Humanitarian Leadership Academy

DATA FRIENDLY SPACE

Ethical AI use and decision-making in humanitarian work

Every humanitarian using artificial intelligence (AI) faces ethical decisions, often without formal guidance. This practical guide introduces key ethical principles to consider when using AI tools.

This guide has been co-created with humanitarian and AI practitioners as follow-up to the 2025-26 joint research initiative between the Humanitarian Leadership Academy and Data Friendly Space: *Artificial intelligence in the humanitarian sector: mapping current practice and future potential.*

Ethical use of AI is not optional

In 2026, 75% of humanitarian practitioners are using AI tools regularly for work. Yet, fewer than one quarter of organisations have a formal AI policy.

AI tools are being used across humanitarian work, including to support decisions about who receives assistance, how data is handled, how programmes are designed, and how impact is measured and reported.

As humanitarian organisational approaches are emerging, practitioners have to use judgement on AI use, often under pressure and in high-stakes contexts.

Ethical AI use is not a technical concern: it is a humanitarian one.

Without ethical guardrails, AI use can, for example:

- Expose sensitive personal data
- Reinforce existing inequalities and exclude marginalised groups
- Influence decisions that are difficult to challenge or reverse

In humanitarian contexts, the consequences of these failures can fall on people who are already in crisis.

Why ethical AI use matters in humanitarian contexts

Humanitarians serve crisis-affected communities. AI use must not add to their vulnerability.

A poorly handled AI decision is not just an error. It can mean, for example, a family is missed in a distribution, an individual seeking safety is incorrectly flagged at a border, or sensitive data reaches actors who could use it to cause harm.

Ethical AI use matters because:

- AI decisions can have serious consequences in humanitarian contexts
- Affected populations often have no way to challenge or correct AI outputs that concern them
- Once data is exposed or misused, the harm cannot always be reversed
- Trust is foundational to humanitarian access, and trust is hard to rebuild

Ethical AI use is how we protect both the people we serve and the integrity of humanitarian action itself. Ethical use must be built into individual and organisational approaches to AI.

The intention of checks and balances is not to slow down processes but to act as essential safeguards, especially when the stakes and the cost of error are high.

What can ethical AI use look like in practice?

Ethical AI involves the practice of using AI tools in ways that protect the rights, safety, and dignity of all people.

Many ethical principles are relevant to humanitarian AI use, drawn from frameworks across the sector, including:

- **Data minimisation and purpose limitation:** Collect and retain only the data needed for a defined humanitarian purpose.
- **Human-in-the-loop oversight and accountability:** Humans remain responsible for decisions supported by AI.
- **Informed consent:** People understand what data is collected, how it will be used, and the risks.
- **Representational bias in data:** Data may underrepresent certain groups, leading to less fair outcomes.
- **Allocative harm:** AI can unfairly influence how resources, services, or opportunities are distributed.
- **Do no harm (in relation to AI):** AI should not create, worsen, or shift risks onto communities or staff.

Ethical use is not just about avoiding harm from AI outputs, but also about knowing when it may not be appropriate or possible to use AI at all. Build non-AI processes and pathways for both practitioners and communities served.

You are responsible for what you put into AI

Before pasting anything into an AI tool, always remember that you are accountable for the data you enter.

Commercial AI tools may store your inputs, use them to train models, or expose them in a data breach. In humanitarian settings, this may potentially lead to serious harm to the people whose information you handle.

Ask yourself before you input:

- Does this data identify individuals? E.g. names, locations, biometrics, household details.
- Could it endanger someone if exposed? E.g. witnesses, displaced people, staff.
- Is there a data processing agreement in place? If you don't know, treat the answer as no.
- Could I use anonymised or aggregate figures instead?

If you are uncertain, do not paste the data. Use AI for the parts of a task that don't require sensitive data, and handle the rest yourself.

AI can reflect and amplify bias

AI tools learn from the data they are given. When that data underrepresents a community, the tool will too.

AI systems that have been trained on external or global datasets may miss or misinterpret local contexts such as languages, cultural factors, and community dynamics.

Examples of bias and harm include:

- **Representational bias in data:** Tools trained on data that excludes certain groups will reflect those gaps.

- **Allocative harm**: Biased outputs can shape decisions about resources, leaving underrepresented communities further excluded from a response.

Examples of how bias could shape humanitarian action include:

- Satellite imagery may miss informal settlements
- Mobile network data may underrepresent older people and women due to differences in access and usage patterns
- Administrative records may exclude people outside official systems.

When a tool misses a vulnerable community, your knowledge is data. Document your reasoning and flag the gap.

Community knowledge is part of the evidence base

Community voices are not simply a validation step for AI outputs: they are part of the evidence humanitarian decisions are based on.

An algorithm cannot fully understand real-world contexts. A practitioner's job is to bring AI outputs and lived experience together.

- **Triangulate as standard**: Use multiple sources of information to check accuracy and reduce bias. Treat AI outputs and community reports as two inputs, not a hierarchy. Cross-check with other available evidence (e.g. field data, partner inputs) before deciding.
- **Verify in person where possible:** *Ground-truth* verification is the bridge between data and reality.
- **Document the reasoning:** When AI and community knowledge conflict, record how you resolved it.
- **Feed gaps back:** If the AI missed something the community caught, tell the team behind the tool.

This approach protects affected populations, supports better decisions, and creates the accountability trail that ethical practice requires.

Aggregation can turn low risk data into dangerous data

Data that appears low risk in isolation can become high risk at scale.

Cross-referenced reports on location, timing, and movement can become a map, potentially revealing the location of a vulnerable community.

Factors to be aware of include:

- **Patterns**: Aggregated data can identify a community's routine
- **Cross-referencing:** Datasets that appear low risk on their own can expose sensitive information when combined
- **Retention**: The longer dangerous data is held, the greater the risk
- **Sharing:** Even partner organisations introduce new access points

When aggregated data becomes a targeting risk, the obligation to protect people overrides the interest in retaining it. Escalate, do not act alone, and document.

Questions every humanitarian should ask

Before using AI in a decision that affects people's lives, work through questions for your context. Example questions:

- **Data minimisation and purpose limitation:** Am I using the minimum data needed, and only for its original purpose?
- **Human-in-the-loop oversight:** Will a qualified human review this output before action?
- **Accountability:** Could I explain and defend this decision later?
- **Informed consent:** Do affected individuals genuinely understand and agree?
- **Representational bias:** Could the data behind this output miss any affected communities?
- **Allocative harm:** Could acting on this unfairly affect who receives aid or services?
- **Do no harm:** Is AI the right tool for this process or decision, or are there non-AI pathways that might be more appropriate? Could this create or shift risk onto affected people or staff?

Our principles don't pause when we use AI tools

The values that guide humanitarian action - do no harm, impartiality, neutrality, independence - don't stop applying when we use AI.

Ethical use of AI is how we carry those principles into the digital tools that we use in our work.

Resources

The HLA in partnership with Data Friendly Space is leading a global study into how humanitarians are using AI. Access the research products, articles, webinars, podcasts and microlearning guides on our Resources Hub.

kaya

This guide accompanies a Kaya microlearning module: *Introduction to Ethical AI Decision-Making in Humanitarian Work.* Access the module and more free learning at kayaconnect.org. Claim certificates and HPass digital badges and share them on social media to let others know about your learning.

This microlearning guide was produced in May 2026 for informational purposes only. It reflects practice at the time of publication. Please consult your organisation's own AI and data protection policies. Cover image credit: Save the Children. Contact: info@humanitarian.academy or hello@datafriendlyspace.org