



Not child's play: protecting children's data in humanitarian AI ecosystems

December 14, 2023, Artificial Intelligence and Armed Conflict / Children in conflict / Cybersecurity and data protection in humanitarian action / Humanitarian Action / New Technologies / Technology in Humanitarian Action

8 mins read



Roxana Radu

Associate Professor of Digital Technologies and Public Policy at the Blavatnik School of Government, University of Oxford



Eugenia Olliaro

Programme Specialist, UNICEF's Chief Data Office



In times of crisis and conflict, advanced artificial intelligence and machine learning (AI/ML) might become as much part of the problem as the solution. An ever-growing number of children in need risk having their data used as part of AI systems that do not implement safeguards and protections for vulnerable populations. Given the short- and long-term consequences for the first generations to grow up with AI, there is an urgent need to scrutinize AI-powered systems that are not aligned with the rights, needs, and realities of children in humanitarian action.

In this post, part of a series on Cybersecurity and data protection in humanitarian action, Roxana Radu, Associate Professor of Digital Technologies and Public Policy at the University of Oxford's Blavatnik School of Government, and Eugenia Olliaro,

Programme Specialist at UNICEF's Chief Data Office and the global UNICEF lead of the Responsible Data for Children (RD4C) initiative, make three recommendations for the humanitarian sector to become children-centred in the age of AI.

ICRC Humanitarian Law & Policy Blog · Not child's play: protecting children's data in humanitarian AI ecosystems

Artificial intelligence (AI) and machine learning (ML) systems are increasingly being used when providing emergency relief and vital services to children, at a time when *more children than ever before* are affected by different types of crises around the world. Their data will be consequential in as yet unknown ways, since they are the first generations to be brought up with AI.

In a humanitarian context, some of these consequences can be dire, as data can put children and their communities at immediate risk. This holds true not only for “*big data*”, but also for “small data”. Gender-disaggregated information about children above a certain age could allow those who want to prevent girls from going to secondary school to do so. Data about child-headed households in a refugee setting would indicate to malicious actors where the unsupervised children are and put those children at risk of exploitation and abuse. Equally true, the profiling of beneficiary groups can turn into a life-threatening practice within or outside a humanitarian setting. New ways of recombining data can thus shape children's short- and long-term opportunities and life chances in unparalleled ways.

Developed at unprecedented speed and scale, ML/AI-based systems rely on data to produce their valuable output: more data. In the process, information about children might feed advanced AI/ML tools beyond their originally specified intent. Due to the complexity and opacity of these systems, it is difficult to ensure that deep learning on particular types of data is not integrated in evolving forms of ML or in related products.

This is already the case in highly regulated environments, where existing data protection frameworks have proven *insufficient* to prevent the harms associated with the use of personal data in AI. Regulators in Canada, Australia or EU countries race to identify and enforce new rules to regulate AI companies and protect their citizens from “*murky data protection practices*“. If business is difficult to govern in highly regulated countries, *some wonder* how AI would be regulated in places where the rule of law is weak – and it is precisely in those jurisdictions that AI-powered solutions tend to be deployed first.

In ecosystems with weaker protections for children, ML systems are able to collect, explore, and integrate more data, when allowed to learn in unsupervised ways. At scale, both anonymized information and non-personal data about children becomes valuable. In times of emergency, there is pressure to make data available as fast as possible, meaning that there are inherent risks of much of the data about children being captured by systems that are not scrutinized for their vulnerable population(s) safeguards. While safety checks and licensing of AI systems for specific sectors are under discussion, we are witnessing the fast roll-out of systems with minimum or no protections in place for those most at risk. For this reason, machine learning models need to be taught to forget information. Paradoxically, to adapt machine learning to humanitarian contexts, ‘*machine unlearning*’ will be required.

Child-centred AI applications

By design, AI systems are *not developed with conflict in mind*. Nor are they highly sensitive to the needs of populations at risk, who are rarely involved in their creation. While there are emerging calls for children to be (meaningfully) consulted as a key stakeholder group in child-centred AI applications (e.g. UNICEF *Policy Guidance on AI for children*, *Scottish Government and Alan Turing Institute, Beijing AI Principles*), the reality of humanitarian action often limits such engagement. With pressing needs to ensure affected populations' survival and wellbeing, organizations providing services in humanitarian crises struggle to find time to talk about data with the communities they serve — let alone engage in explanations about ML and advanced processing of data. Communities in need of aid might provide a wide range of information about them and consent to its use without grasping the potential implications, or without really having the choice — as they would otherwise simply not receive support.

If this is the case for adults, it is all the more true for children. Children may not be provided with adequate support to assess associated risks and benefits when they participate in programmes that generate personal or non-personal data related to them; or they may not even be asked, because their participation depends on someone else's legal consent. Further risks associated with a poor communication and sensitization of children include loss of trust in both emerging technologies and in the frontline relief organizations adopting them.

Practitioners themselves might have trouble understanding how the ML models work and how they impact decision-making, in the short- and in the long-run. Despite their unresolved “black box problem”, AI systems *already support decisions* in humanitarian contexts, including for vulnerable children. Alongside concerns for *surveillance humanitarianism* and *techno-colonialism*, new forms of bias need to be explored in relation to children and AI, such as the push to serve those better captured in data, which might leave many others behind. Until now, we know *very little about the harms* engendered by the use of AI in humanitarian contexts. As such, when AI is used as part of a toolkit to aid decision-making – as it is often the case, alongside human reviews and checks – data fed to the system will shape future programmes and developments, for both the organizations involved and the children targeted.

Although some AI-powered services currently in use are developed in-house, *limited technical expertise and lack of funding* have driven many non-profits in the humanitarian sector towards partnerships with the private sector. Innovation in this space mostly comes from partnerships with technology companies operating globally, which have market incentives to increase the accuracy of their models by integrating as much data as possible. Though some tech giants have adhered to a set of principles for humanitarian action (see, for example, Microsoft's *AI for Humanitarian Action*), there are no specific commitments to children's data and contextual vulnerability safeguards. Moreover, there is no guidance available as to which ML capabilities from humanitarian action work should not be integrated in global AI systems deployed for commercial purposes.

Three recommendations for the humanitarian AI ecosystem to become children-centred

In the humanitarian sector, the *shift from responsive to anticipatory action is powered by AI*. Many of the international organizations, governments, non-profits and for-profits involved on the ground have started integrating ML/AI practices into their work, but there is insufficient scrutiny for the protections afforded to children, when data that is required for access to services is fed into AI systems.

As we move beyond the ML hype and AI techno-solutionism, urgent changes are needed for the humanitarian AI ecosystem to become children-centred. We make three recommendations for the responsible use of artificial intelligence by humanitarian organizations and their private sector partners:

1. Improve transparency around the children-related data points (personal and non-personal data) that are added to AI/ML models deployed in humanitarian contexts and assessing their long-term consequences;
2. Specify how AI systems are designed and in particular how they are configured to unlearn in situations of high vulnerability; and
3. Develop a public register of tested and vetted AI solutions in use in humanitarian contexts, to set minimum standards of data protection and data use when it comes to data for and about children; strengthen collective learning before humanitarian needs arise; and allow for public scrutiny during and after emergencies.

It is key to set adequate safeguards for children in the humanitarian AI ecosystem now. We owe it to them and to the generations to come, whose life opportunities and wellbeing depend on it.

See also:

- Susanna Acland, Barnaby Willitts-King, *Mobile phones for participation: building responsible public-private humanitarian partnerships*, December 7, 2023
- Ed Millet, *Deploying OSINT in armed conflict settings: law, ethics, and the need for a new theory of harm*, December 5, 2023
- Vincent Cassard, Stuart Campo, Jonas Belina, *Responsible data sharing between humanitarian organizations and donors: towards a common approach*, June 22, 2023

Tags: AI, armed conflict, artificial intelligence, children, children in armed conflict, cyber security, data protection, humanitarian action

You may also be interested in:



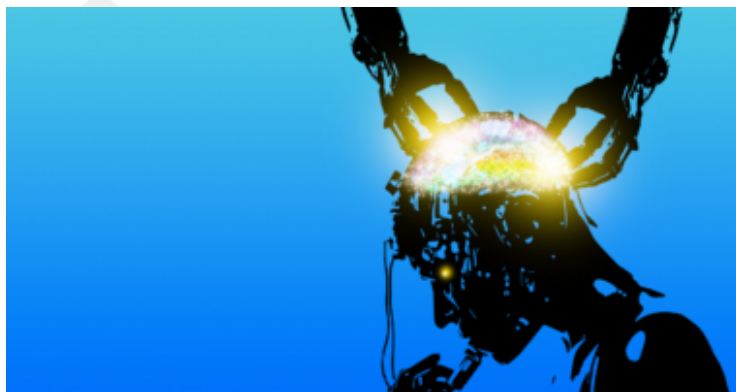
Mobile phones for participation: building responsible public-private humanitarian partnerships

11 mins read

Artificial Intelligence and Armed Conflict / Children in conflict / Cybersecurity and data protection in humanitarian action / Humanitarian Action / New Technologies / Technology in Humanitarian Action

Susanna Acland & Barnaby Willitts-King

Mobile phones are a powerful tool for participation and engagement in humanitarian crises. Responsible public/private ...



Deploying OSINT in armed conflict settings: law, ethics, and the need for a new theory of harm

16 mins read

Artificial Intelligence and Armed Conflict / Children in conflict / Cybersecurity and data protection in humanitarian action / Humanitarian Action / New Technologies / Technology in Humanitarian Action Ed Millet

The deployment of open-source intelligence, or OSINT – information gathered from publicly available data sources ...