash-flow visibility in operations. I have also directed and refined AI-assisted operations prototypes using fictional scenarios. This guide brings that practical operating perspective to decisions about AI.

An AI adoption review

If you have a specific use in mind, I can help work through whether it is worth pursuing, what it should be allowed to do, and what needs to be in place before a trial. You can reach me through the contact form on my website.

Contact Joel Ward

Free to download and use internally. Retain attribution when sharing the unmodified guide. This is operational guidance, not a security audit, certification, or sector-specific compliance assessment. NIST and OWASP have not endorsed it.