

Norm and Value Setting for AI-Adapted Workflows

A guide for teams defining how they want to work with AI

Every AI workflow carries embedded choices about what to prioritize, what to emphasize, and what to produce. Those choices shape how teams are represented, how their work comes across, and whether the output reflects something they would actually stand behind. A workflow can be accurate, efficient, and well-structured while quietly reshaping how your team prioritizes, communicates, and makes decisions. When your team hasn't articulated its own values and ways of working, the tool's assumptions become the default. This guide helps teams define their own norms before the tool defines them, so that outputs reflect how your team actually works and what your work stands for.

Part 1: Questions to Consider

Before your team builds or adopts an AI-adapted workflow, work through these questions together. Not every question will apply to every workflow, but skipping them entirely is how teams end up with something technically useful but misaligned with what your team values about how it does this work. These questions are a starting point. If your team wants to go further and develop documented norms, the process guide and template below will walk you through it.

Terminology and Language

- Does the terminology in the AI outputs match how your team actually talks about this work? If the AI is using different terms than your team uses with each other, with partners, or with people accessing IRC services, the framing needs adjustment.
- Are there terms that carry specific meaning in your context that AI wouldn't know? Words that sound neutral in general use but carry loaded meaning in a specific country, program, or donor relationship.
- Would a colleague reading the AI output recognize it as coming from your team, or does it read like it could have been written by someone without your team's unique perspective?

Prioritization and Weighting

- Does what you've built weight what your team actually values? A prompt or skill or project etc. that prioritizes speed over thoroughness, or compliance over partnership quality, is making a values choice whether or not anyone intended it. Consider what are the key IRC values that are relevant to this task? How do those values show up in the work product?
- Whose perspective is centered in the outputs? If what you've built consistently surfaces one perspective (fiduciary risk, donor alignment, organizational positioning) while downplaying others (community voice, partner capacity, field realities), that's a design choice, whether or not anyone made it intentionally.
- What would be lost if this workflow became the default way of doing this work? Is there productive friction in the current process (e.g. thinking, consulting, wrestling with tradeoffs) that has value beyond efficiency?
- How and when will the values you identified be checked during the workflow? oDoes everyone who will do a values check have the same understanding of what a successful and unsuccessful representation of the values look like?

Decision Making

- Is the AI being asked to recommend things it should only be informing? There's a difference between surfacing options and ranking them or making a recommendation. There's a difference between AI identifying gaps and AI proposing how to fill them. Where that line sits depends on your domain, but your team should name it explicitly.
- Who is accountable for the decision this workflow feeds into? If the answer is unclear, or if the workflow makes it easy to skip the human decision point, the design of the adapted workflow needs to change.
- If the AI gets this wrong in a way that isn't immediately obvious, who would notice and how long would it take? And what are the potential repercussions?
- Is the human checkpoint at each stage a designed process or just an intention? There's a difference between "a human reviews the output" as a stated goal and a process where a named person reviews against clear criteria at a specific point in the workflow. If you can't describe exactly who reviews what, against what standard, and what happens when they flag a problem, the oversight isn't real yet.

Contextual Knowledge

- What contextual knowledge is essential to the work of this team?
- Where does that contextual knowledge reside and can we input it into Claude?
- For knowledge that cannot be input, determine how it will be incorporated into outputs and the checks of those outputs.

Stakeholders

- Who is affected by this workflow beyond the people using it (e.g. people accessing IRC services, partners, communities, other colleagues). Are their interests reflected in how the workflow is designed, or only the interests of the team using it?
- Has the team consulted with the people closest to the work and to the people it affects?
- If this workflow were widely adopted, what changes (positive or negative) would it create for people beyond your team?

Sustainability and Drift

- Does this fit into how your team already works? Workflows that require new meetings, new tools, or new routines face a higher adoption barrier than those that slot into what the team is already doing. If using this means adding a step to an existing process, that's manageable. If it means creating an entirely new touchpoint, think carefully about whether the team will sustain it once the initial enthusiasm fades.
- If someone other than the person who built this will use it, what's the handoff plan? The judgment about when to trust an output and when to question it often lives in the designer's head, not in the written instructions. Users who didn't build the workflow can follow the steps and catch obvious errors, but may not recognize when something is subtly wrong. Watch for hesitation, pauses, or workarounds from new users. These signal the handoff was incomplete.
- What happens to staff expertise if this workflow handles part of their work for an extended period? Will staff maintain the ability to do this work without AI if access were interrupted?
- What happens if this workflow scales beyond your team? Are there guardrails that only work because of your team's specific knowledge, and would those break down if others used the same workflow without that context?
- How will you know if the workflow's outputs have drifted from your norms over time? AI models change, team composition changes, context changes. What's your mechanism for checking that what was aligned six months ago still is?

Part 2: Process for Determining Your Team's Norms

This process is adapted from the methodology used by T4P to develop their AI for Proposal Development Guidance Note. It was designed to surface genuine positions, including areas of disagreement, rather than impose a single view.

Step 1: Start with a survey or pulse check

Before convening the group, gather initial perspectives. A short survey asking staff about their current AI use, where they see potential, and where they have concerns gives you a baseline and surfaces the range of opinion. Keep it brief (5 to 10 questions). The goal is to understand the landscape before the conversation.

Step 2: Host a listening session or sessions

Bring together staff from across the relevant functions, not just the most senior or the most AI-enthusiastic, but the people who do the work and the people whose work is affected by it. For the BD guidance note, this included staff across AMU, RAM, regions, program offices, and country programs. The cross-functional composition matters because different perspectives surface different risks.

In these sessions, map the workflow or the proposed use(s) of AI against different AI collaboration modes.

- AI reviews/checks human-written content
- AI structures, human writes
- Human inputs verified documents, AI drafts, human reviews and substantially rewrites
- AI drafts, human reviews and substantially rewrites

For each component, ask: what is the maximum level of AI involvement the group feels is appropriate, and what conditions need to be in place for that level to be responsible?
Name areas of consensus and areas of genuine disagreement openly. Not every section will produce agreement, and documenting where the team diverges is as valuable as documenting where it converges.

Four Collaboration Modes

AI drafts, human reviews and substantially rewrites

AI produces a starting point that a human substantially revises to ensure accuracy, specificity, and IRC voice.

AI structures, human writes

AI structures, human writes: AI creates an outline or framework; the human writes all substantive content.

AI reviews/checks human-written content

The human writes; AI checks for gaps, inconsistencies, compliance issues, or alignment with donor requirements.

Human Only

No AI involvement; the section requires judgment, relationships, or contextual knowledge that AI cannot access or replicate.

Step 3: Document and validate

After the sessions, synthesize what you heard into a written guidance note. Organize it in two parts: high-level themes that emerged across the conversations (the principles behind the norms), and a component-by-component breakdown covering each part of the workflow. For each component, document the maximum collaboration mode, what the AI role should be, what the human role must be, and any conditions or caveats.

Share the draft back with the group for validation. Ask: did I get this right? What did I miss? What do you disagree with? Revise based on their feedback. The document should reflect where the team actually landed, including areas of tension, not a sanitized version that avoids hard questions.

Step 4: Position it as norms, not rules

The BD guidance note is explicitly positioned as shared norms, not a compliance requirement. This matters. Staff are expected to use judgment, and the document supports that judgment rather than replacing it. Frame your team's document the same way: this is our shared understanding of how we want to work with AI. It represents where we landed. It's a reference point, not a rulebook. Staff are always free to use less AI involvement than what the document suggests.

Step 5: Build in a review mechanism

AI capabilities change, team composition changes, and norms that made sense six months ago may need updating. Build in a review point (semiannually or annually) where the team revisits the guidance note and asks: is this still right? Has anything changed in our work, our tools, or our context that should shift these norms?

Part 3: Template for Documenting Team Norms

Adapt this template for your team's context.

[Team Name] — AI Collaboration Norms

This document represents [team name]'s shared norms for how we collaborate with AI in [area of work]. It is not a compliance requirement. It reflects where staff across [list relevant functions/roles] landed on where AI can add value, where it creates risk, and what needs to be in place for it to be used well.

This guidance is primarily for [target audience]. [Other relevant stakeholders, e.g. reviewers, supervisors] should be aware of these norms and incorporate them into their oversight.

Background

[Brief description of the work, why AI norms are needed, and how this document was developed — who was consulted, what process was used.]

What We Heard: High-Level Themes

[3-10 themes that emerged from the input sessions. For each theme, state the principle and a "what this means in practice" section.]

Quick Reference Table

Workflow Component	Maximum Collaboration Mode	AI Role	Human Role
[Component 1]	[Mode]	[What AI does]	[What the human owns]

Collaboration Modes (define for your context)

- AI reviews/checks human-written content — lowest AI involvement. AI surfaces issues; humans resolve them.
- AI structures, human writes — AI creates the framework or outline; human writes the substance.
- Human inputs verified documents, AI drafts, human reviews and substantially rewrites — human gathers and verifies source material first; AI produces a draft from that material; human reviews and rewrites.
- AI drafts, human reviews and substantially rewrites — AI produces a starting draft; human substantially rewrites. Output must reflect human judgment and voice, not AI's first pass.

Workflow Components: Where We Landed

For each component:
[Component Name]
Maximum Collaboration Mode: [Mode]
[2 to 3 paragraphs explaining: what the group decided, why, any conditions or caveats, and areas of disagreement if relevant.]

What Should Never Be Entered Into AI Tools

[List categories of information that should not be used as AI inputs, adapted from your team's data sensitivity landscape.]

Document History

- Developed by: [Names]
- Input from: [Who was consulted, how many staff, which functions]
- Date: [Date]
- Next review: [Date]