Luiss
School of Government

The Refugees' Crisis and the Ethics of Algorithms

Working Paper Series
SOG-WP5/2024

Authors: Simona Bonini Baldini, Alessandro Picchiarelli
May 2024

LUISS

Table of Contents

1. Abstract

The progressive increase in requests for international protection makes it even more urgent to develop systems capable of processing these large quantities of applications. Algorithmic and artificial intelligence systems represent an answer to this need but, given the advantages they offer and the potential they present, they also bring out ethical challenges and questions. Those cannot be overlooked as they directly affect the dignity of refugees and immigrants, denying or recognizing rights that belong to all human beings and especially to the most disadvantaged.

This digital humanitarianism needs to be understood to ensure that algorithmic systems do not exhibit discrimination, prejudice or bias which can penalize refugees and worsen the humanitarian emergency we are experiencing.

Keywords: Artificial Intelligence – Ethics – Digital humanitarianism – Refugees

2. Introduction

At the end of the Second World War, the Displaced Persons (DPs) in Europe were identified as the protagonists of a problem whose resolution was necessary in order to rebalance an international political order disrupted by the disastrous outcomes of the war[1]. Indeed, at that time, the rehabilitation of the victims of the German occupation was at the heart of the international reconstruction projects promoted by the United Nations. The management of refugee camps and their relocation contributed to redrawing the European geopolitical map, just as the recent refugee crisis, 2015-2016, tested the reception system and highlighted the need to renew the European asylum system. In the new European Migration and Asylum Pact, the management of migration flows is strongly linked to the management of migrant pre-screening procedures and redistribution criteria within Member States. Analyzing the affirmation of new Artificial Intelligence systems in the migratory context therefore means questioning the relationship between the management of new technologies and reception policies aimed at redefining the forces in play both in the territories within member states and at the borders of the European countries most affected by migration routes.

Nowadays, the use of new technologies is a reality in every area of human life. Technology is present in all those contexts in which the analysis of enormous quantities of data makes it necessary to resort to systems that can quickly produce reliable results. Even in the field of management of migration flows, algorithmic and artificial intelligence systems play a fundamental role.

They both control and predict the flows which can reach the national borders of a certain state, to provide appropriate reception policies and allocate sufficient funds to manage these situations. Besides, they play a fundamental role in managing the reception of those refugees and immigrants, who are already in the territory of a state and who may have greater possibilities of integration, if inserted in certain contexts than in others. Understanding the impact that these systems have on the reception policies of these people becomes important to figure out how they work, and what consequences they can have in the lives of refugees and immigrants.

At the same time, however, these systems, due to biases, prejudices or malfunctions, can have a very serious influence on decisions. Even if the final decision is up to humans, at least in the algorithmic systems that will be analysed, it is always true that artificial intelligence systems guide the decisions which humans make, and it is not always possible to understand how this process occurs.

Most of the systems developed aim to support governments in managing the increasingly numerous people requesting protection and reception. Against this backdrop, while recognizing the advantages they

[1] Cf. S. SALVATICI, ***Senza casa e senza paese. Profughi europei nel secondo dopoguerra***, Il Mulino, 2008, P. 37.

present, we must not forget the risks and dangers that these systems can introduce including seriously damaging human dignity and depriving some people of fundamental rights.

All these considerations should help us understand the need for legislative and ethical frameworks which regulate the development, and the use of these systems to guarantee ever greater equity and the pursuit of social justice which protects both immigrants and refugees and the recipient states of these migratory flows.

3. Evaluation of the Migratory Phenomenon

At the end of the Second World War, Europe was a continent of emigration. The enormous presence of refugees and displaced people led to the creation of an international organization for the protection of refugees[2]. Since the creation of United Nations Relief and Rehabilitation Administration (UNRRA), and then the United Nations High Commissioner for Refugees, the task of managing humanitarian interventions towards vulnerable subjects, such as asylum seekers and refugees has opened a wide debate on concepts such as " hospitality", "assistance", "self-determination" and "care". The UNRRA logo, the image of a stylized globe, later taken up by the United Nations, recalled the internationalist spirit and the ambition to take care of all humanity[3]. Despite the ambitions, UNRRA mainly placed the focus of its mission on the old continent, which in September 1945 recorded 1,036, 138 Displaced Persons (DPs) present mainly in Western Germany[4].

The international humanitarianism project, in addition to promoting repatriation through the publication of brochures, posters, flyers, and documentaries screened in the arena and in appeals broadcast on the radio, also proposed the idea of "rehabilitation"[5]. It consisted of restoring the mental, and moral health of the refugees, thus delegating important functions and responsibilities to non-governmental organizations.

Assistance, which is understood not just as relief but as a process of transformation of the conditions of the victims, is present throughout the history of international humanitarianism; from its debut to the contemporary actions promoted by the United Nations High Commissioner for Refugees and organizations non-governmental organizations who are the implementers and spokespersons of this value[6].

Since 1999, the European Union has been working for the creation of a Common European Asylum System (CEAS) which establishes minimum standards for the treatment of all asylum seekers, and all asylum applications in the EU[7]. The new pact on migration and asylum highlights the need for solid and fair management of external borders, including identity, health, and security enquiries; fair and effective asylum rules, and simplification of asylum and return procedures. It also underlines a new solidarity mechanism for search and rescue, pressure, and crises; strengthening crisis forecasting, preparation and response; and an effective return policy coordinated at the European level[8].

[2] Cf. P. Gatrell, ***The making of the modern refugee***, Oxford University Press, 2013, pp. 85-97
[3] S. Salvatici, ***Nel nome degli altri. Storia dell'umanitarismo internazionale***, Il Mulino, 2015, p. 193
[4] P. Gatrell, ***The making of the modern refugee***, Oxford University Press, 2013, p. 97.
[5] S. Salvatici, ***Nel nome degli altri. Storia dell'umanitarismo internazionale***, Il Mulino, 2015, p. 205.
[6] Cf. S. Salvatici, ***Senza casa e senza paese. Profughi europei nel secondo dopoguerra***, Il Mulino, 2008.
[7] Common European Asylum System https://home-affairs.ec.europa.eu/policies/migration-and-asylum/common-european-asylum-system_en
[8] European Commission, ***New Pact on Migration and Asylum***, European Union, 2020.

In 2022, international protection developments have highlighted the importance of having an effective multi-stakeholder protection architecture[9]. Millions of people around the world are forced to move every year due to violent conflict, persecution, human rights abuses, natural disasters, and deterioration of ecosystems. "While the record levels of 2015 and 2016 were driven primarily by applications for international protection from people from Syria, Afghanistan and Iraq, the current increase comes from a much broader range of nationalities. In the biennium 2015-2016, applicants from Syria represented more than a quarter of all applications, while in 2022 a seventh. In 2022, the number of applications submitted by Syrian and Afghan citizens was the highest since 2016, while for several other groups - including Turkish, Venezuelan, and Colombian citizens - it was the highest ever recorded"[10].

Resettlement and complementary reception paths are an expression of international solidarity, which involves various national and international actors. EU resettlement programs are based on national commitments corresponding to the number of third-country nationals that the Member States commit to welcoming the requests. The aim is to manage migration based on predictable timelines and ensure common grounds for reception while carrying out rigorous security controls[11].

In this context, the growing access to the Internet by humanitarian agencies, and the increasing pervasiveness of mobile phone use throughout the migration journey, has highlighted the possibility of tracking the movements of migrants and asylum seekers and increasing the possibility of predicting migratory flows, and consequently, organizing the resources needed to provide adequate reception in the destination countries. The response of governments and humanitarian agencies to the refugee crisis, in the 2015-2016 period, was to apply technological innovation such as artificial intelligence systems to the study and management of migratory flows. This has led to a change of perspective compared to previous reception and rehabilitation practices, introduced since the creation of UNRRA, and opened a broad debate on the role of technology as a socio-economic practice – which at his core involves rethinking the very concept of human beings.

On the one hand, the growing use of the Internet through increased mobile phone use and access to networks by migrants has highlighted the possibility for migrants themselves to express their agency by expanding the possibility of communicating outside and inside their country of origin. On the other hand, however, the increase in the production of digital data produced by migrants during their migratory journey and the spread of biometric data collection practices at the border have reinforced the practice of digitising the migrant's body, risking obscuring the person with his or her history, emotions and motivations that prompted him or her to leave their country of origin. Understanding therefore how technology and new AI systems are used within the migratory context means asking if and how it is possible to employ these technological systems in such a way that the migrant is not de-humanised, but on the contrary placed at the centre of a rethinking of the reception system itself through the use of technology. The technological spaces in which the reception of migrants is organised, together with the preparation of the actors involved in this process, play a key role in determining the process of transformation of the conditions of asylum seekers and refugees, from victims to people capable of

[9] Cf. EUAA, ***Asylum Report 2023. Annual Report on the Situation of Asylum in the European Union***, European Union Agency for Asylum, 2023, p. 3.
[10] EUAA, ***Asylum Report 2023. Annual Report on the Situation of Asylum in the European Union***, European Union Agency for Asylum, 2023, pp. 85-86.
[11] Cf. EUAA, ***Asylum Report 2023. Annual Report on the Situation of Asylum in the European Union***, European Union Agency for Asylum, 2023, p. 277.

expressing their agency, to personally contribute to transforming reception policies from emergency ones into policies centred on human development[12].

[12] For further information on this topic, see A. K. Sen, ***Readings in Human Development***, Oxford University Press, 2003.

4. Digital Humanitarianism: How Humanitarian Reception is Changing in the Artificial Intelligence Era

Over the last decade, the concept of "digital humanitarianism"[13] emerged due to the development of new technologies and the widespread diffusion and collection of digital data on migrants.

Digital humanitarianism uses new media to communicate its services. New media are defined as digitalization processes, whose main channel of communication is the Internet, influenced by the introduction of smartphones and which allow users to express themselves in a variety of ways using written words, sounds, images and moving images. It refers also to the space where social entrepreneurs and technology companies collaborate internationally to generate, share, and spread good practices on digital applications for migrants and refugees. In recent years, applications have been developed that cover all phases of the refugee reception process: arrangement and accommodation, education and school, the course of time and preparation for work, health, cooperation with the local community, reunification with separated family members from conflicts or disasters, education for unaccompanied minors. The applications were created by NGOs, public institutions - such as the German government's ***Ankommen*** meaning "arriving" - and refugees themselves sharing their experiences of being refugees in mobile applications such as ***Bureaucrazy*** and ***Gherbetna***[14].

Digital humanitarianism is also exemplified by «web mapping initiatives such as Humanitarian OpenStreetMap Team and Ushahidi, in which large numbers of volunteers collaboratively process and map data collected by humanitarian agencies. The Humanitarian OpenStreetMap Team[15], for example, is an online community which collaboratively maps humanitarian crisis zones; Ushahidi[16] is a website which collects and maps social media and SMS messages in similar contexts» [17].

Existing research indicates the need to further understand digital humanitarianism as a set of socio-political practices around data and to examine how it plays out in the political economy of humanitarian

[13] M. Benton – A. Glennie, ***Digital Humanitarianism. How Tech Entrepreneurs Are Supporting Refugee Integration***, Migration Policy Institute, 2016, p. 1.
[14] Cf. M. Benton – A. Glennie, ***Digital Humanitarianism. How Tech Entrepreneurs Are Supporting Refugee Integration***, Migration Policy Institute, 2016, pp. 3-5.
[15] https://www.hotosm.org/impact-areas/refugee-response/
[16] https://www.ushahidi.com/
[17] R. Burns***, Digital Humanitarianism and the Geospatial Web. Emerging Modes of Mapping The Transformation of Humanitarian Practices,*** in ***Dissertation thesis of Doctor of Philosophy,*** University of Washington, 2015***,*** p. 3.

interventions. Recent growth in data and available processing power have significantly increased the number of applications of artificial intelligence in daily life and the international migration process[18].

Among the applications aimed at enhancing the control of public opinion, there is the United Nations "Global Pulse" program which in Uganda has experimented with «a toolkit that makes public radio broadcasts readable automatically through the use of speech recognition and translation tools that transform radio content into text. This tool, developed by the Pulse Lab in Kampala, aims to identify trends among different population groups, especially those in rural areas. The mindset behind the initiative is that these trends could provide the government and development partners with a better understanding of public opinion on the country's development needs, which could be considered when implementing development programs» [19].

Another example is the Indian project "National Tracking System for Missing & Vulnerable Children"[20] aimed at identifying and locating missing children. Thanks to the project, almost 3,000 missing children were identified in the four days following the launch of a trial of a facial recognition system that matches the faces of missing individuals to photographs of children living in retirement homes and orphanages. The visual comparison system has also been used to track attacks on civilians and human rights violations.

Amnesty International's "Decode the Difference"[21] project recruited volunteers to compare images of the same location from different periods to identify damaged buildings, which may demonstrate systematic attacks against civilians. In the future, the data could be used to train machine learning tools to analyse images, thus speeding up the process and increasing operational capacity.

Some international organizations already use machine learning and artificial intelligence alongside biometric technology. The IOM has launched the Big Data for Migration Alliance[22], which aims to explore the use of technology in international migration. UNHCR has been using technology for migration management since 2002 when it developed the IT case management tool Profile Global Registration System (proGres)[23].

AI-based technological applications have thus «the potential to revolutionize the way states and international organizations seek to manage international migration and can impact international migration management in different ways»[24]. They can modernize the traditional practices of states and international organizations as well as reinforce contemporary requests for more evidence-based migration and border security management.

On one hand, it has been noticed how artificial intelligence, in the context of digital humanitarianism, can facilitate humanitarian work and related activities by making them more effective and efficient. On the other, however, the risk of making stricter requests for migration and border security management has

[18] For further information on this topic, see United Nations, ***Use of new data sources for measuring international migration***, United Nations, 2022; European Migration Network, ***The use of digitalisation and artificial intelligence in migration management***, European Migration Network, 2022; J. Bither – A. Ziebarth, ***AI, digital identities, biometrics, blockchain. A primer on the use of technology in migration management,*** Migration Strategy Group, 2020; D. Ozkul, ***Automating Immigration and Asylum. The use of New Technologies in Migration and Asylum Governance in Europe,*** University of Oxford, 2023.
[19] C. Kuner, M. Marelli (eds), ***Handbook on data protection in humanitarian action,*** Brussels Privacy Hub and Data Protection Office of the International Committee of the Red Cross, 2020, p. 279.
[20] https://trackthemissingchild.gov.in/trackchild/index.php
[21] https://citizenevidence.org/2020/10/09/designing-amnesty-decoders-how-we-design-data-driven-research-projects/
[22] https://data4migration.org/
[23] https://www.unhcr.org/registration-guidance/chapter3/registration-tools/
[24] A. Beduschi, ***International migration management in the age of artificial intelligence***, in ***Migration studies*** 3, 2021, p. 576.

emerged. These are based mainly on the collection of data, neglecting the person as such, composed not only of a digitized body but also of an ability to express active thought.

5. Artificial Intelligence in Relocation Management: Annie MOORE, GeoMatch, EUMigraTool

When refugees and asylum seekers arrive in a host country, several factors can significantly impact their well-being and integration possibilities. For example, the availability of adequate housing, the possibility of finding employment and taking advantage of educational opportunities, the possibility of relying on existing support networks and of being welcomed into a non-discriminatory living environment. Typically, resettled refugees or asylum seekers are distributed by administrators or reception staff to specific geographical locations, often based on the reception capacity of the existing community, or according to allocation rules predefined by individual states[25]. Depending on the host country's situation and the distribution system used, assignments can be a very lengthy process in what is already an area with often limited staffing capacity.

To overcome this problem, algorithmic systems have been developed. They match refugees and asylum seekers arriving through resettlement with an "optimal placement" with places or communities where they are most likely to find work, with the main objective to "improve integration" through "algorithmic assignments"[26].

The two experimental projects, aimed at proposing an alternative solution to the usual assignment of places, designed to develop the well-being of refugees and asylum seekers, are the Annie MOORE system used in the United States, and GeoMatch of the Immigration Policy Lab, tested both in Switzerland and in the United States. The Annie MOORE (Matching and Outcome Optimization for Refugee Empowerment) system "is named after the first immigrant to pass through Ellis Island in the United States. It was started in 2018 as a collaboration between academia, the non-profit resettlement agency HIAS, and the U.S. Department of State"[27] It recommends locations where newly arrived refugees are most likely to find work based on their profile. The system is based on open-source software, and the developers claim that it can be updated, adapted, and replicated in other contexts[28].

The GeoMatch algorithm, developed by the Immigration Policy Lab and tested since 2020 in the United States and Switzerland, works similarly. GeoMatch predicts the probability of finding a job in different locations in the host country and recommends the "optimal" location for the newcomer. A human operator can confirm or modify this recommendation. GeoMatch in Switzerland is jointly coordinated by the Immigration Policy Lab, stakeholders in the placement process and the Swiss State Secretariat for

[25] Cf. J. Bither – A. Ziebarth, ***Automating decision-making in migration policy. A navigate guide***, Migration Strategy Group, 2021, p. 21.
[26] J. Bither – A. Ziebarth, Automating decision-making in migration policy. A navigate guide, Migration Strategy Group, 2021, p. 21.
[27] J. Bither – A. Ziebarth, Automating decision-making in migration policy. A navigate guide, Migration Strategy Group, 2021, p. 23.
[28] Cf. A. C. Trapp – al., ***Placement Optimization in Refugee Resettlement***, in ***Working Paper*** 23, 2018.

Migration (SEM). The GeoMatch code is public, and the data can be requested from the SEM for research purposes.

Annie MOORE users claim that it arrives at an optimal location result/preference six times faster than a human. Karen Monken, director of reception and initial resettlement at HIAS, says that "the effectiveness of my operations has increased dramatically. I now spend 80% less time on routine matching and can focus my time and energy on the most difficult cases, such as those with significant medical conditions, ensuring their placement is as good as possible"[29]. At the same time, employment opportunities have increased by 30% compared to manual placement. According to GeoMatch's developers, the algorithm produced results that improved the employment prospects of refugees in the United States by about 40% and in Switzerland by 75%. The Swiss model involves an evaluation and has included a randomized control trial that will provide further results[30].

The ITFLOWS project, funded by Horizon 2020, "was born as a response to the challenge of using computational predictive tools to anticipate and develop responses to the arrival of migrants in the European Union (EUMigraTool - EMT). The project worked closely with those promoting the successful integration of newly arrived migrants, asylum seekers and refugees who decide to stay in a particular EU country. It is an evidence-based ICT solution, using simulation and artificial intelligence models previously approved and under constant review by members of the ITFLOWS User Board"[31]. One of the main objectives of ITFLOWS is to provide a useful tool to produce accurate forecasts of the number of arrivals in Europe. These forecasts can also be used by associations and first and second reception agencies to more accurately organize services for migrants and asylum seekers once they arrive in their new destination country. The added value that this project intends to bring, within the framework of migration flow forecasting systems, is to involve civil society and organizations responsible for managing the reception of migrants and asylum seekers, in evaluating the actual functioning of the system. The active participation of civil society in the evaluation of this predictive system is useful not only to verify its accuracy but also to activate a shared debate on migration policies and the use of new technological systems to improve conditions, both for the receiving society and for new arrivals.

[29] J. BITHER – A. ZIEBARTH, ***Automating decision-making in migration policy. A navigate guide***, Migration Strategy Group, 2021, p. 25.
[30] Cf. K. BANSAK – al., ***Improving refugee integration through data-driven algorithmic assignment***, in ***Social Science*** 359, 2018, pp. 1-4.
[31] D. M. GONZÁLEZ – N. GKEVREKIS, ***EUMigra Tool Report 3,*** in ***ITFLOWS Project***, 2023, pp. 4-8.

6. Outstanding Questions on the Use of Artificial Intelligence in the Humanitarian Context

The considerations made allow us to understand that artificial intelligence algorithmic systems are nowadays a great opportunity for European States to control who crosses their borders and to manage those who have had access to their territories. There are multiple advantages of the use of these technologies: decision-making processes are speeded up, large quantities of data can be managed, and opportunities for programming and managing migratory flows can be obtained which would otherwise be obtained with great difficulty.

However, there are many questions and ethical challenges which arise from the use of these artificial intelligence systems. How does a human being get involved in the decision-making process? Human operators have the last word on the placement of a refugee or asylum seeker but, as in other ADM systems, "decision fatigue" and "automation biases"[32] must be taken into account, while humans rely more or less blindly on the recommendations of the algorithmic model. Prejudices and discrimination can corrupt the quality of the data being processed, so it is necessary to ask: is the model controlled to identify any prejudices and/or could it perpetuate systemic discrimination?

Furthermore, in the case of the use of algorithmic systems, all subjects involved must have the right level of information in the use of their data. The subjects, whose data have been used for the training model, have been informed or were they asked for permission even though their data was anonymized? Have the refugees or asylum seekers been informed of the use of an algorithmic support system for a placement suggestion?

Algorithmic decision-making systems in the migration space are set to increase, while there is still no significant oversight and understanding of the long-term effects on the world of migration[33]. Questioning the long-term effects of the use of AI in the migration context means considering the results from the use of such systems as generating further data, which if extrapolated from the socio-economic context can give the impression of providing neutral information about people, whereas we know that no data is neutral and that it needs to be linked to a cultural and historical context.

These and many other questions must guide the ethical reflection on these systems.

[32] J. Bither – A. Ziebarth, ***Automating decision-making in migration policy. A navigate guide***, Migration Strategy Group, 2021, p. 25.
[33] Cf. J. Bither – A. Ziebarth, ***Automating decision-making in migration policy. A navigate guide***, Migration Strategy Group, 2021, p. 36.

7. Some Ethical Considerations on the Use of Artificial Intelligence

To understand the impact that artificial intelligence systems have on human life, we must keep in mind how «technology and its artefacts are not just elements of the human world that question our actions but are anthropological places where our very constitution is revealed to us»[34]: technological artefacts appear, therefore, as cultural products of a hermeneutic process of understanding the world, of one's nature and of how man adapts to the context in which he lives, shaping it and allowing himself to be shaped by it[35].

The algorithmic systems introduced in the previous paragraphs belong to the category of profiling algorithms. They are algorithms that characterize the typical behaviour of a person or a group of people, starting from the data known about the subject under examination. This data can be of different types: tastes, political, sexual, religious orientation, places frequented, friendships and so on. Starting from the knowledge of these parameters it is possible, within certain limits, to predict the behaviour of some subjects or to classify and divide the subjects into groups with similar characteristics. This type of algorithm can work in three different ways. In the first case, we try to analyse the relationships existing between the behaviour of the sample under examination, and the characteristics of a known sample. The goal is to understand whether the sample under examination will have similar behaviour to that of the known sample and therefore whether it can be classified and assigned to a certain category or another. In the second case, we try to investigate an event and the relationships between the causes that provoked it, so that it will be possible to statistically predict when this event will repeat itself. Finally, in the third case, a large amount of data and complex analysis algorithms are combined to predict the behaviour of a certain group of subjects[36].

Algorithmic systems for managing immigrants and refugees predominantly work according to the first mode. This is the case of Annie MOORE and GeoMatch while ITFLOWS brings together the characteristics of the three operating models.

We can immediately highlight a few problems. Acting on a user's sensitive data to provide a profile and to make hypotheses about his actions can create discrimination against some subjects who statistically could be more at risk than others or can generate the denial of services and rights which are instead theirs, such as reception in a state or assignment to a specific territorial area. Since these systems are protected by intellectual property rights, transparency, which is necessary to understand why a user is classified in a

[34] P. Benanti, ***La condizione tecno-umana. Domande di senso nell'era della tecnologia***, EDB, 2017, p. 131.
[35] Further insights into the historical evolution of the technological artefact and the automation of human capabilities can be found in I. McNeil, ***An Encyclopedia of the History of Technology***, Routledge, 1990, p. 1-43.
[36] Cf. A. Zhang, ***Data Analytics***, CreateSpace Independent Publishing Platform, 2017, p. 106-110.

certain way compared to others or why he is assigned to a certain territory instead of another, is not always guaranteed. Furthermore, from a strictly technical point of view, the preparation of the data to be provided to the algorithm could be affected by a series of errors that could compromise the correct functioning of the profiling and decision-making process. One can think about, for example, the databases that are too small. They do not allow accurate predictions or the presence of many borderline or particular cases within the available records. Think about the unavailability or partiality of some information which makes profiling or prediction less objective[37] or the malfunction of these systems which could recognize a subject belonging to a certain group, which does not necessarily coincide with the one to which he belongs. Furthermore, computer algorithms inevitably influence the decisions made by human, in a way in which he is often not even aware: the pervasiveness and speed with which computer algorithms have entered the lives of human beings makes this reality an integral part of his experience in the world, not only in an instrumental way, but as an essential part of his ability to interpret the reality which surrounds him. This property of the computer algorithm can be defined as empathy and implies a second property, the reliability, that computer algorithms understood not so much in operational terms, but as the ability to trust and rely on this artificial reality.

Lastly, by acting on human decision-making processes, computer algorithms, in addition to touching personal individuality, also act on a social level as they generate economic, political, and financial control mechanisms that regulate human social living and that determine the understanding that the community has of the common good.

Given these quick considerations and the analysis of the three algorithmic systems that were previously introduced, four ethical issues appear evident: algorithmic transparency, prejudices, system biases, recognition of responsibility for any incorrect decisions.

[37] Cf. A. Rezzani, ***Big data analytics***, Maggioli Editore, 2017, p. 312-314.

8. Algorithmic Transparency

The complexity of the algorithms that have been produced and presented makes it increasingly difficult to control, monitor and possibly correct, the functioning of these systems. In general, transparency is related to the availability of operating information and the possibility of accessing it[38]. This means that in an algorithm with a high degree of transparency, it is possible to sufficiently control the type of input, the intermediate steps and the results that are produced. While, in a non-transparent, also called opaque algorithm, this is not possible.

However, today this information is less and less available, both for competitiveness and security and privacy reasons[39], as well as for the increased degree of complexity that algorithms have reached. This reduced transparency generates a certain level of uncertainty, and, because of that, the behaviour of the computer algorithm is no longer perfectly predictable. Within this space of indeterminacy, where multiple operational choices are possible for the algorithm due both to the rules established by the programmer and to those that the algorithms give themselves by learning from the data they have available, come into play values and biases. They are not always immediately traceable by humans and are partly not linked to the developers but to the learning of the algorithms themselves[40] and therefore to the data provided in the training phase.

Transparency should not be confused with the possibility of being able to manipulate the algorithm itself. In fact, a good algorithm must always guarantee a certain level of robustness against external intervention. This discussion can be applied to the three systems presented: at least Annie MOORE and GeoMatch are open-source systems, so we manage to have a good level of transparency. In the case of GeoMatch, it is also possible to access the data used and produced for research and analysis purposes. However, it is not always as immediate as it is for other types of systems where the degree of transparency is much lower.

Algorithmic transparency allows us to discover most biases that may be present in algorithms and prevent disadvantageous situations for humans.

[38] Cf. I. RAHWAN – ***al.***, ***Machine behaviour***, in ***Nature*** 568 (2019), p. 478.
[39] Cf. B.D. MITTELSTADT – al., The ethics of algorithms: mapping the debate, in Big Data & Society 3 (2016) p. 6-7.
[40] Cf. C. ACCOTO, ***Il mondo dato: cinque brevi lezioni di filosofia digitale***, Egea, 2017, p. 64-68.

9. Prejudices in Algorithmic Systems

In general, algorithms are characterized by a certain degree of objectivity. This is true in classical algorithms, while it becomes less and less evident in machine learning or profiling algorithms and artificial intelligent systems[41]. In fact, algorithms are inevitably affected by the prejudices and sensitivity of those who developed and designed them and of those who use them, but also by the social prejudices of the context in which they work[42]. A recent MIT study has also shown that artificial intelligences are not only affected by programmers' prejudices but are able to independently develop new ones that undermine the fairness, robustness, and objectivity of the implemented algorithm[43]. This is because, if learning is not sufficiently equitable, human prejudices are amplified.

The learning phase via training set, therefore, must be designed in a way to avoid biases and prejudices in the data. For this reason, approaches which try to make algorithms increasingly fair have been defined. In this context, fair does not actually indicate being fair or equitable, but is used as an acronym to outline four fundamental properties which a training set should have in order not to be intentionally discriminatory. The first property is that the training set system must be ***Fundamentally sound*** because it must be representative of the entire set to which the data refers. In statistical terms, it is said that the data set must be independent and identically distributed (IID). Otherwise, a certain group of data is favoured over others, and this jeopardizes the fair functioning of the algorithm. In the case of the systems under examination, with all the difficulties that may be encountered, linked both to the lack of exhaustive information on the identity and status of the migrant or refugee and the verification of the goodness and truthfulness of these data, it may happen that the results products are falsified. This non-adherence to reality will constitute a learning basis for the algorithm itself and this will amplify the non-objectivity of the data and any prediction and profiling errors that may be made by the system.

The second property is that the system must be ***Assessable*** which means that it must be possible to evaluate how the system makes its decisions. In other words, the system must guarantee a certain level of algorithmic transparency. We have already seen that, within certain limits, this property is quite respected by the algorithms we have analysed.

The third property is that the system must be ***Inclusive*** because it must not discriminate some types of data over others. It may happen, in fact, in some social and cultural contexts that some information

[41] For further information on this topic, see A. Picchiarelli, ***Tra profilazione e discernimento. La teologia Morale nel tempo dell'algoritmo***, Cittadella Editrice, 2021.
[42] Cf. B.D. Mittelstadt – ***al.***, ***The ethics of algorithms: mapping the debate***, in ***Big Data & Society*** 3 (2016), p. 7-9.
[43] For further information on the topic, see the article R.M. Whitaker – G.B. Colombo – D.G. Rand, ***Indirect Reciprocity and the Evolution of Prejudicial Groups***, in ***Scientific Reports*** 8 (2018).

regarding the status of a refugee or immigrant may preclude certain rights or certain reception mechanisms that must instead be recognized. This must be avoided by trying to define a training set that tries to be as inclusive as possible.

Lastly, it must be ***Reversible***, namely, it must be able to be re-evaluated if the exclusion of a certain group of data compared to others is perceived. Also in this case, in the examples shown, human intervention is essential, and the final decision is made by humans to guarantee this additional control. However, it remains significant to inform people of the use of an algorithmic decision-making system and, in case of doubts about the validity of the decision taken by the artificial intelligence system, to resort to verification and control procedures so as not to damage human dignity. In this sense, the system must be able to be considered reliable by its users and by those who are managed using this type of procedure[44].

A final aspect to consider in the analysis of algorithmic discrimination is the fact that data, in addition to being influenced by possible biases and by the algorithms that produce or process them, are also affected by the effects and consequences of human choices. In fact, they can be seen as a history of past choices: every piece of data always refers to a decision that someone has made, and any choice that man makes is always the result of a decision-making process that is not independent of the past choices that were made. These choices are political, social, cultural, and economic and therefore they are never neutral: it follows that the data are never just informative elements but reflect this individual and social complexity that every decision brings with it, and which must be taken into account especially when it has to do with decisions that affect the future of human beings and the recognition or denial of their rights[45].

[44] Cf. B. PURCELL, ***The ethics of AI***, 2018, https://www.forrester.com/report/The-Ethics-Of-AI-How-To-Avoid-Harmful-Bias-And-Discrimination/RES130023.
[45] Cf. M. AIROLDI – D. GAMBETTA, ***Sul mito della neutralità algoritmica***, in ***The Lab's Quarterly*** XX (2018), p. 27.

10. System Biases

From the previous paragraph, we understand that systems trained with data sets in which non-objective information is present can show biases if questioned: this generates system bias. As a result, data collection and segmentation, as well as its truthfulness, are a critical aspect of machine learning and artificial intelligence systems.

A computer algorithm can present three types of bias. The first type consists of prejudices that arise from values and attitudes which exist from the design of the artefact itself. It can be individual, namely, linked to the person who designs the algorithm or who uses it, or social, linked to the social context, political and cultural in which the system operates.

Then, there are the technical biases which depend on the technological limits that today's development allows us to reach.

Finally, there are emerging biases that do not derive from the intentions of the designer, but from the users and the experience that the artefact acquires. This last type of bias is often difficult to identify because how an algorithmic system is understood by its users does not always coincide with the intention of those who designed it and we do not always, in the different contexts in which the same system is used, have the same pre-understandings and the same understanding of man, of his rights and his dignity.

For this reason, once again, guaranteeing a good level of transparency allows us to understand what values are at stake and to verify that these are respected and promoted[46]. It is no coincidence that the algorithmic systems analysed are open source: this, within certain limits, should guarantee greater adaptation to the various contexts and above all to the local legislation applied. However, this can obviously also have negative implications, especially when the adaptation of the functioning of the system introduces social inequalities and amplifies the already existing discrimination between those who ask for hospitality and those who should welcome these people.

[46] Cf. P. BREY, ***Values in technology and disclosive computer ethics***, in L. FLORIDI, ed., ***The Cambridge handbook of information and computer ethics,*** Cambridge**,** 2010, p. 49-51.

11. Recognition of Responsibility

Traditionally, when an algorithm failed to perform its function, the responsibility was immediately placed on its producer or developer. This argument, valid for classical algorithms, no longer automatically applies to all algorithms. Complex algorithmic systems, such as artificial intelligence systems, in fact, do not allow sufficient control of the machine's actions to guarantee that the responsibility belongs entirely to the manufacturer. If we refer to systems for the management of immigrants and refugees, responsibility becomes a very important issue. This is because, establishing who can be welcomed into a certain state and who must be rejected, necessarily involves the recognition or denial of some rights that belong to these people. It is therefore necessary to understand who can be responsible for this type of choice, keeping in mind that the final decision in these systems is up to a human.

The distance between the intention of the producer and the behaviour of the algorithm creates an ***accountability gap*** in which it is possible to reflect on the type of responsibility that belongs to the computer algorithm. This has led some philosophers to talk about a certain degree of moral responsibility to be associated with the computer algorithm[47]. This reflection refers to the classic concept of ***voluntarium in causa*** according to which the effect of an action is caused even though it is not the true purpose of the action itself, but only an unwanted and often unpredictable consequence[48]. From this principle, it follows that the agent is responsible for the effect of the action only if the bad effect was reasonably foreseeable and if it was morally possible to avoid bringing in the cause.

In the examined cases, where a non-personal agent cooperates with personal agents, the situation is further complicated because the unpredictability of the algorithm's behaviour makes it very difficult to establish whether an effect could have occurred or not, and therefore it is not immediate to be able to attribute full responsibility to some of the agents who carried out the action.

For this reason, it is possible to define a type of responsibility in which the agent influences the choices of another agent. This does not diminish the responsibility of the agent who performed the action but determines some responsibility also for the agent who influenced him in his behaviour and choices[49]. This

[47] Cf. B.D. MITTELSTADT – ***al.***, ***The ethics of algorithms: mapping the debate***, in ***Big Data & Society*** 3 (2016), p. 10-12.
[48] Referring to I-II, q. 77, a. 7 of T. D'AQUINO, Somma di teologia, II, Città Nuova, 2018, p. 786-788, we can state that we speak of voluntary action when the will acts on the cause but not on the effect. This means that the negative effect is not deliberately sought and caused but derives accidentally and unpredictably from the cause put in place. According to Herbert Kramer, speaking of an accidental effect means speaking of all those effects that do not necessarily occur when an action is posed that the will directly seeks since an action does not always correspond to only the hoped-for effect. This uncertainty can be linked, according to Kramer, both to the indeterminate nature of the tools used in the action and to natural causes that could not be known a priori in an objective manner (Cf. H.G. KRAMER, ***The indirect voluntary or voluntarium in causa***, Universitas Catholica America Washingtonii, 1938, p. 62-63). The indeterminacy of computer algorithms fully falls into this discussion.
[49] Cf. C.F.R. ILLIES – A. MEIJERS, ***Artefacts, Agency, and Action Schemes***, in P. KROES – P.P. VERBEEK, ed., ***The Moral Status of Technical Artefacts***, XVII, Springer, 2014, p. 171-174.

type of responsibility can be associated with the computer algorithm. In fact, at the current state of technological development, it cannot yet be recognised full responsibility for an artificial agent, but at least it can be recognized accountability for the influence it has had on the human agent. This approach helps to reflect on the fact that it is not so immediate to be able to associate full responsibility only with human beings but it may happen that human responsibility can be attenuated by the influence that artificial agents have in the human decision-making process. Clearly, this does not imply holding an artificial intelligence system guilty for an effect caused by an action but it means attributing to it a non-negligible role in the way in which man understands reality and responds accordingly to the stimuli he receives from it. Furthermore, in this approach, it becomes increasingly urgent to reflect on this relationship of ever closer cooperation between man and artificial intelligence, so much so that new typologies of subjectivity are starting to be introduced. These subjectivities try to describe this reality, taking into account the characteristics of each agent and the degree of responsibility that today, and in the near future, will be associated with each of them[50].

[50] An example of subjectivity is ***Homo Algorithmus*** described in A. PICCHIARELLI, ***Tra profilazione e discernimento. La teologia morale nel tempo dell'algoritmo***, Cittadella Editrice, 2021, 151-154.

12. Conclusions

The widespread use of artificial intelligence systems and the enormous impact they have on human life makes clear the need to adequately understand them and develop ethical and regulatory systems that protect human dignity. This urgency becomes even more evident when it comes to disadvantaged categories of people, such as immigrants or asylum seekers. What appears clear is that in every phase of their existence, development, implementation and use, algorithmic systems must be designed in such a way as to avoid any type of discrimination and prejudice. It is key to try to eliminate system malfunctions or biases both in the data and in the algorithms themselves. In fact, these vulnerabilities can facilitate mistakes that jeopardize an objective analysis of the situation of each individual person requesting hospitality and also the type of placement that each person welcomed must have to facilitate their social integration. All this can be achieved by ensuring a sufficient level of transparency which helps all people involved to adequately understand the decision-making process carried out by the algorithmic system.

What must never be forgotten is that, if technological development is not adequately understood and regulated, the technological artefact can become an instrument of power, even more relevant than the political one[51], of strong discrimination and dehumanization, instead of a means for achieving authentic integral human development. Technological artefacts, in fact, express man's attitudes towards the world and the way in which these attitudes are embodied in every culture, meaning what the vision of the common good that every society possesses, thus determining its life orientation, the relationships between the subjects and the expectations experienced[52]. Technological development is never an independent force that influences society from the outside but is a social activity that reflects the particularities of its being located in a place and time[53]. All of this must be guided and supported by an understanding of the values involved not as something to be exchanged for technological efficiency, but as potentials that can guide future development[54].

In these situations, therefore, what really matters is not so much the instrumental vision of the algorithmic system but rather the vision of the future, the understanding of what is human and the direction we want to give to society: digital humanitarianism must rediscover that. When technical progress is not guided by a search for the common good and by the assumption of morally qualified values, it cannot be considered

[51] Cf. A. Feenberg, ***Tecnologia in discussione. Filosofia e politica della moderna società tecnologica***, Rizzoli, 2008, p. 155.
[52] Cf. P. Benanti, ***The cyborg: corpo e corporeità nell'epoca del post-umano***, Cittadella Editrice, 2012, p. 417-421. See also A. Picchiarelli, ***Tra profilazione e discernimento. La teologia Morale nel tempo dell'algoritmo***, Cittadella Editrice, 2021, p. 281.
[53] Cf. P. Benanti, ***The cyborg: corpo e corporeità nell'epoca del post-umano***, Cittadella Editrice, 2012, p. 338.
[54] Cf. A. Feenberg, ***Tecnologia in discussione. Filosofia e politica della moderna società tecnologica***, Rizzoli, 2008, p. 263.

development as it exposes man to a blind arbitrariness that dehumanizes him and makes him a problem to be faced rather than an opportunity to discover[55].

Acknowledgements

Il presente paper è stato realizzato nell'ambito del progetto "Geopolitica del Digitale", promosso dalla Fondazione Med-Or, in collaborazione con il Center for International and Strategic Studies (CISS) della Luiss Guido Carli, grazie al sostegno della Fondazione Compagnia di San Paolo all'interno del bando "Geopolitica e tecnologia.

[55] Cf. A. Picchiarelli, ***Tra profilazione e discernimento. La teologia Morale nel tempo dell'algoritmo***, Cittadella Editrice, 2021, p. 284.

13. Additional Information

Simona Bonini Baldini, Alessandro Picchiarelli – Paper: The refugees' crisis and the ethics of algorithms.

Information about the author:

Simona Bonini Baldini, Independent researcher

Alessandro Picchiarelli, Independent researcher and Professor

Institutional affiliation:

Pontifical Urbaniana University

Full postal and email address:

Via B. Cairoli, 38 – 06034 Foligno (PG)

Via Urbano VIII, 16 – 00165 Roma

bonini.simona@gmail.com

a.picchiarelli@urbaniana.edu, apicchio2003@gmail.com

About Luiss School of Government

The Luiss School of Government (SoG) is a graduate school training high-level public and private officials to handle political and government decision-making processes. It is committed to provide theoretical and hands-on skills of good government to the future heads of the legislative, governmental and administrative institutions, industry, special-interest associations, non-governmental groups, political parties, consultancy firms, public policy research institutions, foundations and public affairs institutions. The SoG provides its students with the skills needed to respond to current and future public policy challenges. While public policy was enclosed within the state throughout most of the last century, the same thing cannot be said for the new century. Public policy is now actively conducted outside and beyond the state. Not only in Europe but also around the world, states do not have total control over those public political processes that influence their decisions. While markets are Europeanised and globalised, the same cannot be said for the state.

The educational contents of the SoG reflect the need to grasp this evolving scenario since it combines the theoretical aspects of political studies (such as political science, international relations, economics, law, history, sociology, organisation and management) with the practical components of government (such as those connected with the analysis and evaluation of public policies, public opinion, interests' representation, advocacy and organizational leadership).

For more information about the Luiss School of Government and its academic and research activities visit. www.sog.luiss.it

Luiss
School of Government

Via di Villa Emiliani 14
00197 Roma
T +39 85 225052
sog@luiss.it