

The ChangemakerXchange  
Mindful AI  
Manifesto & Policy

This document was last revised  
in November 2025.

# The ChangemakerXchange Mindful AI Manifesto

## Our Compass in the Age of AI

Human connection lies at the heart of everything we do at ChangemakerXchange, as we support individuals and groups through programmes that help them sustain, scale, or deepen their impact in the world.

The advent of Artificial Intelligence is already profoundly affecting how people connect and relate to themselves, to each other, and to the world around them. This transformation carries both potential and great risk for the social fabric we care so deeply about.

The term 'AI' has become a clunky catch-all, obscuring the vast differences in each tool's ethical and environmental impact.

We see powerful potential in specific AI applications used by changemakers—from using image recognition to make second-hand shopping effortless, to powering digital interpreters for the deaf community, or discovering new materials for clean water through data modelling.

Yet we are deeply critical, especially of Generative AI, recognising its risks: the immense environmental toll of its data centers energy and water use, its tendency to amplify bias by being trained on an unequal past, and its potential to flatten the unique human creativity our work depends on.

We must also challenge the emerging 'AI Empire', the concentration of power and resources in a few hands, often built on hidden labour and environmental extraction, and ask who truly benefits.

Rooted in our Team Manifesto and Community Values, this AI Manifesto serves as our 'why', a compass, a guide to help us navigate the world of AI with intentionality, conscience, and integrity.

It is the foundation for our **Mindful AI Policy**, a separate document which details the 'How', the practical implications of AI for our governance, processes, and daily decision-making.

We believe the social impact sector must share the onus of shaping the present and future of AI, to ensure it serves humanity and the planet. To claim our role in steering this technology towards more just and regenerative present and future realities, we explore a different set of questions:

✨ *Should we? not just Do we?*

✨ *What exactly are we optimising for? Are we unleashing or unlocking potential? not just How do we become more efficient?*

✨ *How do we lead with integrity? not just How do we comply?*

We invite others to take this Manifesto, use what serves them, and adapt it to their own journeys as we collectively learn to navigate this time of complexity and uncertainty with care and intention.

## Our Mindful AI Principles

These principles guide how we engage with AI in our daily work, with curiosity, care, and a constant reminder of the human connections at the heart of all we do.

### 1. Mindful Use

AI is not the default. Before every potential use, particularly with generative AI, we try to pause and ask ourselves:

- *Is AI (really) needed for this task?*
- *Is it (really) aligned with our values?*
- *What do I lose when I use this tool for this task? (e.g. creativity, authenticity or genuine human connection)*

Some tasks call for our full presence and care, such as writing a heartfelt message to members, developing a theory of change, or designing a session for one of our gatherings. When such things are over-delegated to AI, they risk losing depth, authenticity or meaning.

Some of these tasks may benefit from AI assistance — for example, refining language, improving clarity, or offering new perspectives. And others may be better suited to AI altogether, such as financial forecasting or identifying connections across large data sets.

Whatever the case, we assess each use not only for its practical benefits but also for its effects on human connection and ensuring we stay true to our values.

## 2. Awareness and Responsibility

We acknowledge AI's broader social and environmental footprint and aim to stay informed about its real impacts.

We seek to understand which tools, models and companies are more transparent and socially and environmentally responsible.

We will choose, where possible, options with lower footprints or stronger ethical practices and aim to share what we learn with others in our network and community.

## 3. Human at the Core

We celebrate the spark of human creativity and the beautiful imperfections of authentic expression. While AI can be a powerful assistant, we believe the ultimate decision-making must remain firmly in our hands and not just 'a human in the loop' but retaining agency and oversight.

In communications, we guard against 'AI slop' (content that feels too generic or lacking in substance) and ensure our voice, stories, and connections remain recognisably human. AI may assist in tone or clarity, but cannot replace the empathy or humour that give our words life.

While AI can support our work in numerous ways, it mustn't replace our responsibility as ultimate decision-makers. Whether we are using AI tools to help us in a selection process, event flow design, or problem-solving session, we are the ones ultimately responsible for the outcomes and outputs of our work.

## 4. Data Privacy & Trust

We treat the data people share with us with the respect and care it deserves. We are committed to protecting privacy, and ensuring that any use of AI does not compromise the trust placed in us.

We strive to use tools responsibly, understand where and how data is handled, and obtain consent whenever personal or sensitive information may be involved. As a non-negotiable rule, **confidential or sensitive community and member data must not be input into public, third-party generative AI models.**

## 5. Speaking Up for What's Right

We will not shy away from holding those developing AI technologies accountable and we recognise that stepping into this space as ChangemakerXchange also means questioning how the technology is developed, governed, and deployed.

Responsible leadership in the changemaking ecosystem we are part of entails speaking out for justice, transparency and accountability.

# Our Journey Forward

Ultimately, we aspire to use AI to amplify human connection, creativity, and changemaking power, never to replace them. This manifesto is not a static document; it is a living and evolving commitment. We will formally review and adapt these principles and their associated policies on a regular basis, ensuring our 'compass' remains true as we learn and the technology evolves.

# The ChangemakerXchange Mindful AI Policy

A Living Document in Permanent Beta and reviewed every 6 months

## 1. Our Starting Point: The Framework

This document is the heart of our organizational approach to AI. It is our ‘how’, our practical guide for daily decisions about the technology.

It does not stand alone, but alongside **The CXC Mindful AI Manifesto**: Our ‘why’, the higher-level compass and principles which underpin this policy,




**Who it’s for:** This policy was written for all CXC core team members and contractors. We share it externally as an open-source resource for other organizations to adapt and use in their own journeys.

## 2. Using AI: The Traffic Light Check

This check is designed to be run the first time you use a **new workflow, tool, or data type**. For example:

- Using a **new AI tool** (e.g., trying a new image generator).
- Using an existing tool for a **new purpose** (e.g., using an LLM for selections when you’ve only used it for blog posts).
- Using an existing tool with a **new type of data** (e.g., using confidential community data for the first time).

We trust our team to be mindful in their more regular uses of AI (covered in section 3).

We use a traffic light system for AI usage. If the result is  Green, you’re cleared for that task. If it’s  Yellow, you get guidance from the AI Stewardship Circle (our internal AI review team, detailed in Sec 4.1). If it’s  Red, you should NOT use AI for this task.

# The Check: Find Your Action

Go through these steps in order. The first one you answer “YES” to determine your action.

## CHECK 1: THE RED LIGHTS (STOP)

Does your intended new use violate any of these red light rules? These are our non-negotiable red lines, directly reflecting our core Manifesto principles.

- 1. Safety & Harm** (from Principles 1, 2 & 3 in our manifesto) Could I be using AI to create content that could **directly cause harm, discriminate, or promote dangerous misinformation** (e.g., generating hateful content, giving unqualified medical/legal advice for a community programme, or creating propaganda)?
- 2. Human Agency** (from Principle 3 in our manifesto) Am I using AI to make a **final, un-reviewed decision** about a person (e.g., in a selection process)?
- 3. Deception & Misrepresentation** (from Principle 3 in our manifesto) Could I be using AI to deceive, such as creating a ‘deepfake’ (audio, video, or image) of a real person without their consent, or passing off generated content as a ‘direct quote’ or ‘personal story’ from a real community member?
- 4. Personal & Organisational Privacy** (from Principle 4 in our manifesto) Am I about to paste ‘**un-scrubbed**’ **confidential or sensitive** community/member data (e.g., a personal story, name, or email) into a **public, third-party** AI model?

**How to Fix This Red Light: The Data ‘Scrub’ (and Our Long-Term Plan)** We recognise that ‘scrubbing’ every document is a significant effort. It should be a last resort, not the daily rule.

As of November 2025, our organisational priority is to move our confidential AI work to secure, **solutions backed by a Data Processing Agreement (DPA)** (e.g. ‘Google for Nonprofits’ program or other low-cost enterprise-grade tools). This will allow us to handle confidential data safely *without* needing to scrub it, turning this “Red Light” workflow into a “Yellow Light” one.

In the meantime, if you *must* use a public tool, scrubbing is the only non-negotiable technique for protecting our people: \* **Duplicate**: Never work on the original document. \* **Replace Names & Identifiers**: Replace all names, emails, orgs, and locations with generic placeholders (e.g., [Community Member], [Their City]). \* **Generalise**: Broaden any unique details (e.g., ‘The only solar-powered bakery in rural Colombia’ becomes ‘A social enterprise in South America’). \* **The Litmus Test**: ‘If this scrubbed text leaked

online, could it be traced back to a specific person?' If no, it is now 'anonymised data', and you can proceed to Check 2.

**1. Informed Consent** (from Principle 4) Am I deploying an AI note-taker in an intimate community gathering without explicit consent?

➔ **If YES to any:** ● **RED LIGHT. STOP. ACTION: Do not proceed** (or, for the Privacy rule, Scrub Your Data First).

## CHECK 2: THE YELLOW LIGHTS (PAUSE)

If the answer to all Red checks was 'NO' (or you fixed the Privacy Red by scrubbing), continue:

This check is for **new workflows** or **new tools** that aren't 'Hard Red' but are too complex for a simple 'Green Light'.

- **Flag (High-Stakes Decisions / Critical Documents):** Am I using AI to help make or communicate a **high-stakes decision** about an external person (e.g., a selection assistant, *summarising applications, drafting rejection emails*) or to generate a **critical document** (e.g., a *grant proposal* or *donor report*)?
- **Flag (Data & Security Check):** Am I about to use **any confidential or sensitive data** (even 'scrubbed' personal data or internal org data) in **any** AI tool, and I'm **not 100% sure** if it has a DPA or is secure?
- **Flag (New Tool 'Gut Check'):** Am I about to use a **new AI tool** that isn't on our **AI Tools Inventory** (our internal list of vetted tools, *defined in Sec 4.2*) and it just *feels* like it might be misaligned with our Manifesto (e.g., its values, its energy use)?

➔ **If YES to any:** ● **YELLOW LIGHT. PAUSE & ASK. ACTION: This requires consultation. Stop and contact the AI Stewardship Circle.** This is *not* a "no"—it is a "let's check together".

## CHECK 3: THE GREEN LIGHT (GO)

If the answer to all Red and Yellow checks was "NO", you are Green.

This means your new workflow is: Using **public** or **fully scrubbed** data for an **internal** task (like a blog post draft or social media idea) OR for a **low-stakes** external task.

➔ **Action:** ● **GREEN LIGHT. You are good to go with this workflow.**

### 3. Our Rules of the Road (For ALL AI Use)

These are our basic rules of the road that apply to all AI use, even Green Light workflows. This is the part to have top of mind for your daily prompts.

- 1. Be Mindful (Prompt Critically & Sparingly):** Protect your time and creativity. Avoid 'doom-prompting' (endless, low-value prompt spirals). Be aware that LLMs are designed to be agreeable and will tell you what you want to hear.
  - **Anti-Sycophant Prompts:** To get more critical feedback, use prompt chains like:
    - **Prompt 1:** "Critically examine the core assumptions, unstated premises, and potential cognitive biases in my request above."
    - ...Wait for the answer, then...
    - **Prompt 2:** "Now, based *only* on logic and established facts, analyse my idea. Focus *also* on its limitations, counter-arguments, and potential downsides. What is being overlooked?"
- 2. Be Authentic (No 'AI Slop'):** If you use AI to help you write, ensure the final output is recognisably human, reflects our 'soul', and pay particular attention to content that feels generic, hollow, 'business speak' that is lacking a human touch and using many cliches and buzzwords. You are the author, not just the editor, definitely do NOT just copy-paste content over from an LLM without giving it your own personal touch!
- 3. Be Proportional (Environmental Impact):** AI is not virtual; it runs on massive, energy- and water-guzzling data centers. The full, often hidden, environmental footprint of this technology is still being uncovered, but we know the cost is significant. Therefore we also ask that you have in mind a mindful 'proportionality check'. As a general rule we look to avoid **high-energy uses for low-value tasks**.
  - **Relatively Low Energy:** Text generation (drafting, summarising).
  - **Medium Energy:** Simple image generation.
  - **High Energy:** Complex/photorealistic image generation.
  - **Very High Energy:** Any video, 3D model, or audio generation. This is a judgment call: Is the high energy cost of generating a video justified by the high value of the task?
- 4. Be Transparent (Disclose Use):**
  - **Substantive Use:** Disclose any AI use that **substantively generated or shaped** the final content. This does not apply to simple grammar or spelling checks.
  - **Evaluations:** AI use in selection or evaluation processes must be disclosed to applicants.

## 4. Bringing This Policy to Life

### 4.1 The AI Stewardship Circle

There is a small dedicated team within our core team (potentially including a community member / expert on AI) which acts as the keepers of our mindful AI manifesto and policy, ensuring it is kept up to date and applied consistently. Their role is to:

- Ensure the team is aligned with our approach to AI and incorporates all feedback.
- Ensure twice-yearly updates of all documents.
- Guide the team through an impact assessment of new tools / workflows.
- Escalate major or contentious decisions to the entire team.

### 4.2 Tools Inventory

To make life easier, the **AI Stewardship Circle** maintains the **CXC Dynamic AI Inventory**.

- **What it is:** A living list of all AI tools and *workflows* the team has tested, with their assessments, known risks, and usage rules.
- **How to use it:** You can check this list for recommendations and to see what workflows are already cleared.

### 4.3 When Things Go Wrong (Our Learning Process)

Mistakes and unexpected issues are opportunities to learn. If you hit a snag or an AI tool produces a biased or weird result, we recommend establishing a dedicated, simple channel (we use a chat thread) for your team to post issues. The AI Stewardship Circle (or your equivalent) should review this channel to see if the policy needs adapting.

## 5. Yellow Light Check-in (A Guide for the Stewardship Circle)

When a new workflow is flagged as  **Yellow Light**, the AI Stewardship Circle facilitates a simple, **collaborative check-in**. This is not a formal audit; it's a guided conversation.

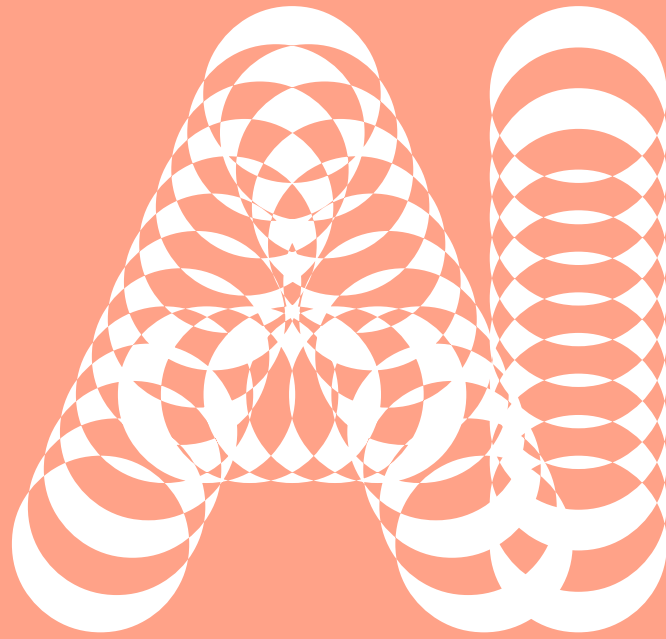
- i. The Initial Conversation (with the team member) This is a quick, informal chat to understand the request:
  - What is the tool and context? (What is it? Why do we want to use it? Who is the “owner” or champion for this workflow?)
  - What are the opportunities? (How does this align with our values? How does it help our team or community?)
  - What are the immediate risks? (What could go wrong? What’s your gut feeling?)
  - What is our decision? (Is this a clear “Go” or “No-Go”? Or does the Circle need to do a deeper review?)
- ii. **The Circle’s Deeper Check-in** If a deeper review is needed, the Circle is responsible for it. The goal is to be the guardian of our Manifesto principles for workflows that are too complex for a simple Red/Green check.

Here is the Circle’s checklist, organised by our Manifesto principles:

- Principle 2: Awareness and Responsibility
  - **The Check:** Do the provider’s ethics or footprint violate our values?
  - **The Action:** We’ll do a quick check on the provider.
    - **Ethics:** Are there *obvious* values conflicts (e.g., human rights, military contracts)? We’ll consult resources like the **AI Now Institute** or **Mozilla’s *Privacy Not Included***.
    - **Footprint:** What’s the *energy cost* (e.g., **Hugging Face’s** “Carbon Footprint” tag)? This helps us uphold our “proportionality” principle.
- Principle 3: Human at the Core
  - **The Check:** If the “Yellow Flag” was for a *high-stakes* task (like selections or grant writing), how do we keep this process “human-at-the-core”?
  - **The Action:** We will discuss the *real* impact on the person this is for (the applicant, the donor). The goal is to ensure the workflow has **clear human oversight** and doesn’t feel cold or impersonal to the recipient.
- Principle 4: Data Privacy & Trust
  - **The Check:** If the “Yellow Flag” was for *confidential data*, how do we *guarantee* this data is not public and not used for training?
  - **The Action:** This is where we must go beyond a simple policy scan. The best solution is to use a paid API or enterprise version of the tool where we can sign a **Data Processing Agreement (DPA)**. This is the legal contract that *prohibits* the provider from training on our data (it’s exactly what we did for *an* OpenAI selection assistant). If a DPA isn’t possible, we must find a tool that *explicitly* guarantees (in its business terms) that our data is private.

- Principle 5: Collective Stewardship
  - **The Check:** Does this new workflow need a new “Rule of the Road” for *everyone*?
  - **The Action:** We will ask: “If we approve this, what new rule must we add to Section 3?” For example: if we approve AI for selection summaries, we will add a *new* “Rule of the Road” that says, “AI use in selection processes must be disclosed to applicants”.

Based on this review, the Circle will add the tool and the workflow to the AI Tools Inventory (as defined in Sec 4.2). They will document the ‘Known Risks’ and ‘Usage Rules’ in the inventory’s columns. This entry is our simple, living risk register for this workflow, so the team doesn’t have to ask again.



[changemakerxchange.org/mindful-ai](https://changemakerxchange.org/mindful-ai)