



REINVENTING HUMANITARIAN AID PROCUREMENT FOR THE AGE OF AI

accessnow.org



accessnow

Access Now defends and extends the digital rights of people and communities at risk. As a grassroots-to-global organization, we partner with local actors to bring a human rights agenda to the use, development, and governance of digital technologies, and to intervene where technologies adversely impact our human rights. By combining direct technical support, strategic advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

REINVENTING HUMANITARIAN AID PROCUREMENT FOR THE AGE OF AI



This report is an Access Now publication,

written by **Giulio Coppi**, Senior Humanitarian Officer at Access Now. This publication has been made possible thanks to generous and consistent support from the German Federal Foreign Office, which provided essential funding and accompaniment throughout the research process.

The authors would like to thank the Access Now team members who provided support, including Arzu Geybulla, Donna Wentworth, Méabh Maguire, Milica Pandzic, and Natalia Krapiva. We appreciate the help of all the interviewees, researchers, and external organisations who provided us with key insights, feedback, and information for this publication.

Design by Loren Giordano

Published in March 2026

Licensed under [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)



For more information, please visit:

<https://www.accessnow.org>

Contact: **Giulio Coppi** | Senior Humanitarian Officer, Access Now

giulio@accessnow.org

Access Now defends and extends the digital rights of people and communities at risk. As a grassroots-to-global organisation, we partner with local actors to bring a human rights agenda to the use, development, and governance of digital technologies, and to intervene where technologies adversely impact our human rights. By combining direct technical support, strategic advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

Table of contents

Executive summary

Methodology and guidance: author's note

1. Introduction

1.1. Current implementations of algorithmic tools in humanitarian programming

Case study 1: Signpost AI, a platform to de-risk humanitarian AI

1.2. Humanitarian procurement in the digital era

Case study 2: OpenAI and the nonprofit-to-defence contractor journey

Case study 3: Microsoft and heightened human rights due diligence for business in conflict-affected contexts (hHRDD+)

1.3. Why does this matter for human rights and humanitarian actors

Case study 4: The hidden geopolitics of generative AI

Case study 5: Wikimedia Foundation human rights impact assessment on AI and machine learning

1.4. Main trends in humanitarian procurement of algorithmic systems

1.5. Taxonomy of emerging practices in the integration of algorithmic systems

2. Conclusions

2.1. Money cannot buy human rights compliance

2.2. Algorithmic capture: The AI-poisoned gift for local actors and communities

2.3. Procurement is becoming heavier, not smarter nor safer

Case study 6: Erin, Tara, and Copilot enter the Irish Department of Justice

2.4. Beyond bells and whistles: from product price-based selection to continuous environmental scanning

Case study 7: Clearview AI

2.5. Bringing the zero trust environment in procurement

3. Suggested foundations for a digital service framework in the algorithmic age (see Appendix)

4. Recommendations

4.1. Recommendations for states and donors

4.2. Recommendations for the humanitarian community

4.3. Recommendations for tech companies

4.4. Recommendations for local aid actors and communities

4.5. Recommendations for research groups and cyber experts

Executive summary

In our 2024 report “Mapping humanitarian tech: exposing protection gaps in digital transformation programmes,”¹ we examined claims and consequences behind the partnerships between humanitarian actors and private corporations. We showed how these often-opaque partnerships impact the digital rights of affected communities, who often rely on a small group of humanitarian actors, trusted because of their claims of neutrality, impartiality, and independence from warring parties.

Building on some of the conclusions and recommendations made in 2024, this new report focuses on how **the rise of artificial intelligence (AI) and algorithmic tools is changing the role of humanitarian actors; from trust-holders into mediators of trust between affected communities and tech companies.** This comes with unprecedented risks for both aid actors and individuals, and requires new ways of imagining and governing the relationship between humanitarian actors and technology.

This investigation highlights how algorithmic systems are entering the humanitarian space mostly by the backdoor, often without safeguards. In a global landscape signalling the upcoming end of the AI hype cycle for large companies,² humanitarian organisations are still reporting very slow and limited institutional adoption, despite growing individual use by staff. In short, **AI adoption seems to be still driven by individual curiosity and needs, rather than by organisational policy, strategy, or measurable added value.**

Our research reveals that the organisational use of AI and algorithmic systems often happens through a process of unstructured integration rather than formal adoption. More precisely, cloud-based, algorithmic-enhanced functionalities and processes are making their way into humanitarian entities’ tech stack through updates and add-ons to preexisting products, slowly but inexorably seeping into most internal systems. **A price-based procurement process designed around the static idea of digital asset purchase is inadequate to ensure adequate protection from hyperdynamic tech product development.** This is especially evident looking at the growing power imbalances between private actors and nonprofits, and among nonprofits themselves.

Digital technology is often heralded as key to democratise and equalise the playing field between humanitarian actors, by finally unlocking local actors’ access to resources traditionally reserved for international entities. But **emerging trends seem to disprove such claims.** The lack of adequate legal frameworks and safeguards, lagging investments in humanitarian funding for information and

¹ Coppi, G. (2024, February 13). *Private tech, humanitarian problems: how to ensure digital transformation does no harm*. Access Now. <https://www.accessnow.org/private-tech-humanitarian-mapping/>

² Ropek, L. (2025, September 9). *AI Use at Large Companies Is in Decline, Census Bureau Says*. Gizmodo. <https://gizmodo.com/ai-adoption-declining-at-large-companies-2000655693>

communication technology (ICT) resources and connectivity, and inadequate dedicated funding for digital programmes are thwarting these opportunities.³ More broadly, **virtuous and responsible procurement practices are proving inadequate for protecting human rights alone**, as tech companies' commitments to human rights are replaced by defence-oriented pitches and the size of their human rights teams keeps shrinking.

Preexisting digital divides are enhanced by several deep and subtle changes in traditional tech business models, resulting from the gradual cloudification of most common digital systems and the fast-paced, distributed nature of modern digital development. Even more worryingly, **we are seeing new forms of digital divide well beyond the traditional gaps in digital capacity between local and small non-governmental organisations (NGOs) and international non-governmental organisations (INGOs) or UN entities.**⁴ Cloud-based companies' takeover of generic coordination platforms creates a constantly broadening chasm between the few organisations that are digitally mature, those who are barely retrofitting their processes to fit the new reality, and the broad majority forced to access the promised land of digital humanitarianism through the predatory gates of algorithmic and cloud services' free tech licensing.

Our findings highlight how organisation-led digital development, building on open-source systems (despite some “openwashing” practices)⁵ and trusted external providers, remains the best option for risk reduction when it comes to algorithmic systems. However, in the short to medium term, this might not be financially and technically feasible for the majority of humanitarian actors, especially considering the catastrophic funding environment.

To fill this gap, **we propose a toolkit in an appendix to this report. This is a foundational governance model for transforming digital procurement from a transactional into a strategic business function,**⁶ integrated across key organisational departments. By doing so, protection-mandated aid groups could move from running purchase-centred processes, to adopting a dynamic, proactive, and rights-based approach to tech management. **This report also contains a set of recommendations for states and donors, the humanitarian community, tech companies, local aid actors and communities, research centres, and cyber experts.**

³ CARE International & Accenture. (2024). *AI and the Global South: Exploring the Role of Civil Society in AI Decision-Making*.

https://careinternationaluk.ams3.cdn.digitaloceanspaces.com/media/documents/Althe_Global_South_Exploring_the_Role_of_Civil_Society_in_AI_Decision-Making.pdf

⁴ Rebollo, C. (2025, September 25). *Gaza Aislada: El Asedio a la Tecnología Humanitaria Agudiza El Genocidio*.

Wired. <https://es.wired.com/articulos/gaza-aislada-el-asedio-a-la-tecnologia-humanitaria-agudiza-el-genocidio>

⁵ Liesenfeld, A., & Dingemans, M. (2024). Rethinking open source generative AI: Open washing and the EU AI act. *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 1774–1787.

<https://doi.org/10.1145/3630106.3659005>

⁶ Thakur-Weigold, B., Parsa, I., Balcik, B., & Van Wassenhove, L. N. (2025). Purchasing and supply management in humanitarian settings. *Journal of Purchasing and Supply Management*, 31(3), 100975.

<https://doi.org/10.1016/j.pursup.2024.100975>

Methodology and guidance: author's note

This report builds on conclusions and recommendations made in our 2024 report,⁷ notably in the section on cloud systems, advanced analytics, and AI. This work is the result of several months of desk research, two workshops, several private feedback sessions, and 70 interviews with actors from across the humanitarian, public, academic, private, and social impact sector. Throughout this report, any reference to opinions or commentary collected during the research process without an external reference should be understood as having been collected during these interviews.

In the absence of a universally agreed definition of AI⁸ and considering the rapid convergence of most digital systems and services towards a cloud-first model involving some type of algorithmic process, this document privileges the terms “algorithmic tools” or “algorithmic systems”. We consider this to be software that automatically generates outputs, makes predictions, makes decisions, and/or gives advice by using data analysis, statistics, and/or self-learning logic.

Whenever the term AI is used, it should be understood within this broader meaning, unless presented in quotation marks, as an external citation. **Our report findings include the suggestion to break away from the unhelpful obsession with the AI label.** In line with this recommendation, we have aimed to use the term AI as little as possible. Any reference to digital tools and systems should be assumed to incorporate or make use of one or more algorithmic tools of some kind defined by their users or developers as AI-like.

Our report shows that aid organisations should start treating any external digital system — especially if proprietary and non-auditable — as possibly involving algorithmic tools. It also suggests how they can adjust their own IT, cybersecurity, procurement, legal, and digital protection governance accordingly.

⁷ Coppi (n. 1).

⁸ While a definition of AI as a general field is still lacking, various definitions have been developed for the concept of “AI system”. For example, for the OECD an “AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment” (Perset, K. (2024). *Explanatory memorandum on the updated OECD definition of an AI system* | OECD. OECD.

https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html). For a non exhaustive collection of technical definitions and standards see *Supplement ITU-T Y suppl. 72 (11/2022) - ITU-T Y.3000-series – artificial intelligence standardization roadmap*. International Telecommunication Union. (2022, November 24). <https://www.itu.int/epublications/es/publication/itu-t-y-suppl-72-2022-11-itu-t-y-3000-series-artificial-intelligence-standardization-roadmap/en>

Any assessment of the appropriateness, usefulness, legality, or ethical nature⁹ of certain algorithmic systems is outside of this report's scope, and should be made on a case-by-case basis based on existing guidance and research. This includes, for example, guidance on the usefulness of existing AI solutions for nonprofits,¹⁰ and extensive mapping of ethics- or rights-based approaches to AI principles.¹¹ Under no circumstances should any analysis of a specific algorithmic system for a specific use presented in this report be taken as an endorsement of its adoption, or a suggestion for broader deployment.

Finally, we stress that this report's recommendations and the integrated digital service framework model proposed in the appendix are not intended as a silver bullet or checklist for solving all of the humanitarian sector's challenges related to AI adoption. We hope this document is seen as a foundational step for influencing how digital solutions are procured by humanitarians and other protection-mandated actors, to help them set their digital transformation engagement on a better path. We also hope to provoke a surge of solidarity across the sector, **breaking the siloed approach that is generating new specific forms of digital divide¹² between few actors in the Global Minority with privileged, on-call access to Big Tech,¹³ and the majority of frontline organisations and communities struggling to protect their principled approach, and their own rights, in the digital age.**

⁹ Robinson, T. (2025, January 21). *Ten reasons not to use AI for development and ten routes to more responsible use*. Institute of Development Studies.

<https://www.ids.ac.uk/opinions/ten-reasons-not-to-use-ai-for-development-and-ten-routes-to-more-responsible-use/>

¹⁰ *Guide to AI Solutions for Nonprofits*. NetHope. (2025, November 5).

<https://nethope.org/toolkits/the-guide-to-usefulness-of-existing-ai-solutions-in-nonprofit-organisations/>

¹¹ Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. *SSRN Electronic Journal*.

<https://doi.org/10.2139/ssrn.3518482>

¹² CARE International & Accenture (n. 3). See also Rebollo (n. 4).

¹³ Rotter, J., & Bailkoski, W. (2025, October 17). *AI Adoption in NGOs: A Systematic Literature Review*. arXiv.

<https://arxiv.org/html/2510.15509v1>

1. Introduction

1.1. Current implementations of algorithmic tools in humanitarian programming

Analysis of available documentation and direct feedback from interviewees shows that the humanitarian sector is prioritising narrow or function-specific systems involving single-task applications of artificial intelligence (AI).¹⁴ These seem to promise better performance gains, compared with purely human interventions. The sector is showing lukewarm interest towards general applications requiring deeper integration among internal systems, databases, and units, which are in some cases sandbox-tested but not entirely trusted.

On the other hand, there seems to be broad agreement on the irrelevance of the debate on so-called artificial general intelligence (AGI). Humanitarian experts consider that this hypothetical technology – involving systems that exhibit intelligent behaviour across a range of cognitive tasks¹⁵ – is unachievable in a foreseeable timeline, if at all. Most importantly, it would be hardly compatible with humanitarian principles and would thus be unusable.

NGOs' current implementation of algorithmic systems has clustered across four operational domains: external communication; internal operations; monitoring, evaluation and insight generation; and mission delivery.¹⁶ The main types of algorithmic applications tested and sometimes deployed by humanitarian organisations may be tentatively grouped into three main categories:

- **Standalone AI systems**, i.e. generative or other AI systems where the algorithmic functionality is the main process expected by the user;
- **Enhancing features** integrated within data and information management systems, where the main process expected by the user is also supported by an algorithmic system, even if the user is unaware; and
- **Critical backend systems** where algorithmic functionalities connected to a given cloud environment are separated from the user experience by one or more layers of the tech stack.

Beyond these broad trends lies a rich and diverse set of technical and governance approaches in the way these solutions are selected, customised, and managed. This speaks to **the growing maturity of a sector whose tech landscape is increasingly animated by smaller, more agile actors adopting non-traditional operational, governance, and sustainability approaches, often inspired by a**

¹⁴ *Human rights in the age of artificial intelligence*. Access Now. (2018, November).

<https://www.accessnow.org/wp-content/uploads/2018/11/AI-and-Human-Rights.pdf>

¹⁵ Ibid.

¹⁶ Ahmed, M., Khan, A., Wang, H., Kim, S. S. Y., Muchhala, M., Merkle, A., Papakyriakopoulos, O., & Ali, D. (2025). *AI adoption across mission-driven organisations*. TUM think tank.

<https://tumthinktank.de/en/output/ai-adoption-across-mission-driven-organisations/>

data-first strategic model. While digitally native humanitarian organisations appear to be more at ease and less concerned about the implications and challenges of integrating algorithmic systems, they too face traditional hurdles as soon as discussions on digital transformation touch on core governance positions or require leadership to make strategic decisions. **Internal reflections in particular tend to stall when proposed interventions require multi-year planning or considerable upfront investment.**

Despite the many challenges, **we recorded an unexpected upside to the ongoing AI hype. Forced out of their inertia, humanitarian entities have been pushed to explore alternatives to the cloud-first and Big Tech-dependent algorithmic tools, notably in the form of small language models or dedicated hosting through trusted providers.** For example, the United Nations International Computing Centre (UNICC) AI Hub¹⁷ recently started offering UN agencies various hosting options and services, including an AI sandbox for safe testing.

Recent years have also seen explosive growth in actors and ecosystems building algorithmic tech in Global Majority regions,¹⁸ covering the whole ecosystem including accelerators and investors, startups, universities, and the social sector. Unfortunately, this emerging process is already threatened by the downturn in funding, which might push organisations into even deeper tech dependency. **The amplifying effects of algorithmic systems are widening the digital divide between those organisations that can afford to plot a journey towards digital independence, and those barred from even considering such options.**

Multiple surveys indicate that, while many humanitarian organisations still face challenges in formally adopting algorithmic tools and the mass adoption claimed in the prevalent public discourse may still be lagging, humanitarian professionals are indeed increasingly experimenting with AI tools on an individual basis.¹⁹ In a Data Friendly Space (DFS) survey of 2,500

¹⁷ UNICC AI hub. UNICC. (n.d.). <https://www.unicc.org/artificial-intelligence/>

¹⁸ *Global South map of emerging areas in artificial intelligence.* Knowledge 4 All Foundation Ltd. (n.d.). <https://k4all.org/project/aiecosystem/>

¹⁹ Düchting, A., & Pellowska, D. (2025, August 31). *Balancing innovation, efficiency, and principled humanitarian action.* Centre for Humanitarian Action. <https://www.chaberlin.org/en/publications/balancing-innovationefficiency-and-principled-humanitarian-action-2/> See also Parkinson, K. M., Johnson, M., & Hall, L. (2025, August 7). *How are humanitarians using artificial intelligence in 2025? Mapping current practice and future potential: Initial insights report.* <https://reliefweb.int/report/world/how-are-humanitarians-using-artificial-intelligence-2025-mapping-current-practice-and-future-potential-initial-insights-report-august-2025>; and Johnson, M. (2025, August 5). *Global study highlights “Shadow AI” as humanitarian workers outpace organisations in technology adoption.* Data Friendly Space. <https://www.humanitarianleadershipacademy.org/news/global-study-highlights-shadow-ai-as-humanitarian-workers-outpace-organizations-in-technology-adoption/>

aid workers, 93% of individual respondents said they had tried AI tools, yet only 8% of organisations have integrated this technology.²⁰

A majority of the organisations assessed for this report seem to be uncertain about how to regulate individuals' use of algorithmic tools, whether for professional purposes or for personal purposes but using devices connected to professional accounts. When humanitarian workers use digital tools without clear ethical frameworks or guidance, they risk inadvertently introducing new harms for the populations they are meant to protect. **This risk is compounded by the fact that when users engage with AI systems on a personal basis, they often tend to rely on free or low-cost versions which offer worse performances than their expensive, higher-tier counterparts.**²¹

Aid groups' responses involve various nuances and combinations of blacklisting, whitelisting, policy, and guidance. When appropriate and possible, organisations limited or blocked access to or download of certain tools, and took steps to pre-install whitelisted applications on professional devices. The most common intervention, however, is introducing general AI policies and guidance aimed at instructing staff on acceptable usage and prohibited actions, accompanied by internal awareness-raising campaigns and training.

Humanitarians face the specific challenge of harmonising efficiency, low resources, and humanitarian principles,²² while striving to keep pace with rapidly evolving technologies and their trade-offs,²³ despite limited practical experience with their use, let alone with scaling. It is hardly surprising that in a Centre for Humanitarian Action (CHA) survey, 81% of respondents expressed ethical concerns about the use of AI in humanitarian action due to “highlighted ethical risks, data protection response to rapidly evolving technology, expectations, accountability frameworks, and regulatory pressures.”²⁴

The considerations presented thus far might help explain why the majority of systems mapped during our research remain in the experimental or pilot phase, and why their access is often subjected to a whitelisting process and restricted to selected users. Respondents to the CHA survey cited²⁵ data protection and privacy concerns, lack of trust, and uncertainty about the reliability of AI tools as the main barriers to organisational adoption. In contrast, respondents to another survey led by Sphere identified a lack of technical expertise as the top barrier, followed by concerns over data protection,

²⁰ Data Friendly Space (n. 19)..

²¹ Dyuzhov, A. (2025, September 19). *ChatGPT free vs. paid features: Is premium worth it?*. Overchat AI. <https://overchat.ai/ai-hub/chatgpt-free-vs-paid-features>

²² Devidal, P. (2024). Lost in digital translation? The humanitarian principles in the digital age. *International Review of the Red Cross*, 106(925), 120–154. <https://doi.org/10.1017/s1816383124000080>

²³ DÜchting & Pellowska (n. 19).

<https://www.chaberlin.org/en/publications/balancing-innovationefficiency-and-principled-humanitarian-action-2/>

²⁴ Ibid.

²⁵ Ibid.

privacy, and tool reliability.²⁶ Empirical evidence from the earliest adopters of algorithmic systems also confirms the persistent struggle²⁷ in finding ways to engage in meaningful conversations with local communities²⁸ about the tools used by aid providers, often worsened by the lack of visibility or understanding of the upstream supply chain.

In our interviews, **senior humanitarian managers instead framed the challenges in further integrating algorithmic tools as being mostly related to financial and strategic reasons. In fact, significant upfront investments are required to quickly and safely scale up new tools requiring deep integration with other internal systems while offering unclear added value, building on mostly unsubstantiated claims of future efficiency gains.** These considerations, in combination with digital literacy gaps in decision-making spaces, are reportedly chilling the enthusiasm of senior managers who find it difficult to justify investing in demanding digital transformation projects to their leadership.

In terms of areas of adoption, advanced analytics and algorithmic tools present a more mature stage of deployment in humanitarian fields that are historically data-friendly, such as in disaster and weather forecasting to support preparedness measures, early warning systems and early action. Preexisting data readiness and these teams' stronger digital capacities compared with the humanitarian average makes it easier for them to identify potential efficiencies, and to seamlessly identify and integrate the most appropriate algorithmic tools for the task at hand. Experts also confided that the relatively recent addition of these areas of work into the humanitarian digital services portfolio allowed them to escape from the growing grip of Big Tech actors, offering them more flexibility in vetting and selecting their stack. That said, interviewees also stressed how such maturity has yet to be effectively integrated into decision-making processes.

Another area that has been experimenting with algorithmic tools since well before the AI hype is the field of humanitarian information management. Pioneered by DFS and a coalition of partners during the 2015 Nepal earthquake crisis, the data entry and exploration platform (DEEP) has left an invaluable legacy, even if it was discontinued²⁹ in 2025 due to funding constraints. This unusually long life has given the information management field a considerable advantage in learning how to navigate the complexities of harnessing machine learning and algorithmic tools. Today, DEEP's legacy is carried on

²⁶ Sphere. (2024, December 19). *How can humanitarian organisations use AI safely?*
<https://spherestandards.org/how-can-humanitarian-organisations-use-ai-safely/>

²⁷ Desai, N., & Lummis, V. (2025, September 4). *How to explain AI/ML across the digital divide.* GiveDirectly.
https://www.givedirectly.org/explain_ai/

²⁸ ECNL & SocietyInside. (2023, September 29). *Framework for meaningful engagement in human rights impact assessments of AI.* OECD.AI.
<https://oecd.ai/en/catalogue/tools/framework-for-meaningful-engagement-in-human-rights-impact-assessments-of-ai>

²⁹ Johnson, M. (2025, May 28). *A farewell to DEEP.* Data Friendly Space.
<https://www.datafriendlyspace.org/post/farewell-to-deep>

by several text analysis tools.³⁰ These include GANNET,³¹ which aims to help organisations sift through vast amounts of crisis information to extract actionable insights, the Lebanese Red Cross' ICT offline AI assistant for operational security,³² or the Norwegian Refugee Council internal knowledge management and retrieval system,³³ among others.

CASE STUDY 1: SIGNPOST AI, A PLATFORM TO DE-RISK HUMANITARIAN AI

Signpost AI,³⁴ a project led by the International Rescue Committee (IRC), starts from the assumption that LLMs are not fit for use without safeguards. The project aims to deliver information and knowledge for people facing crises in multiple countries through an open-source and open-access orchestration system.³⁵ Their suite of products includes a mapping hub, and automated multilanguage chatbots (only for low-risk questions) for customer relationship management (CRM) integration. This project joins a growing movement of actors trying to build digital humanitarian infrastructure for the whole sector by allowing all actors, including local or less resourced ones, access to the most relevant foundational models. This includes an agents ecosystem to create AI agents through no-code processes. Like most AI chatbots, the Signpost AI information assistant relies on foundational models designed, trained, built, and maintained by private tech companies, and faces the common challenge of being only as ethical, good, and effective as these models are. To overcome this problem, the Signpost team is testing protocols³⁶ based on technical and human enhancement features, including constant training through direct feedback by human evaluators. The team also runs retrieval-augmented generation (RAG) processes³⁷ to make sure the chosen LLM only refers to authoritative knowledge bases pre-vetted and whitelisted by their staff, and ensures human-in-the-loop moderation, evaluation, training, and review of the AI agents in their system. Quite unique in the current landscape of humanitarian tech solutions is the team's approach to transparency, offering access to a broad range of resources including a research

³⁰ Rocca, R., Tamagnone, N., Fekih, S., Contla, X., & Rekabsaz, N. (2023). Natural language processing for humanitarian action: Opportunities, challenges, and the path toward humanitarian NLP. *Frontiers in Big Data*, 6. <https://doi.org/10.3389/fdata.2023.1082787>

³¹ *AI-powered tools for Social Impact*. GANNET. (n.d.). <https://gannet.ai/>

³² Toro, M. (2025, November 4). *The human in the loop: How oversight turns AI into a humanitarian ally*. ReliefWeb. <https://reliefweb.int/report/world/human-loop-how-oversight-turns-ai-humanitarian-ally>

³³ Mandefro, M. T. (2025, June 19). *AI-powered knowledge chatbot (Norwegian Refugee Council)*. UKHIH. <https://www.ukhih.org/news/ai-powered-knowledge-chatbot-norwegian-refugee-council/>

³⁴ *Signpost AI*. (n.d.). <https://www.signpostai.org/>

³⁵ Nicoll, L. (2025, October 15). *Behind the scenes: Orchestrating AI for humanitarian response*. Signpost AI. <https://www.signpostai.org/insights/behind-the-scenes-orchestrating-ai-for-humanitarian-response>

³⁶ Ali, A. (2025, July 15). *Signpost AI Information assistant: A pilot case study*. Signpost AI. <https://www.signpostai.org/research/signpost-ai-a-case-study>

³⁷ *What is Rag? - Retrieval-augmented generation AI explained*. AWS. (n.d.). <https://aws.amazon.com/what-is/retrieval-augmented-generation/>

roadmap,³⁸ its repository,³⁹ testing reports,⁴⁰ and accompanying policies and frameworks.⁴¹ The Signpost AI repository is mostly available on Github⁴² under GPL-3.0,⁴³ a licence ensuring that the original code is open for contribution and that derivative versions must also remain open source.

Several organisations are also testing ways to connect their analytical work with programming activities. The World Food Programme (WFP) is planning to use algorithmic systems to combat global food insecurity,⁴⁴ while the Red Cross Red Crescent network and the UN Office for Coordination of Humanitarian Affairs (OCHA) are testing predictive models to assess the impact of disasters before they happen.⁴⁵ GiveDirectly is piloting a flood forecast system triggering early cash payments to help communities to prepare and increase their resilience.⁴⁶ And the International Committee of the Red Cross (ICRC) has developed an algorithmic tool for improving efficacy of medical supply delivery, is exploring a tool forecasting the delivery of medical supplies to health facilities in war zones,⁴⁷ and multimodal open-weight AI models to support medical diagnosis in low-resource contexts.⁴⁸

The full list of ongoing or closed activities harnessing algorithmic tools is long and remains outside the scope of this research. Attempts at classifying the current landscape of AI projects have been made by others,⁴⁹ and a more advanced repository of AI-enabled humanitarian projects — most of them at pilot

³⁸ Ali, A. (2024, August 16). *Charting the course: Signpost AI Research Roadmap*. Signpost AI.

<https://www.signpostai.org/research/signpost-ai-agent-technology-a-research-roadmap>

³⁹ Welcome to Signpost AI. (n.d.-b). <https://doc.signpostai.org/>

⁴⁰ Ali, A. (2025, June 27). *Pilot report: Signpost AI Information assistant*. Signpost AI.

<https://www.signpostai.org/research/pilot-report-signpost-ai-information-assistant>

⁴¹ Nicoll, L. (2024, October 4). *Introducing SignpostAI: An AI lab for humanitarian aid*. Signpost AI.

<https://www.signpostai.org/insights/blog-post-title-two-6a63d>

⁴² IRC. (n.d.). *THEIRC/signpost-AI: SignpostAI is a repository to power the digital infrastructure of humanitarian aid*.

GitHub. <https://github.com/theirc/signpost-ai>

⁴³ Licence Publique Générale GNU, v3.0 - Projet GNU. Free Software Foundation. (n.d.).

<https://www.gnu.org/licenses/gpl-3.0.fr.html>

⁴⁴ *Artificial Intelligence*. WFP Innovation. (n.d.). <https://innovation.wfp.org/artificial-intelligence>

⁴⁵ Breker, S. (2024, November 20). *Embracing data-driven early action protocols*. 510.

<https://510.global/2024/11/embracing-data-driven-early-action-protocols/>

⁴⁶ Mandefro, M. T. (2025, July 23). *AI supported triggers for cash transfers, GiveDirectly*. NetHope.

<https://nethope.org/case-studies/ai-supported-triggers-for-cash-transfers-givedirectly-2/>

⁴⁷ Baranyk, F. (2024, July 11). *Projects to develop tech for humanitarian goals*. ETH for Development.

<https://eth4d.ethz.ch/news-and-events/eth4d-news/2024/07/projects-to-develop-tech-for-humanitarian-goals.html>

⁴⁸ *Meditron: An LLM suite for low-resource medical settings leveraging Meta Llama*. Meta. (2024, April 25).

<https://ai.meta.com/blog/llama-2-3-meditron-yale-medicine-epfl-open-source-llm/>

⁴⁹ Salaj, R. (2024). *Artificial Intelligence in Development and Humanitarian Work*. ImpactSkills.

<https://impactskills.it/wp-content/uploads/2025/01/Artificial-Intelligence-in-Development-and-Humanitarian-Work.pdf>

stage — is available on the United Kingdom Humanitarian Innovation Hub (UKHIH) website,⁵⁰ while the Centre for Humanitarian Data houses the largest catalogue of predictive analytics models in the humanitarian sector⁵¹ inclusive of a dedicated peer review framework.⁵²

On the programmatic side, our research confirms findings from the above-mentioned surveys pointing at organisations integrating algorithmic systems in their stack as tools for natural language processing, as well as project cycle management tasks like proposal writing, monitoring and evaluation, data analysis, or knowledge management. The resulting applications include chatbots for information sharing, biometric verification systems for fraud detection, or generative tools used for proposal writing, monitoring, and evaluation purposes. This trend points at a consistently marginal role for algorithmic systems in the overall humanitarian tech ecosystem, mostly to enhance preexisting functions or processes. So far, while some experts consider this technology to be still in its early days, the vast majority expressed scepticism with regard to its real transformative potential for meaningful organisational change.

Algorithmic systems are also experiencing slow rates of development and deployment due to the technical challenges involved, and the need for enhanced precautionary measures required for implementation. A notable exception to this trend is biometric technology, whose adoption rate in the humanitarian field is currently plateauing after spreading early and quickly over the past decade, and sparking significant controversy. As noted in our 2024 report,⁵³ the implementation of biometric systems in humanitarian response has already gone through an intense growth cycle that further entrenched lingering public mistrust of aid organisations for their alleged inability to prevent waste and the diversion of aid. Support for this kind of solution has now become less common outside of the health and vaccine sectors. The most common exceptions are seen via a handful of global humanitarian entities, and those specific local actors operating where the use of biometrics is required by donors, accountability, or security policies.

During our interviews, experts flagged the spread of chatbots and similar assistive tools as another exception to the slow integration of algorithmic tools. However, as we previously documented, most past experiences in pushing chatbots in frontline services have been sporadic and short-lived, and

⁵⁰ *Directory of AI Solutions in the Humanitarian Sector*. UKHIH. (n.d.).

<https://www.ukhih.org/humanitarian-directory-of-ai-enabled-projects/>

⁵¹ *Catalogue of predictive models in the humanitarian sector*. The Centre for Humanitarian Data. (n.d.).

<https://centre.humdata.org/catalogue-for-predictive-models-in-the-humanitarian-sector/>

⁵² *Peer Review Framework for Predictive Analytics in Humanitarian Response*. The Centre for Humanitarian Data. (2022, August 11).

<https://centre.humdata.org/peer-review-framework-for-predictive-analytics-in-humanitarian-response/>

⁵³ Coppi (n. 1).

none used AI functionalities.⁵⁴ For example, ‘Hakeem’⁵⁵ and ‘Karim’⁵⁶ are two chatbots developed by NGOs with support by Microsoft and Silicon Valley start-up X2AI respectively. Both aimed to support out-of-school refugees or displaced youth in Lebanon, yet neither moved beyond the prototype stage. More importantly, both used rules-based algorithms instead of AI-enhanced approaches.⁵⁷

The journey of ‘Karim’ offers a critical cautionary tale to humanitarian entities. After the startup X2AI rebranded as Cass⁵⁸ and renamed ‘Karim’ to ‘Tessa’, the chatbot was adopted by the US National Eating Disorder Association (NEDA), allegedly to replace helpline staff.⁵⁹ Without notifying or warning NEDA, the Cass team transformed ‘Tessa’ from a strictly rule-based chatbot⁶⁰ with pre-programmed responses, into a generative AI (GenAI) version with “enhanced questions and answers feature”⁶¹ as part of a system upgrade. Suddenly, the chatbot started giving users potentially harmful answers that were outside of what was programmed by researchers in the previous rule-based system.⁶² The chatbot was ultimately taken down,⁶³ but there is no clarity on where responsibility for what happened lies: Cass says the system upgrade was within the boundaries of NEDA’s contract, but NEDA’s leadership contend that they were “never advised of these changes and did not and would not have approved them.”⁶⁴

The surge in available funding and public attention for GenAI is lowering barriers to accessing these systems, facilitating their embedding in preexisting platforms and tools, and their

⁵⁴ Ibid.

⁵⁵ Toplic, L. (2020, January 22). *AI and learning: Insights from Global EdTech event*. NetHope. <https://nethope.org/articles/ai-and-the-future-of-learning-what-we-talked-about-at-the-worlds-largest-edtech-event/>

⁵⁶ Madianou, M. (2021). Nonhuman humanitarianism: When “AI for good” can be harmful. *Information, Communication & Society*, 24(6), 850–868. <https://doi.org/10.1080/1369118x.2021.1909100>

⁵⁷ Caballé, M. (2023, July 25). *Rule-based chatbots vs. AI chatbots: Key differences: Hubtype blog*. Hubtype. <https://www.hubtype.com/blog/rule-based-chatbots-vs-ai-chatbots>

⁵⁸ X2AI announces rebrand to Cass. Cass. (2022, August 29). <https://www.cass.ai/blog/x2ai-announces-rebrand-to-cass>

⁵⁹ Morris, C. (2023, May 26). *AI chatbot will replace human helpline workers at National Eating Disorder Association*. Fortune. <https://fortune.com/well/2023/05/26/national-eating-disorder-association-ai-chatbot-tessa/>

⁶⁰ Shah, J., DePietro, B., D’Adamo, L., Firebaugh, M., Laing, O., Fowler, L. A., Smolar, L., Sadeh-Sharvit, S., Taylor, C. B., Wilfley, D. E., & Fitzsimmons-Craft, E. E. (2022). Development and usability testing of a chatbot to promote mental health services use among individuals with eating disorders following screening. *International Journal of Eating Disorders*, 55(9), 1229–1244. <https://doi.org/10.1002/eat.23798>

⁶¹ Tsagas, M. (2023, October 30). *Chatbot Tessa serves as a cautionary tale against replacing frontline workers with AI*. SmartCompany. <https://www.smartcompany.com.au/health/chatbot-tessa-cautionary-tale-replacing-frontline-workers-ai/>

⁶² Ibid.

⁶³ Morris, C. (2023, May 31). *National Eating Disorder Association shuts down A.I. chatbot it planned to use to replace humans saying it ‘may have given’ harmful information*. Fortune. <https://fortune.com/well/2023/05/31/neda-ai-chatbot-harmful-advice/>

⁶⁴ Wells, K. (2023, June 9). *An eating disorders chatbot offered dieting advice, raising fears about AI in health*. NPR. <https://www.npr.org/sections/health-shots/2023/06/08/1180838096/an-eating-disorders-chatbot-offered-dieting-advice-raising-fears-about-ai-in-hea>

potential for escaping internal tracking and oversight mechanisms. As shown by case study 2, humanitarians lack any meaningful control over GenAI tools owned by third parties. This has led to “intensified calls to revisit humanitarian standards, strengthen accountability mechanisms, and integrate existing guidelines.”⁶⁵ **Once more, while most of the debate on the role of AI in aid has focused on the need for improved frameworks to use AI safely, considerations on due diligence and procurement have been mostly absent from the general discourse.**

1.2. Humanitarian procurement in the digital era

After almost two decades of stable partnership building with tech companies, the humanitarian community is now facing a series of seismic shifts calling into question advances made so far. While the sector was previously able to cautiously monitor and control the integration of emerging technologies, the progressive consolidation of IT systems within externally hosted cloud environments and the rapid introduction of increasingly complex algorithmic systems and obscure digital supply chains are testing the limits of many humanitarian actors’ internal governance mechanisms. Moreover, aid groups are facing an unprecedented culling of human rights and humanitarian teams at tech companies,⁶⁶ in conjunction with increased pandering to defence authorities even by those private actors which were the main drivers in the humanitarian tech partnerships landscape.⁶⁷

While these trends raise serious questions about the truly principled nature of existing private-humanitarian partnerships, **the collapse of the humanitarian fundraising landscape in 2025 following the shutdown of all USAID funding⁶⁸ further worsened this power dynamic. Faced with the challenge of harnessing the complexity of AI, humanitarian actors find themselves with fewer resources and allies, and increasingly dependent on their tech supporters and providers.**

Evidence collected by the UN working group on the issue of human rights and transnational corporations and other business enterprises (UN WG) has shown that there are still significant gaps when it comes to legislative frameworks on rights-respecting procurement and deployment of AI systems by states and businesses.⁶⁹ Also because of these normative gaps, many businesses are reportedly designing, procuring, and deploying algorithmic systems without conducting proper

⁶⁵ DÜchting (n. 19).

⁶⁶ De Vynck, G., & Oremus, W. (2023, March 30). *As AI booms, tech firms are laying off their ethicists*. Washington Post. <https://www.washingtonpost.com/technology/2023/03/30/tech-companies-cut-ai-ethics/>

⁶⁷ Pascual, M. G. (2025, July 21). *Big Tech enters the war business: How Silicon Valley is becoming militarized*. EL PAÍS English. <https://english.elpais.com/economy-and-business/2025-07-21/big-tech-enters-the-war-business-how-silicon-valley-is-becoming-militarized.html>

⁶⁸ Konyndik, J. (2026, February 9). *What We Lost When We Lost U.S.A.I.D.* The New York Times. <https://www.nytimes.com/2026/02/08/opinion/usa-id-humanitarian-aid-america.html/>

⁶⁹ A/HRC/59/53 - *Artificial intelligence procurement and deployment: ensuring alignment with the Guiding Principles on Business and Human Rights*. UN Human Rights Council. (2025, May 14). <https://docs.un.org/en/A/HRC/59/53>

human rights due diligence (HRDD),⁷⁰ let alone the heightened process (hHRDD+)⁷¹ required to deploy information and communication technologies in conflict⁷² or disaster-affected areas.

On the other hand, humanitarian organisations have often been found slacking⁷³ in implementing their commitment under existing humanitarian data responsibility frameworks, which are often considered as second or third-tier priorities. In our interviews, several experts who shared their experience in supporting tech partnerships or contracting reported feeling like they were asked to run box-checking exercises where each actor trusts the other to do the due diligence and conduct appropriate, tech-specific human rights or protection impact assessments.⁷⁴ In practice this arrangement lacks oversight, since nobody is ultimately responsible for verifying the existence and quality of these assessments, or checking what the emerging risks and mitigation measures might be.

The trust-based relationship is especially controversial because of the asymmetric power dynamics between private and aid entities. Humanitarian actors, including those with most influence and resources, shared their lack of ease in dealing with emerging risks from system updates or product upgrades beyond the initial scope, often because of contractual documents and non-disclosure agreements (NDAs). The persistence of these commercial practices is also hardly compatible with the aid groups' fiduciary responsibility towards affected communities and their donors.

The aid sector's dependence on unenforceable, ethics-based corporate policies is becoming even harder to justify and rationalise, in the face of some tech leaders' blatant and very public rejection of core human rights and humanitarian values.⁷⁵ This trend raised red flags even before the well-known partnership between Palantir and the WFP where public criticism on ethical and rights-based grounds was dismissed by both parties, who shielded themselves behind alleged respect

⁷⁰ *Corporate human rights due diligence – identifying and leveraging emerging practices*. Working Group on Business and Human Rights. (2018).

<https://www.ohchr.org/en/special-procedures/wg-business/corporate-human-rights-due-diligence-identifying-and-leveraging-emerging-practices>

⁷¹ Pachoud, G., & Milatović *Heightened human rights due diligence for business in conflict-affected contexts: A guide*. UNDP. (2022, June 16).

<https://www.undp.org/publications/heightened-human-rights-due-diligence-business-conflict-affected-contexts-guide>

⁷² *Conflict-sensitive human rights due diligence for ICT companies*. BSR. (2022, December 7).

<https://www.bsr.org/en/reports/conflict-sensitive-human-rights-due-diligence-for-ict-companies-guidelines-and-toolkit-for-corporate-human-rights-practitioners>

⁷³ Lidén, K. (2025). Ethics and the governance of digital data in humanitarian action. *Disasters*, 50(1).

<https://doi.org/10.1111/disa.70018>

⁷⁴ *Conflict-Sensitive Human Rights Due Diligence For ICT Companies*. BRS & JustPeace Labs. (2022).

https://www.bsr.org/reports/eHRDD_Toolkit_Presentation_Dec_2022.pdf

⁷⁵ Soave, R. (2025, September 12). *Tech CEO proposes public execution to deter crime*. The Hill.

<https://thehill.com/opinion/robby-radar/5640692-public-execution-crime-deterrence/>

for privacy safeguards.⁷⁶ **In such a complex paradigm, the selection of a specific product and the associated provider becomes crucial not only for the organisation’s long-term outcomes, but even just for its immediate survival.**

CASE STUDY 2: OPENAI AND THE NONPROFIT-TO-DEFENCE CONTRACTOR JOURNEY

Ahead of the 2024 AI for Good Summit, the International Rescue Committee (IRC) announced a new humanitarian partnership with OpenAI “to improve learning outcomes and reduce disruptions in learning for children in crisis-affected communities.”⁷⁷ Since then, ChatGPT — the GenAI chatbot trained by OpenAI to interact in a conversational way — has been growing in popularity with several national and international NGOs testing its system, mostly for internal knowledge management purposes. But in early 2026, less than two years after the announcement of the IRC partnership, OpenAI’s own staff called out the deal struck by their company to supply the US Department of War (DoW) with AI, with the chief of the OpenAI Robotics team resigning in protest.⁷⁸ This controversial deal, defined as “opportunistic and sloppy” even by OpenAI’s chief executive officer Sam Altman, allows OpenAI to take over an important contract previously held by Anthropic. This competitor fell out of favour with the DoW following the company’s refusal to allow its AI to be used for mass surveillance or for autonomous weapons.⁷⁹ While this might have surprised some less attentive observers, the company’s transformation into a defence contractor did not happen overnight and should have raised red flags well before the recent Anthropic/OpenAI spat. Originally registered as a nonprofit that “benefits all of humanity,” OpenAI shifted to a “capped-profit” model in 2019⁸⁰ and completed its transition to a for-profit company in 2025 through the registration as a public benefit corporation (OpenAI Group PBC) and an important recapitalization⁸¹ process. This change allowed the company to attract more investment, strike deals with the US Department of Defence including a USD \$200 million contract to

⁷⁶ Martin, A. (2025, July 25). *Why sovereignty matters for humanitarian data*. Big Data & Society.

<https://journals.sagepub.com/doi/10.1177/20539517251361109>

⁷⁷ *OpenAI x International Rescue Committee: Leveraging AI to Scale Ed-Tech in Crisis Affected Settings*. (2024, May 29). International Rescue Committee.

<https://www.rescue.org/press-release/openai-x-international-rescue-committee-leveraging-ai-scale-ed-tech-crisis-affected>

⁷⁸ *OpenAI robotics chief quits over AI’s potential use for war and surveillance*. (2026, March 8). France24.

<https://www.france24.com/en/americas/20260308-openai-robotics-chief-quits-over-ai-potential-use-for-war-and-surveillance-artificial-intelligence-pentagon>

⁷⁹ Field, J. (2026, March 4). *OpenAI signs US defense contract after Anthropic drops out*. CoinGeek.

<https://coingeek.com/openai-signs-us-defense-contract-after-anthropic-drops-out/>

⁸⁰ Niemeyer, K. (2024, September 15). *OpenAI’s mission to develop AI that “benefits all of humanity” is at risk as investors flood the company with cash*. Business Insider.

<https://www.businessinsider.com/sam-altman-openai-mission-drift-for-profit-nonprofit-structure-investment-2024-9>

⁸¹ Taylor, B. (2025, October 28). *Built to benefit everyone*. OpenAI.

<https://openai.com/index/built-to-benefit-everyone/>

develop “prototype frontier AI capabilities,” and partnerships with defence contractors such as Anduril,⁸² to develop AI and military robotics and drones. OpenAI is far from alone in their quest to strike defence deals: xAI,⁸³ Anthropic,⁸⁴ Scale AI⁸⁵ and more recently Google’s Gemini⁸⁶ have also signed contracts with the US Pentagon. Despite this complex and controversial trajectory, OpenAI remains active in the debate on how to embed AI solutions into humanitarian response⁸⁷ and into the nonprofit sector in general,⁸⁸ supported by both technical support and financial grants through its foundation.⁸⁹ OpenAI’s rapid transition from a nonprofit partnering with humanitarian protection actors into a defence-oriented company is an ominous sign for a sector that has not yet fully developed adequate procurement and legal mechanisms to deal with emerging misalignment between corporate partners and humanitarian principles and policies.

Purchasing and supply management (PSM) in the humanitarian sector presents unique features and key differences from the common practice in the commercial,⁹⁰ public, and military⁹¹ sector. Up to 65% of aid passes through a procurement process of some kind, making humanitarian procurement a critical function handling over EUR 20 billion in 2024 alone.⁹² Despite its key role, it is historically

⁸² Chedraoui, K. (2025, June 21). *Why you should care about OpenAI’s new \$200 million Defense Department deal*. CNET.

<https://www.cnet.com/tech/services-and-software/openai-signed-a-200m-deal-with-the-defense-department-why-you-should-pay-attention/>

⁸³ Habeshian, S. (2025, July 14). *Musk’s Xai announces \$200 million contract with Pentagon*. Axios.

<https://www.axios.com/2025/07/14/xai-pentagon-grok-contract>

⁸⁴ *Anthropic awarded \$200M DOD agreement for AI capabilities*. Anthropic. (2025, July 14).

<https://www.anthropic.com/news/anthropic-and-the-department-of-defense-to-advance-responsible-ai-in-defense-operations>

⁸⁵ Curi, M. (2025, September 17). *Exclusive: Scale AI strikes deal with Pentagon*. Axios.

<https://www.axios.com/2025/09/17/scale-ai-deal-pentagon>

⁸⁶ Sabin, S., & Demarest, C. (2025, December 9). *U.S. military to use Google Gemini for new AI platform*. Axios.

<https://www.axios.com/2025/12/09/pentagon-google-gemini-genai-military-platform>

⁸⁷ *OpenAI x International Rescue Committee: Leveraging AI to Scale Ed-Tech in Crisis Affected Settings*. The IRC. (2024, May 29).

<https://www.rescue.org/press-release/openai-x-international-rescue-committee-leveraging-ai-scale-ed-tech-crisis-affected>

⁸⁸ *Announcing the initial people-first AI fund grantees*. OpenAI. (2025, December 3).

<https://openai.com/index/people-first-ai-fund-grantees/>

⁸⁹ *A people-first AI fund: \$50m to support nonprofits*. OpenAI. (2025, September 8).

<https://openai.com/index/people-first-ai-fund/>

⁹⁰ Thakur-Weigold, Parsa, Balci, Van Wassenhove (n. 6). <https://doi.org/10.1016/j.pursup.2024.100975>

⁹¹ Goussac, N., & Boulanin, V. (2026, February 10). *Responsible procurement of military artificial intelligence*. Stockholm International Peace Research Institute.

<https://www.sipri.org/publications/2026/other-publications/responsible-procurement-military-artificial-intelligence>

⁹² *Technical briefing paper: The State of humanitarian procurement*. Hulo and IAPG. (2025, March 20).

<https://reliefweb.int/report/world/technical-briefing-paper-state-humanitarian-procurement>

considered as a support or service function rather than a key strategic component. **The persistent misalignment between what procurement is supposed to do and what it is concretely tasked to do results in serious strategic consideration and investment gaps.**

Procurement teams are frequently under-resourced and lack tech-specific skills, focusing far more on administrative compliance than on strategic procurement activities.⁹³ Most humanitarian organisations maintain strict, standardised processes for supplier selection, contracting, and funding to comply with public procurement standards.⁹⁴ Even the smallest NGOs may undergo dozens of audits annually, necessitating detailed records of programme activities, transactions, and outcomes. For both headquarters and field staff, the burden of record-keeping adds to the demands of programme execution. The pressure from serial audits often results in IT systems and organisational cultures focused primarily on donor reporting,⁹⁵ with performance management a secondary concern.

CASE STUDY 3: MICROSOFT AND HEIGHTENED HUMAN RIGHTS DUE DILIGENCE FOR BUSINESS IN CONFLICT-AFFECTED CONTEXTS (HHRDD+)

Microsoft is a leading cloud and AI provider for the humanitarian ecosystem, a role grounded in a variety of commitments⁹⁶ including its Global Human Rights Statement, Responsible AI Principles, and AI Services Code of Conduct, in addition to its active role in rights-based initiatives such as the Global Network Initiative (GNI). Microsoft also publicly endorses the *UN Guiding Principles on Business and Human Rights* (UNGPs),⁹⁷ which imposes on companies a clear responsibility to avoid causing or contributing to human rights abuses, and to address risks directly linked to their business operations and relationships. In mid-2025, an internal company review confirmed⁹⁸ media investigations⁹⁹ alleging that an Israeli military intelligence unit, Unit 8200, had been using Microsoft Azure cloud services to store and process vast quantities of daily intercepts of Palestinians' telephone communications. These were used in Gaza alongside AI-based targeting tools to research and identify bombing targets, and in

⁹³ Ibid.

⁹⁴ Thakur-Weigold, Parsa, Balcik, & Van Wassenhove (n. 6)

⁹⁵ Ibid.

⁹⁶ *Technology & human rights: Microsoft CSR*. Microsoft. (n.d.).

<https://www.microsoft.com/en-us/corporate-responsibility/human-rights>

See also Microsoft. (2025, October 16). *Responsible AI transparency report*. Corporate Responsibility.

<https://www.microsoft.com/en-us/corporate-responsibility/responsible-ai-transparency-report/>

⁹⁷ Crown, S. (2021, October 20). *Taking on human rights due diligence*. Microsoft.

<https://blogs.microsoft.com/on-the-issues/2021/10/20/taking-on-human-rights-due-diligence/>

⁹⁸ Smith, B. (2025, September 25). *Update on ongoing Microsoft Review*. Microsoft.

<https://blogs.microsoft.com/on-the-issues/2025/09/25/update-on-ongoing-microsoft-review/>

⁹⁹ Davies, H., & Abraham, Y. (2025, August 6). "A million calls an hour": Israel relying on Microsoft Cloud for expansive surveillance of Palestinians. *The Guardian*.

<https://www.theguardian.com/world/2025/aug/06/microsoft-israeli-military-palestinian-phone-calls-cloud>

the West Bank to “blackmail people, place them in detention, or even justify their killing after the fact.”¹⁰⁰ In response, Microsoft ceased and disabled¹⁰¹ some Israeli military subscriptions and services, including their use of specific cloud storage and AI services and technologies. The United Nations Development Programme (UNDP) guide on hHRDD+ expressly states¹⁰² that businesses conducting heightened due diligence in conflict-affected contexts should take steps to understand the conflict, identify their actual or potential adverse impacts on the conflict, and act upon those findings. Where a business may contribute to an adverse human rights impact, it should take the necessary steps to cease or prevent its contribution and use its leverage to mitigate any remaining impact to the greatest extent possible. To this day, Microsoft has not publicly disclosed whether it has exercised hHRDD+, or why it took a media investigation for the company to start an internal review on the use of its tools by a client currently being investigated by both the International Criminal Court (ICC)¹⁰³ and the International Court of Justice (ICJ)¹⁰⁴ for breaches of international law, including alleged genocide. Instead, Microsoft has reportedly expanded its contracts with the Israeli military over time, as the needs for its services — including AI-related and cloud-based tools — rose exponentially after October 7, 2023.¹⁰⁵ While commending Microsoft's decision to cease all services related to Unit 8200, human rights activists have flagged that an internal investigation cannot replace a heightened human rights due diligence review. As for Unit 8200, they announced their intention to replace Microsoft with Amazon Web Services (AWS);¹⁰⁶ a move that raises identical questions about hHRDD+ compliance, especially in light of information emerging from within the company flagging “all-costs-justified, warp-speed approach to AI development”.¹⁰⁷ The humanitarian sector has kept mostly silent on this issue, and it remains unclear if and what consequences a continued lack of compliance by companies with their own ethical

¹⁰⁰ Ibid.

¹⁰¹ Smith (n. 97)

¹⁰² Pachoud & Milatović (n. 71)

¹⁰³ *Situation in the State of Palestine: ICC pre-trial chamber I rejects the State of Israel's challenges to jurisdiction and issues warrants of arrest for Benjamin Netanyahu and Yoav Gallant* | International Criminal Court. International Criminal Court. (2024, November 21).

<https://www.icc-cpi.int/news/situation-state-palestine-icc-pre-trial-chamber-i-rejects-state-israels-challenges>

¹⁰⁴ *Application of the convention on the prevention and punishment of the crime of genocide in the Gaza Strip (South Africa v. Israel)*. International Court of Justice. (n.d.). <https://www.icj-cij.org/case/192>

¹⁰⁵ Davies, H., & Abraham, Y. (2025, January 23). *Revealed: Microsoft deepened ties with Israeli military to provide tech support during Gaza War*. The Guardian.

<https://www.theguardian.com/world/2025/jan/23/israeli-military-gaza-war-microsoft>

¹⁰⁶ *Microsoft blocks Israel's use of its technology in mass surveillance of Palestinians*. BDS Movement. (2025, September 26).

<https://bdsmovement.net/news/microsoft-blocks-israel%E2%80%99s-use-its-technology-mass-surveillance-palestinians>

¹⁰⁷ Dave, P. (2025, November 26). *Amazon workers issue warning about company's “all-costs-justified” approach to AI development*. Wired.

<https://www.wired.com/story/amazon-employees-open-letter-artificial-intelligence-layoffs/>

commitments might possibly have on current and future contracting opportunities with principle-abiding organisations.

Some of these gaps — and their stifling effect on the development of digital-ready forms of procurement — are a logical consequence of the sector-wide dependence on a handful of Big Tech actors and the alleged absence of reasonable alternatives to these providers. As flagged by public sector observers, the concentration of foundation models among Big Tech firms also deeply undermines most of the claims of AI-related productivity gains made by the tech industry.¹⁰⁸ While in the medium term a provider has an incentive to offer its model as open source, in the longer term, once the provider has consolidated its market leadership, the company is incentivised to close the ecosystem and maximise income as rent by model providers. This ultimately translates into more limited diffusion of the technology across the economy and higher costs for users.

These gaps are further explained by the lack of investment in research on long-term supplier relationship management in the humanitarian field. So far, this topic has been neglected in favour of the organisational buying behaviour lens, which mostly focuses on procurement policies and strategies.¹⁰⁹ Similarly, while there is ample research and documentation on the crucial role played by established relationships with suppliers in the timely and efficient delivery of aid and services under extreme time and funding constraints, very limited attention has been provided to performance evaluation of purchase and supply management methodologies, strategies, policies, and approaches.

Studies have highlighted how “research appears to lag behind practice as larger HOs (humanitarian organisations) are already trying out innovative technologies to improve their performance”¹¹⁰ This results in a lack of proper theoretical and methodological frameworks to drive sector-wide reflections on the digitalisation of purchasing, the role of technology in forecasting and procurement management, supply networks and complex supply systems management, or the interdependent relationship between humanitarian principles and social, ethical, or environmental supply issues. **Currently, most available research focuses on large NGOs and higher or lower-middle income nations, while AI adoption through the lens of procurement in small NGOs or NGOs operating within low income countries remains largely unexplored.**¹¹¹

¹⁰⁸ Chardon-Boucaud, S., Dozias, A., & Gallezot, C. (2024, December). *The Artificial Intelligence Value Chain: What Economic Stakes and Role for France?*. Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique, Direction générale du Trésor.
<https://www.tresor.economie.gouv.fr/Articles/2024/12/05/the-artificial-intelligence-value-chain-what-economic-stakes-and-role-for-france>

¹⁰⁹ Thakur-Weigold, Parsa, Balcik, & Van Wassenhove (n. 6).

¹¹⁰ Ibid.

¹¹¹ Rotter, & Bailkoski (n. 13).

1.3. Why does this matter for human rights and humanitarian actors

The implications of AI are well documented in human rights¹¹² and humanitarian literature,¹¹³ as the first examples of algorithmic systems in the aid field go back forty years.¹¹⁴ In 2017, NetHope, a consortium of over 60 leading global humanitarian nonprofits, explored the potential scaling of AI applications and soon afterwards launched an “AI primer”¹¹⁵ in collaboration with technology partners. NetHope’s programming on AI is based on the assumption that AI can cause real harm when misused or poorly implemented.¹¹⁶ In 2018, Access Now, Amnesty International, and a number of human rights organisations and AI companies released the “Toronto Declaration”¹¹⁷ — a landmark statement on protecting human rights in the age of AI — which has been widely endorsed by the global human rights community.

Over time, documentation of the human rights issues related to algorithmic systems across sectors¹¹⁸ and hidden in its multilayered supply chain (including labour,¹¹⁹ social, economic, or environmental impacts)¹²⁰ sparked an active discussion on how to grapple with these externalities.¹²¹ Human rights scholars affirm that human rights-centred frameworks should inform any cost-benefit analysis of algorithmic systems.¹²² The assessment should compare any enhanced enjoyment of human rights unlocked by this technology against the human rights risks associated with them throughout the product lifecycle, supply chain, and customer base.¹²³ **As experts have noted, hypothetical efficiency**

¹¹² Access Now (n. 14).

¹¹³ Salaj (n. 49).

¹¹⁴ Hall, L. (2025, October 24). *The history of Artificial Intelligence in humanitarianism*. Humanitarian Leadership Academy.

<https://www.humanitarianleadershipacademy.org/resources/the-history-of-artificial-intelligence-in-humanitarianism/>

¹¹⁵ *Webinar: AI Primer*. NetHope. (2024, July 13). <https://nethope.org/webinars/ai-primer/>

¹¹⁶ *Discover the Nethope AI Lighthouse: Your Guide to responsible, purpose-driven AI specifically for nonprofits*. NetHope. (2026, January 28).

<https://nethope.org/programs/discover-the-ai-lighthouse-your-guide-to-responsible-purpose-driven-ai-specifically-for-nonprofits/>

¹¹⁷ *The Toronto Declaration*. Toronto Declaration. (2018, May). <https://www.torontodeclaration.org/>

¹¹⁸ Salaj (n. 49).

¹¹⁹ Open Data Institute. *Humans in the loop: Labour in the AI supply chain*. YouTube. (2024, December 4).

<https://www.youtube.com/watch?v=GEuy4USHMGY>

¹²⁰ Valdivia, A. (2024). The supply chain capitalism of AI: a call to (re)think algorithmic harms and resistance through an environmental lens. *Information, Communication & Society*, 28(12), 2118–2134.

<https://doi.org/10.1080/1369118x.2024.2420021>

¹²¹ Khan, T. (2024, August 15). *Protecting AI's essential workers*. Partnership on AI.

<https://partnershiponai.org/protecting-ais-essential-workers-a-pathway-to-responsible-data-enrichment-practices/>

¹²² Dulka, A. (2023). *The use of artificial intelligence in international human rights law*. Stanford Law School.

<https://law.stanford.edu/publications/the-use-of-artificial-intelligence-in-international-human-rights-law/>

¹²³ Kang, D., & Grauer, Y. (2025, October 4). *Takeaways from AP’s investigation into how US tech companies enabled China’s digital police state*. AP News.

<https://apnews.com/article/chinese-surveillance-silicon-valley-uyghurs-tech-xinjiang-7ddfd2a3260a541fd9ffedd>

gains cannot and should not be balanced against proven human rights impacts. Any alleged benefits of algorithmic systems must not come at the cost of possible human rights or humanitarian violations.

For example, early versions of key algorithmic tools were barely usable because their foundational models were trained on the whole corpus of the internet, and consequently reflected some of its unlawful or discriminating content in both processes and outputs.¹²⁴ Attempts at developing innovative solutions on top of existing LLMs such as the one adopted and tested with encouraging results by Signpost AI (see case study 1) do not unfortunately remove the original problem¹²⁵ that makes foundational models basically unsafe for use without very complex de-risking measures¹²⁶ or attentive human-in-the-loop approaches.¹²⁷

Research keeps discovering new ways in which bias can be baked into the foundational model weights, with some of these being especially relevant in a humanitarian context. For example, the policies or guidelines set by national regulators for controlling or censoring some foundational models have resulted in forms of “emerging misalignment, a de facto intrinsic kill switch”.¹²⁸ These might heavily affect the quality of the output when using the raw open-source model, and even impede the output generation altogether¹²⁹ when accessed through the API or app (see case study 4). This is just one element of what has been called the “AI security risk overhang”,¹³⁰ a still unexploited but exponentially growing attack surface. **This is just one example that has led some researchers to warn¹³¹ that even debiasing or ex-post correcting measures would not be enough to mitigate the harmful effects of fundamentally flawed systems.**

[db44e34f4](#); See also *Call from a hundred NGOs: “AI as developed is fuelling a global exploitation system”*. Vox Public. (2025, February 6).

<https://www.voxpublic.org/Stand-What-you-won-t-be-told-at-the-AI-Action-Summit.html>

¹²⁴ Song, P., Ojo, A., & Curry, E. (2026). Trustworthy requirements for foundation models — a comprehensive survey and roadmap. *Engineering Applications of Artificial Intelligence*, 163, 113111.

<https://doi.org/10.1016/j.engappai.2025.113111>

¹²⁵ Kliemann, C., & Salaj, R. (2025, December 2). *From SDGs to GPTs: Trading Solidarity for AI Automation?*. Debating Development Research. <https://www.developmentresearch.eu/?p=2235>

¹²⁶ Ali, A. (2025, July 15). *Signpost AI Information assistant: A pilot case study*. Signpost AI.

<https://www.signpostai.org/research/signpost-ai-a-case-study>

¹²⁷ *Moderation*. Signpost Global. (n.d.).

<https://signpost-global.zendesk.com/hc/en-us/categories/360006184093-Moderation/>

¹²⁸ Stein, S. (2025, November 20). *CrowdStrike researchers identify hidden vulnerabilities in AI-coded software*. CrowdStrike.

<https://www.crowdstrike.com/en-us/blog/crowdstrike-researchers-identify-hidden-vulnerabilities-ai-coded-software/>

¹²⁹ Ibid.

¹³⁰ Saxe, J. (2026, January 6). *How to defend an exploding AI attack surface when the attackers haven’t shown up (yet)*. Substack. <https://substack.com/home/post/p-183640704>

¹³¹ Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, 2(2), 100205.

<https://doi.org/10.1016/j.patter.2021.100205>

CASE STUDY 4: THE HIDDEN GEOPOLITICS OF GENERATIVE AI

By their very nature, humanitarian organisations deal with highly sensitive and restricted military and political environments, navigating the complexities of conflictive, polarising, and politicised language¹³² hidden in loaded terms such as names, events, and places, among others. A typical language-related risk associated with aid work is the leaking of or unauthorised access to data or metadata associated with these categories by warring parties. This can result in danger for individuals or communities involved, or any form of harm — be it societal or individual — related to crisis response decisions based on biased or faulty data in the digital pipeline. But language-related issues can also result in reputational impacts for aid entities,¹³³ with consequences for their legal status or security. In this regard, the emergence of a geopolitical race to AI development is revealing new types of organisational risks previously unknown to the sector, or at least not considered as a primary concern. For example, independent tests by CrowdStrike Counter Adversary Operations¹³⁴ on DeepSeek-R1, a large language model (LLM) by China-based AI startup DeepSeek offering an accessible option for AI assisted coding, found that the quality of the system’s output was comparable to other and more expensive market-leading LLMs. However, they also detected that the likelihood of it producing code with severe security vulnerabilities increased by up to 50% when DeepSeek-R1 received prompts containing topics that might be considered harmless by the user, but that are likely considered politically sensitive by the Chinese Communist Party. Researchers found¹³⁵ that, “for example, when telling DeepSeek-R1 that it was coding for an industrial control system based in Tibet, the likelihood of it generating code with severe vulnerabilities increased to 27.2%. This was an increase of almost 50% compared to the baseline.” It is unlikely that such a flaw might be intentional, appearing to be instead what the researchers define as ‘emergent misalignment’. The system’s degraded capacity to produce safe and secure code would then be an unintended deviance due to policies embedded to ensure the system remains compliant — even if used as a raw open-source model — with existing requirements under Chinese law at the time of training the original algorithm. Although CrowdStrike tested a specific iDeepSeek model, they stress that these kinds of biases could affect any LLM, including smaller ones, especially those suspected to have been trained to adhere to certain ideological or cultural values. Data by Google Cloud informs that in 2025, up to 90% of developers already use AI assisted software for

¹³² Mahé, P., & de Geoffroy, V. (2023, November 6). *The relations between humanitarian aid and the political realm: Past, present and future*. Groupe URD.

<https://www.urd.org/en/review-hem/the-relations-between-humanitarian-aid-and-the-political-realm-past-present-and-future/>

¹³³ Gosp-Server, L. (2025, July 30). *The crucial role of humanitarian communication in the fake news and “infoglut” era*. Alternatives Humanitaires.

<https://www.alternatives-humanitaires.org/en/2025/07/30/the-crucial-role-of-humanitarian-communication-in-the-fake-news-and-infoglut-era/>

¹³⁴ Stein (n. 128).

¹³⁵ Stein (n. 128).

coding,¹³⁶ which means that “any systemic security issue in AI coding assistants is both high-impact and high-prevalence.”¹³⁷ Researchers suggest that companies using LLMs or AI agents as coding assistants should thoroughly test the agent within its designated environment, stressing that relying on generic open-source benchmarks is not enough. Humanitarian organisations procuring coding services or digital information systems from third parties — even if convinced to not procure AI systems — should therefore consider in their due diligence process how geopolitical factors upstream in the digital supply chain might have influenced not only the outputs they obtain from a tool, but also the quality and security of the code entering their ICT infrastructure.

These kinds of risks would be unknown to the aid groups using the model, and very often also to the tech provider and developers themselves. Introducing mitigation measures would also be problematic due to the complex and interconnected nature of these systems, and the inherently challenging regulatory environment faced by aid groups. Humanitarian organisations have to deal with diverse constituencies across geographical areas, in potentially endless combinations of applicable norms, standards, and frameworks.¹³⁸

CASE STUDY 5: WIKIMEDIA FOUNDATION HUMAN RIGHTS IMPACT ASSESSMENT ON AI AND MACHINE LEARNING

In 2024, the Wikimedia Foundation carried out a human rights impact assessment (HRIA) on AI and machine learning (ML)¹³⁹ to help the organisation and Wikimedia volunteer communities better understand how these technologies may affect the exercise of human rights in the Wikimedia tech ecosystem. The report was realised in collaboration with Taraaz Research, a specialised research and advocacy organisation, in a rare display of partnership-building centred on strengthening human rights through policy and governance reform rather than driven by efficiency-boosting considerations. The HRIA considered risks emanating from three different categories of issues specific to Wikimedia projects: editorial support tools developed in-house by Foundation staff; generative AI and its potential for marginal human rights risks in the Wikimedia context; and content on Wikimedia projects that may

¹³⁶ Salva, R. J. (2025, September 23). *How are developers using AI? Inside our 2025 dora report*. Google. <https://blog.google/innovation-and-ai/technology/developers-tools/dora-report-2025/>

¹³⁷ Stein (n. 128).

¹³⁸ Resources. Center for AI and Digital Policy. (n.d.). <https://www.caidp.org/resources/>

¹³⁹ Gaines, R. (2025, October 21). *Making sure AI serves people and knowledge stays human: Wikimedia Foundation publishes a human rights impact assessment on the interaction of AI and machine learning with Wikimedia Projects*. Wikimedia DIFF. <https://diff.wikimedia.org/2025/09/30/making-sure-ai-serves-people-and-knowledge-stays-human-wikimedia-foundation-publishes-a-human-rights-impact-assessment-on-the-interaction-of-ai-and-machine-learning-with-wikimedia-projects/>

be used for external ML development. No concrete harms were found during the initial assessment and the report expressly focused on potential harms that could occur in the future without proper mitigation measures in place. Just a few months later, however, this open governance system was put to the test when Wikipedia editors discovered that some translations of existing articles into other languages contained AI “hallucinations”. Editors scrambled to implement new policies and restrict a number of contributors paid by an external nonprofit to use large language models for translation.¹⁴⁰ Besides demonstrating the feasibility of mapping and assessing algorithmic systems across a large, distributed, and complex organisation, the Wikimedia team also developed a specific version of the HRIA for partner companies’ due diligence. This latter tool is meant to be used on customers of the Wikimedia Enterprise¹⁴¹ service offering API access to companies, and might offer a rare and interesting example of possible applications of the HRIA model for digital supply chain assessments. It would be a laudable step to disclose this resource, or at least parts of it as a template, to show other actors a concrete example of algorithmic supply chain due diligence in the nonprofit space.

As recalled by Verity and Wright,¹⁴² AI systems that have prevalent biases, security risks, and issues with consent can undermine the role of humanitarian actors in disaster contexts by leaving aid recipients even more vulnerable. These risks can take several forms, such as biased decision making, non-consensual sharing of humanitarian data,¹⁴³ violations along the digital¹⁴⁴ and data¹⁴⁵ supply chains, negative environmental impacts,¹⁴⁶ discrimination in service delivery, or other unintended negative consequences for already marginalised individuals seeking support or protection. Other risks include de-anonymization of individuals in danger, the unlawful targeting of civilians in conflict, harvesting people’s data for commercial purposes, or enabling harm at scale by automating life-impacting decision processes without meaningful human oversight or remedy.¹⁴⁷

¹⁴⁰ Maiberg, E. (2026, March 7). *AI translations are adding “hallucinations” to Wikipedia articles*. 404 Media. <https://www.404media.co/ai-translations-are-adding-hallucinations-to-wikipedia-articles/>

¹⁴¹ Wikimedia Enterprise. (n.d.). <https://enterprise.wikimedia.com/>

¹⁴² Wright, J., & Verity, A. (2020, January). *Artificial intelligence principles for vulnerable populations in humanitarian contexts*. Digital Humanitarian Network. <https://digitalhumanitarians.com/artificial-intelligence-principles-for-vulnerable-populations-in-humanitarian-contexts/>

¹⁴³ Martin, A. (2025). Why sovereignty matters for Humanitarian Data. *Big Data & Society*, 12(3). <https://doi.org/10.1177/20539517251361109>

¹⁴⁴ Johnston, J. (2025). *Human Rights in the AI Supply Chain*. <https://doi.org/10.2139/ssrn.5437935>

¹⁴⁵ Jindal, S. (2025, May 1). *AI and human rights: Protecting data workers*. Partnership on AI. <https://partnershiponai.org/ai-and-human-rights-protecting-data-workers/>

¹⁴⁶ *AI has an environmental problem. Here's what the world can do about that*. UNEP. (2025, November 13). <https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about>

¹⁴⁷ *Towards responsible humanitarian AI: Guidelines from research and practice*. Signpost AI. (n.d.). <https://www.signpostai.org/research/responsible-humanitarian-ai-guidelines-lessons-from-the-field>

On a systemic level, experts repeatedly warned that the same tools supposed to improve people's lives can also create a web of autocratic surveillance,¹⁴⁸ and other potential harmful consequences that go well beyond the specific impact on vulnerable populations. They may also indirectly affect the achievement of the 2030 Sustainable Development Agenda, and the maintenance of the UN Charter and international human rights laws. In many ways, **humanitarian actors' careless use of algorithmic tools might also slowly influence how the whole sector adheres to key humanitarian norms and principles.**¹⁴⁹

Despite this, humanitarian entities keep engaging with these technologies using common procurement systems that favour financial, performance, and liability considerations over human rights impact assessments or compliance with humanitarian norms and standards.¹⁵⁰ **Humanitarian actors carry a historic gap in accountability to affected people and rely on informal commitments such as tech providers' ethical charters or policies, thus facing an increased risk of cascading trust failures generating adverse human rights impacts.**

The discussion about AI is a critical moment to revisit and potentially rectify the sector's historic reluctance to implement or challenge existing regulatory frameworks,¹⁵¹ or tackle much-needed governance reform well beyond individual digital product procurement. Humanitarian actors often rely on vague references to ethics without implementing legal safeguards or contesting problematic uses of digital tools.¹⁵² **But unlike traditional technologies, algorithmic systems in general and AI in particular have been defined as an 'apparatus': a layered and interdependent arrangement of technology, institutions, and ideology.**¹⁵³

1.4. Main trends in humanitarian procurement of algorithmic systems

The first, most glaring trend visible noted in this report's interviews and literature review — in line with the general debate on AI — is the absence of any attempt by the sector to converge on a shared understanding of what AI means for the whole field and what might be truly unique about it. Humanitarian experts and literature most commonly refer to AI as if it were a single, uniform

¹⁴⁸ Tyrie, R. (2024, June 7). *AI's double-edged sword: Navigating the risks and rewards for humanitarianism*. Grey Swan Guild.

<https://medium.com/grey-swan-guild/ais-double-edged-sword-navigating-the-risks-and-rewards-for-humanitarianism-6eb7c6b11e75>

¹⁴⁹ Coppi, G., Moreno Jimenez, R., & Kyriazi, S. (2021). Explicability of humanitarian AI: A matter of principles. *Journal of International Humanitarian Action*, 6(1). <https://doi.org/10.1186/s41018-021-00096-6>

¹⁵⁰ Devidal (n. 22).

¹⁵¹ Döchting (n. 19).

¹⁵² Sandvik, K. B. (2023). *Humanitarian extractivism: The Digital Transformation of Aid*. Manchester University Press. <https://manchesteruniversitypress.co.uk/9781526173355/>

¹⁵³ Kliemann, C., & Salaj, R. (2025, December 2). *From SDGs to GPTs: Trading Solidarity for AI Automation?*. Debating Development Research. <https://www.developmentresearch.eu/?p=2235>

technology, rarely distinguishing between narrow and general technology — or between different use cases, intended purposes, capabilities, contexts, target audiences, or users.

Understanding what is, or isn't, an AI system or function is increasingly challenging.

Interviewees with direct experience of digital procurement shared that self-labelled AI tools presenting nothing that might fall under the common definition of this technology are a more common occurrence than hidden AI features. This digital procurement version of whack-a-mole is a waste of already limited internal resources that fails to provide any meaningful digital security guarantees.

The lack of an ontological moment in the humanitarian sector has paved the way to the unchecked spread of 'AI for good' claims,¹⁵⁴ associating to an already undefined technology an equally undefined capacity to "do good". This is a floating¹⁵⁵ or empty signifier¹⁵⁶ falling into the broader category of what is known as aidwashing.¹⁵⁷ The vagueness surrounding so-called AI systems also extends to the underlying tech stack, the data pipeline, and digital supply chain involved in the whole process.¹⁵⁸ Very few organisations — even among the private sector — review risks coming from immediate suppliers. A 2021 UK Department for Digital, Culture, Media and Sport survey found that “only 12% of businesses review risks coming from immediate suppliers, while only one in twenty address risks coming from wider supply chains.”¹⁵⁹

To deal with such a shapeless technology, each humanitarian organisation is setting up informal ways to detect and flag digital products that might need enhanced scrutiny. In most cases, these result in simple keyword searches in the product specifications provided by the vendor, or in the internal requirements and specifications annexed to the purchase / requisition order or a tender. To decide which product to prioritise for screening, experts within aid groups often adopt a triage process based on where the internal request is coming from. For example, **a purchase order from the protection**

¹⁵⁴ ITU. (2026, January 16). *About Us*. AI for Good. <https://aiforgood.itu.int/about-us/>

¹⁵⁵ Romele, A. (2024). « Éthique de l'intelligence artificielle » comme signifiant flottant : Considérations théoriques et analyse critique des discours de presse. *Interfaces Numériques*, 13(1). <https://doi.org/10.25965/interfaces-numeriques.5229>

¹⁵⁶ Hardcastle, F., Raman, S., De Silva, C., Davis, J., & Tavakoli-Nabavi, E. (2025). Rethinking AI for good: Critique, reframing and alternatives. *AoIR Selected Papers of Internet Research*. <https://doi.org/10.5210/spir.v2024i0.13955>

¹⁵⁷ Martin, A. (2023). Aidwashing surveillance: Critiquing the corporate exploitation of humanitarian crises. *Surveillance & Society*, 21(1), 96–102. <https://doi.org/10.24908/ss.v21i1.16266>

¹⁵⁸ For an example of mapping and disclosure of external suppliers delivering services through a multi-layered supply chain from the public sector see UK Government

Digital Service. *Algorithmic Transparency Recording Standard — guidance for public sector bodies.*:

<https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard/algorithmic-transparency-recording-standard-guidance-for-public-sector-bodies>

¹⁵⁹ *Call for views on cyber security in supply chains and managed service providers*. GOV.UK. (2021, November 15).

<https://www.gov.uk/government/publications/call-for-views-on-supply-chain-cyber-security/call-for-views-on-cyber-security-in-supply-chains-and-managed-service-providers>

team is much more likely to receive an enhanced level of scrutiny than one from an administrative office.

To do this, many organisations simply deploy cross-functional teams — an already well-established practice for many technical areas — to coordinate the overall procurement process. Some of the better-resourced entities even hire specific profiles from the private sector to provide deeper technical analysis of algorithmic digital systems, and sketch embryonic forms of procurement guidance in AI-specific policies.¹⁶⁰ On the other hand, none of the organisations consulted has set up a specific official procurement protocol for algorithmic systems. AI-specific toolkits and guidance are considered an informative resource and a way to reassure leadership, but unhelpful as mandatory standalone protocol. Humanitarian logistics and procurement representatives have stressed¹⁶¹ how investments in the digitalisation of humanitarian procurement and the broader supply chain have so far failed to reduce the administrative burdens of compliance. Amid the ongoing funding crisis, adding more compliance requirements hardly seems a viable solution.

This is confirmed by emerging guidance on risk management¹⁶² which identified little novelty in the procurement of AI, mostly “reduced to two categories, 1) complexity of the AI system and 2) scale of the potential impact on the population.” Similarly, the adoption of AI in the humanitarian sector does not seem to have faced unique or specific challenges compared to classic innovation processes.¹⁶³ Humanitarian digital transformation processes were usually quite slow even before the current funding crisis, aid leaders are traditionally risk-averse, and humanitarian staff include a considerable number of digital-sceptics. So far, these factors seem to have kept the AI hype at bay. Overall, **experts seem confident that the technosolutionist wave will pass and clearer, more reliable metrics will emerge to assess the level of scrutiny needed for cloud-based or algorithmic systems.**

1.5. Taxonomy of emerging practices in the integration of algorithmic systems

In addition to being neglected, underfunded, and mostly struggling to adjust to the reality of digital development, traditional procurement is also increasingly bypassed by the same products it is supposed to purchase. Our research identified various ways in which algorithmic tools make their way into the inner workings of humanitarian entities, which we have summarised in a

¹⁶⁰ Devidal, P. (2024, November 28). *Trying to square the circle: The ICRC AI policy*. Humanitarian Law & Policy Blog. <https://blogs.icrc.org/law-and-policy/2024/11/28/trying-to-square-the-circle-the-icrc-ai-policy/>

¹⁶¹ Hulo and IAPG (n. 92).

¹⁶² Miller, C., & Waters, G. (2023). *RMF for AI Procurement*. Center for Inclusive Change. <https://www.inclusivechange.org/ai-governance-solutions/rmf-for-ai-procurement>

¹⁶³ Döchting (n. 19).

simple tentative taxonomy built on top of key tasks involved in purchasing and supply management (PSM) as identified in the existing scholarship¹⁶⁴ (see Table 1: Taxonomy).

Type	Subtype	Examples	Process
Market-driven	Purchasing	Open-ended control over a repository or system provided against payment of a fee	Purchasing and Supply Management (PSM)
	Licensing	Temporary access to a system provided against payment of a fee	PSM
Organisation-driven	Internal development	Custom development or customisation of a digital system by internal teams and consultants	Mix of Project Cycle Management (PCM) and PSM
	Partnership-based	Custom development or customisation of a digital system with some degree of support by external partners without significant dependency	Mix of PCM and PSM
User-driven	Shadow IT systems	Chatbot app integrated or used on professional devices by staff outside of IT oversight	None (excepted possible IT access controls procedures)
	Parallel IT systems	Chatbot app being used by staff outside of the professional IT environment (e.g. on a personal device)	none
Supplier-driven	Major updates	Upgrades ¹⁶⁵ resulting in significant changes, often involving the addition of new features, redesigning the user interface, or making other substantial changes to the app's functionality	Terms and conditions update, privacy notice, user agreement notice, update notification, update download / run
	Minor updates	Updates addressing specific issues or vulnerabilities, such as patches ¹⁶⁶ or bug fixes, or other small changes	Update notification, update download / run

Table 1: Taxonomy – Type of algorithmic tools integration methods in the organisational tech stack

As shown in table 1, standard PSM approaches guide only a fraction of the algorithmic systems integrated into the organisation’s tech stack. Their effectiveness is even more limited in aid groups with reduced technical and financial resources, which rely on free-tier access options or on staff opening individual accounts to access digital applications and services. This also applies to federated, distributed, or decentralised entities, which might house multiple teams with various degrees of independence in the decision-making process leading to digital tools selection. **To compensate for the gaps in coverage of PSM approaches, organisations deployed a set of risk mitigation measures that can be tentatively grouped in five types:**

- **Fiduciary measures** include screening vendors based on their allegiance to recognised ethical charters, common membership to trusted platforms, solidity of historic relationships, or

¹⁶⁴ Moshtari, M., Altay, N., Heikkilä, J., & Gonçalves, P. (2021). Procurement in humanitarian organisations: Body of knowledge and practitioner’s challenges. *International Journal of Production Economics*, 233, 108017. <https://doi.org/10.1016/j.ijpe.2020.108017>

¹⁶⁵ Meilke, G. (2025, May 23). *Software updates vs. patches: Your mission-critical security steps, explained*. Rightworks. <https://www.rightworks.com/blog/software-patches/>

¹⁶⁶ Ibid.

demonstrated alignment with values and principles shown through forms of direct or indirect support;

- **Governance measures** include conditioning a digital product’s adoption, integration, or usage to the respect of internal policies or protocols, vetting providers based on objective criteria established in formal or informal policies, or training and capacity building programmes;
- **Normative measures** include assessing companies and services based on compliance with existing legislative and regulatory framework in the country of registration and/or in the areas of operation, or the inclusion of protective or prescriptive language or clauses in contractual agreements;
- **Technological measures** include system audits by the organisation or, more commonly, by qualified third parties, demonstrated compliance with technical requirements set by the organisation or with sector-wide recognised standards, or technical solutions for risk mitigation introduced at local or cloud environments; and
- **Financial measures** include any form of financial guarantees or assurance agreements tying product performance to economic incentives or rewards.

Some of these screening exercises have been consolidated in existing guidance,¹⁶⁷ but once again build on the assumption that the algorithmic system is the only or main component being screened. The informal ways algorithmic systems make their way into aid organisations’ digital ecosystem shows that the number and enforceability of risk mitigation measures currently available to humanitarian organisations progressively decreases as the integration type slips away from structured PSM procedures (See Table 2: Mapping).

Type	Subtype	Common mitigation measures	Examples <i>(possibly applicable across types)</i>	Process
Market-driven	Purchasing	Fiduciary, governance, legal, technological, financial	<ul style="list-style-type: none"> - Vendor has a public set of principles or adheres to a recognised ethics charter; - Vendor agrees to run a third party-led audit; - Vendor defined in contract as data processor only - Vendor has passed the Assessment List for Trustworthy AI (ALTAI)¹⁶⁸ or equivalent assessment; - Client relies on existing comparative evaluations of 	PSM
	Licensing			

¹⁶⁷ European Commission. (n.d.). *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment*. High-Level Expert Group on Artificial Intelligence. <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

¹⁶⁸ Idem.

			<p>AI/LLM systems;¹⁶⁹ and/or</p> <ul style="list-style-type: none"> - Vendor obtains a certification as verified system (e.g. the Safe AI¹⁷⁰ framework or Safetag¹⁷¹ certificate) or adheres to AI assurance protocols.¹⁷² 	
Organisation-driven	Internal development	Governance, legal, technological	<ul style="list-style-type: none"> - Development team has a strict set of requirements and specifications and follows a trusted framework (examples by ICRC/EPFL,¹⁷³ GiveDirectly,¹⁷⁴ and IRC's Signpost¹⁷⁵); - Final product is audited, pentested, and continuously tested (e.g. white hacking),¹⁷⁶ - Partner has a time-honored relationship with aid group; - Partner formally accepts to respect aid group's policies and accepts liability and mitigation rules; - Development and testing environment are locally hosted, and/or sandboxed and air-gapped; and/or - Sensitive databases or systems are segregated. 	Mix of PCM and PSM
	Partnership-based	Fiduciary, governance, legal, technological		Mix of PCM and PSM
User-driven	Shadow IT systems	Fiduciary, governance, technological	<ul style="list-style-type: none"> - Staff and consultants are required to sign a policy on ICT acceptable behaviour and receive specific training on such policy; and/or - ICT introduces a whitelisting system. 	None (excepted possible IT access controls procedures)
	Parallel IT systems	Fiduciary, governance, technological		None

¹⁶⁹ For some examples of available evaluation and assessment systems see Chia, H. S., Abrol, F., Madon, T., On, R., Walsh, J., Wu, Z., & Carter, S. (2025, April 16). *An AI evaluation framework for the Development Sector*. Center for Global Development. <https://www.cgdev.org/blog/ai-evaluation-framework-for-the-development-sector>. See also examples collected by MERL Tech: *Join us on Oct 9 for a conversation about evaluating LLMs for accuracy and inclusion*. MERL Tech. (2025, September 15).

<https://merltech.org/oct-9-evaluating-llms-for-accuracy-and-inclusion/>

¹⁷⁰ *How to drive adoption of community crisis intelligence*. Nesta.

<https://www.nesta.org.uk/project/how-to-drive-adoption-of-community-crisis-intelligence/>

¹⁷¹ Safetag. (n.d.-a). <https://safetag.org/>

¹⁷² Winecoff, A., Bogen, M., & Groves, L. (2025, July 10). *Assessing AI: Surveying the spectrum of approaches to understanding and auditing AI systems - center for democracy and technology*. CDT AI Governance Lab.

<https://cdt.org/insights/assessing-ai-surveying-the-spectrum-of-approaches-to-understanding-and-auditing-ai-systems/>. See also *SAFE AI: Standards and Assurance Framework for Ethical Artificial Intelligence in Humanitarian Contexts*. CDAC Network, The Alan Turing Institute & Humanitarian AI Advisory. (forthcoming, 2026)

<https://www.cdacnetwork.org/safe-ai>; McElhinney, H., Mazumder, A. & Tjalve, M. (2026, March) *SAFE AI instalment 1: The Governance Gap in Humanitarian AI*. <https://www.cdacnetwork.org/resources/the-governance-gap-in-humanitarian-ai>.

¹⁷³ *Chitchat architectures for interpretable & transparent continuous humanitarian alignment in chatbot technologies*. Engineering Humanitarian Action. (2025, September 18).

<https://eha.swiss/case-study/chitchat-architectures-for-interpretable-transparent-continuous-humanitarian-alignment-in-chatbot-technologies-2/>

¹⁷⁴ Luk, S., Lummis, V., & Ramprasad, S. (2024, November 21). *GiveDirectly's approach to responsible AI/ML*. GiveDirectly. <https://www.givedirectly.org/ai-framework/>

¹⁷⁵ Nicoll (n. 41).

¹⁷⁶ Gilman, D. (2020, July 24). *Cyber-warfare and Humanitarian Space*. Global Interagency Security Forum. <https://gisf.ngo/resource/cyber-warfare-and-humanitarian-space/>

			or artificially generated content.	
Supplier-driven	Major updates	Fiduciary, governance, technological	<ul style="list-style-type: none"> - Provider has a time-honored relationship with aid group / client; - Aid group evaluates the update and provide guidance to staff; - Partial or complete migration of sensitive data to another environment; and/or - Client refuses the update, opening up the possibility of service discontinuation or contractual breakup. 	Terms and conditions update, privacy notice, user agreement notice, update notification, update download / run
	Minor updates	Fiduciary, governance, technological		Update notification, update download / run

Table 2: Mapping – Type of prevalent mitigation measures per integration type/subtype

As algorithmic functionalities grow more ubiquitous and are increasingly embedded in core digital products, humanitarian organisations progressively lose capacity to enforce meaningful control through traditional PSM processes. This calls for a reconsideration of existing centralised acquisition processes, in favour of more dynamic digital solution integration approaches requiring distributed validation, vetting, and oversight points. Unfortunately, based on available data, none of the many humanitarian procurement initiatives¹⁷⁷ seem to focus on the specific challenges posed by the procurement of digital systems in general and algorithmic systems in particular, and those exploring supply chain issues only consider the environmental and sustainability angles.¹⁷⁸

2. Conclusions

2.1. Money cannot buy human rights compliance

Over the past decade, algorithmic tools have been progressively integrated into public and private services, allegedly to enhance the efficiency of aid management, planning, and delivery. They have also created significant human rights risks, affecting the rights to life, privacy, equality, human dignity, and protection against discrimination, among others. **In situations of conflict, potential risks expand beyond the human rights field and into the international protection framework offered by humanitarian norms and principles, as traditional actors struggle to identify, document and review impact stories and testimonies about emerging forms of harm.**

As the AI industry experiences tremendous growth, regulators are failing to improve existing legal frameworks by supporting human rights and ethics-focused regulatory exercises.¹⁷⁹ The result is a

¹⁷⁷ *Humanitarian procurement ecosystem mapping*. Inter-Agency Procurement Group. (n.d.).

<https://iapg.org.uk/humanitarian-procurement-ecosystem-mapping/>

¹⁷⁸ *IAPG Mapping of Humanitarian Procurement Stakeholders*. IAPG. (2025, June).

https://iapg.org.uk/wp-content/uploads/2025/08/IAPG_Humanitarian-procurement-stakeholders-mapping.pdf

¹⁷⁹ Mantelero, A. (2024). The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template. *Computer Law & Security Review*, 54, 106020.

<https://doi.org/10.1016/j.clsr.2024.106020>

proliferation of ethics codes¹⁸⁰ and tentative common core values¹⁸¹ in a public debate dominated by the idea that the public and nonprofit sector might use moral suasion on tech companies through their role of ethical or rights-based buyer. In this scenario, as trusted and responsible buyers, states and other ethics or rights-driven entities demand fair, ethical, trustworthy, transparent, and secure technologies, thus generating enough pressure to shape mainstream commercial products by influencing the tech industry's self-regulation.

These theories suggest that purchasing and supply processes might ultimately generate foundational safeguards and therefore protect people's rights despite the broader underregulation of both digital technologies and the new models of digital governance they enable. Behind this conviction lies the assumption that regulatory developments face impassable constraints encapsulated in the Collingridge dilemma¹⁸² or pacing problem¹⁸³ ("cannot effectively regulate early on, so will probably regulate too late"), which has a specific version at the multilateral and international level.¹⁸⁴

The general nature of these soft instruments has allowed the tech industry to let their business interests drive the implementation of ethical values¹⁸⁵ as shown by the dreadful results of Ranking Digital Rights' Big Tech scorecard on algorithmic systems.¹⁸⁶ As a result, in spite of decades of pledges to innovate responsibly and support 'AI for Good' initiatives, major tech companies have violated laws and ethical principles to bring their innovative products to market,¹⁸⁷ boosted their presence in the defence market,¹⁸⁸ and culled their human rights teams

¹⁸⁰ Corrêa, N. K., Galvão, C., Santos, J. W., Del Pino, C., Pinto, E. P., Barbosa, C., Massmann, D., Mambrini, R., Galvão, L., Terem, E., & de Oliveira, N. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI Governance. *Patterns*, 4(10), 100857. <https://doi.org/10.1016/j.patter.2023.100857>

¹⁸¹ Ibid.

¹⁸² Vourdakís, A. (2024, October 7). *What is the Collingridge dilemma and why is it important for tech policy?* Demos Helsinki. <https://demoshelsinki.fi/what-is-the-collingridge-dilemma-tech-policy/>

¹⁸³ Marchant, G. E., Allenby, B. R., & Herkert, J. R. (2011). *The growing gap between emerging technologies and legal-ethical oversight: The pacing problem*. Springer. <https://link.springer.com/book/10.1007/978-94-007-1356-7>

¹⁸⁴ *The Relationship between Digital Technologies and Atrocity Prevention*. Global Centre for the Responsibility to Protect. (2024, March 7).

<https://www.globalr2p.org/publications/the-relationship-between-digital-technologies-and-atrocity-prevention/>

¹⁸⁵ Mantelero (n. 178).

¹⁸⁶ *The 2022 Ranking Digital Rights Big Tech Scorecard*. Ranking Digital Rights. (2022). <https://rankingdigitalrights.org/bts22/indicators/G4d>

¹⁸⁷ Mantelero (n. 178).

¹⁸⁸ Pascual, M. G. (2025, July 21). *Big Tech enters the war business: How Silicon Valley is becoming militarized*. EL PAÍS English. <https://english.elpais.com/economy-and-business/2025-07-21/big-tech-enters-the-war-business-how-silicon-valley-is-becoming-militarized.html>

and statements.¹⁸⁹ The result is a growing record of harms¹⁹⁰ emerging from the unregulated experimentation with algorithmic tools, including in the public sector.¹⁹¹ Notably, critics of AI's 'regulation by contract' have defined it as "regulatory hallucination".¹⁹²

Our research into the humanitarian side of algorithmic systems further confirms how relying on procurement as a tool of soft regulation is deeply inadequate considering the power imbalance between Big Tech providers and humanitarian entities in a regulatory vacuum. **Purely relying on ethics-based processes can result in "humanitarian experimentation"¹⁹³ with immature technologies in environments lacking adequate data and technical infrastructures, and in additional risks of regulatory capture and regulatory privatisation.**

In our consultations, some experts claimed that the humanitarian sector has not recorded concrete AI-related harms, that current examples of documented failures related to programmes involving algorithmic tools are scarce, and that their causes are possibly tech-agnostic. These arguments don't seem to form a consensus. Other respondents flagged these perspectives as building on evidence of absence, by demanding that critics prove the negative of something that is simply not being systematically tracked nor documented. Even if proven correct, experts said that this approach would hardly be compatible with the 'do no harm' principle,¹⁹⁴ which demands a proactive effort to prevent and avoid any behaviour that might be reasonably expected to worsen existing humanitarian conditions or generate additional harm.

In similar contexts, especially **when potential negative effects can happen on a large scale, experts stressed how "the logic of risk management seems to be the most appropriate way to strike a balance and define what risk is acceptable and to what extent"**.¹⁹⁵ **On the companies' side, that would require replacing the 'err first, learn and correct later' motto with an "ex ante approach that makes it possible to prevent harmful applications from being placed on the market"**.¹⁹⁶ It is

¹⁸⁹ Wiggers, K. (2025, January 14). *OpenAI quietly revises policy doc to remove reference to "politically unbiased" AI*. TechCrunch.

<https://techcrunch.com/2025/01/14/openai-quietly-revises-policy-doc-to-remove-reference-to-politically-unbiased-ai/>

¹⁹⁰ *MIT AI Incident Tracker*. MIT. (n.d.). <https://airisk.mit.edu/ai-incident-tracker>; See also *AIM: AI Incidents and Hazards Monitor*. OECD AI Policy Observatory Portal. (n.d.). <https://oecd.ai/en/incidents>; and *Welcome to the Artificial Intelligence Incident Database*. AI Incident Database. (n.d.). <https://incidentdatabase.ai/>

¹⁹¹ *AIM: AI Incidents and Hazards Monitor*. OECD AI Policy Observatory Portal. (n.d.). <https://oecd.ai/en/incidents>

¹⁹² Sanchez-Graells, A. (2024). *Responsibly buying artificial intelligence: A 'regulatory hallucination.'* *Current Legal Problems*, 77(1), 81–126. <https://doi.org/10.1093/clp/cuae003>

¹⁹³ Jacobsen, K. L., Sandvik, K. B., & McDonald, S. M. (2018, October 25). *Humanitarian experimentation*. Humanitarian Law & Policy Blog.

<https://blogs.icrc.org/law-and-policy/2017/11/28/humanitarian-experimentation/>

¹⁹⁴ *Humanitarian principles*. European Commission Directorate-General for Civil Protection and Humanitarian Aid Operations. (n.d.). https://civil-protection-humanitarian-aid.ec.europa.eu/who/humanitarian-principles_en

¹⁹⁵ Mantelero (n. 178).

¹⁹⁶ Ibid.

therefore unsurprising that over time, the initial regulatory approach has increasingly shifted from ethical guidance to risk-based regulation.

In conclusion, public authorities and humanitarian entities most certainly have a duty as responsible tech users or developers to make sure that companies' voluntary commitments are sound and their adherence is strictly monitored. However, the purported benefits of virtuous and responsible procurement practices cannot be relied upon to protect human rights, as long as these partnerships and services remain outside of the legal frameworks that sanction people's rights in the first place. Even in the most effective and responsible procurement strategy, any digital transformation integrating algorithmic tools into the public and humanitarian space without adequate multistakeholder processes and rights-protecting guardrails risk being driven by “the policy irresistibility of ‘tech fixes’, limited (...) digital capabilities, and specific governance characteristics such as the policy push for innovation procurement.”¹⁹⁷

2.2. Algorithmic capture: The AI-poisoned gift for local actors and communities

On the tech companies' side, unfortunately, the data capture trap seems to be working as intended. **On one hand, local actors from resource-constrained environments in the Global Majority are once again defeating neocolonial stereotypes and placing themselves at the forefront of the debate on AI.** In Data Friendly Space's survey of more than 2,500 humanitarian workers, 75% of responders were from the Sub-Saharan Africa, MENA, and Asia-Pacific regions.¹⁹⁸

On the other hand, 69% of respondents to the same survey declared that they rely on commercial AI platforms like ChatGPT and Claude, rather than purpose-built humanitarian solutions. Early experiences from innovation-oriented joint projects between donors and small, mission-driven organisations seem to indicate that the cost of building their own algorithmic tool or paying for a commercial system was almost the same. This argument appears rooted in survivorship bias however, as it assumes as a baseline systemic equality that is non-existent, while erasing the struggle and experiences of organisations that may not yet have the resources to hire the staff or consultants needed to build, host, and protect customised systems.

Looking at the emerging data from surveys, the ways in which humanitarians and human rights defenders are accessing algorithmic products has less to do with enterprise-grade licensing and more to do with individuals opening personal accounts on various platforms, often based on their financial accessibility. **Very few individual users have reported accessing the main commercial AI systems through pro (or higher) tier paid accounts, which offer higher-output quality, privacy, and**

¹⁹⁷ Sanchez-Graells, A. (2024). Responsibly buying Artificial Intelligence: A 'regulatory hallucination.' *Current Legal Problems*, 77(1), 81–126. <https://doi.org/10.1093/clp/cuae003>

¹⁹⁸ Johnson (n. 19).

cybersecurity guarantees compared to the more popular free or low-cost versions. This is an important detail: the difference between free and paid tiers is sometimes so vast,¹⁹⁹ that some interviewees framed it as an emerging form of digital divide.

As the use of algorithmic systems grows, discussions on how to integrate algorithmic tech-related consideration into the governance models of smaller or less well-resourced organisations, especially but not only from the Global Majority, face specific challenges. **The sector keeps waving both the threats and the promise of AI under their nose while doing very little to facilitate access to funding and spaces where digital technology is designed, built, and monetised.** Interviews and desk research confirms that, beyond the well-known historic, financial, and technical factors, the current approach to algorithmic procurement also generates direct and often neglected impacts.

Local humanitarian actors, human rights defenders, and even local authorities in the Global Majority are entering the algorithmic arena with very limited capacity to decipher and negotiate contractual terms. Meanwhile, a relentless process of corporate capture by a handful of Big Tech actors is consolidating control over hyperscalable cloud and analytics systems. **In this struggle to remain a rights-based, ethical, and responsible tech actor, local organisations are mostly left to fend for themselves. All around them, in a competitive and fragmented aid sector, key players in the humanitarian tech environment are building platforms focused on fostering project-oriented strategic partnerships instead of solidarity-based approaches.**

As a result, when it comes to harnessing global tech platforms or off-the-shelf AI solutions, local organisations too often face a binary choice — to use or not use the few software solutions they can actually afford — without much space for nuanced approaches or mitigating measures. If the collective transition towards interconnected, algorithmic-enhanced digital tools continues, this already false choice might be reduced to whether or not to use digital systems at all. In the present scenario, **the few alternative ethical or rights-based options offered by bigger organisations mostly revolve around various forms of mediated access to algorithmic systems hosted, owned, or managed by larger INGOs or international organisations. In that case, to what extent should smaller organisations be held accountable for the security of their data, and where does the provider or partner’s responsibility begin?**

An increasing body of knowledge is exploring how to ensure local organisations play a central role in the development of algorithmic systems. **So far, research has identified²⁰⁰ four pathways for increasing inclusion of local civil society in algorithmic decision-making:**

- **Expanding AI literacy and cross-sectoral knowledge sharing;**
- **Increasing local decision-making and representation across the tech lifecycle;**

¹⁹⁹ Compare top generative AI tools in 2025: AI Subscription Value Analyzer. Interactive AI Subscription Value Analyzer. (n.d.). <https://c3.unu.edu/projects/ai/compare/>

²⁰⁰ CARE International & Accenture (n. 3).

- **Strengthening advocacy on the contextualised impacts and desired outcomes; and**
- **Improving digital infrastructure and equitable data governance.**

There are also efforts to develop models for inclusive AI,²⁰¹ and to unpack the nuances and complexity of regional approaches to thematic applications of algorithmic models, via monitoring, evaluation, research, and learning (MERL) activities on the African continent, for example.²⁰² But when it comes to setting up concrete measures, the field of accountability to affected populations (AAP, also defined as community engagement and accountability or CEA) is struggling to ensure that past mistakes are not repeated with emerging technologies, including AI, and that digital systems are built *with* affected communities instead of for them.

As we have previously already stressed,²⁰³ the humanitarian sector is struggling to build actionable AAP approaches from the individual recipient's side, as aid group's policies are often omitting legal pathways for compliance and remedy. Public use technologies are subject to different — often higher — compliance and legal criteria than most private use technologies, including by nonprofits. For example, in 2020, the Dutch court of The Hague ordered the immediate halt of SyRI²⁰⁴ — an algorithmic system cross-referencing personal data from citizens in various databases to prevent and fight fraud — because it violated article 8 of the *European Convention on Human Rights* (ECHR), which protects the right to respect for private and family life.²⁰⁵ In other cases, the use of algorithmic decision-making systems has been found to be in violation of due process law, in particular the use of allocative algorithms to determine Medicare and Medicaid benefits in the US.²⁰⁶ **Despite still relying mainly on public or institutional funding (with a handful of notable exceptions), most humanitarian organisations keep dodging any commitment to open procurement and transparency on their digital and data stack practices, and are still to define which avenues for recourse and remedy might be available to downstream partners and affected individuals.**

As argued by Sandvik,²⁰⁷ many challenges in humanitarian AI arise from the complex intersection of policy, programming, protection, and digital transformation. Solving the “humanitarian AI dilemma”²⁰⁸

²⁰¹ See for example the *Inclusive AI Group*. Utrecht University. (2025, January 20).

<https://www.uu.nl/en/organisation/centre-for-global-challenges/research/inclusive-ai>

²⁰² Matimba, V. (2025, October 31). *Launching our Made in Africa AI for Merl Landscape Study*. MERL Tech.

<https://merltech.org/africa-ai-merl-landscape-study/>

²⁰³ Coppi (n. 1).

²⁰⁴ *Case number C-09-550982-HA ZA 18-388*. The Hague District Court. (2020, March 6).

<https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI%3ANL%3ARBDHA%3A2020%3A1878>

²⁰⁵ *How Dutch activists got an invasive fraud detection algorithm banned*. AlgorithmWatch. (2020, April 6).

<https://algorithmwatch.org/en/syri-netherlands-algorithm/>

²⁰⁶ Sloane, M., Chowdhury, R., Havens, J. C., Lazovich, T., & Rincon Alba, L. (2021, June 28). *AI and procurement - A Primer*. NYU. <https://doi.org/10.17609/bxzf-df18>

²⁰⁷ Sandvik, K. B. (2024). *Framing humanitarian AI conversations: What do we talk about when we talk about ethics?* Peace Research Institute Oslo (PRIO). <https://www.prio.org/publications/14207>

²⁰⁸ DÜchting (n. 19).

would then require defining why algorithmic technology is used, by whom, and for whom, just as much as assessing its technical performance. In our interviews, experts flagged that protection considerations should take precedence over technical results, departing from the current commercial approach. They also stressed the need for normalizing the rejection of algorithmic solutions incompatible with the ‘do no harm’ imperative even if technically feasible or effective. But the inputs received in our research adds another key element on top of that list: agency by local communities and actors. There is an urgent need to consider what people and their local groups and communities can concretely do to decide or influence which algorithmic technology is used, and how to act against it, possibly through judicial or paralegal mechanisms.

2.3. Procurement is becoming heavier, not smarter nor safer

Some responders and most published guidance on the topic of humanitarian AI call for the development of AI-ready internal governance models and processes, often building on past experiences introducing considerations on legal compliance,²⁰⁹ or on rights and ethics-driven issues such as environmental and sustainability requirements into standard operating procedures for procurement.

For example, European Union (EU) guidance defines the preliminary eligibility vetting based on qualifying criteria, before any financial consideration, as “one of the most powerful methods”.²¹⁰ This approach requires the definition of appropriate selection criteria and their inclusion in the technical specifications of the items, followed by adequate publicity to ensure awareness among bidders. Such a process is purported to allow the selection committee to ensure that all qualifying requirements and technical specifications are met and that the bidder is eligible.

In this way, EU guidance for humanitarian procurement aims to ensure that the “price competition happens between sustainable items rather than between sustainable and unsustainable, the latter of which might be cheaper.”²¹¹ The environmental lens of the original model is very narrow and considers only the revised technical specifications of physical products, by focusing on the impacts of packaging, transport/location of production facilities, the materials used and the processes of production, and finally the product’s end of life management.

However, experts consider that, by adjusting the environmental, human rights, and humanitarian criteria in the assessment framework, this process could be incorporated in a broader lifecycle analysis

²⁰⁹ See for example *Litig releases Legal IT AI Due Diligence Questionnaire*. Litig - Legal IT Innovators Group. (2024, May 16). <https://www.litig.org/news/l4zy6wkxepan9gvajgfgeotlsw5opn>

²¹⁰ European Commission Directorate-General for European Civil Protection and Humanitarian Aid Operations (DG ECHO). *Guidance on the operationalization of the minimum environmental requirements and recommendations for EU-funded humanitarian aid operations*. (P. 33). Publications Office of the European Union. (2022). <https://data.europa.eu/doi/10.2795/467817>

²¹¹ *Ibid.*

exercise inclusive of digital systems. **Pre-pricing vetting might then include considerations such as the impact of servers and data centres on the environment and surrounding communities, human rights due diligence and compliance for the full upstream and downstream supply chain, and respect for core humanitarian tenets and principles, including company policies and practices for ensuring their systems aren't used by actors suspected of international humanitarian law violations or atrocities.**

While such an approach might be appropriate to integrate into the lifecycle monitoring for general digital systems, experts remain sceptical about introducing new tech-specific tools, and consider building AI-dedicated internal processes to be a rather pointless exercise. A reason for this lies in the extreme difficulty associated with identifying specific mandatory standards for human rights or humanitarian compliance, which has led the UN WG to suggest embedding such considerations into existing assessments and frameworks.²¹²

A similar reasoning drives a diffused scepticism towards standalone AI assurance processes — “a set of practices that measure, evaluate and communicate the trustworthiness of AI systems, such as AI or algorithm audits, red teaming, conformity assessments, or impact assessments.”²¹³ The reality is that **all evidence points in the direction of algorithmic systems skirting not only new dedicated assessments, but also well-established procurement processes, by blending into existing ICT contracts**, as shown in case study 6.

CASE STUDY 6: ERIN, TARA, AND COPILOT ENTER THE IRISH DEPARTMENT OF JUSTICE

In December 2025, human rights activists flagged²¹⁴ that the Irish Department of Justice had deployed up to three chatbots, all of them misleading users²¹⁵ about sensitive legal and administrative procedures, and all without running a single procurement process. A first chatbot named ‘Erin’ — now discontinued — answered queries from asylum seekers. A second one, ‘Tara,’ responded to citizenship queries. While the

²¹² Working Group on the issue of human rights and transnational corporations and other business enterprises. (2025, May 15). *A/HRC/59/53: Artificial Intelligence Procurement and deployment: Ensuring alignment with the guiding principles on business and human rights - report of the Working Group on the issue of human rights and transnational corporations and other business enterprises*. United Nations.

<https://www.ohchr.org/en/documents/thematic-reports/ahrc5953-artificial-intelligence-procurement-and-deployment-ensuring>

²¹³ Winecoff et al (n. 172).

²¹⁴ Shrishak, K. (2025, November 19). *Department of Justice chatbots mislead people seeking information*. Irish Council for Civil Liberties.

<https://www.iccl.ie/news/irish-department-of-justice-chatbots-mislead-people-seeking-information/>

²¹⁵ Malekmian, S. (2025, March 19). *For “urgent” questions, the Department of Justice directs asylum seekers to a chatbot. Is it trustworthy?* Dublin Inquirer.

<https://www.dublininquirer.com/for-urgent-questions-the-department-of-justice-directs-asylum-seekers-to-a-chatbot-is-it-trustworthy/>

institutional website included a disclaimer that neither chatbot might be accurate, asylum seekers had no choice but to use it because automated email responses from the department pointed them in that direction, while the phone line had been closed since August 2024. Both ‘Tara’ and ‘Erin’ resulted from a special pre-commercial innovation process, but should have required a proper tender to be adopted for official use. Regarding Microsoft Copilot, Ireland’s Minister for Justice, Jim O’Callaghan, declared²¹⁶ that the chatbot was included in a larger tender process covering “the overall cloud based customer service solution”. As a consequence, not only did the AI system not undergo any specific procurement process, but its cost was also blended, and thus hidden, within the general agreement. Activists claim that this approach leads to vendor lock-in, putting other companies that provide comparable solutions without being cloud service providers, out of the running. In addition to lacking proper procurement documentation, the department was also unable to provide the human rights group investigating the issue with a data protection impact assessment or environmental impact assessment. Minister O’Callaghan declared that the deployment of Copilot was preceded by an internal assurance process, but when external experts replicated said process²¹⁷ (an internal testing on a set of customer service team members’ questions), the outputs were mostly erroneous or misleading.

Experts also question the usefulness of adding more burden onto already overstretched and underfunded procurement teams, in a process that is not guaranteed to entirely mitigate the main risks it is supposed to prevent. In their view, multiplying the assessments, forms, and processes would bring few concrete benefits while stifling initiatives to experiment and innovate. Others have directly questioned the capacity of humanitarian organisations, already dependent on Big Tech providers, to be able to obtain this type of information or do anything useful with it.

On the technical side, it has been flagged that the dynamic nature of algorithmic models and cloud based systems would quickly outpace any nonprofit organisation’s capacity to maintain an up-to-date and effective vetting or assessment process to select which algorithmic solution to use and how. This recalls the frustrating experience of academic researchers who often see their work questioning version systems that are already outdated or even discontinued by the time their article is published.

Procurement-stage technical vetting and assessment of all prospective technologies might also prove unrealistic if an organisation aims to use digital technologies at scale: algorithmic functionalities are increasingly enmeshed in complex digital systems with long supply and value chains possibly involving several providers, some of them unknown to the client. Along the same lines, respondents also pointed at the speed rate for introducing new algorithmic tools and newly algorithmic enhanced systems to suggest that there may soon be no option about whether or not to use an algorithmic model-enhanced product.

²¹⁶ Houses of the Oireachtas. (2025, November 4). *Dáil Éireann Debate, Tuesday - 4 November 2025*. Artificial Intelligence. https://www.oireachtas.ie/en/debates/question/2025-11-04/1255/#pq_1255

²¹⁷ Shrishak (n. 214).

This calls into question the feasibility of applying the many decision trees, toolboxes, and guidance whose objective is to support organisations in assessing whether AI-like capabilities might be needed for a specific function. If the global trend is to be trusted, in an environment where the very concept of AI cannot be properly defined and where, AI-free tools are possibly going extinct in some technical areas, continuing to add further extra steps to the procurement process would not only be impractical and cumbersome, but also performative and ultimately ineffective.

Some respondents have suggested considering this phenomenon pragmatically. Apart from the emerging trend of localised partnerships or non-Big Tech algorithmic development, it is very unlikely in the short term that large humanitarian actors will escape the corporate capture spiral they have been caught in for the past decade. Despite all the studies, guidance, and vetting adopted or promoted by these aid groups and their donors, their access to AI and algorithmic tools is likely to remain heavily mediated or enabled by a handful of tech actors. **Pragmatists seem to converge in recommending that humanitarians focus their risk management efforts on securing this core environment, instead of raising impossibly high access barriers for emerging work with innovation-focused partners.**

More broadly, some of the experts we consulted see this emerging discussion on digital governance and policies as a welcome and much-needed moment in the humanitarian space, regardless of the measures' concrete capacity to solve the problems attributed to algorithmic tools. In their view, **the ongoing obsession to frame everything as AI-centred is creating a useful pretext to pressure their own leadership to care about data governance, and forcing the whole sector to rethink its approach to digital and procurement.**

This might start with halting the race-to-the-bottom in endlessly expanding the amount of paperwork and extra steps involved in the procurement of digital systems, to focus instead on a key question: if procurement staff and tools are becoming a minority component in a lengthy cross-functional process, is it still purely a traditional procurement issue?

2.4. Beyond bells and whistles: from product price-based selection to continuous environmental scanning

All evidence points to the need for the aid sector to stop playing catch up and instead adopt a dynamic and forward-looking perspective with regard to algorithmic tools, and digital architecture in general. Doing so starts from the assumption that the challenges explored in this report apply to the whole stack of algorithm-enhanced systems and cloud-based technologies, especially when hosted and managed by tech providers running key models and systems. In the corporate sector, the same realization led to the creation of dedicated services offering AI governance

platforms, such as VerifyWise²¹⁸ or Modulos,²¹⁹ often including due diligence in AI procurement²²⁰ within their offering.

Firstly, this requires rethinking the way that tech contracts and partnerships are formed and managed in the humanitarian sector. As mentioned, the ‘ethical buyer’ approach seems to not be succeeding in steering technology development in the right direction. Besides that, it is also uncertain whether the current competitive bidding process and vendor evaluations are structured in a way that makes it possible to even detect any eventual misalignments between aid groups and vendors goals. What is certain, is that they do not seem to offer any safeguard against future critical divergences, or any guarantee of carefully assessing vendor risk in a highly dynamic environment.

Just as the continual auditing of AI systems²²¹ is crucial to prevent or detect flawed models, continuous monitoring of tech companies is essential to make sure the services provided are aligned with the original purpose of the partnership. In particular, while some humanitarian actors already have screening mechanisms in place to make sure companies and their leadership aren’t directly or indirectly managed or controlled by defence actors, these mechanisms do not account for the specific governance, business, and service models of digital tech companies. In particular, vendors’ assessments do not commonly include upstream and downstream supply chain relationships, tech dependencies, or their investors and shareholders’ portfolio composition.

The rise of dynamic digital commercial services on the frontlines of conflict or crisis has led to an emergence of ‘bait and switch’ approaches, a particularly insidious form of “aidwashing”²²² aimed at performing a market grab. In this scenario, ambitious digital tech actors aim to move in a sensitive emergency context, exploiting NGOs and humanitarian actors’ capacity to cut corners on local regulation in order to obtain exemptions intended purely for humanitarian causes, only to launch a commercial service — or even a defence or security vertical — once the system has gained a solid-enough local foothold.

CASE STUDY 7: CLEARVIEW AI

The inherent tendency towards facilitating sprawling surveillance by algorithmic systems is plain to see in the case of Clearview AI and how the company has angled to pursue business, humanitarian, and military objectives simultaneously. In Ukraine, Clearview AI initially offered the government free access

²¹⁸ *AI Governance Platform: Enterprise AI compliance*. VerifyWise. (n.d.). <https://verifywise.ai/>

²¹⁹ *AI Governance & EU AI act compliance platform*. Modulos. (n.d.). <https://www.modulos.ai/>

²²⁰ *Due diligence in AI procurement: AI Governance Lexicon*. VerifyWise. (n.d.-b). <https://verifywise.ai/lexicon/due-diligence-in-ai-procurement>

²²¹ Tyrie (n. 148)

²²² Martin (n. 157).

to its facial recognition system in order to identify fallen Russian soldiers,²²³ help to reunite refugees with their families, identify war victims, and counter Russian propaganda.²²⁴ However, to identify a person, their photo must be uploaded to Clearview's biometric database to find a match. The database contains about 30 billion images previously harvested from online social networks such as Facebook, Twitter, and Vkontakte, and search engines such as Google.²²⁵ While its 'humanitarian' contribution has remained limited, Clearview AI is actively selling this data to different authorities, mainly law enforcement and military departments.²²⁶ In Ukraine, this is used "to detect infiltrators at checkpoints, process citizens who lost their IDs [and] identify and prosecute members of pro-Russia militias and Ukrainian collaborators". This approach raises concerns due to its potential for mass surveillance, misidentification, and privacy violations and because its unregulated use threatens fundamental rights, as shown by several legal actions²²⁷ facing Clearview AI breaching data protection laws in different countries. While the company's attempt to aidwash its product was ultimately unsuccessful, as no aid group was actively involved in this specific project,²²⁸ this type of scenario represents an open challenge for the humanitarian sector. As documented in 2024²²⁹ and in this report, there are myriad examples of how "reputational damage can occur if a private sector partner has been linked to human rights violations in a previous project or is seen as whitewashing by working with a humanitarian partner."²³⁰

As highlighted by the UN WG,²³¹ businesses have only just started acknowledging the potential or current adverse effect of their AI products. Nor is the humanitarian sector, which has barely begun documenting digital harms in general, pushing the corporate sector to define and categorise algorithmic-related harms. And when humanitarian and private actors strike a partnership, there is often no common understanding around the distribution of data management responsibilities.²³²

²²³ Hill, K. (2022, April 7). *Clearview AI used in Ukraine to identify deceased soldiers, despite concerns about use of facial recognition technology during conflict*. Business and Human Rights Centre.

<https://www.business-humanrights.org/en/latest-news/clearview-ai-used-in-ukraine-to-identify-deceased-soldiers-despite-concerns-about-use-of-facial-recognition-technology-during-conflict/>

²²⁴ Mysyshyn, A. (2024, September 30). *Advanced technologies in the war in Ukraine: Risks for democracy and human rights*.

<https://www.gmfus.org/news/advanced-technologies-war-ukraine-risks-democracy-and-human-rights>

²²⁵ Ibid.

²²⁶ Coppi (n. 1).

²²⁷ Mysyshyn (n. 224).

²²⁸ Kurpiewska-Korbut, R. (2024). *Digital Technologies in the Global Humanitarian Sector: A Case Study of Ukraine*. *Journal of Modern Science*, 60(6), 730–743. <https://doi.org/10.13166/jms/199490>

²²⁹ Coppi (n. 1).

²³⁰ *Data responsibility in public-private partnerships*. The Centre for Humanitarian Data. (2020, January).

<https://centre.humdata.org/guidance-note-data-responsibility-in-public-private-partnerships/>

²³¹ Working Group on the issue of human rights and transnational corporations and other business enterprises. (n. 212).

²³² The Centre for Humanitarian Data (n. 230).

Among the many gaps in our existing evidence base on algorithmic technologies, a full mapping of these systems inclusive of their impacts and related requirements for data responsibility with support from a multiple range of humanitarian stakeholders definitely tops the list.

The fact that companies themselves lack a clear understanding of how the systems being offered — or any future changes to them — might result in harm to people²³³ undermines the previous trust-based approach adopted by most humanitarian organisations when building tech relationships with partners or vendors. From this perspective, fiduciary mitigation measures typically adopted by the aid sector need to shift away from providers' capabilities to deliver a quality product. Instead, aid actors should assess each company's record and their ability to effectively ensure the compliance of their product development and supply chain with the agreed set of commercial clauses and principles. Considering that trust is often exploited in supply chain attacks,²³⁴ humanitarians must move past the naive concept of trusted tech provider and rather expand their procurement governance to encompass the concept of zero trust architecture²³⁵ typical of hostile cyber environments.

Aid groups should prioritise providers based on their capability and intention to work with the client on building adequate monitoring and preventative measures to avoid unilateral or unexpected transformations in or across their systems. We might tentatively define this as an emerging form of expectation for constant due diligence falling on the provider for future performances of any system which is either experimental or unpredictable by design, or susceptible to becoming one. Algorithmic errors may be hard to detect due to a lack of organisational or technical transparency or to specific digital literacy skills. These errors may have cascading effects as algorithmic systems interact with other systems within and beyond the units or even organisations that deploy them.²³⁶

Secondly, **while function creep was previously decried as a pathology of digital transformation, it is now a core feature of most commercially available algorithmic systems, especially those commonly identified as AI.**²³⁷ A growing body of evidence points to the existence of a clear pipeline leading most research work on algorithmic systems towards surveillance applications.²³⁸

²³³ Landymore, F. (2025, November 13). *AI-powered toys caught telling 5-year-olds how to find knives and start fires with matches*. Futurism. <https://futurism.com/artificial-intelligence/ai-toys-danger>

²³⁴ Aarland, M. (2024). *A resilience framework for managing cyber risks in digital supply chains in the Norwegian power industry*. Nasjonalbiblioteket. <https://www.nb.no/items/317aedcb29d4c6e9176a8f97f92dca42?page=0>

²³⁵ Zero trust (ZT) is the term for an evolving set of cybersecurity paradigms that move defences from static, network-based perimeters to focus on users, assets, and resources. See Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture*. National Institute of Standards and Technology. <https://doi.org/10.6028/nist.sp.800-207>

²³⁶ Sloane et al (n. 206).

²³⁷ *A new AI lexicon: Function creep*. AI Now Institute. (2025, April 29).

<https://ainowinstitute.org/publications/collection/a-new-ai-lexicon-function-creep>

²³⁸ Kalluri, P. R., Agnew, W., Cheng, M., Owens, K., Soldaini, L., & Birhane, A. (2023, October 17). *The surveillance AI pipeline*. arXiv.org. <https://arxiv.org/abs/2309.15084>

Purpose-specific software design, data minimisation, data anonymisation, and other core tenets of responsible humanitarian data management²³⁹ are antithetical to the massive data harvesting needs and unpredictable nature of most AI products.²⁴⁰ Whenever they are not deliberately designed and implemented to keep a narrow scope within a contained environment, they tend to expand, change, and grow.

When dealing with similar systems, be they new or newly enhanced, an ethical and rights-based procurement process must assess how they fit within the broader tech stack, the dependencies they might establish, or the possible future functionalities that might impact how the overall digital ecosystem operates and its compliance with laws and policies. As public procurement experts have observed, “AI is not just a tool, it acts as an infrastructure for future action, with long-term societal and environmental consequences.”²⁴¹

Thirdly, as noted in our 2024 report²⁴² and confirmed by researchers and journalists, Big Tech giants and AI startups are in a race to scoop up elusive real-world data, “forging alliances across industries and regions to help gather real-world data that can’t be scraped from the internet.”²⁴³ This, of course, does not respond to a humanitarian imperative but rather frames a product value proposition that relies on these companies’ ability to obtain constant streams of training signals unavailable in the existent public data environment. Continuous access to rare and diverse datasets to keep training their algorithmic models is key for keeping an edge in their hyper-customised and hyper-personalised offerings²⁴⁴ on one side, and the unlikely quest for the holy grail of artificial general intelligence on the other.²⁴⁵

In this context, humanitarian entities purchasing, licensing, or building an algorithmic system stop being a purely occasional actor engaged in a transactional negotiation on a service agreement. **In the dynamic field of algorithmic models, roles are often determined by obscure concepts and metrics such as the amount of cloud compute deployed for a model.**²⁴⁶ **Aid groups and their partners**

²³⁹ *Humanitarian Data Ethics*. The Centre for Humanitarian Data. (2020, January).

<https://centre.humdata.org/guidance-note-humanitarian-data-ethics/>

²⁴⁰ Coppi, Moreno Jimenez, & Kyriazi. (n. 149).

²⁴¹ *Principles for public participation in AI procurement*. Connected by Data. (n.d.).

<https://connectedbydata.org/projects/2025-procurement-principles>

²⁴² Coppi (n. 1).

²⁴³ Sharma Punit, I. (2025, August 21). *AI giants race to scoop up elusive real-world data*. Rest of World.

<https://restofworld.org/2025/ai-data-collection-global-deals/>

²⁴⁴ Ibid.

²⁴⁵ Sarikaya, R. (2025). Path to artificial general intelligence: Past, present, and future. *Annual Reviews in Control*, 60, 101021. <https://doi.org/10.1016/j.arcontrol.2025.101021>

²⁴⁶ *Providers of general-purpose AI models - what we know about who will qualify*. EU Artificial Intelligence Act. (2025, April 25).

<https://artificialintelligenceact.eu/providers-of-general-purpose-ai-models-what-we-know-about-who-will-qualify/>

partaking in a lucrative data supply chain mechanism might inadvertently qualify as model modifiers or providers,²⁴⁷ data custodians, controllers,²⁴⁸ and possibly data brokers,²⁴⁹ resulting in obligations to run complex upstream and downstream due diligence exercises. To avoid partaking in humanitarian extractivism, aid organisations need to transform not only the way they work, but also be deliberate about the way they see (and position) themselves.²⁵⁰

Even if algorithmic technology can no longer be seen only as a commodity,²⁵¹ most humanitarian procurement protocols and policies still consider it as such. As highlighted in the NYU primer on AI and procurement²⁵² and prevalent literature, any design choices that vendors make about an algorithmic system and even the framing of the algorithm’s decision “all encode policy choices”.²⁵³ Thus, **if the procurement of AI is treated “with the same off-the-shelf purchasing philosophy as other IT systems, the result is policy making by third party design, ultimately resulting in a lack of (...) accountability for these decisions.”**

There are additional reasons to conclude that the price-based approach to digital procurement needs to be reformed. Considering the dynamic nature of algorithmic systems, their constantly changing parameters, pricing, and even core functionalities, and their impact on an increasing number of connected processes, is it still possible to identify a clear distinction between the ‘pre-contract and ‘contract execution’ phase in the algorithmic procurement pipeline? In the past, innovative approaches to major procurement strategies for infrastructure projects in the public sector tried to tackle this challenge by experimenting with pre-contract uncertainty reduction.²⁵⁴ Is this approach still reasonable in a risk environment that remains highly dynamic even long after a product has entered the execution stage?

Instead, **our research highlights how uncertainty should now be considered an inherent and persistent risk factor throughout the contract lifecycle. Considering their peculiar nature, digital systems should rather be framed as highly dynamic contractual processes where every change to a digital product involving an algorithmic tool or function might revive the procurement process and possibly open the way to a reassessment of the appropriateness, sustainability, and responsibility criteria.** This also clearly impacts the governance of the procurement process, and of

²⁴⁷ Ibid.

²⁴⁸ Chapter 4 – Controller and Processor. General Data Protection Regulation (GDPR). (2018, October 5). <https://gdpr-info.eu/chapter-4/>

²⁴⁹ Firoz, M. (2024). Quantifying vulnerability: Humanitarian Datafication and the Neophilia of Integrated Power. *Cultural Anthropology*, 39(3). <https://doi.org/10.14506/ca39.3.02>

²⁵⁰ Sandvik (n. 152).

²⁵¹ ICRC Technology strategy. ICRC. (2024). <https://shop.icrc.org/icrc-technology-strategy-summary-pdf-en.html>

²⁵² Sloane et al (n. 206).

²⁵³ Mulligan, D. K., & Bamberger, K. A. (2019). *Procurement as policy: Administrative Process for Machine Learning*. Berkeley Technology Law Journal. <https://doi.org/10.15779/Z38RN30793>

²⁵⁴ Procurement strategy in major infrastructure projects. (2021). *OECD Public Governance Policy Papers*, (06). <https://doi.org/10.1787/38996343-en>

the associated partnerships-building models. In a way, this reflects the need for what the Inter-Agency Procurement Group (IAPG) has called the need “for a shift from compliance-focused to strategic procurement.”²⁵⁵

Algorithmic systems degrade reliable, trust-based, private-public relationships as they make any human rights impact assessment possibly outdated upon any product update. Instead, constant monitoring and evaluation of all algorithmic systems’ behaviour and performances now seems necessary. Some emerging theories seem particularly well placed to help in restructuring traditional procurement approaches. Some of these find their origin in social and behavioural sciences, such as the dynamic capabilities theory with their focus on continuous environmental scanning.²⁵⁶ **Others stem from cybersecurity and require us to expand our focus from the mere vendor assessment to include cyber risks across the digital supply chains.**²⁵⁷ **This is where our analysis identifies the stronger potential for reform.**

2.5. Bringing the zero trust environment in procurement

Just as procurement is traditionally seen as a static exercise for purchasing assets, the traditional view of cybersecurity in the humanitarian sector is limited to operational areas such as account management, internal system control, firewalls, or threat detection systems. While maintaining this approach might still make sense for non-digital items, bridging the gap between digital procurement and cybersecurity is essential to ensure responsible digital transformation in the age of algorithmic systems. The need to add a cybersecurity lens to procurement becomes even more evident if we consider that Gartner’s top 10 strategic technology trends for 2025 identified cloud, AI and algorithmic systems, and supply chain interdependencies as the main factors expanding an organisation's attack surface. Analysts notably suggested reframing the role of “procurement, security architects, security engineers, and other stakeholders”.²⁵⁸

Our analysis points at the need to develop a sector-wide and inclusive resilience framework for managing cyber risks in the humanitarian digital supply chain, possibly taking inspiration from other industries providing core civilian services.²⁵⁹ Cyber risk management in digital supply chains involves “identifying, assessing, and mitigating potential cyber threats that could disrupt the supply of

²⁵⁵ IAPG Position Paper on Improving Humanitarian Procurement Efficiency. Inter-Agency Procurement Group. (2025, June). <https://iapg.org.uk/position-paper/>

²⁵⁶ Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), 1319–1350. <https://doi.org/10.1002/smj.640>

²⁵⁷ The term “digital supply chain” is intended in its broader meaning, including other variations as explained by Aarland (n. 234).

²⁵⁸ Gartner identifies the top cybersecurity trends for 2025. Gartner. (2025, March 3). <https://www.gartner.com/en/newsroom/press-releases/2025-03-03-gartner-identifiesthe-top-cybersecurity-trends-for-2025>

²⁵⁹ Aarland (n. 234).

services or products.”²⁶⁰ The constantly growing degree of interconnectivity among organisations means that cyber-attacks or harmful events often originate from partners within the supply chain,²⁶¹ and possibly from entities whose existence might be completely unknown to the humanitarian actor running a procurement exercise.

This new reality requires holistic approaches beyond any single organisation line of sight. The speed, variety, and virality of some threats, including spyware or unpredictable changes in algorithmic tools behaviour, makes it so that previously trusted actors within the organisation network or authentication gates might be flagged as compromised sometimes weeks or months after a problem has occurred. This turns trust in algorithmic systems inside any organisational ICT environment into a liability.

More broadly, **this essential process faces specific challenges due to a variety of factors including digital supply chains’ inherent unpredictability, fluctuation over time, and transnational dependencies.**²⁶² The unpredictability makes it so that errors can appear anywhere across a very complex network of systems, and propagate in ways that might generate disproportionate effects on seemingly unrelated functions. The constant fluctuation over time undermines the static vendor vetting process. Even if an NGO identifies and evaluates all suppliers and sub-suppliers at the time of procurement, technical, financial or legal events that occur during the lifecycle of the contractual relationship might result in different or additional unknown suppliers and sub-suppliers being introduced into the digital supply chain. Finally, digital supply chains invariably create transnational dependencies, since suppliers are likely to fall under different jurisdictions. For example, an NGO using biometric systems in Myanmar may rely on digital services such as cloud storage from the US and scanners produced in China, while their software license may be stored on a server outside the country, potentially creating a series of critical cross-border liabilities.²⁶³

Given so much is out of a single organisation’s exclusive control, experts suggest that a wholesale redesign of existing standards, forms, and procedures is needed, beyond just the simple integration of AI-specific tool evaluations or checklists. By working as one instead of just coming together as a cross-functional team on a case-by-case basis, procurement, ICT, legal, humanitarian policy, and cybersecurity departments can become an integrated business function, essential for preventing and eventually responding to cyber threats or digital risks originating from within the network or the broader supply chain. Even in a less resourced and distributed organisation, this integrated team could better assist across the full spectrum of

²⁶⁰ Ibid.

²⁶¹ Pergande, R. (2025). Bridging Procurement and cybersecurity: A systematic literature review and research agenda. *Advanced Studies in Supply Management*, 123–154. https://doi.org/10.1007/978-3-658-48750-8_7

²⁶² Aarland (n. 234).

²⁶³ Dias, T. (2025, September 16). *AI Agents & Global Governance*. Partnership on AI. <https://partnershiponai.org/resource/ai-agents-global-governance-analyzing-foundational-legal-policy-and-accountability-tools/>

opportunities, challenges, and risks that might potentially reverberate across the whole organisation.

This integrated team would need to develop new approaches beyond policymaking and capacity building; for example, by introducing procurement-led horizon scanning exercises to detect trends and changes across the digital supply chain and to spot anomalies and deviances in the internal systems' signature or behaviour. Ideally, this would happen on a rolling basis, starting well before the traditional pre-contracting phase of any digital procurement process, to maximise comparative data and detection capabilities. **This approach would be key to securing organisations and supply chains through the design and deployment of the aforementioned zero trust digital architecture²⁶⁴ in procurement or, less utopically, the trust-but-verify²⁶⁵ principle.**

A few experts downplayed the need for systemic change in the face of algorithmic growth and adopted a cynical approach to this problem: if an NGO can't secure their ICT environment and protect their data, it should just simply abstain from using digital tools for their work. This radical position — often verbalised by international experts while referring to local NGOs — is hardly feasible in practice, and most of all, is ethically tainted. Even when removing the neocolonial undertones from the core argument and disregarding the resulting humanitarian impacts of excluding most of frontline aid from algorithmic systems, such an argument simply cannot be translated into policy.

How would an entity frame a set of indicators to define when to stop using algorithmic tools? Some humanitarian experts point to an even more serious flaw: in this age when most application systems are transitioning to integrate some forms of algorithmic functionalities, the only ethical decision for any organisation without a solid IT department would be to walk away from digital technology. Based on this reasoning, given the pace of technology, the overwhelming majority of the humanitarian system would have no other choice but to give up on digital programming and cloud systems altogether, and rewind to the analogue era.

But leaving behind fiduciary relationships with vendors and their machines doesn't have to mean letting go of trust (or technology) altogether. Experts indicate that the AI hype is breaking the illusion of the friendly techno-optimist private partner, and opening up opportunities for rebuilding trust among aid groups, and between them and local communities. In order to do so, we have highlighted some questions for the sector to consider, in order to change how it works:

- Can the humanitarian sector come together to push for open procurement policies and resist non-disclosure agreements (NDAs)?

²⁶⁴ Rose et al (n. 235).

²⁶⁵ Pace, T. (2024, September 12). *Trust but verify*. Security Today.
<https://securitytoday.com/articles/2024/09/12/trust-but-verify.aspx>

- Can aid sector and tech companies agree on standard mechanisms and metrics to verify if future system upgrades are managed with respect to data security and responsibility principles?
- Can procurement and legal experts develop adequate contractual terms for enforcing such clauses?
- Can parties to tech agreements and NGOs agree on new approaches for continual auditing of AI systems,²⁶⁶ in order to detect unintended consequences or biases infiltrating models trained on messy real-world data, or to spot unruly or unknown AI agents?²⁶⁷
- Is there space and capacity to negotiate more responsible, tailored solutions with providers? Can these terms be made publicly available?
- Can risk mapping and assessment become an exercise spanning across the value chain through cross-sectorial cooperation, reducing the burden on individual companies or actors?²⁶⁸

None of this is doable without strengthening sector-wide trust, cohesion, coordination, and solidarity.

3. Suggested foundations for a digital service framework in the algorithmic age (see Appendix)

Our research confirms the impossibility of developing one-size-fits-all guidance to the challenges facing humanitarian organisations in the era of cloud and AI technologies. The slightest variation within the broad spectrum of governance and compliance options with the humanitarian principles and human rights norms have significant impacts in the way aid groups and affected communities experience, engage with, and suffer from algorithmic systems.

Furthermore, some of the key decisions taken by aid groups on core ICT systems are often influenced by the principle of resource scarcity, and are strictly dependent on the leadership team's current priorities. Building a standard checklist is likely to result in a rubber-stamping exercise by the team required to run a process in which they have limited leeway. **The endless stream of guidance tools seems to add little practical added value, while putting additional burden on both humanitarian teams and vendors when trying to build a partnership.**

²⁶⁶ Tyrie (n. 148).

²⁶⁷ Boulanin, V., Blanchard, A., & Lopes da Silva, D. (2025, October 1). *Before it's too late: Why a world of interacting AI agents demands new safeguards*. SIPRI.
<https://www.sipri.org/commentary/essay/2025/its-too-late-why-world-interacting-ai-agents-demands-new-safe-guards>

²⁶⁸ Working Group on the issue of human rights and transnational corporations and other business enterprises. (n. 212).

Also, as repeatedly mentioned, focusing on developing solutions to the ‘AI problem’ — while useful in the early days of reckoning with these technologies — does not seem to be the best long-term investment in times when resources are stretched thin across the board. Algorithmic systems — including those currently associated with AI functionalities — are very likely to blend and seep into the majority of commercial digital tools, and become inextricably fused with their main application. Once the AI hype has passed and users stop caring about the AI label²⁶⁹ there may be no public notice nor identifiable marker for detection, unless and only to the extent expressly required by applicable laws.

Instead, our research shows that any digital tool might, at some point, potentially become or integrate an algorithmic functionality or component, or even present agentic behaviour or characteristics involving a higher degree of autonomy,²⁷⁰ which may be invisible to humanitarian organisations but is not without danger.²⁷¹ **We have identified a set of core questions to help aid groups question the very basis of their procurement approach, and identify ways to reform how we build a digital ecosystem. This stems from one essential premise: what would we want our digital governance to look like, if we were building our organisation from scratch?** The main objective is to move away from treating tech products as commodities and — recalling the words of the UN WG — make sure that “efforts shift from regulating machines to protecting people”.²⁷²

The guiding schema proposed in this report (see Appendix) is a humanitarian digital service framework; a strategic blueprint for planning and managing risk-driven plans considerate of future algorithmic developments and their impacts on people and systems. This foundational framework is designed to be resources, scale, and tech-agnostic. It can support immediate change, but also inform further research by operational and academic partners involved in humanitarian logistics and supply chain innovation, and internal reflection by aid groups, digital transformation teams, and their boards.

The framework is an adaptation of the following existing procurement guidance models relevant for algorithmic environments:

- Civic Machines Lab’s whitepaper on AI adoption across mission-driven organisations;²⁷³
- The Center for Inclusive Change’s risk management framework for AI procurement;²⁷⁴

²⁶⁹ James, D. (2026, January 6). *Dell seems to be the first to realise we don’t actually care about AI PCs*. PC Gamer. <https://www.pcgamer.com/hardware/dells-ces-2026-chat-was-the-most-pleasingly-un-ai-briefing-ive-had-in-mybe-5-years/>

²⁷⁰ Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). AI agents vs. Agentic AI: A conceptual taxonomy, applications and challenges. *Information Fusion*, 126, 103599. <https://doi.org/10.1016/j.inffus.2025.103599>

²⁷¹ Boulanin, Blanchard, & Lopes da Silva (n. 267).

²⁷² Working Group on the issue of human rights and transnational corporations and other business enterprises. (n. 212).

²⁷³ Ahmed et al.(n. 16).

²⁷⁴ Miller, C., & Waters, G. (2023). *Risk Management Framework for AI Procurement*. Center for Inclusive Change. <https://www.inclusivechange.org/ai-governance-solutions/rmf-for-ai-procurement>

- MERL Tech’s tool for assessing AI vendors;²⁷⁵
- WEF’s AI procurement in a box tool;²⁷⁶
- WEF’s guidelines for private sector procurement of AI solutions ;²⁷⁷
- New York University’s AI and Procurement primer;²⁷⁸
- The Responsible AI procurement section of the Scottish AI Playbook;²⁷⁹
- UK Local Government Association’s guide on responsibly buying AI;²⁸⁰
- The relevant sections of the NIST cybersecurity standard;²⁸¹ and
- NetHope’s Humanitarian AI code of conduct.²⁸²

We have processed these based on the findings presented above and framed them through a customised version of the adaptive risk management (ARM) principles.

We hope this framework will help humanitarian groups of all kinds and sizes to establish relationships on most algorithmic-based systems, regardless of an AI label, based on shared values and mutual trust with their tech providers, partners, and communities. **To work effectively, these requirements would of course need the implementation of the largest number of recommendations suggested in this research. In the meantime, this framework can already provide a common ground for the whole humanitarian sector to build on, to ensure that the benefits of the digital era are spread in a fair and transparent manner without leaving behind smaller or local organisations with limited negotiating or purchase power.**

The proposed framework also has an operational component which is far from conclusive or final, and may never be. This roadmap offers building blocks that any organisation can modify, customise, and

²⁷⁵ Higdon, G. L. (2025, April). *Tool for assessing AI vendors a resource for decision-makers*. MERL Tech. <https://merltech.org/wp-content/uploads/2025/05/Assessment-Tool--Evaluating-AI-Enabled-Service-Providers-May-2025.pdf>

²⁷⁶ *AI Procurement in a Box*. World Economic Forum. (2020, June 11). <https://www.weforum.org/publications/ai-procurement-in-a-box/>

²⁷⁷ *Adopting AI Responsibly: Guidelines for Procurement of AI by the Private Sector*. World Economic Forum. (2023, June). https://www3.weforum.org/docs/WEF_Adopting_AI_Responsibly_Guidelines_for_Procurement_of_AI_Solutions_by_the_Private_Sector_2023.pdf

²⁷⁸ Sloane et al (n. 206).

²⁷⁹ *Responsible AI procurement*. Scottish AI Playbook. (n.d.). <https://www.scottishaiplaybook.com/responsible-ai-procurement>

²⁸⁰ *Responsibly buying AI*. Local Government Association. (2025, April 16). <https://www.local.gov.uk/publications/responsible-buying-how-build-equality-data-protection-your-ai-commissioning>

²⁸¹ Mahn, A. (2018, October 23). *Identify, protect, detect, respond and recover: The NIST Cybersecurity Framework*. NIST. <https://www.nist.gov/blogs/taking-measure/identify-protect-detect-respond-and-recover-nist-cybersecurity-framework>

²⁸² *NetHope Humanitarian AI Code of Conduct 2024*. NetHope AI Working Group. (2024). <https://nethope.org/toolkits/humanitarian-ai-code-of-conduct/>

develop through an inclusive multistakeholder process. **To ensure this research allows any organisation to take the first steps in the right direction, the language in the framework is presented in a format compatible with immediate integration into training design,²⁸³ job descriptions, and terms of reference templates. It might also offer a stepping stone for academic researchers interested in further shaping an ARM-inspired humanitarian digital service framework for volatile and high-risk environments** by developing the aforementioned dynamic capabilities and microfoundations theories on top of this concept.

4. Recommendations

This section contains practical recommendations aimed at different actors, suggesting ways to implement our research's core findings. To avoid repetition and in the spirit of continuity with our prior analysis, **we invite readers to consider these as building upon the previous recommendations made in our 2024 report.**²⁸⁴

4.1. Recommendations for states and donors

We call on donors to:

- **Establish and support a forum dedicated to reforming digital procurement and developing harmonised and shared guidance on responsible procurement of emerging technologies** in the humanitarian sector, inclusive of evidence base gathering and transparent reporting on progress similarly to the reporting framework of the Hiroshima AI process code of conduct;²⁸⁵
- **Promote and support the transformation of digital procurement into a strategic business function**, and the required upskilling of existing procurement profiles, including through training on humanitarian digital supply chain risks;²⁸⁶
- **Privilege the development of small language models, custom built AI systems, and locally owned and hosted open-source solutions** to better meet the local organisations' needs²⁸⁷ and reduce the risk of vendor lock-in, mission creep of external tools, ethical misalignments due to changing business models and practices, or mishandling of product lifecycle processes by tech vendors;

²⁸³ See for example *EPFLX: Understanding the digital supply chain and its stakes for humanitarian actors*. edX. (n.d.).

<https://www.edx.org/learn/ethics/ecole-polytechnique-federale-de-lausanne-understanding-the-digital-supply-chain-and-its-stakes-for-humanitarian-actors>

²⁸⁴ Coppi (n. 1)

²⁸⁵ Perset, K., & Esposito, S. F. (2025). *How are AI developers managing risks?* OECD Artificial Intelligence Papers. https://www.oecd.org/en/publications/how-are-ai-developers-managing-risks_658c2ad6-en.html

²⁸⁶ edX. (n. 283).

²⁸⁷ Verhulst, S. G., & Sloane, M. (2020, February 7). *Realizing the potential of AI localism*. Project Syndicate. <https://www.project-syndicate.org/commentary/local-regulation-of-artificial-intelligence-uses-by-stefaan-g-verhulst-1-and-mona-sloane-2020-02>

- **Fund foundational modules for AI fairness** such as API access to publicly available data hosting, **and cost-sharing mechanisms with trusted data providers** to allow a broader range of organisations to access audited algorithmic models;
- **Demand that recipient organisations adopt open procurement policies and introduce key requirements such as explainability requirements in contracting,**²⁸⁸ or minimal use of NDAs for tech systems and services, with very limited exceptions;
- **Facilitate and support the development of human rights-centred frameworks as a requirement to inform any cost-benefit analysis of algorithmic systems and tools,** moving to a thorough assessment comparing the enhanced enjoyment of human rights unlocked by this technology with the associated human rights risks throughout the product's lifecycle, its supply chain, and customer base;
- **Support processes aimed including humanitarian organisations in computer emergency response team (CERT) mechanisms** to ensure they are promptly informed in case of emerging vulnerabilities or other threats to their digital integrity, and the establishment of a common humanitarian reporting framework for algorithmic incidents,²⁸⁹ and
- **Reconsider blanket requirements for open data policies as a condition for funding.** In the age of cloud based systems and algorithmic models, such requirements might hamper or pose significant barriers to the sustainability of projects aiming to reform core internal processes such as procurement, legal, accounting, or human resources, or external areas such as health or protection.

We call on regulators to:

- **Strengthen transparency and due diligence requirements across the whole supply chain for all algorithmic tools, with a special requirement for hHRDD+²⁹⁰ and remedy mechanisms** through legal clauses, standardised procurement processes, and mandatory impact assessments, among other things;
- **Harmonise ethical and regulatory instruments covering procurement in the public and nonprofit sector based on a human rights framework to ensure consistency,**²⁹¹
- **Identify uses of algorithmic systems incompatible with respect for human rights,**²⁹² without introducing broad national security or defence carveouts in reporting and implementation mechanisms; and

²⁸⁸ Miller, C., & Waters, G. (2023a). *AI Explainability*. Center for Inclusive Change.

<https://www.inclusivechange.org/ai-governance-solutions/ai-explainability>

²⁸⁹ *Towards a common reporting framework for AI incidents*. OECD. (2025, February 28).

https://www.oecd.org/en/publications/towards-a-common-reporting-framework-for-ai-incidents_f326d4ac-en.html

²⁹⁰ Pachoud & Milatović (n. 71).

²⁹¹ Working Group on the issue of human rights and transnational corporations and other business enterprises. (n. 212).

²⁹² *Ibid.*

- **Address emerging gaps in our existing evidence base on algorithmic technologies**, support a mapping exercise of these systems inclusive of their impacts and related requirements for data responsibility with support from a range of humanitarian stakeholders.

We also invite the the European Commission’s Directorate-Generale for Civil Protection and Humanitarian Aid Operations (ECHO) to consider this report’s findings in the action points following up on workshops run in 2025 on humanitarian procurement,²⁹³ and related implications on preparedness, digitalisation, localisation and environmental sustainability.

4.2. Recommendations for the humanitarian community

We call on the humanitarian community to:

- **Rethink their governance model, moving from a transactional vision of procurement towards a strategic one**, and allocate adequate resources for upskilling or recruitment of qualified personnel by harnessing existing procurement reform coordination fora,²⁹⁴ or establishing new ones;
- **Invest in upskilling PSM decision-makers and all decision trees related to digital procurement**, or to the procurement of digital solutions;
- **Update common contractual terms to include clauses on legal and financial penalties in case of breaches of trust or excessive reputational/operational risk** emerging from audits of digital systems detecting unintended consequences or biases infiltrating models;
- **Identify uses of algorithmic systems incompatible with the respect for human rights²⁹⁵ or humanitarian principles**, and constantly demand tech partners to run and disclose their hHRDD+ assessments;
- **Partner with research institutions to develop technical interpretations and benchmarking tools to assess the compliance of algorithmic systems and their developers**, against national and international human rights²⁹⁶ and humanitarian legal frameworks, rather than vague principles;²⁹⁷
- **Develop standards and processes for a humanitarian resilience framework for managing cyber risks including in the digital supply chain**, both upstream and downstream, and **develop and enforce internal principles and guidance for risk mitigation for the whole**

²⁹³ Hulo and IAPG (n. 92).

²⁹⁴ See for example the *ProcureNext Tech*. UN System Chief Executives Board for Coordination. (n.d.). <https://unsceb.org/procurenex-tech>

²⁹⁵ Working Group on the issue of human rights and transnational corporations and other business enterprises. (n. 212).

²⁹⁶ Meyer, F. (2024). *How law-abiding is AI? ETH researchers put it to the test*. ETH Zurich.

<https://ethz.ch/en/news-and-events/eth-news/news/2024/10/how-law-abiding-is-ai-eth-researchers-put-it-to-the-test.html>

²⁹⁷ Mavi, J., Găitan, D. T., & Coronado, S. (2025, June 5). *From rogue to safe AI: The role of explicit refusals in aligning LLMs with international humanitarian law*. arXiv. <https://arxiv.org/html/2506.06391v1>

digital ecosystem beyond baseline GDPR rules, including, for example, third-party minimisation, continuous system testing, or context-specific labelling;

- **Ensure that the Humanitarian Reset roadmap²⁹⁸ includes considerations and actionable steps to protect humanitarian actors, especially local and small ones, from falling into tech corporate capture** as a result of reduced resources and increased responsibilities and expectations: **there cannot be true localisation without adequate distribution of resources and agency over the tech stack**;
- **Take example from existing practices of open and transparent participation in algorithmic solutions procurement²⁹⁹** to develop similar platforms for the humanitarian sector; and
- **Develop red lines and compliance rules associated with clear criteria for vetting and evaluating tech partners or providers** also providing services or products to warring parties. This includes:
 - Negotiating strategic vendor agreements³⁰⁰ and updating procurement rules to include checks for political and legal risks specific to algorithmic systems;³⁰¹
 - **Requiring vendors to disclose third parties involved in their AI systems**, and how they would protect their systems and supply chain from undue access or interference during crises or conflicts;
- **Reduce centralised digital solution acquisition processes, in favour of more dynamic digital solution integration approaches** requiring more distributed validation, vetting, and oversight points. This avoids one-size-fits-all mandates and creates flexibility for local contexts and regulatory environments. A centralised oversight role can be maintained by developing a set of overarching digital governance principles covering human rights, ethics, data use, and transparency³⁰² that can be adopted or adapted by country offices, and even organisation-wide AI assurance protocols to ensure consistent testing;
- **Ensure that internal processes and contractual agreements include specific provisions and obligations to adequately cover algorithmic systems' end-of-life³⁰³** including

²⁹⁸ *Humanitarian Reset*. IASC. (2025, November 10).

<https://interagencystandingcommittee.org/humanitarian-reset>

²⁹⁹ *Principles for public participation in AI procurement*. Connected by Data. (n.d.).

<https://connectedbydata.org/projects/2025-procurement-principles>

³⁰⁰ Ahmed et al.(n. 16).

³⁰¹ Rosenauer, P. (2024, November 7). *Legal considerations when procuring AI tools* . PWC Switzerland.

<https://www.pwc.ch/en/insights/regulation/legal-considerations-AI-tools.html>

³⁰² Ahmed at al.(n. 16).

³⁰³ Farrell October 6, R. (2025, October 6). *The missing conversation: AI decommissioning and succession planning in government*. GovLoop.

<https://www.govloop.com/community/blog/the-missing-conversation-ai-decommissioning-and-succession-planning-in-government/>

safeguards to prevent harm during system retirement, and considering algorithmic imprints³⁰⁴ as part of the mitigation and reparation for algorithmic harms;

- **Collaborate for shared AI solutions**,³⁰⁵ forming partnerships with peer organisations to develop common tools, share implementation costs, and maintain collective control over mission-critical systems. This isn't about every organisation building everything independently, but about the sector working together to reduce dependence on commercial providers. Ensure Global Majority voices are central to these collaborations to rebalance power in the global AI ecosystem; and
- **Proactively support and engage CERT mechanisms** to remain promptly informed in case of emerging vulnerabilities or other threats to digital integrity; and
- **Humanitarian platforms and aggregators to facilitate collection of contractual best practices** to contribute to the fulfilment of the commitment taken through the Humanitarian AI code of conduct³⁰⁶ designed by NetHope, stating that members will “align on the approach to due diligence, assessment, and procurement of AI vendors and tools, as well as limits on the commercial use of data”, and “where we have different regulatory environments and uphold the highest available standards and practice of data protection, privacy, and security”.

We also recommend that **humanitarian platforms and aggregators facilitate collection of contractual best practices** to help fulfill their commitment made through NetHope's Humanitarian AI Code of Conduct,³⁰⁷ stating that members will “align on the approach to due diligence, assessment, and procurement of AI vendors and tools, as well as limits on the commercial use of data”, and “uphold the highest available standards and practice of data protection, privacy, and security” in case of multiple applicable regulatory environments.

4.3. Recommendations for tech companies

We call on tech companies to:

- **Replace opt-out with opt-in as the default option for all algorithmic functionalities integrated or added into systems provided to humanitarian actors, restore adequate resources for their human rights teams**, and establish stable and reactive communication channels with their humanitarian clients;
- **Adopt open and inclusive design practices when developing algorithmic systems** by, for example, implementing multistakeholder approaches when building tools and policies,³⁰⁸

³⁰⁴ Ehsan, U., Singh, R., Metcalf, J., & Riedl, M. (2022). The algorithmic imprint. *2022 ACM Conference on Fairness, Accountability and Transparency*, 1305–1317. <https://doi.org/10.1145/3531146.3533186>

³⁰⁵ Ahmed et al. (n. 16).

³⁰⁶ NetHope (n. 282).

³⁰⁷ *Ibid.*

³⁰⁸ Working Group on the issue of human rights and transnational corporations and other business enterprises. (n. 212).

- **Adopt a risk-based approach to the development of AI, especially for systems involving high-risk users or deployments**,³⁰⁹ designed around an “ex ante approach that makes it possible to prevent harmful applications from being placed on the market”,³¹⁰
- **Mandate regular hHRDD+ assessments** for any tech system that is reasonably expected to support or manage humanitarian data or services, and ensure their disclosure to relevant humanitarian clients or partners;
- **Adopt an open approach to the product roadmap of any process affecting core functions of those systems used by humanitarian actors or managing protected or restricted data**, and create appropriate communication and engagement mechanisms to ensure organisations are informed about significant changes affecting the internal compliance of their systems (e.g. product roadmap disclosure or inclusive product lifecycle monitoring);
- **Engage with the nonprofit sector to identify ways to ensure that humanitarian personnel can always access enterprise-grade safety and security features**, possibly also outside of an organisational framework agreement through trusted partner mechanisms;
- **Always include adequate, transparent, and effective complaint mechanisms and remedy options** for individuals or groups harmed by the deployment of algorithmic systems, including from the so called “algorithmic imprints”,³¹¹
- **Collaborate with digital rights organisations and investor groups to develop transparent and reliable indicators** for assessing the company’s exposure to salient impacts on people during conflict, and engage in regular reporting;
- **Standardise the disclosure of upstream and downstream supply chain relationships**, tech dependencies, or their investors and shareholders’ portfolio composition to humanitarian clients; and
- **Clearly define and disclose policy, technical, and governance distinctions between verticals to ensure defence-oriented systems and capabilities are walled off from those parts of the ecosystem dedicated to the humanitarian and human rights sector.**

4.4. Recommendations for local aid actors and communities

We call on local aid actors and communities to:

- **Continue raising the need to prioritise funding for local tech development with donors and the international community**, possibly through collaborative or public-private initiatives when applying for funding;
- **Pressure INGOs and regional actors into supporting the local tech environment** instead of imposing Big Tech stack by default;

³⁰⁹ Ibid.

³¹⁰ Mantelero (n. 178).

³¹¹ Ehsan et al (n. 304).

- **Increase relationship with local tech community and research centres to identify non-Big Tech** partnership options when exploring the purchase design, or development of customised algorithmic systems; and
- **Run customised audits, similar to those outlined in Amnesty Tech’s algorithmic accountability toolkit,³¹² on humanitarian tech partnerships** before accessing their services, **whenever possible and with support from digital rights associations if needed.**

4.5. Recommendations for research groups and cyber experts

We call on research groups and cyber experts to:

- **Analyse tech adoption through the lens of procurement in small NGOs or NGOs operating within low-income countries,³¹³** instead of only focusing on on large NGOs and those based in higher or lower-middle income nations;
- **Develop standards and frameworks for assessing tech companies’ exposure to conflict-related risks,** to enhance transparency when reporting to communities, institutional actors, investors, and shareholders;
- **Research models and practices for enhanced due diligence for business in conflict** with a focus on the digital technologies sector, possibly assessing emerging risks and documenting harms and impacts;
- **Analyse legal language used in contractual documents for tech services and products related to the humanitarian sector** to identify emerging standards and practices that might affect compatibility with internal policies and sector-wide regulations;
- **Include humanitarian organisations in CERT mechanisms³¹⁴** to ensure they are promptly informed in case of emerging vulnerabilities or other threats to their digital integrity; and
- **Develop and run customised audits similar to those outlined in Amnesty Tech’s algorithmic accountability toolkit,³¹⁵** but for humanitarian tech partnerships.

³¹² *Algorithmic accountability toolkit*. Amnesty International. (2025, December 9).

<https://www.amnesty.org/en/latest/research/2025/12/algorithmic-accountability-toolkit/#h-introduction>

³¹³ Rotter, & Bailkoski (n. 13).

³¹⁴ For example, Access Now’s Digital Security Helpline is a member of the Computer Incident Response Center for Civil Society (CiviCERT), an accredited CERT focused on improving the incident response capabilities of civil society groups and individuals around the world. For more info see <https://www.civicer.org/>

³¹⁵ Amnesty Tech (n. 312).