



UNAUTOMATED
THE HUMAN INTELLIGENCE PROJECT

PRACTICAL GUIDE 004

THE HUMAN IN THE LOOP PLAYBOOK

A leadership guide for deciding what AI may assist with, what requires human review, and what should never be delegated.

FREE EDUCATIONAL RESOURCE

Plain language. Practical tools. Human judgment.

unautomated.org

The Human in the Loop Playbook

Putting a human 'in the loop' sounds reassuring, but the phrase can become meaningless if nobody defines what the human is expected to do. A person who automatically approves hundreds of AI recommendations may be present in the workflow without exercising meaningful judgment.

HOW TO USE THIS RESOURCE

Use this playbook to classify AI supported work into three categories: AI MAY ASSIST, HUMAN REVIEW REQUIRED, and DO NOT DELEGATE. Then define what meaningful review looks like for your organization.

The three zones

AI MAY ASSIST	Low consequence, reversible, administrative, creative, or exploratory work where errors can be noticed and corrected without materially harming a person. Examples include brainstorming, formatting, summarizing nonconfidential material, first drafts, routine scheduling, and internal organization.
HUMAN REVIEW REQUIRED	Work where an AI output can materially affect money, rights, opportunities, safety, employment, reputation, customers, official records, or important organizational decisions. AI may inform the decision, but a qualified person must review the relevant information and remain responsible.
DO NOT DELEGATE	Decisions or acts that should remain human because delegation would remove meaningful accountability, dignity, consent, due process, or moral responsibility. The exact boundary depends on context, but an organization should define prohibited uses before pressure or convenience makes the decision for it.

What meaningful human review requires

Authority

The reviewer can disagree with the AI and change the outcome.

Time

The reviewer has enough time to examine the matter rather than rubber stamp a recommendation.

Information

The reviewer can see the relevant facts, uncertainty, source material, and important limitations.

Competence

The reviewer understands the subject well enough to recognize an implausible or harmful recommendation.

Independence

The workflow does not punish or discourage people for overriding the system when judgment requires it.

Accountability

A person or institution can explain the final decision and remains responsible for its consequences.

Appeal

When a decision materially affects a person, there is a meaningful way to ask for review, correction, or reconsideration.

Decision matrix

ACTIVITY	ZONE	WHY / CONTROL
Meeting summaries	ASSIST	Human checks important commitments, assignments, and official records.
Routine customer drafts	ASSIST	Human review when the message creates commitments, resolves disputes, or affects money.
Marketing concepts	ASSIST	Check factual claims, rights, brand risk, and deceptive synthetic media.
Hiring administration	REVIEW	Do not allow AI to become the sole basis for selecting or rejecting a person.
Employee discipline	REVIEW	A qualified leader reviews context, evidence, policy, and fairness.
Housing, lending, insurance, or access	REVIEW	High consequence decisions require meaningful review, explanation, and applicable legal safeguards.
Medical or safety decisions	REVIEW	Qualified professionals remain

		responsible. AI output does not replace clinical or safety judgment.
Legal strategy or rights	REVIEW	Appropriate professionals verify law, facts, jurisdiction, and consequences.
Autonomous impersonation	DO NOT DELEGATE	Do not use AI to deceive people about whether they are interacting with a person when the distinction matters.
Fabricating evidence	DO NOT DELEGATE	Synthetic evidence should not be created or presented as authentic.
Irreversible high consequence action without review	DO NOT DELEGATE	Do not permit an AI system to make a final consequential decision merely because automation is possible.

A seven question test before delegating

1. How serious is the consequence if the AI is wrong?

Notes: _____

2. Can the decision be reversed or corrected later?

Notes: _____

3. Does the person affected know AI is involved when that matters?

Notes: _____

4. Can a qualified reviewer understand the basis for the recommendation?

Notes: _____

5. Can that reviewer realistically disagree with the system?

Notes: _____

6. Would we be comfortable explaining this workflow to the person most affected by it?

Notes: _____

7. Who remains responsible when something goes wrong?

Notes: _____

THE HUMAN QUESTION

If we remove the human from this decision, are we removing unnecessary labor or are we removing judgment, responsibility, dignity, or the ability to appeal?

Workshop exercise

Choose five AI uses already occurring or being considered in your organization. Classify each one, identify the human owner, and describe what review actually means.

AI USE	ZONE	HUMAN OWNER	REVIEW STANDARD	PROHIBITED / ESCALATE IF
A PRACTICAL PRINCIPLE Human oversight should become stronger as the consequence of being wrong becomes greater. Presence is not enough. Review must be meaningful.				