

The Jonesi Consultancy | Houston, TX

EXECUTIVE AI ADOPTION · OVERSIGHT GUIDE

HUMAN-IN-THE-LOOP



Where Human Judgment Still Matters, and Why It Cannot Be Delegated

Accountability does not transfer to a system. It stays with the person who signed off.

HOW MUCH CHECKING IS ENOUGH?

Every organization using AI eventually arrives at the same question. Check everything and the efficiency disappears. Check nothing and you are one confident-sounding error away from a client conversation you do not want to have. The answer is not a percentage. It is a judgment about stakes, made in advance, so that nobody has to make it under time pressure at four o'clock on a Friday. This guide gives you a way to make that judgment consistently.

1 The Part That Does Not Move

THE PRINCIPLE

AI assists work. It does not approve work. When something goes wrong, the question asked will be who sent it, not what produced the first draft. No regulator, client, board, or court has yet accepted a tool as the responsible party, and none is likely to.

This is worth stating plainly inside your organization, because in the absence of a statement people develop private theories about where responsibility sits. The most common private theory is that responsibility is shared, which in practice means nobody feels like the last line.

2 What Human-in-the-Loop Requires

People remain actively involved at the points where context, verification, judgment, accountability, or consequences require human review. The goal is not to slow useful automation. It is to place human attention where it matters most.

DIRECTION

A person defines the objective, scope, constraints, and acceptable use of AI.

VERIFICATION

A person checks whether important claims, calculations, interpretations, and recommendations are accurate enough to use.

CONTEXT

A person applies organizational knowledge, professional standards, and situational factors the model may not know.

DECISION OWNERSHIP

A named human remains accountable for the final action, communication, approval, or decision.

A USEFUL RULE

As consequences increase, human review should become more deliberate, more qualified, and more visible.

3

3 Levels of Stakes

Most work sorts cleanly into one of three levels. The point of naming them is that review effort should be wildly different across them. In most organizations it is uniform, which means the low-stakes work is over-checked and the high-stakes work is under-checked.

LEVEL 1

LOW-STAKES ASSISTANCE

The output stays inside your own head or your own team, and a mistake costs you a few minutes.

TYPICALLY Brainstorming, first drafts you will rewrite anyway, summarizing something you already read, reformatting, explaining a concept to yourself, generating options you will discard.

WHAT REVIEW MEANS HERE Read it. Notice if something looks wrong. Move on.

LEVEL 2

REVIEW REQUIRED

The output leaves your hands, reaches a colleague or a client, or informs a decision somebody else will make.

TYPICALLY Client correspondence, internal analysis others will rely on, meeting summaries that become the record, proposals, reports, training material, anything that will be quoted back to you.

WHAT REVIEW MEANS HERE Verify every factual claim, figure, name, date, and citation. Read the reasoning as a skeptic. Confirm it reflects your actual situation rather than a generic answer.

LEVEL 3

HIGH CONSEQUENCE

The output affects someone's money, employment, health, legal position, safety, or rights, or it is going somewhere it cannot be taken back from.

TYPICALLY Employment decisions, financial reporting and filings, clinical or case documentation, legal and regulatory submissions, safety procedures, public statements, anything with a compliance obligation attached.

WHAT REVIEW MEANS HERE A qualified human does the substantive work and AI assists at the margins. A second reviewer signs. The AI contribution is documented. If you would not let a capable but unaccountable stranger draft it, do not let a model draft it either.

THE CLASSIFICATION IS THE DECISION

Deciding which level a piece of work sits in is the whole exercise. Once that is settled, what to do is obvious. Organizations that struggle with AI review have almost always skipped the classification and gone straight to arguing about how careful to be.

6 Areas Where Judgment Cannot Be Delegated

General organizational guidance, not legal, clinical, financial, or regulatory advice. The pattern across all six: AI expands the information available to people. Human judgment interprets consequences, context, competing values, and responsibility.

1 FINANCIAL AND RESOURCE DECISIONS

AI CAN SUPPORT

Model scenarios, summarize financial information, compare options, or identify anomalies.

HUMAN JUDGMENT MUST OWN

Assumptions, materiality judgments, approvals, resource tradeoffs, and final decisions that affect budgets, investments, controls, or stakeholders.

2 EMPLOYMENT AND PEOPLE DECISIONS

AI CAN SUPPORT

Organize information, draft communications, identify themes, or support administrative tasks.

HUMAN JUDGMENT MUST OWN

Context, fairness, evidence, interpersonal consequences, and accountability when decisions affect hiring, performance, discipline, advancement, or termination.

3 HEALTHCARE AND CLIENT SERVICE DECISIONS

AI CAN SUPPORT

Non-sensitive drafting, information organization, education, or workflow support within approved boundaries.

HUMAN JUDGMENT MUST OWN

Qualified professionals retain responsibility for decisions involving care, safety, eligibility, clinical judgment, protected information, or services with meaningful consequences.

4 LEGAL, COMPLIANCE AND REGULATORY WORK

AI CAN SUPPORT

Summarize source material, organize requirements, draft questions, or surface issues for review.

HUMAN JUDGMENT MUST OWN

People with appropriate expertise verify current authority, jurisdiction, applicability, interpretation, and the consequences of acting on the output.

5 CLIENT-FACING AND PUBLIC COMMUNICATION

AI CAN SUPPORT

Create drafts, alternatives, summaries, and formatting support.

HUMAN JUDGMENT MUST OWN

Accuracy, tone, confidentiality, disclosure where appropriate, brand alignment, and readiness to stand behind what is communicated.

6 STRATEGIC AND OPERATIONAL DECISIONS

AI CAN SUPPORT

Analyze alternatives, identify patterns, challenge assumptions, and accelerate synthesis.

HUMAN JUDGMENT MUST OWN

Leaders interpret tradeoffs, understand organizational context, weigh uncertainty, preserve accountability, and decide what should actually change.

4 Questions Before You Act on Any Output

Use these whenever an AI response will inform a decision, leave the organization, affect another person, or be treated as fact.

1 What is at stake if this is wrong?

Not how likely it is to be wrong. How bad it is if it is. Consider size, reversibility, and human impact. Low probability and high consequence still means Level 3.

2 Who is affected, and do they know AI was involved?

Identify the people, teams, clients, or communities who may experience the consequences. Higher human impact requires stronger review.

3 What could plausibly be fabricated here?

Models fail most reliably at specifics: citations, statistics, quotations, proper names, dates, URLs, policy and regulatory detail, and anything recent. Check those first.

4 Who is accountable, and do they know they are?

Name the person with authority to approve the work and responsibility for the result. If the answer is unclear to you, it is certainly unclear to everyone else, and unclear accountability is how work reaches a client unchecked.

DO NOT CONFUSE CONFIDENCE WITH CORRECTNESS

AI systems produce fluent, persuasive answers even when important parts are wrong. Review should be based on stakes and evidence, not on how confident the output sounds.

WHERE THE ETHICAL WEIGHT SITS

The people carrying the consequence are frequently not the people who chose to use the tool. That asymmetry is the reason question two is not a formality.

Run CHECK on Anything at Level 2 or Above

CHECK is a widely used community verification mnemonic. It takes under a minute once it becomes habit, and it is the single highest-return behavior in everyday AI use.

✓	STEP	ASK YOURSELF
☐	CLAIMS	What factual claims does this make? List them before you defend them.
☐	HALLUCINATIONS	Could any citation, statistic, quotation, name, or date have been invented?
☐	ERRORS IN REASONING	Does the logic actually hold, or does it only sound like it does?
☐	CONTEXT	Did the model have what it needed, or is this generic where it should be specific?
☐	KNOWLEDGE GAPS	Is this time-sensitive, proprietary, or particular to your industry in a way the model could not know?

ESCALATE RATHER THAN DECIDE WHEN ANY OF THESE IS TRUE

- › You cannot tell whether the work is Level 2 or Level 3.
- › The output contradicts something you believe to be true and you cannot resolve which is right.
- › Verifying it properly would take longer than doing the work yourself would have.
- › The model appears to be filling gaps in missing organizational context in ways that materially affect the recommendation.
- › No qualified person is clearly responsible for approving the decision or standing behind the final output.
- › You are being asked to sign off on something you are not qualified to evaluate.

THE ONE THAT MATTERS MOST

Being asked to sign off on something you are not qualified to evaluate. The pressure to approve work you cannot assess is the characteristic failure mode of AI in organizations, because the output looks finished and competent regardless of whether it is either.

What This Looks Like Across Functions

Examples, not rules. Your obligations may place work at a different level than the one shown here, and where they do, your obligations win.

FUNCTION	USUALLY LEVEL 1	LEVEL 2 OR 3
FINANCE AND ACCOUNTING	Drafting commentary, explaining a variance to yourself, formatting a schedule.	Anything entering the financial statements, a filing, a covenant calculation, or a board pack. Judgment about estimates and treatment stays with a qualified person.
LEGAL AND CONTRACTS	Summarizing a clause for your own understanding, generating a first list of questions.	Any advice, any filing, any executed document, any interpretation someone will rely on. Citation fabrication has already ended careers here.
HUMAN RESOURCES	Drafting a job description, rewording a policy, preparing interview questions.	Hiring decisions, discipline, termination, performance evaluation, accommodation, compensation. Documented human judgment is required, and in some jurisdictions so is disclosure.
HEALTHCARE AND CLINICAL	Explaining a general concept, drafting non-clinical correspondence.	Anything touching diagnosis, treatment, clinical documentation, or protected health information. This is Level 3 by default and by regulation.
CLIENT-FACING WORK	Internal drafts, structuring your own thinking, preparing questions.	Deliverables, correspondence that commits you to something, anything quoted back to you later, and anything a client will act on.
OPERATIONS AND SAFETY	Reformatting a procedure, drafting a training outline.	Safety procedures, incident findings, and anything that changes how a physical task gets done.
PUBLIC AND STRATEGIC	Brainstorming, testing how an argument sounds.	Public statements, funder and investor communications, board material, and anything that becomes the organization's position.

This is general guidance rather than legal, medical, or compliance advice. Regulated organizations should have counsel and their compliance function set the lines that apply to them.

Human oversight works best when it is designed into the process instead of added after something goes wrong. A checkpoint belongs wherever work changes hands or leaves the organization, because those are the moments after which correction becomes expensive.

1

IDENTIFY THE DECISION POINT

Locate the moment where AI output begins to influence a communication, recommendation, approval, or action.

2

DEFINE THE REVIEW STANDARD

Specify what the reviewer must check: accuracy, source support, confidentiality, bias, policy alignment, or professional appropriateness.

3

NAME THE DECISION OWNER

Assign a specific person with authority to approve, reject, revise, or escalate the work.

4

DOCUMENT WHAT MATTERS

For consequential uses, preserve what was reviewed, which sources were relied on, and who approved the final action.

5

CREATE AN ESCALATION PATH

Define what happens when the output is uncertain, contradictory, or beyond the reviewer's expertise.

GOOD HUMAN OVERSIGHT

The reviewer knows what to check, has enough context, has authority to act, and is expected to challenge the AI when needed.

WEAK HUMAN OVERSIGHT

The reviewer is asked to "look it over" after the decision is effectively made, lacks the expertise or time to challenge it, or assumes the AI is probably right.

THREE RULES FOR THE CHECKPOINT ITSELF

- › **Name the reviewer.** One named person per checkpoint. Shared responsibility reliably produces none.
- › **Bind sending to checking.** The person who sends it is the person who checked it. Splitting those two roles is how unchecked work ships.
- › **Reserve the second reviewer.** A second reviewer belongs at Level 3 and almost nowhere else. Applied everywhere it becomes a formality people learn to skip.

WHAT TO DOCUMENT

For Level 3 work, one line is usually enough: what the AI contributed, who verified it, and when. Keep it with the work product rather than in a separate log nobody will find.

WHEN TO DISCLOSE

Where a reasonable person would expect to know: external publications, client deliverables, funder submissions, and anything involving a likeness or voice. Check your contracts first; several probably already address it.

A sixty-second human-review checklist for meaningful AI output. Print this one and put it where people can see it.

✓	CHECK	ASK YOURSELF
☐	PURPOSE	Is the AI being used for an appropriate task, with a clear objective and scope?
☐	CONTEXT	Did the model receive the information and constraints it needed, without exposing sensitive or restricted data?
☐	FACTS	Have important claims, numbers, names, dates, quotations, and sources been independently checked?
☐	LOGIC	Does the conclusion actually follow from the evidence, or is the model making unsupported leaps?
☐	FIT	Does the output reflect this organization, audience, and situation rather than generic assumptions?
☐	BIAS AND IMPACT	Could the output disadvantage, misrepresent, or create avoidable harm for a person or group?
☐	CONFIDENTIALITY	Is this information appropriate to share with the tool being used, under organizational rules and applicable obligations?
☐	AUTHORITY	Is a qualified person reviewing the work before it informs a consequential decision or external communication?
☐	OWNERSHIP	Is it clear who approves the final output and accepts responsibility for the action taken?
☐	ESCALATION	If uncertainty remains, is there a defined person, source, or process to consult before proceeding?

WHY THIS IS AN OPERATIONAL QUESTION

Models produce wrong information with complete confidence. That is not a defect awaiting a patch in the next version, it is a property of how they work. Organizations that treat verification as a values statement tend to have inconsistent verification. Organizations that treat it as a step in a process, assigned to a named person at a defined point, tend to have reliable verification and, as it happens, a defensible ethical position as well.

Keep human judgment, authority, and accountability visible at the moments where they matter.

General organizational guidance. Not legal, medical, clinical, financial, compliance, or regulatory advice.

© 2026 The Jonesi Consultancy. All rights reserved.

Deciding where the checkpoints belong in your operation.

We help organizations set the lines, train to them, and put governance behind them.

