Aid and AI: The Challenge of Reconciling Humanitarian Principles and Data Protection

Júlia Zomignani Barboza, Lina Jasmontaitė-Zaniewicz, Laurence Diver

▶ To cite this version:

Júlia Zomignani Barboza, Lina Jasmontaitė-Zaniewicz, Laurence Diver. Aid and AI: The Challenge of Reconciling Humanitarian Principles and Data Protection. 14th IFIP International Summer School on Privacy and Identity Management (Privacy and Identity), Aug 2019, Windisch, Switzerland. pp.161-176, 10.1007/978-3-030-42504-3_11 . hal-03378953

HAL Id: hal-03378953

https://inria.hal.science/hal-03378953

Submitted on 14 Oct 2021

Aid and AI: the challenge of reconciling humanitarian principles and data protection

Júlia Zomignani Barboza, Lina Jasmontaitė-Zaniewicz and
Laurence Diver

Vrije Universiteit Brussel, Belgium

Abstract

Artificial intelligence systems have become ubiquitous in everyday life, and their potential to improve efficiency in a broad range of activities that involve finding patterns or making predictions have made them an attractive technology for the humanitarian sector. However, concerns over their intrusion on the right to privacy and their possible incompatibility with data protection principles may pose a challenge to their deployment. Furthermore, in the humanitarian sector, compliance with data protection principles is not enough, because organisations providing humanitarian assistance also need to comply with humanitarian principles to ensure the provision of impartial and neutral aid that does not harm beneficiaries in any way. In view of this, the present contribution analyses a hypothetical facial recognition system based on artificial intelligence that could assist humanitarian organisations in their efforts to identify missing persons. Recognising that such a system could create risks by providing information on missing persons that could potentially be used by harmful actors to identify and target vulnerable groups, such a system ought only to be deployed after a holistic impact assessment has been made, to ensure its adherence to both data protection and humanitarian principles.

Keywords: humanitarian action, artificial intelligence, facial recognition, data protection, humanitarian principles

1 Introduction

The use of artificial intelligence (hereafter AI) and biometrics systems is no longer the preserve of science fiction. On the contrary, a combination of advances in computing power and the vast amounts of data being generated by Internet-connected devices means that AI systems are now a truism in our everyday lives [1]. These systems are present, for example, in voice-activated digital assistants, biometric and facial recognition systems that unlock smartphones or allow access to buildings, traffic routing applications, purchase or viewing recommendations on online platforms, and many other features of smart devices. It is not surprising, therefore, to see that the humanitarian sector is also exploring how AI and biometrics tools can be applied to further the provision of humanitarian aid.

The use of AI and biometrics in the humanitarian sector, as in any other field, comes with many challenges. Some of these are specific to AI, while others are inextricably linked with the use of biometrics, particularly in relation to the protection of personal data. The most frequently cited challenge of AI systems relates to machine bias, exemplified by the controversial COMPAS algorithm used in the US to predict recidivism rates in criminal cases in order to assist judges in determining bail. The algorithm attracted criticism on the basis that it predicts black defendants as being almost twice as likely to reoffend as white defendants [2]. These predictions can be considered biased, depending on the technique for measuring fairness that is adopted [3]. Furthermore, the use of AI can have serious implications for individuals' rights to privacy and personal data protection, enshrined in international law,[1] especially considering the systems' increasing "capability of linking data or recognising patterns of data [that] may render non-personal data identifiable" [1; p. 11].

Apart from these general concerns about AI, its use in the humanitarian sector must also be reconciled with the humanitarian principles that govern how humanitarian organisations are supposed to act. In particular, organisations ought to comply with the principles of humanity and impartiality, of 'do no harm', of accountability, and the principled goals both of facilitating participation on the part of humanitarian beneficiaries and of building on local capacities [4].

Complying with such principles when deploying AI is particularly challenging given that humanitarian organisations usually lack the technical knowledge and resources to develop their own AI and biometric systems and must, therefore, rely on partnerships with private, for-profit technology companies. Such partners may have incentives which are incompatible with the humanitarian principles and which may in turn cause reputational damage to humanitarian organisations. As an illustrative example, the World Food Program (WFP) recently partnered with controversial data analytics firm Palantir to improve its food delivery and cash-based assistance programs. The WFP was heavily criticised as a result, leading it to issue a defensive public statement explaining how the partnership complied with the organisation's principles [5].

The use of AI systems in humanitarian aid raises numerous complementary issues from both the humanitarian and data protection perspectives, not least as to whether such use is compatible with either the principles pertaining to humanitarian assistance or the data protection principles implicated by the use of such technologies. Bearing these in mind, the challenge is how to deploy AI systems that comply with both data protection requirements and humanitarian principles. Taking into account the limited space available and the broadness of the field, we focus on one specific AI application, namely the use of facial recognition to identify missing persons.

The next section provides a brief overview of the evolution of humanitarian action, followed by section 3 which analyses the use of AI and biometrics in the humanitarian sector. Sections 4 and 5 consider the challenges for privacy and data protection and the implementation of humanitarian principles, respectively.

[1] See Article 12 of the Universal Declaration of Human Rights, Article 17 of the International Covenant on Civil and Political Rights, and the Council of Europe Convention for the protection of individuals with regard to automatic processing of personal data, no.108.

1.1. Terminology

For the purposes of this paper, we consider **artificial intelligence** to be "[a] set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being." [6] It includes the development of algorithms that improve their performance when completing a certain task with experience in the form of machine-readable data. The latter is usually referred to as 'machine learning', a subset of artificial intelligence more generally – this is the form of AI we focus on in this paper. We use the terms '**humanitarian action'**, 'humanitarian assistance' and 'humanitarian aid' interchangeably to refer to "any activity undertaken on an impartial basis to carry out assistance, relief and protection operations in response to a Humanitarian Emergency." [7; p. 8].

2 The evolution of humanitarian action

Traditionally, humanitarian action was carried out by only a handful of dedicated organisations. In the present day, however, humanitarian action is provided by numerous actors, including states, international organisations, non-governmental organisations (NGOs), private companies, and private and technology-related philanthropic foundations (e.g. The Bill & Melinda Gates Foundation) [8; p. 7]. This shift toward the involvement of multiple parties has been partially driven by the increasing use of digital technologies to speed up the delivery of aid. The strive for (technology-based) efficiency has drawn in a range of new actors, e.g. mobile network operators and financial institutions involved in cash transfer programmes. This has increased the complexity of governing humanitarian action and has created more intricate data flows that include non-personal and personal data concerning both beneficiaries and humanitarian staff.

The presence of multiple actors in humanitarian action is not problematic per se, but the lack of a 'core set of shared values' can make such cooperation challenging [8], especially when the processing of beneficiaries' personal data is involved. To address these challenges, leading humanitarian organisations (e.g. the International Federation of Red Cross and Red Crescent Societies (IFRC), the International Committee of the Red Cross, and the United Nations High Commissioner for Refugees (UNHCR)) as well as smaller actors in the field (e.g. Terre des Hommes) have developed internal and external privacy policies. Many of these policies and their more recent revisions are intended to align such instruments with principles stemming from the EU's General Data Protection Regulation 2016/679 (GDPR). For example, the revised IFRC privacy policy allows data subjects to object to data processing, and provides for the reasonable expectation of notification should their personal data be disclosed to an unauthorised third party. It also foresees that in situations where "processing operations appear likely to result in a high risk to the rights or freedoms of a data subject", a data protection impact assessment should be carried out.

Taking into account the plurality of actors involved in humanitarian action and the fact that not all of them stand on the same footing, some organisations, such as the Harvard Humanitarian Initiative (HHI), have put forward guidelines and codes of ethics for those providing humanitarian assistance that are based largely on the principles of the EU data protection framework.[2] Building on these observations it might be said that the GDPR represents the 'gold standard' for data protection, not least in the humanitarian sector. This is in spite of the fact that many humanitarian organisations are, due to their status in international law, not bound by the GDPR or any other national or regional data protection legislation. It should be noted that many of these organisations are based in Western countries that fall within the EU's sphere of influence and, therefore, the choice to follow EU standards may not be entirely neutral (this phenomenon is sometimes referred to as the 'Brussels effect' [9]). Regardless of the exact reasons behind the extraterritorial influence of the GDPR, it has in practice become the standard for data protection and, consequently, we rely on the principles and definitions that it sets out.

3 AI and biometrics in humanitarian action

The use of AI and biometrics is not new to humanitarian organisations. For example, AI tools are being used to identify patterns and make predictions based on the analysis of social media platforms to detect disasters and identify the needs of affected populations [10]. Systems like the Artificial Intelligence for Disaster Response (AIDR) platform use AI "to automatically identify informative content on Twitter during disasters" [11].[3] In emergencies, AI also enables the automated mapping of disaster areas using satellite and aerial images [12]. AI can also assist in diagnosing disease, as well as in the early detection of pathogens, helping to avoid outbreaks and pandemics that may occur in the context of humanitarian crises (e.g. Microsoft's Project Premonition [13]).

Biometric systems, particularly those based on fingerprints and iris scans, have been used by humanitarian organisations "as part of their identification systems because of the benefits it can bring in efficiently identifying individuals and preventing fraud and/or misuse of humanitarian aid" [7; p. 98]. Such systems are often chosen by humanitarian organisations in the belief that they are typically "more difficult to counterfeit and, being digitally produced and stored, facilitate the efficient management of humanitarian aid in the field" [7; p. 99]. Some even argue that the use of biometric systems in humanitarian action "reveals a determined humanitarian focus on the individual" [14].

Despite the perceived advantages of using AI and biometrics in humanitarian action, the combination of the two technologies to facilitate the next generation of AI-based biometric systems (including facial recognition) raises important legal and ethical concerns. When combined, the two technologies could be (re)used for purposes and in contexts that humanitarian organisations may not fully anticipate or be able to control.

[2] For example, the HHI Signal Code includes the Right to Privacy and Security, according to which data of affected individuals should be (i) processed fairly and lawfully, and not further processed in a way incompatible with that purpose, (ii) adequate, relevant, and not excessive in relation to that purpose, (iii) accurate and, where necessary, kept up-to-date, and (iv) not kept longer than necessary to achieve the stated purpose under which informed consent and/or participation was obtained.

[3] This platform goes beyond simple keyword search, which its makers say can fail to identify over 50% of the textual content posted on Twitter that is relevant to disaster response.

In this regard, the next subsection reflects on general trends emerging in the area of AI and facial recognition, with the aim of unpacking some of the issues that these pose in the broader context.

3.1 General trends in AI-based facial recognition

Facial recognition from photographs is widely used by governments and private companies, and can be performed to a high degree of accuracy.[4] Such systems allow full identification of individuals and, with the rapid pace at which the technology is developing, facial recognition is possible even from live video (such examples are seen in China) in addition to static images.

A critical literature on facial recognition is beginning to emerge in tandem with these technological advancements, with a number of arguments being developed both in favour and against the technology. Some argue that more balanced and representative training datasets can facilitate accurate recognition across racial and gender boundaries, avoiding discrimination [16], while others warn that facial recognition is inherently dangerous and ought to be banned outright for any and all purposes (see [17, 18]).

There are multiple examples of facial recognition systems being used in practice and for various purposes. Traffic Jam [19], for example, uses facial recognition to assist law enforcement to locate victims of human trafficking and was estimated to have identified 3,000 victims of sex trafficking in 2018. Samsung provides facial recognition as a mechanism both for users to unlock their smartphones [20] and, in tandem with Diebold Nixdorf, to authenticate identity at ATMs [21].

Other uses of the technology currently being explored include identifying criminals, tracking school attendance, and seamless border crossing, as well as the identification of missing persons, which is further explained in the next section.

Considering the multiple and sometimes life-saving potential uses of the technology, it is to be expected that various public and private actors would seek to invest in the development of facial recognition systems. Recent revelations show, however, that some are willing to employ questionable methods in the development of facial recognition. The Ever smartphone application for example, which offered users free storage for photos, used the uploaded images to train and improve the company's facial recognition system. The only notification users received of these activities was a short and vague statement in the company's lengthy privacy policy [22]. Similar concerns have been expressed about other photo applications such as FaceApp (see [23, 24, 25]).

Besides concerns over consent and the longstanding debate over facial recognition systems' interference with privacy (see [26, 27, 28, 29, 30]), the potential of the technology to profile specific ethnic groups [31] and even purportedly to identify homosexuals [32] are worrisome, particularly if such systems are relied upon despite their inaccuracies. The harms that could arise from the use of facial recognition systems impact humanitarian organisations' compliance with humanitarian principles, as explained below in section 5.

[4] Taigman et al. claim, for example, that Facebook facial recognition using the Deep Face method "reaches an accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset, reducing the error of the current state of the art by more than 27%, closely approaching human-level performance". See [15].

3.2 The prospective use of facial recognition to identify missing persons in the humanitarian sector

As mentioned in the introduction, taking into account the limits of this paper and the broadness of the field, we focus on the use of a hypothetical AI-based facial recognition system to identify missing persons. Various public actors have already deployed such systems: examples such as India's National Tracking System for Missing & Vulnerable Children are positively regarded; the system identified nearly 3,000 missing children within four days of launching a trial that matched photos of missing children with the faces of children throughout New Delhi [33]. To achieve that result, however, the technology processed the data of 45,000 children, raising concerns about interference with their privacy. Similarly, the British police is also preparing to use facial recognition systems to identify missing people in the UK by scanning CCTV footage [34].

In the humanitarian sector, the use of AI-based facial recognition will take a different shape. Humanitarian actors such as the International Committee of the Red Cross (ICRC) have a long history of identifying the fate of missing persons, restoring contact between family members, and facilitating family reunification [35]. In view of the significant inward migration to Europe in recent years, the ICRC together with National Red Cross Societies in Europe have developed the Trace the Face program [36], where those looking for a family member manually search an online database containing photos posted by individuals who are looking to be found.

For the purposes of this paper, we predict that in the near future AI-based facial recognition systems will be used to automate this search. In practice, this would mean that instead of manually going through the database, someone looking for a family member would upload a photo of their relative into the system to try and locate their family member in the database. This would be done upon entering into a contract with a humanitarian organisation providing access to such software. The system would then map the facial features in the photo uploaded, such as the distance between the eyes and the distance from forehead to chin, thus creating a facial 'signature'. This would then be compared to a database of known faces to look for a match (i.e. a photograph that contains the same facial features) [37].

In the posited case, then, the photo of the relative uploaded by the person seeking their family member would not be published publicly, but would instead pass through the facial recognition system in order to be compared with the public photos on the database. In the next sections, we analyse the potential challenges such use of facial recognition would pose to humanitarian organisations.

4 Privacy and data protection concerns

As pointed out earlier, the deployment of AI-based facial recognition systems for identifying missing persons raises privacy and data protection concerns. As is the case with any technology with disruptive potential, especially when they are intended to process large amounts of personal data, the deployment of such systems should be preceded by a data protection impact assessment (DPIA), which is used to identify, evaluate, and address the risks to individuals and their personal data arising from a specific system.

Such assessments are typically done on the basis of principles stemming from the data protection framework, and include questions such as:

- Is the processing fair and lawful?
- Are the data subjects able fully to exercise their rights? If not, are the restrictions lawful and proportionate?
- How does the system operate?
- How is responsibility attributed among the involved parties?
- What are the risks posed by the chosen system to data subjects? How will each risk be treated?

Although some steps and characteristics are common to all impact assessments, these need to be tailored to the specificity and needs of a given project and its context. Humanitarian organisations, therefore, should develop DPIA guidance (if necessary with the assistance of external advisors and support of national data protection authorities) that takes the specific characteristics of humanitarian action into account. One might ultimately envision a standard emerging from the DPIA guidance documents of various humanitarian organisations.

In this regard, it is important to note that in the humanitarian sector compliance with data protection principles alone is not sufficient because of certain contradictions embedded in them. Take, for example, the principle of data minimisation. It can be said that the principle of data minimisation opposes the principle of accuracy, requiring data to be updated [38]. By minimising the amount of data processed, the controller can end up with applications that provide discriminatory results. Similarly, it is suggested that "data minimisation is not always be able to exclude privacy violating or discriminatory results given the redlining effect." In this regard, "[d]ata minimisation not only offers no adequate solution in this respect, it might also make it difficult to assess whether a rule is indirectly discriminatory or privacy violating" [38; p. 162].

Considering this background, we suggest that such assessments should be enriched by including humanitarian principles, which we consider further below in section 5. For now, we maintain our focus on privacy and data protection. More specifically, the next sections will address the principles of purpose limitation, data security and fairness, as these principles are some of the most challenging ones to implement in the application foreseen by this paper.

4.1 Purpose limitation

According to the purpose limitation principle, data must be collected for a clear and specific purpose and cannot be further processed for any purpose unrelated to the original. When it comes to AI, it is important to note that these systems 'learn' from the data that passes through them to improve their outcomes [39]. In the case of facial recognition, an AI system will learn from every photo that passes through it to better identify facial features in future photos. Thus it can be argued that despite the original purpose being the identification of missing persons, the photos will potentially be further processed to improve the system itself.

At this point, it is important to note that AI systems can be static or dynamic.[5] The former process (personal) data only to perform the task which was assigned to them (in this case, matching photos), while the latter process data both to perform the task (i.e. matching photos) and to refine their internal model to improve accuracy.

Consequently, even if humanitarian organisations deploy off-the-shelf AI systems developed by technology companies instead of developing their own, they may still have to deal with the fact that the data passing through the system (i.e. the photos of missing persons) is used to improve it. It is, therefore, essential that humanitarian organisations are aware of which type of system they are using and what the data protection implications are.

In this regard, it is worth noting that while it could be also argued that further processing the data to improve the system for humanitarian use is compatible with the initial purpose, since a better trained system has a better chance of identifying missing persons, humanitarian organisations should also be aware of the possibility of their partners using these data for commercial purposes. For example, when purchasing an off-the-shelf system, the technology company behind the AI may request that the data collected by the humanitarian organisation be used to improve the company's non-humanitarian systems that will later be commercialised by private or governmental entities.

4.2 Data security

As the facial recognition system we envision processes biometric data – the features of an individual's face – to identify missing persons, such data must be subjected to a high level of security. This is because biometric data that is used to identify someone is considered to be a special category of personal data deserving of enhanced protection in a variety of legal regimes. According to the GDPR, for example, biometric data, which is defined as personal data relating to the physical, physiological or behavioural characteristics of a natural person which allow or confirm the unique identification of that natural person [41], will be considered a special category of data when collected "for the purpose of uniquely identifying a natural person" [42]. Uniquely identifying someone, especially someone belonging to a vulnerable group (which is often likely to be the case for those who have gone missing during emergency situations) can lead to stigmatisation and discrimination.

In China, for example, the government is said to be using facial recognition systems to track down and target Uighur Muslims [31]. While the system foreseen by this paper would not aim at categorising members of specific minorities, attacks against such systems (mentioned below) could alter them or retrieve the biometric data that passes through them, which could later be used to identify members of certain groups. Furthermore, if technology companies manage to use data gathered by humanitarian organisations to improve their commercial systems (as foreseen above), such systems

[5] Static models will not change over time and will always apply the model developed with the training data. This allows the programmer to maintain full control of the model but stops the system from refining itself over time. Dynamic models, on the other hand, avail themselves of input data to adjust to changes and refine their outputs, for more see [40; p. 10].

could be trained to accurately identify the traits of certain ethnicities and later be sold to harmful actors that wish to target them. In this regard, it is important to note that when using biometrics,

> [g]iven the potentially harmful consequences for the persons concerned, more stringent requirements will have to be met in the impact assessment process of any measure interfering with an individual's dignity in terms of questioning the necessity and proportionality as well as the possibilities of the individual to exercise his right to data protection in order for that measure to be deemed admissible. [43; p. 15]

When determining which security measures should be put in place, humanitarian organisations should take into consideration that harmful actors may try to conduct deliberate attacks that aim at (i) revealing information about the data that passed through the system (model inversion), (ii) undermining the utility of the system by adding noise to the input data or inserting bad data that will induce the system to misread the information or emphasise the wrong features (poisoning attack) or (iii) gaining unauthorised access to the system (backdoor attack) and modifying it after it has been trained. Such deliberate attacks can further decrease the quality of outcomes, sometimes leading to false positives and negatives. This is particularly relevant because even data that would not be publicly available (such as the photos of missing relatives uploaded to the system, as in the example above) might potentially be revealed through these types of attacks. The need for a high level of security is thus evident.

4.3 Fairness and bias

The principle of fairness requires that all processing activities respect data subjects' interests and that data controllers take action to prevent arbitrary discrimination against individuals [40]. The risk of discriminatory bias in AI systems is widely known and may be a result of, for example, using biased datasets to train the system, systemic biases in society that are reflected in the data, or even choices of the programmer when deciding which features to assign more value to in each dataset. Without delving into the intricacies of information theory and machine learning research design, it can at a minimum be said that AI developers should assess the quality, nature and origin of the personal data used to develop the system as well as considering the potential risks to individuals and groups of using de-contextualised data, which can create de-contextualised results [44].

In the case at hand, the matching of the photo of a relative to that of a missing person in a database should not determine the fate of an individual in the same way the recidivism algorithm mentioned in the introduction did; however, when one considers that most facial recognition systems perform better with male faces than female faces and on lighter skinned faces than on darker skinned faces [16], there is a high probability that the system will not perform as well on certain minority groups, inevitably leading to mismatches. Such results are likely to exacerbate an already stressful situation for vulnerable beneficiaries who are relying on such systems to find their missing relatives. Conversely, systems that are very efficient in identifying those belonging to a certain minority can be used by harmful actors to target them, as is the concern with the Uighur in China.

It is essential therefore that those deploying such systems (considered to be data controllers) carry out frequent assessments on both the system itself and the data used to develop and improve it, in order to address any possible form of bias or discrimination [45]. The consequences of not taking such measures are not only legal, but also have an impact on humanitarian principles, as will be explained below.

5 Application of humanitarian principles

As mentioned above, data protection impact assessments can be used to identify risks associated with the processing of individuals' personal data, including the potential negative effects of a specific technology on a variety of fundamental rights as well as the ethical and social consequences of the data processing [46]. Considering humanitarian organisations' mandate of providing humanitarian assistance, it is essential that humanitarian principles are also considered and, therefore, included in such assessments. Furthermore, it should be noted that principles setting requirements for personal data protection and humanitarian action are complementary and reinforce each other (see [47]). Indeed, the two frameworks are built with individuals' dignity and empowerment at their core. In the following sub-sections we elaborate on some the most topical humanitarian principles. In this regard, humanity and impartiality are the core two principles of humanitarian action. They embody the idea that humanitarian action should aim to prevent and alleviate human suffering wherever it may be found, and that it should be provided to anyone in need, regardless of nationality, race, religious beliefs, class, or political opinion [4].

In this regard, the concept of discriminatory bias in AI is once again relevant. Studies have shown that facial recognition systems perform better in the population of the region where they were developed [16]. Considering that many of these systems are developed in Western countries, this is one of the explanations of why they tend to perform better on light skinned males (see above in section 4.2). Databases such as the ICRC's Trace the Face are based in Europe and, most likely, will make use of Western-developed systems, even though the database itself contains photos from different regions and ethnic groups. It would be necessary therefore to scrutinise whether a facial recognition system applied in such a case is able correctly to identify those coming from other regions or belonging to certain minority groups that are underrepresented in the data.

Considering that the impartiality principle requires that humanitarian action benefits everyone in need in a non-discriminatory manner, deploying a system that is likely to benefit only those belonging to a majority Western group may not be appropriate. Furthermore, because such systems can be less accurate within minority groups, using them to identify these groups might lead to a high number of false positives or false negatives, magnifying the detrimental emotional effects on an already-vulnerable class of user.

It is also important to consider that deploying ineffective systems may negatively affect an organisation's reputation and local acceptability, hindering their access to local population and, potentially, their ability to operate in a certain area, thus obstructing the humanitarian mission.

5.1 'Do no harm'

The 'do no harm' principle is part of humanitarian organisations' commitment to ensure "[c]ommunities and people affected by crisis are not negatively affected and are more prepared, resilient and less at-risk as a result of humanitarian action" [48; p. 59]. In other words, humanitarian organisations should not leave a negative footprint in the contexts where they act.

Organisations may face various dilemmas in which the consequences of their actions might be unclear. For example, whether they would have to share data they gather to identify missing persons with law enforcement authorities and other public or state actors. This is particularly relevant considering that if harmful actors were somehow to acquire such data it may allow them to identify vulnerable groups with the intention of harming them. The ICRC highlighted this potential issue in its Restoring Family Links (RFL) strategy for 2020-2025:

> "The sharing of potentially sensitive information on affected people with other entities, such as States, that wish to use such data for non-humanitarian purposes might expose individuals to new risks, such as profiling, discrimination, arrest or even exclusion from humanitarian assistance. This would negatively impact the safety of the very people humanitarian action is trying to help, contradict the "do no harm" principle and be incompatible with the Fundamental Principles, especially neutrality, impartiality and independence" [49; p. 5]

When making decisions around such dilemmas, it is recommended that humanitarian organisations make "sufficient enquiries about the ethics, interests, risks and professional reliability of individuals and organizations in your agency's political, commercial and humanitarian supply chain and delivery network; and acting upon information received" [4; p. 109].

5.2 Participation by beneficiaries and building on local capacity

In essence, the principle of building on local capacity implies that humanitarian organisations should act in a way that brings beneficiaries into the process as far as possible, such that disaster responses are not simply 'imposed' upon them [4; p. 82]. This principle, however, extends beyond participation. As rightly noted by former ICRC President Jakob Kellenberger, humanitarians' work should allow the re-establishment of physical and mental resilience in affected populations. In particular, it should aim to assist those affected to regain their autonomy, in order to better "cope with the shock and trauma caused" [50; p. 987]. This notion of autonomy is also present in the data protection framework, linking closely to the notion of informational self-determination [51; pp. 8-9]. We recognise that the two notions are being used in different contexts, but ultimately they each serve to respect and empower the individual.

At the same time, Western individualistic notions of autonomy and rational decision-making may not be appropriate across the spectrum of geographical and cultural contexts in which humanitarian assistance is provided. At any rate, if the goal of humanitarian action is to save and protect life in order to facilitate its subsequent flourishing, it is important that this goal is not undermined by data processing that denies the individual similar opportunities. Squaring that circle may prove to be extremely challenging in practice, where AI systems that rely on personal data are used in contexts where beneficiaries have little power to contest the processing, and perhaps not even the correct frame of reference to conceptualise it.[6]

In the hypothetical system introduced above, following the rationale of the building on local capacity principle, several questions arise. One might ask to what extent the "technical wisdom that local knowledge often knows best" could be applied within the scope of such a project. Furthermore, it might be queried whether the skills and understanding of the local population are such that the application is in practice useful to them [4; p. 80]. Finally, it must be considered whether or not the affected beneficiaries can provide meaningful consent to the processing of personal data and, perhaps more importantly, whether they comprehend the risks that the processing may constitute in the case of an AI-facilitated biometric system. In the end, it may well be that respecting the principles of beneficiary participation and local capacities entails the conclusion that AI ought not to be used.

5.3 Accountability: value for money or humanitarian effectiveness

The Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief [52] states that "[w]e [those who adhere to it] hold ourselves accountable to both those we seek to assist and those from whom we accept resources." This commitment can prove challenging, however, when resources come from national or regional authorities which, as mentioned above, may pressure humanitarian organisations or NGOs into sharing the biometric data sets they have collected (in this case, the photos of missing persons and facial patterns found by the system), "with the risk of the data being used for purposes other than strictly humanitarian purposes (e.g. law enforcement, security, border control or monitoring migration flows)" [7; p. 99]. In this case, being accountable to donors can compromise accountability to beneficiaries, since acquiescing to donors' requests to share data might endanger beneficiaries (e.g. in case such data is used to target minorities such as the Uighur, mentioned above).

Furthermore, beyond addressing the risks of their own activities (including the sharing of beneficiaries' data), humanitarian organisations also need to consider the actions of their technology partners, who might have divergent interests and whose commercial imperative might be at odds with the humanitarian programs they engage with, for example the improvement of their technology for later re-use in non-humanitarian contexts (recall the example of Ever above). Vulnerable groups and individuals whose data are processed for the purposes of providing aid might find themselves inducted into

[6] But compare [51], detailing how the WFP halted aid where local beneficiaries refused to permit their biometric data to be harvested, reportedly for reasons of sovereignty (as opposed to data protection).

surveillance-capitalist paradigm that is not of their choosing, particularly where the social and legal norms of (Western) data protection law do not easily translate across cultural and developmental boundaries. These concerns are often raised when humanitarian organisations partner with the private sector, as was the case in WFP's partnership with Palantir, mentioned above. When partnering with the private sector, therefore, it is essential that humanitarian organisations can clearly present to partners what is needed and what is expected in the project, especially given that technology companies might have different ways of working in terms of budget, timelines, reporting etc. Moreover, humanitarian organisations need to assess the risks of working with external partners, possibly through an impact assessment, to ensure that affected populations and their data are not endangered by the partnership.

Such risks reinforce the utmost importance of humanitarian organisations taking a precautionary approach. They should conduct thorough impact assessments before opting to use facial-recognition techniques in their programs and should use such technologies only after taking all feasible precautions to protect those they seek to help, in the humanitarian spirit that is their very *raison d' être*. In this regard, and similar to the accountability principle in data protection, accountability requires evidence. More specifically, "[a]gencies need to be able to 'know and show' what they have intended, decided, done, and the results that have flowed from their actions." [4; p. 99].

6 Outlook for future engagement

Compliance with data protection principles by themselves does not suffice for the legitimate deployment of new technologies by humanitarian organisations. Due to the humanitarian sector's mission to carry out assistance, relief and protection operations in an impartial manner that does not cause harm to beneficiaries, it is essential that the humanitarian principles are also complied with. Before deciding on the use of a certain technology, therefore, humanitarian organisations should conduct a holistic assessment of compliance of a specific system with both legal rules and humanitarian principles. In the case of using AI-based facial recognition to identify missing persons, the principles identified above are particularly relevant and unless all of them can be complied with, such a system should not be implemented.

Further research and analysis will be necessary to fully tease out the interplay between the humanitarian and data protection principles. However, by drawing attention to the complementariness of the two regimes, humanitarian organisations can be sensitised to the multidimensional concerns that AI-based facial recognition poses in the context of providing aid. The laudable aim of saving and protecting vulnerable human lives must not be inadvertently undermined by the use of AI systems that infringe fundamental rights and/or disrespect local culture and custom.

References

1. Centre for Information Policy Leadership: First Report: Artificial Intelligence and Data Protection in Tension (2018).
2. Angwin, J., Larson, J., Mattu, S. and Kirchner, L.: Machine Bias. In ProPublica (2016). Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing, last accessed 2019/8/2.
3. S. Corbett-Davies, E. Pierson, A. Feller and S. Goel: A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. In: Washington Post (2016), https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/, last accesses 2019/10/23.
4. Slim, H.: Humanitarian Ethics - a guide to the morality of aid in war and disaster. 1st edn. Oxford University Press, New York (2015).
5. World Food Program: A statement on the WFP-Palantir partnership (2019), https://insight.wfp.org/a-statement-on-the-wfp-palantir-partnership-2bfab806340c, last accessed 2019/8/2.
6. Council of Europe, Glossary on Artificial Intelligence: https://www.coe.int/en/web/artificial-intelligence/glossary, last accessed 2019/8/5.
7. Kuner, C., Marelli, M.: Handbook on Data Protection in Humanitarian Action. 1st edition, International Committee of the Red Cross, Geneva (2017).
8. Scott, R.: Imagining More Effective Humanitarian Aid A Donor Perspective. In: 18 OECD Development Co-operation Directorate 34 (2014).
9. Bradford, A.: The Brussels Effect. Northwestern University Law Review 107(1), 1-68 (2012).
10. Hashtag Standards for Emergencies (2014). Available at: https://www.unocha.org/sites/unocha/files/Hashtag%20Standards%20for%20Emergencies.pdf, last accessed 2019/8/2.
11. iRevolutions: AIDR: Artificial Intelligence for Disaster Response (2013), https://irevolutions.org/2013/10/01/aidr-artificial-intelligence-for-disaster-response/, last accessed 2019/8/2.
12. Meier, P.: Digital Humanitarians: How Big Data is Changing the Face of Humanitarian Response. 1st edn. CRC Press, Boca Raton (2015).
13. Project Premonition website, https://www.microsoft.com/en-us/research/project/project-premonition/, last accessed 2019/8/2.
14. Slim, H.: Eye Scan Therefore I am: The Individualization of Humanitarian Aid. In: European University Institute (2015), https://iow.eui.eu/2015/03/15/eye-scan-therefore-i-am-the-individualization-of-humanitarian-aid/, last accessed 2019/8/2.
15. Taigman, Y., Yang, M., Ranzato, M. A., Wold, F..: DeepFace: Closing the Gap to Human-Level Performance in Face Verification. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1701–1708. IEEE, Columbus (2014), https://research.fb.com/wp-content/uploads/2016/11/deepface-closing-the-gap-to-human-level-performance-in-face-verification.pdf, last accessed 2019/10/23.
16. Buolamwini, J., Gebru, T.: Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research 81, 1–15 (2018).

17. Hartzog, W.: Facial Recognition is the Perfect Tool for Oppression. In: Medium (2018), https://medium.com/s/story/facial-recognition-is-the-perfect-tool-for-oppression-bc2a08f0fe66, last accessed 2019/5/27.
18. Golumbia, D.: Do You Oppose Bad Technology, or Democracy? In: Medium (2019). https://medium.com/@davidgolumbia/do-you-oppose-bad-technology-or-democracy-c8bab5e53b32, last accessed 2019/10/23.
19. Traffic Jam website, https://www.marinusanalytics.com/traffic-jam, last accessed 2019/8/2.
20. Samsung: Use Facial recognition security on your Galaxy phone, https://www.samsung.com/us/support/answer/ANS00062630/, last accessed 2019/8/2.
21. Samsung: Diebold ATM Samsung SDS Nexsign – Digital Banking (2017), https://www.samsungsds.com/global/en/about/news/1196788_1373.html, last accessed 2019/8/2.
22. Vincent, J.: A photo storage app used customers' private snaps to train facial recognition AI – Photo app Ever pivoted its business without informing users. In: The Verge (2019), https://www.theverge.com/2019/5/10/18564043/photo-storage-app-ever-facial-recognition-secretly-trained-ai, last accessed 2019/8/2.
23. Collie, M.: 'Just walk away from it': The scary things companies like FaceApp can do with your data. In: GlobalNews (2019), https://globalnews.ca/news/5653531/faceapp-data-mining/, last accessed 2019/8/5.
24. Carman, A.: FaceApp is back and so are privacy concerns. In: The Verge (2019), https://www.theverge.com/2019/7/17/20697771/faceapp-privacy-concerns-ios-android-old-age-filter-russia, last accessed 2019/8/5.
25. Olavario, D.: FaceApp: Are security concerns around viral app founded? In: Euronews (2019), https://www.euronews.com/2019/07/17/faceapp-are-security-concerns-around-viral-app-founded-thecube, last accessed 2019/8/5.
26. Milligan, C. S.: Facial Recognition Technology, Video Surveillance, and Privacy. Southern California Interdisciplinary Law Joiner 9, 295-334 (1999).
27. Bowyer, K. W.: Face recognition technology: security versus privacy. IEEE Technology and Society Magazine 23(1), 9-19 (2004).
28. Privacy International: The police are increasingly using facial recognition cameras in public to spy on us (2019), https://privacyinternational.org/feature/2726/police-are-increasingly-using-facial-recognition-cameras-public-spy-us, last accessed 2019/8/1.
29. Frew, J.: How Facial Recognition Search Is Destroying Your Privacy. Make Use Of (2019), https://www.makeuseof.com/tag/facial-recognition-invading-privacy/, last accessed 2019/8/1.
30. Curran, D.: Facial recognition will soon be everywhere. Are we prepared? In: The Guardian (2019), https://www.theguardian.com/commentisfree/2019/may/21/facial-recognition-privacy-prepared-regulation, last accessed 2019/8/1.
31. Mozur, P.: One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority. In: The New York Times (2019), https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html, last accessed 2019/8/2.
32. Lewis, P.: 'I was shocked it was so easy': meet the professor who says facial recognition can tell if you're gay. In: The Guardian (2018), https://www.theguardian.com/technology/2018/jul/07/artificial-intelligence-can-tell-your-sexuality-politics-surveillance-paul-lewis, last accessed 2019/8/2.
33. Cuthbertson, A.:Indian Police Trace 3,000 Missing Children in Just Four Days Using Facial Recognition Technology. In: Independent (2018), https://www.independent.co.uk/life-style/gadgets-and-tech/news/india-police-missing-children-facial-recognition-tech-trace-find-reunite-a8320406.html, last accessed 2019/8/2.

34. Bernal, N.: Facial recognition to be used by UK police to find missing people. In: The Telegraph (2019), https://www.telegraph.co.uk/technology/2019/07/16/facial-recognition-technology-used-uk-police-find-missing-people/, last accessed 2019/8/2.
35. Restoring Family Links website, https://familylinks.icrc.org/en/Pages/home.aspx, last accessed 2019/8/2.
36. Trace the Face – Migrants in Europe website, https://familylinks.icrc.org/europe/en/Pages/Home.aspx, last accessed 2019/8/2.
37. Norton: How does facial recognition work?, https://us.norton.com/internetsecurity-iot-how-facial-recognition-software-works.html, last accessed 2019/8/5.
38. Gellert, R.: Understanding the risk based approach to data protection: An analysis of the links between law, regulation, and risk. PhD Thesis, Vrije Universiteit Brussel (2017).
39. Burrell J.: How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms. Big Data & Society, 1-12 (2016).
40. The Norwegian Data Protection Authority: Artificial intelligence and privacy (2018), https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf, last accessed 2019/8/2.
41. GDPR, Art. 4(14).
42. GDPR, Art. 9.
43. Article 29 Data Protection Working Party: Opinion 3/2012 on developments in biometric technologies (2012), https://www.pdpjournals.com/docs/87998.pdf, last accessed 2019/8/2.
44. Council of Europe: Guidelines on artificial intelligence and data protection (2019), https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8, last accessed 2019/8/2.
45. Article 29 Data Protection Working Party: Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (2018), https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053, last accessed 2019/8/2.
46. Mantelero, A.: Artificial Intelligence and Big Data: A blueprint for a human rights, social and ethical impact assessment. Computer Law & Security Review, 34(4), 754-772 (2018).
47. Kuner, C., Svantesson, D. J. B., Cate, F. H., Lynskey, O., Millard, C.: Data protection and humanitarian emergencies. International Data Privacy Law, 7(3), pp. 147-148 (2017).
48. Sphere: The Sphere Handbook – Humanitarian Charter and Minimum Standardsin Humanitarian Response (2018), https://spherestandards.org/handbook/editions/, last accessed 2019/8/2.
49. ICRC: Restoring Family Links and Data Protection – background document (2019), https://rcrcconference.org/app/uploads/2019/06/33IC-RFL-background-document_en.pdf, last accessed 2019/8/2.
50. McGoldrick, C.: The future of humanitarian action: An ICRC perspective. International Review of the Red Cross 93(884), 965-991 (2011).
51. Martin, A. and Taylor, L.: Biometric Ultimata — what the Yemen conflict can tell us about the politics of digital ID systems. In: Global Data Justice (2019), https://globaldatajustice.org/2019-06-21-biometrics-WFP/.
52. Code of Conduct, https://media.ifrc.org/ifrc/who-we-are/the-movement/code-of-conduct/, last accessed 2019/8/5.