AI Ethical Guidelines for NGOs

This guideline provides practical principles and actions to support the ethical, responsible, and sustainable use of AI in NGO environments. It aims to ensure that technology serves people, protects rights, promotes equity, and aligns with the humanitarian and social values of our work.

TABLE OF CONTENT

Co-funded by the European Union

1. ETHICS IN GENERAL

AI is becoming an increasingly important tool for NGOs. It helps with data analysis, communication, improving project efficiency, and monitoring human rights compliance. However, its use raises numerous ethical questions that cannot be ignored.

Privacy and Data Protection

NGOs often work with sensitive information. Organizations have an ethical obligation to ensure data anonymization, transparent processes, and compliance with legislation (e.g., GDPR).

Algorithmic Bias

AI models learn from historical data, which may contain biases. NGOs must check that algorithms do not harm vulnerable groups and that decisions remain fair.

Transparency and Accountability

AI decisions are often a “black box.” For NGOs, it is crucial to explain the basis of any decision made by the system. A lack of transparency can undermine public and donor trust. Therefore, it is advisable to use explainable models and clearly communicate the technology’s limitations.

Risk of Manipulation and Disinformation

AI can generate content that appears credible but is false. For NGOs fighting disinformation, this is a double-edged sword. Organizations must set ethical rules for content creation and fact-checking to avoid becoming part of the problem.

2. Transparency and Communication in AI Use for NGOs

Transparency builds trust — the foundation of humanitarian and social-impact work. When NGOs use AI, they must be open and clear to ensure fairness, responsibility, and credibility. Here are seven key practices:

1

INFORM WHEN AI IS USED

- Tell people when AI supports tasks like data analysis, chatbots, or risk prediction.
- Explain what it does, why it's used, and what decisions it influences.

2

EXPLAIN DATA HANDLING

- Share how data is collected, used, stored, and protected.
- Clarify who can access it and security measures.

3

OFFER OPT-OUT CHOICES

- Allow declining AI-supported programs.
- Provide human review and data deletion options.

4 BE HONEST ABOUT AI LIMITS

- AI can make mistakes.
- It supports—not replaces—human judgment.

5 SHOW ACCOUNTABILITY

- Explain monitoring and evaluation.
- Clarify responsibility for errors or bias.

6 USE CLEAR COMMUNICATION

- Allow declining AI-supported programs.
- Provide human review and data deletion options.

7 INVITE FEEDBACK

- Encourage questions and complaints.
- Provide channels for suggestions and dialogue.

3. ALGORITHM BIAS, FAIRNESS & DISCRIMINATION

The Problem

Machine learning algorithms learn from data: if the data contains social or historical bias, the algorithm replicates it.
Biases are distortions of reality that influence mental and decision-making processes, and they can transfer to AI.

Examples:
a) A recruitment system trained on data where most leaders are men, may favor male candidates.
b) A facial recognition system trained mostly on images of light-skinned people may perform poorly on darker-skinned faces.
Consequences:
Since the training data is unbalanced, the algorithm becomes less accurate for certain groups and ends up reproducing the bias present in the data.

The Risk

Biased systems can generate systematic discrimination, reinforcing social prejudices. Biases — human, statistical, computational — can emerge at any stage: source selection, processing, data grouping, learning, or from third-party systems.

A flawed algorithm produces decisions that are not fair or impartial.

Interventions to prevent or reduce algorithmic bias can occur on 2 levels:

The Remedies

1 TECHNICAL LEVEL

– Train algorithms on diverse and representative datasets to avoid skewed learning.

– Continuously monitor model outputs to detect, measure, and correct emerging biases.

– Use fairness-aware machine learning techniques (e.g., rebalancing data, adjusting model weights, or applying bias-mitigation algorithms).

– Ensure transparency and documentation throughout the model's lifecycle.

2 LEGAL & REGULATORY LEVEL

– Apply European and national anti-discrimination laws that protect individuals from unfair treatment in automated decision-making.

– Follow the principles of the EU AI Act, which require risk assessment, human oversight, and accountability in high-risk AI systems.

– General goal: prevent single, stereotyped views and promote cultural and identity pluralism, ensuring that automated systems respect fundamental rights.

WHAT CAN YOU DO AS A YOUTH WORKER?

1. Check Incoming Data

Review the quality, representativeness, and sources of the data.

Example: health data from a single ethnic group or urban area does not correctly represent other groups.

2. Test the System on Different Groups

Perform fairness tests with sets representing different groups. Analyze whether the system favors one group.

Example: scholarship selection software that prefers men with equal performance.

3. Request Transparency From the AI Provider

Ask for explanations about how the model works. Request independent audits.

Example: check how an educational AI handles socio-demographic data and whether it integrates anti-discrimination measures.

4. Use Open-Source AI Models

Prefer transparent and modifiable tools. Collaborate with the community to correct bias.

Example: improve an algorithm used to analyze applications for educational labs by involving other NGOs and developers.

5. Educate the Team on AI Ethics

Provide ongoing training on recognizing bias. Create internal ethical guidelines.

Example: workshops to identify signs of discrimination in data.

6. Regularly Monitor Results

Continuous feedback collection. Update models when distortions emerge.

Example: a system that consistently favors youth from a specific region must be corrected and recalibrated

4. HUMAN OVERSIGHT WITH AI — 5 SIMPLE STEPS FOR NGOS

1

PEOPLE STAY IN CHARGE

Decide which staff members are responsible for checking what the AI suggests or decides. AI can assist — but people make the final call.

2

KEEP IT TRANSPARENT

Write down how and why you're using AI. Make it clear to staff, donors, and beneficiaries. No hidden systems, no hidden decisions.

3

LISTEN AND ALLOW QUESTIONS

Give communities and beneficiaries a way to ask questions or say, "This decision feels wrong." Make sure they know how to reach a real person.

4

TRAIN YOUR TEAM

Help staff understand what AI can and cannot do. Teach them to double-check results rather than trust them automatically.

5

CHECK REGULARLY AND ADJUST

Plan regular reviews to see if the AI is doing more good than harm. If something isn't working — change it or stop it.

HUMAN OVERSIGHT

- AI must support—not replace—human judgment.
- NGOs should: - Ensure final decisions are made by trained staff.
- Use AI as a tool, not an authority.
- Train staff about risks and limitations.

Human oversight means that people — not algorithms — remain accountable for decisions that affect rights, resources, or vulnerable populations.

In NGO contexts, this ensures **AI does not make final judgments** on eligibility, aid distribution, or risk assessments without human review, especially when these decisions impact safety, dignity, or access to services.

Relevant NGO cases:
WHY HUMAN OVERSIGHT MATTERS

AI suggests which family receives limited food assistance

AI looks at data like family size, income, and housing to decide who needs food support.

However:

- Some families earn cash informally, so the system may think they are poorer or richer than they are.
- A family caring for a disabled relative may have high medical costs that are not shown in the data.

➡ Why human oversight matters:
A staff member can understand personal circumstances and make fair decisions that data alone cannot capture.

AI translates beneficiary statements automatically

Translation tools help NGOs communicate across languages, especially in emergencies.

However:

- Trauma survivors might express themselves using broken grammar, which AI misinterprets.
- A mistranslated safety plan could create dangerous misunderstandings.

➡ Why human oversight matters:
Human interpreters understand tone, urgency, and emotional meaning — something machines cannot fully capture.

AI predicts which children are “at risk” of dropping out of school

NGOs use data from attendance records, grades, or household income.

However:

- A child with temporarily low attendance due to illness or caring for siblings may be wrongly labeled.
- AI may amplify pre-existing biases — poorer children receive more scrutiny and stigma.

➡ Why human oversight matters:
Labels can shape futures. A teacher or caseworker must review individual situations before actions are taken.

5: SAFETY AND DATA PROTECTION

It is paramount for NGOs to protect the privacy and security of their stakeholders' data and ensure the ethical use of AI technologies, because they often deal with sensitive information and serve vulnerable populations. To prioritize safety and data protection in AI NGOs can implement the following best practices:

Conduct a comprehensive risk assessment

Identify potential risks associated with the AI system, including data breaches, privacy violations, and algorithmic bias. This will help in understanding the potential threats and vulnerabilities that need to be addressed.

Implement strong data protection measures

Utilize encryption, access controls, and data anonymization techniques to protect sensitive information. Adopt secure data storage and transfer protocols to prevent unauthorized access to data.

Ensure compliance with data protection regulations

Stay up-to-date with data protection laws and regulations, such as the GDPR, and ensure compliance with these regulations in the development and deployment of AI systems.

Prioritize transparency and accountability

Provide clear information to users about how their data is being used and give them the option to opt-out or delete their data if they wish. Implement mechanisms for auditing and accountability to ensure that AI systems are operating ethically.

Conduct regular security audits and assessments

Regularly assess the security posture of AI systems and conduct penetration testing to identify and address any vulnerabilities. Implement a process for promptly addressing any security incidents or breaches.

Train staff on data protection and security best practices

Provide training to employees on data protection policies, security measures, and ethical considerations in using AI systems. Foster a culture of data protection and security awareness within the organization.

Engage with stakeholders and users

Build trust with stakeholders and users by engaging in transparent communication about data protection measures and security practices. Solicit feedback from users on their concerns and preferences related to data privacy and security.

This will help in upholding the highest standards of data protection and safety in their work and fulfilling their social responsibility to protect individuals' data and privacy. Furthermore, by promoting ethical use of AI technologies, NGOs can set a positive example for other organizations and contribute to shaping a more responsible and sustainable AI landscape.

6. Sustainability and the Ethical Use of AI in NGOs

For NGOs, **ethical AI use is not only about fairness and safety — it also includes environmental and social responsibility.** AI models require vast amounts of electricity, rare minerals, and human labour across global supply chains. When NGOs advocate for climate justice, workers' rights, or equitable development, it becomes ethically essential that AI adoption does not harm the same communities they aim to protect.

Ethical AI means choosing solutions that **support long-term human and environmental wellbeing,** not just efficiency or innovation.

What NGOs Can Do (Practical Ethical Actions)

1

CHOOSE ENERGY-EFFICIENT OR SMALLER AI SYSTEMS

→ Ethical AI balances need with impact, using the least resource-intensive option that still serves its purpose.: Whisper Small or Google's on-device tools, DistilBERT etc.

2

AVOID UNNECESSARY AUTOMATION OR COMPUTATION

→ Ask: Is AI solving a real problem, or are we using it because it's available? Don't add unnecessary prompts and create unnecessary pictures and videos.

3

SELECT VENDORS COMMITTED TO SUSTAINABILITY AND FAIR LABOR STANDARDS

→ Request transparency on renewable energy use, supply chains, labor practices, and recycling plans.

SUSTAINABILITY and the ETHICAL USE of AI in NGOs

Why This Is an Ethical Issue for NGOs?

Environmental impact:

High energy-consuming AI tools increase emissions, worsening climate impacts often felt most by vulnerable communities.

Resource extraction:

Hardware relies on minerals often mined in regions facing exploitation or conflict.

Invisible labor:

Large AI projects sometimes depend on low-paid workers for data labeling and content moderation.

Reputational integrity:

NGOs must align their digital practices with their public values.

Justice and equity:

The negative impacts of AI are uneven — communities with the least power often bear the greatest cost.

Conclusions

NGOs can reduce bias by following clear technical and ethical practices. This guide offers practical tools to use AI responsibly, with verified and adapted content to avoid stereotypes or inappropriate examples.

With the collaboration of:

SiR

AMPERSAND

High on Life

consciousness & environment

HAWK STARS NGO

Co-funded by the European Union