



Responsible AI Use Guidelines

A policy-style framework your team can adopt to use AI tools safely, consistently and with clear accountability.

WORKS WITH THE ASSISTANTS TEAMS ALREADY USE



EDITION 01 / EMBER & ICE

EXECUTIVE WORKING EDITION - 08 OCT 2026

01

AT A GLANCE

A policy-style framework your team can adopt to use AI tools safely, consistently and with clear accountability.

These guidelines turn responsible AI from a values statement into daily rules people can follow. They define what is approved, what needs approval and what is not allowed, how to handle information by sensitivity tier, when a human must review output, how to disclose AI use, and what to do when something goes wrong. Each section ends with a working question, and the document includes checklists, an intake form, an incident note format and a one-page quick reference. Adapt the text to your organization and have counsel confirm it against current law. This is general guidance, not legal advice.

WHO IT IS FOR	Executives, department heads, IT, security, privacy and HR leads, and team managers who must publish and enforce rules for everyday AI use across a company of 20 to 2,000 people.
TIME TO USE	60 minutes to read; one half-day working session to adapt and approve a first version

What you will have at the end

- A short, adoptable policy covering purpose, scope, principles and roles
- Three-column rules for approved, approval-required and prohibited AI uses
- Four information handling tiers that tell staff exactly what may enter which tool
- Human review rules by risk level, plus disclosure and attribution norms
- An incident process, intake form and quick reference your team can use the same week

Contents

01	Why these guidelines exist and who they cover	03
02	Six principles that decide the gray areas	04
03	What is approved, what needs approval, and what is not allowed	06
04	Four tiers that tell people what may enter which tool	08
05	Review rules by risk level, plus disclosure and attribution	09
06	What to do when something goes wrong	11
07	Owners, approvers, reporters, and what everyone must learn	12
08	Review cadence, intake form and the one-page reference	13
09	A first quarter that produces evidence	15
10	Terms used in this publication	16

PURPOSE AND SCOPE

Why these guidelines exist and who they cover

01



A good AI use policy gives people a clear yes for common work and a clear route for everything else.

Most staff already use AI, whether the company has a policy or not. They draft emails, summarize documents, write code, brainstorm and translate. The question is whether they do it inside boundaries you chose or outside boundaries nobody set. These guidelines exist to move use into the first category. They protect customers, colleagues and the organization, and they protect the individual employee who would otherwise guess where the line is.

Scope matters because AI now appears in many forms. It includes standalone chat assistants, AI features inside office software, coding assistants, meeting transcription tools, image and audio generators, automation platforms and agents that take actions on a user's behalf. The guidelines apply to all of them, whether the tool is paid for by the company, offered free to the public, or built into a product someone already licenses.

Who and what is covered

- Every employee, contractor, intern and temporary worker who does work on behalf of the organization.
- Every AI tool, model, feature or agent used for work, including personal accounts used for work tasks.
- Every kind of work output touched by AI: text, code, analysis, images, audio, decisions and recommendations.
- AI used inside products and services you deliver to customers, which also follows the implementation guidelines and any contract terms.



NOT LEGAL ADVICE

These guidelines are an operating template. Laws on data protection, employment, consumer protection, intellectual property and AI itself differ by place and change over time. Have qualified counsel review the adapted version against the rules where you operate before you publish it.

WORKING QUESTION

If a new hire asked today which AI tools they may use and for what, would they get one consistent answer?

PRINCIPLES

Six principles that decide the gray areas

02

Rules cover what you foresee; principles cover everything else, so keep them short and make them testable.

No policy can list every situation. People will meet new tools and new tasks faster than you can update a table. A small set of principles gives them a way to reason when the rules are silent. These six are written as commitments that can be checked. They draw on the common themes in frameworks such as the NIST AI Risk Management Framework and ISO/IEC 42001, expressed in plain working language rather than abstract values.

PRINCIPLE	WHAT IT MEANS IN PRACTICE	TEST QUESTION
Accountable	A named person owns every AI-assisted output and every AI system. The tool never owns the result.	Whose name is on this?
Proportionate	Review effort rises with the harm a wrong output could cause.	What happens if this is wrong?

PRINCIPLE	WHAT IT MEANS IN PRACTICE	TEST QUESTION
Protective of data	Information goes only into tools approved for its sensitivity tier.	Am I allowed to give this tool this data?
Transparent	People know when they are dealing with AI or AI-generated material that matters to them.	Would the reader feel misled if they knew?
Fair	AI is not used in ways that disadvantage people unlawfully without a way to contest the outcome.	Who could be treated worse because of this?
Verified	Facts, figures, citations and code are checked before anyone relies on them.	How do I know this is true?

Train managers to use the test questions in everyday conversation. A manager who asks 'whose name is on this?' when a team presents AI-drafted work does more for accountability than a page of policy text. The questions are deliberately short so they can be asked in a meeting without ceremony.

COMMON MISTAKE

Writing principles that cannot be violated. 'We use AI ethically' is not testable. 'A named person verifies any figure before it leaves the company' is. If nobody could ever be found in breach, the principle is decoration.

WORKING QUESTION

Which of these six principles would your people find hardest to follow on a busy day, and why?

PERMITTED USE

What is approved, what needs approval, and what is not allowed

03

Sort every use into one of three lanes so most work moves fast and only the few risky uses slow down.

The core of any usable policy is a short list in three lanes. Approved uses need nothing more than good judgment and the information rules in this document. Approval-required uses go through the intake form in the next section. Prohibited uses are off the table unless the policy owner grants a written exception. Publish the lists and update them as the tool set and your confidence change. The examples below are a starting point to edit, not a final answer.

Approved without further approval

USE	CONDITIONS
Drafting and editing internal Approved tool; Tier 1 or Tier 2 information only; author reviews and text such as emails, notes and outlines	owns the result
Summarizing public or non-sensitive documents	Check the summary against the source for anything that will be relied on
Brainstorming, outlining, naming and structuring ideas	No confidential detail in prompts
Explaining concepts, learning and practice exercises	Treat explanations as leads to verify, not authority
Writing and explaining non-sensitive code with an approved coding assistant	Standard code review and testing still apply
Translating non-sensitive text for internal understanding	Do not publish translations without a fluent human review

Needs approval before use

USE	WHY APPROVAL IS NEEDED
Any use with Tier 3 information	Needs a tool and contract that restrict data use, plus an owner

USE	WHY APPROVAL IS NEEDED
Customer-facing AI output, chatbots or auto-generated communications	Wrong output reaches customers; testing, disclosure and monitoring required
Any AI-assisted input to decisions about hiring, pay, promotion, credit, insurance, benefits or health	Affects individuals; legal and fairness review required
Connecting AI to company systems, or letting it send, buy, delete or change anything	Actions can be irreversible; permissions and approval gates required
Meeting recording, transcription or voice and video generation	Consent, privacy and recording laws vary by location
Adopting a new AI tool or enabling an AI feature in existing software	Vendor terms, security and data flows need review

Not allowed

- Entering Tier 4 information into any AI tool, whatever its settings.
- Using personal or unapproved AI accounts for company confidential, client or personal data.
- Presenting AI-generated work as personal expertise or as verified fact when it is neither.
- Creating deceptive content, impersonating a real person without consent, or fabricating evidence, reviews or records.
- Using AI to make a final decision about a person's employment, credit, benefits or similar significant matter with no human review.
- Attempting to bypass security controls, filters or approval gates with AI assistance.

WORKING QUESTION

Which uses are people in your team already doing that this three-lane list would classify as needs approval or not allowed?

INFORMATION HANDLING

Four tiers that tell people what may enter which tool

04

Staff should never have to guess whether data is safe to paste; a four-tier scheme makes the answer a lookup.

Most real AI incidents begin with data placed in the wrong tool. Prevent this with a simple tier scheme linked to a list of approved tools. If your organization already has a data classification policy, map these four tiers onto it rather than creating a second scheme. What matters is that every person can name the tier of what they are about to paste, and knows which tools are cleared for it.

TIER	EXAMPLES	WHERE IT MAY GO
Tier 1: Public	Published website copy, marketing already released, public regulations	Any approved tool, and public tools at the user's discretion
Tier 2: Internal	Internal memos, non-sensitive plans, generic process documents	Approved company tools only
Tier 3: Confidential	Client work, contracts, financials, source code, unreleased plans, personal data of customers or staff	Approved tools with data-use restrictions and a recorded owner
Tier 4: Restricted	Credentials and keys, health records, government identifiers, payment card data, privileged legal advice, regulated data	No AI tool, unless the policy owner and counsel have approved a specific system

Rules that apply to every tier

- Use the minimum necessary: remove names, identifiers and numbers the task does not need before pasting.
- Check the account type. A company-managed account with data-use restrictions is not the same as a personal account on the same product.
- Do not upload files you do not have the right to share, including third-party material under confidentiality terms.

- Treat prompts and outputs as business records that may be retained, reviewed or requested in legal or regulatory proceedings.
- If you realize you pasted something into the wrong tool, report it immediately; do not delete and hope.



CHECK THE CONTRACT, NOT THE MARKETING

Whether a tool may be used for Tier 3 depends on its terms: whether it trains on your inputs, how long it retains them, where it processes them and who its sub-processors are. Procurement, security and privacy should confirm these points before the tool is listed as approved for that tier.

WORKING QUESTION

Can you point to the list that says which approved tool may hold which tier, and does every employee know where it is?



HUMAN REVIEW

Review rules by risk level, plus disclosure and attribution

05



A person must stay responsible for what AI produces, and the depth of their check should match the harm of an error.

Models produce fluent text whether or not it is true. They can invent figures, sources and legal citations, misread instructions and reflect bias in their training data. The remedy is not distrust of everything, which would waste the benefit. It is a review rule scaled to risk. Low-risk drafts need a read-through. High-risk outputs need independent verification against source material by someone competent in the subject, with the check recorded.



RISK LEVEL	EXAMPLE OUTPUTS	MINIMUM HUMAN REVIEW
Low	Internal drafts, brainstorming, meeting summaries for the attendees	Author reads in full and corrects errors before use

RISK LEVEL	EXAMPLE OUTPUTS	MINIMUM HUMAN REVIEW
Medium	External communications, reports, analysis, code going to production	Author verifies facts and figures against sources; a second person reviews before release
High	Legal, financial, medical or safety content; anything affecting an individual's rights or opportunities	Qualified expert verifies against primary sources; decision-maker is a human; record kept
Critical	Automated actions that are hard to reverse, or decisions at scale	Approved design with approval gates, monitoring and a tested way to pause

Disclosure and attribution norms

- Tell people when they are interacting with an AI system, such as a chatbot, and how to reach a person.
- Disclose AI assistance when it matters to the reader: in formal deliverables for clients, regulated filings, academic or credentialed work, and wherever a contract or law requires it.
- Do not claim AI-written work as the unaided product of a named expert; the expert must have reviewed and adopted it.
- Credit and respect source material. Do not present summaries of copyrighted content as original, and check licensing before publishing generated images or text at scale.
- Follow platform and market rules on labeling synthetic media, which are changing quickly in many jurisdictions.

EXAMPLE

An analyst uses an assistant to draft a market summary. She checks every figure against the cited source, finds one number invented by the tool, corrects it, and her manager reviews the final version before it goes to a client with a note that AI assisted the first draft.

WORKING QUESTION

For your most important external document last month, who verified the facts, and could they show how?

INCIDENTS AND ESCALATION

What to do when something goes wrong

06

K'

People report problems quickly only when reporting is easy, blame-free and clearly worth their time.

An AI incident is any event where AI use caused or could have caused harm, a breach of these guidelines, a data exposure, a materially wrong output being relied upon, a biased or offensive result, or a tool behaving outside its intended scope. Near misses count. The aim is learning and containment, not punishment. Reserve disciplinary action for deliberate or repeated violations and for concealment. Say so explicitly, because fear of blame is the main reason incidents are reported late.

01 Stop and preserve

Stop using the tool for the task. Keep the prompt, output and any screenshots. Do not delete anything.

02 Report

Tell your manager and the policy owner through the designated channel as soon as you notice, ideally within one working day. Use the incident note format below.

03 Triage

The policy owner classifies severity, decides who else to involve (security, privacy, legal, communications, HR) and orders containment such as disabling access or rotating credentials.

04 Notify where required

Counsel decides whether customers, regulators or other parties must be told, and by when. Many data protection laws set short notification windows, so involve legal early.

05 Fix and learn

Correct the harm, fix the cause, update guidance or training, and record the lesson. Share the pattern, not the person's name.

Incident note format

- Date, time and name of the reporter, and the name of the system or tool involved.

- What happened, in plain words, and how it was found.
- Information involved, with its tier, and who could have seen or relied on it.
- Immediate action taken and by whom.
- Suggested cause and suggested prevention.
- Severity (low, medium, high, critical) set by the policy owner, and the date the incident was closed.

WORKING QUESTION

If someone pasted a client contract into a personal account this afternoon, would they know whom to tell and feel safe doing it?

ROLES AND TRAINING

Owners, approvers, reporters, and what everyone must learn

07

Policies work when each rule has a named person behind it and each person has had a short, practical briefing.

ROLE	RESPONSIBILITY	TYPICAL HOLDER
Executive sponsor	Approves the policy, sets risk appetite, resolves disputes	COO, CIO, CRO or general counsel
Policy owner	Maintains these guidelines, the approved tool list and the registered decisions	Risk, compliance or an AI program lead
Approvers	Decide approval-required requests within the stated service level	Policy owner plus security, privacy and legal as needed
Managers	Apply the rules in their teams, check review practice, escalate issues	Team and department leads
Users and reporters	Follow the rules, verify output, report incidents and near misses	Every person covered by scope

Set a service level for approvals, such as a decision within five working days for medium-risk requests, and publish it. A slow or silent approval path is the largest single driver of unapproved use. Track how long decisions take and report it to the sponsor along with the number of requests approved, changed and declined.



Training and acknowledgement

- All staff: a short briefing (about 45 minutes) at onboarding and annually covering the three lanes, the data tiers, review rules and how to report.
- Managers: additional guidance on coaching review habits and handling reports without blame.
- Roles with higher exposure such as HR, finance, legal, customer support and engineering: scenario-based training using their own workflows.
- Every person acknowledges the guidelines in writing or through a recorded click at onboarding and after each material update.
- Keep records of who completed training and when, because they are evidence that the policy operates in practice.



WORKING QUESTION

Who owns the approved tool list today, and when did they last update it?

REVIEW AND QUICK REFERENCE

Review cadence, intake form and the one-page reference

08



Policies decay quickly in a fast-moving field, so schedule the review and give people a page they will actually keep open.



Review these guidelines at least twice in the first year and then annually, with extra reviews when a major tool is adopted, a law changes materially or an incident shows a gap. Each review should consider incident reports, questions staff keep asking, requests that were declined, and any changes to vendor terms. Version the document, list what changed and

communicate the changes in plain language.

01 Intake form: describe the use

Requester, team, tool or system name, and a plain description of the task and who benefits.

02 Intake form: data and people

Highest information tier involved, whether personal data is included, and who is affected by the output.

03 Intake form: consequence

What happens if the output is wrong, and the proposed review step, owner and fallback if the tool is unavailable.

04 Intake form: decision

Approver, date, conditions, review date and the entry made in the register.

Before you use AI for a work task

- The tool is on the approved list for the information tier I am about to use
- I removed names, identifiers and detail the task does not need
- I know who owns the result and it is me or someone I named
- I will verify facts, figures, quotes and code before anyone relies on them
- I know whether this output needs a second reviewer or disclosure
- I know how to report a problem

QUICK REFERENCE	RULE
Always	Verify before relying; own the result; use approved tools; minimize data
Ask first	Customer-facing output; decisions about people; system access; new tools; Tier 3 data
Never	Tier 4 data in AI; deception or impersonation; automated final decisions about people
If it goes wrong	Stop, keep evidence, report within one working day, no blame for honest reports
Questions	Ask the policy owner; answer within five working days

WORKING QUESTION

What would have to change in your organization for people to open the quick reference before they act rather than after?



30 / 60 / 90

A first quarter that produces evidence

09



Roll the guidelines out in three steps: publish a usable version, make approval and reporting work, then prove the policy operates.

First 30 days: ground it

- Name the executive sponsor and the policy owner
- Adapt these guidelines to your tools and data classes, and have counsel review them
- Publish the approved tool list with the tier each tool is cleared for
- Send the one-page quick reference to all staff

**Days 31 to 60: test it**

- Run the 45-minute briefing for all staff and a deeper session for managers
- Open the intake form and meet the five-day approval service level
- Collect written or recorded acknowledgement from every person in scope
- Set up the incident channel and run one tabletop exercise

Days 61 to 90: decide

- Review every approval and incident so far and update the guidelines
- Report approvals, declines, training completion and incidents to the sponsor
- Sample reviewed outputs from two teams to check that verification happens
- Set the annual review date and training calendar



GLOSSARY

Terms used in this publication

10

TERM	MEANING
Acceptable use	The set of rules describing what people may, may not and must ask permission to do with a tool.
Approved tool	An AI product or feature the organization has reviewed and listed as cleared for stated information tiers.
Information tier	A level of data sensitivity, from public to restricted, that determines which tools may receive it.
Hallucination	A fluent output that is false or unsupported, such as an invented figure, source or legal citation.
Human in the loop	A design where a person reviews or approves AI output before it has an effect.
Automation bias	The tendency to accept machine output without enough checking.
Shadow AI	AI tools used for work without the organization's knowledge or approval.
Policy owner	The person who maintains the guidelines, the approved tool list and the record of decisions.
Near miss	An event that could have caused harm but did not, reported so the cause can be fixed.
NIST AI RMF	A voluntary US framework for managing AI risk, organized around govern, map, measure and manage.
ISO/IEC 42001	An international standard for an AI management system that organizations can implement and be audited against.
Tabletop exercise	A discussion-based rehearsal of an incident to test roles, decisions and communication.



A note on use

These guidelines are general operating guidance and not legal advice. Align the adapted version with qualified counsel and the laws and contracts that currently apply to your organization.

TAKE THE NEXT STEP

Take the free Halden assessment at halden.company/assessments, or talk to us at halden.company/contact about turning this into a plan for your organization.

