OPINIO JURIS inComparatione

Studies in Comparative and National Law

CONTENTS

Articles

Section 1

Section 2

ARTIFICIAL INTELLIGENCE AND THE GOVERNANCE OF MIGRATION: POTENTIALITIES AND PITFALLS BETWEEN TECHNOLOGICAL NEUTRALITY AND POLITICAL DESIGN

Simone Penasa[*]

Abstract

The article provides a general framework, in a comparative and multilevel perspective, of the current technological and legal state of the art of the use of systems equipped with AI in the sphere of migration policies. Taking the case of "smart borders" as a paradigmatic sphere, the potentials and risks deriving from such an intertwining are analysed, proposing, in the light of the current normative initiatives adopted at a comparative level, some possible regulatory tools, with the general objective of guaranteeing the integration between the efficiency of technological tools implemented and international standards of protection of fundamental rights.

Indice Contributo

[*] Simone Penasa is Assistant Professor of Comparative Public Law at the Department of Legal Sciences – University of Trento. simone.penasa@unitn.it.

Keywords

Artificial intelligence – Smart borders – Immigration – Fundamental rights – European Union

1. AI-based tools vis-a-vis migration management: the need for a critical and 'neutral' assessment

The use of autonomous systems based on artificial intelligence has now become part of everyday life in contemporary societies, as well as of the concrete execution of the functions that are traditionally associated with the State. Legal scholarship has begun to critically analyse the possible positive effects, as well as the criticalities, which can be produced by this intertwining, focusing in particular on the effectiveness of safeguards for fundamental rights of the people involved. When they are linked to the public dimension, AI-based systems are usually adopted in order to make more efficient, fast and less expensive the concrete implementation of functions such as the administrative, judicial, social and policing ones.

In recent years, the use of autonomous systems and algorithms, often associated with big data, has also affected the immigration sector, finding application in various contexts, such as the control of national borders and the management of the first contacts between migrants and national authorities, the formalisation and assessment of applications for visas and international protection, and even the organisation of reception and integration activities and the evaluation of applications for social assistance.[1] Even if it is a relatively new and recent field of application, it is possible to carry out an initial assessment, albeit provisional, of the effects caused by the use of systems based on artificial intelligence in the context of migration management.[2] Although, as will be discussed in the following paragraphs, the application of artificial intelligence to immigration management expresses the potentialities and risks that are constantly associated with the use by the public administration of such 'intelligent' systems, it is possible to immediately underline that the specificity of this context – and in particular the condition of particular fragility that typically characterises the people involved – determines the opportunity to carry out a specific study, in terms of legitimacy of the legal basis, technological reliability of the systems implemented, their adequacy with respect to the purposes associated with immigration policies and effective guarantee of the rights of the people involved.

Although often characterised by an experimental and temporary nature, the adoption of autonomous systems based on algorithms in the implementation of migration regulations and policies is taking on a dimension, as well as an impact in legal terms, which deserves careful consideration by the doctrine. From that perspective, it is significant that the World Migration Report 2022[3] of the International Organization for Migration dedicates a chapter to 'Artificial intelligence, migration and mobility:

---

[1] For a general overview, among others, see P. Molnar, 'Technology on the margins: AI and global migration management from a human rights perspective'(2019) 8 *Cambridge International Law Journal* 2, pp. 310. With regard to the European Union dimension, see C. Dumbrava, 'Artificial intelligence at EU borders. Overview of applications and key issues' European Parliamentary Research Service. In-Depth Analysis (July 2021), available at https://www.europarl.europa.eu/thinktank/en/document/EPRS_IDA(2021)690706.

[2] See A. Beduschi, 'International migration management in the age of artificial intelligence' (2020) *Migration Studies*, 3.

[3] A. Beduschi, 'Artificial intelligence, migration and mobility: Implications for policy and practice' (2021) International Organisation for Migration, World Migration Report 2022, pp. 281-300, available at: https://publications.iom.int/system/files/pdf/WMR-2022-EN-CH-11_0.pdf.

Implications for policy and practice'. In the IOM's Report, main pros and cons are highlighted, through the analysis of concrete cases and specific technologies that may be implemented in all phases traditionally characterising the migration cycle, thus for example the pre-departure identity checks, forecasting of migration trends and lodgement, applications and compliance with visa conditions.[4]

In fact, the intertwining of policies in the field of immigration and the use of devices based on autonomous systems are destined to produce potentially virtuous effects, to which, however, are associated some critical issues deriving in particular from the nature and degree of development of the technologies available. These criticalities are expressed both in terms of the technological feasibility of the systems actually adopted, on the one hand, and the effectiveness of the safeguards provided in relation to the rights of the persons involved, on the other hand.

In general terms, many studies have highlighted the positive potentials, not only in terms of effectiveness of immigration management policies but also of a more reasonable and impartial implementation of the relevant regulations, associated with the use of 'intelligent' systems in areas such as the presentation of applications for international protection[5] or the management of migrants' reception.[6] Reception management represents a particularly useful area for expressing the potential inherent in the use of intelligent systems, as well as the concomitant need to provide adequate guarantees in order to avoid, or at least minimise, the possible risks deriving from such use. Some countries, such as the United States, Sweden and Switzerland, have already implemented AI-based tools, able to match asylum seekers or refugees' arrivals with 'optimal placement', that is, locations or communities in which they are most likely to find employment, with the main aim to 'improve integration' of refugees through 'algorithmic assignments'.[7] As far as it is technically reliable, the function performed in this area by systems based on algorithms seems to be consistent with the declared objective of the relevant legislation, in particular the

---

[4] Ivi, p. 297.

[5] Critically, L. Jasmintaite-Zaniewicz and J. Zomignani Barboza, 'Disproportionate Surveillance: Technology-Assisted and Automated Decisions in Asylum Applications in the EU?' (2021) 33 *International Journal of Refugee Law*, 1, pp. 89-110.

[6] A. Beduschi, 'International migration management in the age of artificial intelligence', cit., p. 6.

[7] J. Bither, A. Ziebarth, 'Automating decision-making in migration policy: a navigation guide' (November 2021) Migration Strategy Group on International Cooperation and Development, p. 21.

legislation on reception in the European context. In fact, the ordinary and non-emergency management of reception, as well as the integration of individuals seeking international protection, represents declared goals, which are difficult to reach in a context in which, also due to the size and the usually unpredictable nature of migratory flows, an emergency approach to reception tends to prevail. The availability of tools, capable of processing elements and data useful for maximising the organisation of the reception, not only in order to guarantee greater efficiency from the point of view of the competent administration, but also the acceptability of the presence of foreigners within a concrete social community and the chances of integration and regaining autonomy for asylum seekers and refugees, certainly represents a useful and suitable resource.

At the same time, decision-makers must be aware of possible risks embodied in such algorithmic tools, which derive directly from the technical characteristics of their design and functioning, such as errors and biases linked to the nature of data processed in order to achieve their outcomes. In particular, it has been stressed that, since these systems are designed to optimise the future employment rates of migrants participating in integration and reception projects, they risk not adequately taking into account other factors, such as for example the will of the applicant for international protection, on the one hand, or to reproduce, or even increase, the inequalities existing between applicants for international protection,[8] as 'this use of automated systems can also reinforce and exacerbate inequalities by placing those individuals with the least prospect of success into under-resourced areas, perpetuating cycles of poverty and potentially justifying negative attitudes towards refugee integration'.[9] Thus, even if it has been proven to be statistically very efficient and effective in achieving the expected goal,[10] the complexity of the migration context, combined with the duty for public authorities to avoid discriminatory or stigmatising effects that may be possibly provoked by political choices, calls for a more in-depth analysis and

---

[8] Ivi, p. 22.

[9] P. Molnar and L. Gill, 'Boats at the gate. A human rights analysis of automated decision-making in Canada's immigration and refugee system' (2018) International Human Rights Program, University of Toronto, p. 39.

[10] Ibidem («The Lab found that if the algorithm had selected the resettlement locations for refugees, the average employment rate would have been about 41% and 73% higher than their current rates in the United States and Switzerland, respectively»).

assessment of the effects that can be produced,[11] even unintentionally, in terms of protection of fundamental rights and respect for the principle of non-discrimination.

2. 'Smart borders': the risk of the 'politicisation' of AI-based tools and possible remedies

The management of migratory flows at States' borders represents an area in which this complexity is destined to express itself in a paradigmatic way. It is in this geographical area where AI-based systems interact with the pre-existing need to balance between the efficient management of migration fluxes and the effective protection of migrants' fundamental rights. The set of various systems based on AI and big data that are increasingly used in order to manage the flow of foreign people at States' borders is identified through the concept of 'smart borders', which have been defined, with specific regard to the EU context, as those 'automated systems to speed up and facilitate the border check procedure of the majority of travellers, and to hinder and stop those immigrants that pose a threat to the security of the EU through their status as irregular immigrants, criminals or terrorists'.[12]

The implementation of technologies connected to the concept of smart borders (drones, facial recognition, big data, tracking apps, chatbot) is unquestionably able to increase the efficiency of the activities carried out, for example in terms of timeframe for carrying out procedures, predictability of the concrete characteristics of the various migratory flows, and the number of visa and international protection applications evaluated. On the other hand, however, it is necessary to carefully assess what is, or may be, the impact of this use in terms of effectiveness of legal safeguards

---

[11] We will focus on the idea of risk assessment applied to AI-based tools in the following paragraphs, here we refer to an innovative approach provided by N. Ioannidis, S. Casiraghi, A. Calvi and D. Kloza, 'A tailored method for the process of integrated impact assessment on border control technologies in the European Union and the Schengen Area' in J. P. Burgess and D. Kloza (eds.), *Border Control and New Technologies. Addressing Integrated Impact Assessment* (ASP, 2021), pp. 143-160.

[12] J. P. Burgess, D. Kloza (eds.), *Border Control and New Technologies. Addressing Integrated Impact Assessment*, cit., p. 25.

that in this context – the border – must be guaranteed in the light of international and national law.[13]

As anticipated, AI-based systems are content-neutral per se, but they can be designed and implemented in such a way as to be functional to certain migration policy objectives, thus not only being able to reduce the risks in terms of protection of rights and the effectiveness of the procedures carried out at the border, but on the contrary to contribute to consolidating no-entry policies already adopted at a regulatory or practice level. With regard to the Canadian legal framework (see infra), the idea of 'invisible border walls' has been proposed, in order to emphasise the fact that 'introducing AI into the decision-making on immigration and border control has the potential to supplement (…) non-entrée policies, such as visa control and extradition practices'.[14] The risk of a 'politicisation' of AI-based tools, especially in the direction of a strengthening of 'state practices that are aimed at curbing international migration and preventing certain individuals from reaching state territories',[15] must be taken into account at the time of proposing the use of such technologies: a necessary precondition is the enactment of an ad hoc regulatory framework, capable of guaranteeing the ex ante assessment of the technical trustworthiness of these tools and their compatible use – and functional to – with the purposes and safeguards already provided for in terms of migration management and international protection. From that perspective (see next paragraph for more details), two issues may come into view: on the one hand, the ability of existing regulations, and in particular of the standards and indicators that characterise them,[16] to be adequately understood and processed by systems based on algorithms; on the other hand, the opportunity to provide regulations that are able to metabolise and manage the specificity,

---

[13] P. Molnar, 'Technology on the margins: AI and global migration management from a human rights perspective', cit., 314, with regard to the US-Mexico border, highlights that «While so-called 'smart-border' technologies have been called a more 'humane' alternative to the Trump Administration's calls for a physical wall», concretely «The use of these technologies by border enforcement is only likely to increase in the 'militarised technological regime' of border spaces, without appropriate public consultation, accountability frameworks and oversights mechanisms».

[14] R. Akhmetova, 'Efficient Discrimination: On How Governments Use Artificial Intelligence in the Immigration Sphere to Create and Fortify 'Invisible Border Walls'' (2020) Centre on Migration, Policy and Society, Working Paper n. 149, University of Oxford, p. 14.

[15] Ivi, p. 16.

[16] Ivi, p. 8. For example, the concept of safe country, or migrant's reliability.

technological but also normative, social, and political in the broad sense, of these instruments, and the current widespread lack of legislative initiatives.[17]

Those technologies that employ facial recognition, geolocation or that may favour the prediction of future migratory flows become particularly relevant. There are currently many projects at both the international and national level, which aim to enable responsible authorities to better understand and predict migrants' movements, by matching and processing data of different types, such as the ones developed by IOM, UNHCR,[18] or EASO. The algorithm developed by EASO is particularly relevant, as it tries to predict pressures up to four weeks in advance and suggest possible future medium-term scenarios using historical and current data[19]; but, in so doing, it makes use also of migrants' social media data, which can raise relevant concerns in terms of fundamental rights, especially privacy and possible discrimination, and protection.[20]

As a matter of fact, the ability to locate arrivals in advance, also foreseeing their size, can allow national authorities to prepare, adapt and modify the resources – human, organisational, economic, infrastructural – necessary to manage the first phases of arrivals, the subsequent activity reception of persons requesting international protection, as well as the resources and procedures necessary for the expulsion of persons who are not entitled to remain in the State's territory. Potentially, this may help to overcome the traditional declination in terms of emergency of migration policies, justified by the unpredictability and concentration of flows, as the availability of technologies that exploit big data and autonomous systems based on machine

---

[17] Among others, see T. Krügel, B. Schützea and J. Stoklas, 'Legal, ethical and social impact on the use of computational intelligence based systems for land border crossings' (2018) International Joint Conference on Neural Networks (IJCNN); P. Molnar, 'Technological Testing Grounds and Surveillance Sandboxes: Migration and Border Technology at the Frontiers' (2021) 45 *The Fletcher Forum of World Affairs* 2, p. 112, according to whom «It is therefore not surprising that the regulatory and legal space around the use of these technologies remains murky and underdeveloped, full of discretionary decision making, privatized development, and uncertain legal ramifications».

[18] N. Kinchin, 'Technology, Displaced? The Risks and Potential of Artificial Intelligence for Fair, Effective, and Efficient Refugee Status Determination' (2021) 37 *Law in Context* 3, p. 6.

[19] T. Bircan and E.E. Korkmaz, 'Big data for whose sake? Governing migration through artificial intelligence' (2021) 8 *Humanities & Social Sciences Communications* 241, p. 2.

[20] Ibidem. See also L. Jasmintaite-Zaniewicz and J. Zomignani Barboza,' Disproportionate Surveillance: Technology-Assisted and Automated Decisions in Asylum Applications in the EU?', cit.,p. 91, on other possible uses of personal data derived from social media.

learning could allow a State – or the European Union – to predict, albeit only in probabilistic terms, the short-medium-term trend of these events. This may facilitate an ordinary management of flows and a greater capacity of the single State, or of the European Union, to absorb in a physiological way the impact on national reception and asylum systems of arrivals of migrants at the borders. Eventually, if properly implemented, these tools may support national and international authorities to 'better discharge their legal and moral obligations towards refugees and asylum-seekers by providing them with a more efficient system of humanitarian protection'.[21]

From a European Union perspective, the implementation of migration fluxes' predictive tools may be seen as coherent and functional for the concrete building of the Migration Preparedness and Crisis Blueprint,[22] which aims to establish 'A structured migration management mechanism (…), with real-time monitoring, early warning and a centralised, coordinated EU response to mobilise structures, tools, human and financial resources as needed, across EU institutions and agencies and in cooperation with Member States' (§ 7). Within the regulatory package foreseen by the New European Pact on Migration and Asylum,[23] in the crisis management stage the Blueprint 'should ensure that up-to-date comprehensive information on the migratory situation is available to all actors allowing to take timely decisions and that the implementation of those decisions is monitored and coordinated properly' (§ 15). A feasible, reliable and data-driven capacity to predict migration fluxes and movements may also reduce States' margin of discretion in declaring a crisis in the field of migration and asylum, as the availability of evidence-based data and information shall put States in the condition to be able to adequately and timely strengthen their situational awareness and organisational and humanitarian response capacity.

---

[21] J. Napierala, et al., 'Toward an Early Warning System for Monitoring Asylum-Related Migration Flows in Europe' (2021) 4 *International Migration Review*.

[22] Commission Recommendation, 23 September 2020, on an EU mechanism for Preparedness and Management of Crises related to Migration (Migration Preparedness and Crisis Blueprint, C(2020) 6469 final.

[23] See EU Immigration and Asylum Law and Policy, D. Thym (ed.), Special Collection on the 'New' Migration and Asylum Pact, available at https://eumigrationlawblog.eu/series-on-the-migration-pact-published-under-the-supervision-of-daniel-thym/.

3. How to avoid the risk of AI-based 'invisible border walls'? The thick trail of the 'algorithmic impact assessment'

However, the same tools can be used also to consolidate and make even more efficient administrative practices that end up hindering or preventing the arrival of migrants at national borders,[24] thus apparently allowing States to legitimately avoid the duty to respect internationally recognised principles, such as the prohibition of refoulement, the right to have access to the international protection procedure, as well as the right to an effective judicial remedy against decisions taken by national authorities. From this perspective, the context of immigration does not escape the consideration according to which the use reserved for a technology cannot be considered neutral,[25] when it is functional to the realisation of political purposes and it is in any case interacting with certain and consolidated implementation practices. As vividly pointed out by the legal doctrine, which specifically refers to the raising of an approach towards the building of 'smart borders' in the European Union context, 'these technological experiments also play up the "us" vs "them" mentality at the centre of migration management policy' and the risk to '(…) only exacerbate deterrence mechanisms already so deeply embedded in the EU's migration strategy'.[26] Probably, it is appropriate to clarify that it is not the technologies actually used that express a specific political or legal objective, but rather the methods through which and the purposes with respect to which the political decision-makers, holders of the functions implemented through these systems, must be accountable to the concrete use – and related purposes – they make of such tools. Also in light of the recent case-law of the European Court of Human Rights (ECHR) regarding the prohibition of inhuman and

---

[24] A. Beduschi, 'International migration management in the age of artificial intelligence', cit., p. 6.

[25] See P. Molnar, 'Technology on the margins: AI and global migration management from a human rights perspective', cit., p. 306.

[26] P. Molnar, 'Technological Testing Grounds and Surveillance Sandboxes: Migration and Border Technology at the Frontiers', cit., p. 115.

degrading treatment[27] and the prohibition of collective expulsions,[28] it appears evident that the use of these technologies, especially when targeted towards people who are in particularly vulnerable and fragile conditions, requires an adequate legal basis, on the one hand, and a consistent ex ante assessment of the possible impact in terms of respect for fundamental rights involved, on the other hand.

In this direction, the case of Canada seems to be paradigmatic. Canada has long been experimenting with risk assessment tools, based on algorithms, for evaluating applications for protection of migrants in order to identify potential indications of fraudulent or illegal application, as well as applications that may put at risk national security.[29] While the characteristics and safeguards set forth within the triage of temporary visa applications by using 'risk assessment' algorithms have been extensively and critically analysed by the Canadian legal scholarship,[30] it must be underlined that the Canadian federal government approved a Directive on Automated Systems,[31] which provides that the use by the public administration of autonomous systems based on algorithms is previously subjected to an 'Algorithmic Impact Assessment', in which different standards are required according to the foreseeable level of impact on the rights, health and well-being, or economic interests of individuals or communities. Requirements and safeguards provided by the Directive, which entered into force in 2021, will apply also to automated systems implemented in the immigration field, thus calling federal authorities to guarantee – through the fulfilment of a comprehensive check-list enumerated by the Directive – principles such as transparency, accountability, legality and procedural fairness, while achieving by the implementation of AI-based tools 'more efficient, accurate, consistent, and

---

[27] Article 3 of the ECHR.

[28] Protocol 4, Article 4, ECHR. According to A. Beduschi, 'International migration management in the age of artificial intelligence', cit., p. 13, «AI technologies could be used to assist states in maritime interventions aiming to return migrants and asylum-seekers to unsafe countries and territories».

[29] M. Forti, 'AI-driven migration management procedures: fundamental rights issues and regulatory answers' (2021) 2 *BioLaw Journal-Rivista di BioDiritto*, p. 442.

[30] See L. Nalbandian, 'Using Machine-Learning to Triage Canada's Temporary Resident Visa Application' (2021) 9 Working Papers Series, Ryerson Centre for Immigration and Settlement (RCIS) and CERC in Migration and Integration.

[31] T. Scassa, 'Administrative Law and the Governance of Automated Decision-Making: A Critical Look at Canada's Directive on Automated Decision-Making' (2021) 54 *University of British Columbia Law Review* 1, pp. 251-298, available at: https://ssrn.com/abstract=3722192.

interpretable decisions' (§ 4.1.). Particularly relevant is the reference to the right of the public administration to have access to and to test the system used, if this is also necessary in the context of a trial, guaranteeing against unauthorised disclosures but providing for the possibility of authorising also third parties to review and verify these elements (§ 6.2.5.2 and 3). However, this right is related to a concomitant duty to test the development processes and the data used by the autonomous systems, before the start of production, in order to verify the absence of involuntary biases attributable to the data used and other factors that may affect the results in an unfair way, before the start of the production (§ 6.3.1), and to develop monitoring tools in order to identify any unintended results and to verify compliance with the law on the matter (§ 6.3.2). The Directive also provides the duty to carry out an 'Algorithmic Impact Assessment' before the start of production of any autonomous systems (§ 6.1.1), which could prove decisive in the area of immigration.[32]

To underline the need to guarantee an adequate and formal legal basis for the use of intelligent systems in the field of immigration, it is necessary to report the decision of the British administration to suspend the use of a risk assessment system for the evaluation of applications for entry into the country, following the appeal presented before the UK High Court by the Joint Council for the Welfare of Immigrants (JCWI). The appeal is grounded on the supposed discriminatory nature based on the nationality of applicants of the visa algorithms used by the UK Home Office. The Home Office clarified that the algorithm only had the aim of categorising different applications with reference to how much scrutiny each application needed, thus leaving the final decision to a human operator and that, among the indicators and data assessment, the applicant's nationality was not taken into consideration. According to the appellants, there may be the risk that some predetermined nationalities will automatically be attributed the greatest risk, thus resulting in more complex, lengthy procedures and – on the basis of the statistics available[33] – with a much lower percentage of being accepted than the others. As a consequence of the appeal, the Home Office decided to suspend the use of the so-called 'streaming tool' and to

---

[32] L. Nalbandian, 'Using Machine-Learning to Triage Canada's Temporary Resident Visa Applications', cit., p. 13.

[33] «Applications with a red rating were much less likely to be successful than those rated green, with around 99.5% of green being successful but only 48.59% of red», according to data referred by R. Jennings, 'Government Scraps Immigration "Streaming Tool" before Judicial Review' (2020) *UK Human Rights Blog*, available at https://ukhumanrightsblog.com/2020/08/06/government-scraps-immigration-streaming-tool-before-judicial-review/.

perform an in-depth and thoughtful assessment of the algorithm, in order to detect and correct possible unconscious bias and discriminatory outcomes.

It must be recalled that, in other sectors of public administration activity, the use of algorithms aimed at assessing and predicting citizens' behaviour[34] was considered illegitimate, as the transparency and verifiability of the system were not adequately ensured. In particular, there is a need to make public the risk or evaluation models used and the indicators introduced in the algorithm, as well as its actual functioning. Moreover, there is a need to avoid the possibility that the use of such assessment systems could lead to the risk of discrimination or stigmatisation of certain categories of people, based for example on their ethnicity or personal characteristics. In this perspective, from the albeit small jurisprudence on algorithmic assessment tools, some standards of protection can be obtained that are also relevant in the context of migration policies, such as the need to minimise the risk of cognitive biases, placing on the competent authorities a duty to verify ex ante that the system does not suffer from 'training biases' that could lead to discriminatory results, through the conduct of independent research and assessment.

An experiment was also carried out within the European Union that envisaged, along borders particularly affected by constant migratory flows (Greece and Hungary in particular), the use of an algorithm – the project iBorderCtrl[35] – able to detect a person's emotions through facial recognition systems, in order to identify false or contradictory statements made to border authorities.[36] In this regard, it is useful to recall a recent judgment of the Court of Justice of the European Union (T-158/19, 15 December 2021). Without going deeply into the argumentative process, the Court partially accepted the request to make public the documents relating to the authorisation procedure of the project, which was financed entirely by European

---

[34] See I. Leijten, 'The Dutch SyRI Case: Some Thoughts on Indivisible Interferences and the Status of Social Rights' (2020) *IACL-IADC Blog*, available at: https://blog-iacl-aidc.org/social-rights/2020/5/19/the-dutch-syri-case-some-thoughts-on-indivisible-interferences-and-the-status-of-social-rights; M. van Bekkum and F. Zuiderveen Borgesius, 'Digital welfare fraud detection and the Dutch SyRI judgment' (2021) 23 *European Journal of Social Security* 4, pp. 323-340.

[35] For a comment, J. Sánchez- Monedero and L. Dencik, 'The politics of deceptive borders: 'biomarkers of deceit' and the case of iBorderCtrl' (2020) *Information, Communication & Society*, pp. 1-18.

[36] L. Hall, 'Programming the machine: gender, race, sexuality, AI, and the construction of credibility and deceit at the border' (2021) *Internet Policy Review*, p. 9.

public funds. While limiting the right of access to such documents, in particular those relating to ethical and legal assessments, to the parts that are not covered by a commercial interest of the members of the Project's Consortium, the Court recognises the existence of a public interest to participate in an informed and democratic public discussion on the question of whether control technologies are desirable and whether they should be publicly funded, and that this interest should be duly safeguarded (§ 200), as it has happened – according to the Court – in the concrete case.

4. Algorithmic biases and migrants' rights effectiveness: the challenge of integrating AI-based tools' efficiency and human rights' international standards

The cases just mentioned, which are characterised by the use of artificial intelligence systems in an area that can be decisive – 'life changing'[37] – in determining the fate of the people involved, thoroughly express the main criticalities that are generally associated with these technologies: e.g., the margin of error that can be associated with technologies such as facial recognition, which turns out to be particularly high precisely in reference to the 'types' of individuals that traditionally make up the composite category of migrants; the risk of discrimination and stigmatisation towards specific social groups (such as minorities and women), deriving from the presence of cognitive bias that can characterise autonomous systems based on machine learning; the effectiveness of the right not to be subject to a completely autonomous decision; and the need for a human oversight to be ensured within the decision-making process.

Errors and biases directly associated with the technical features of the devices adopted by public administration may provoke a particularly harmful impact in the context of the management of migratory flows, ending up by increasing – rather than reducing – the current criticalities in terms of effective protection of migrants' rights, especially in the early stages of contact with national authorities. In this regard, also the safeguards related to the use of AI-based systems, for example the presence of a human operator who assumes the responsibility for the decision and the right not to be subject to a completely automated decision, risk being inadequate, taking into account the practices adopted by the border authorities to carry out expulsions or

---

[37] World Migration Report 2022, cit., p. 292.

push-backs without giving migrants the concrete possibility to express their desire to request protection or in any case to prevent their entry into the national territory. Accordingly, the adoption of technologies associated with the concept of smart borders may favour, and not prevent, the consolidation of practices that are the expression of policies that are contrary to the safeguards provided by the European Convention of Human Rights (ECHR).[38] As a recent case decided by the European Court of Human Rights (ECtHR) has shown, the practices adopted by the national border authorities may be able to minimise, even exclude, the AI-based tools' potential function to assure the legitimacy, or at least guarantee the possibility of verifying the legitimacy, through the analysis and verification of reliable data, of the actions adopted by national authorities. Thus, 'intelligent' tools might empower the effectiveness of safeguards, even if only procedural, as set forth by the ECHR and EU law. In the M.H. and Others v. Croatia judgment,[39] the ECtHR recalls the principle according to which 'footage of video surveillance may be critical evidence for establishing the circumstances of the relevant events', especially when, as it was in the assessed case, the concerned area is under constant surveillance, including by stationary and thermographic cameras, due to the frequent attempts by migrants to illegally cross the border (§271). Furthermore, the availability of these recordings would also have made it possible to clarify whether the contact with the Croatian authorities took place on Croatian territory. As reported by the Croatian Ombudswoman, also in previous cases in which she had sought to obtain such recordings, the thermographic camera recordings had also not been available owing to technical problems (§ 104). It must be recalled, eventually, that the ECtHR has declared the violation of Article 2 ECHR on the ground, among other reasons, that 'the investigative authorities never verified the police allegation that there were no recordings of the impugned events, and that they had failed to inspect the signals from their mobile telephones and the police car GPS in order to establish the applicants' whereabouts and their contact with the Croatian police (…)' (§ 272).

Without denying the potentialities deriving from the use of these devices in terms of fundamental rights' protection, it is necessary to adequately assess, through a data-

---

[38] In particular, the risk of being subjected to inhuman and degrading in the event of refoulement and the prohibition of collective expulsions, as well as the right to an effective remedy against decisions taken at the administrative level.

[39] See H. Hakiki and D. Rodrik, 'M.H. v. Croatia: Shedding Light on the Pushback Blind Spot' (2021) *VerfBlog*, available at https://verfassungsblog.de/m-h-v-croatia-shedding-light-on-the-pushback-blind-spot/.

driven and comprehensive assessment, the impact on the legal status – and the same existences – of the persons involved by their implementation, in particular when the outcomes may be decisive in order to determine migrants' future.

As is evident, in these cases the use of artificial intelligence can produce very different effects in terms of guarantees, both substantial and procedural, being able to alternatively contribute to reducing practices that are contrary to what is envisaged at the regulatory level (e.g., the principle of non-refoulement, and access to procedure) or, on the contrary, to strengthen its effectiveness. It is therefore necessary to raise the general question of the conditions – technological, legal but also organisational and institutional – that may allow a legitimate use of these tools, at least in order to assess their technical trustworthiness (in order to reduce the risk of errors or discriminatory effects), to guarantee their transparency and explicability for both public authorities and the public, as well as to select at the legislative level the concrete functions that they can legitimately perform[40] (for instance, avoiding the complete replacement of the human operator). In general terms, the public administration's duty to provide a comprehensive, independent, evidence-based, multidisciplinary, and rights-centred ex ante assessment of AI-based devices may become a shared standard for achieving a trustworthy use of AI-based systems in the migration policy context.

In addition, a series of related questions arises, such as the possibility of declaring the jurisdiction of a State that provides for the use, pursuant to Article 1 of the ECHR, of such measures as the following: the visual contact between migrants and thermographic cameras or cameras installed on drones, which may be used to locate groups of people who are approaching a state border, in order to intercept them in advance; the existence of adequate judicial remedies against decisions taken by competent authorities regarding the level of risk for national security or public order associable to a particular subject; the reliability of an application for international protection, which is also based on data obtained from the use of predictive algorithms and assessment tools; or, again, the legitimacy of implementing concepts such as 'safe country', or assessing the concrete social, political and cultural situation existing in a

---

[40] J. Purshouse and L. Campbell, 'Automated facial recognition and policing: A Bridge too far?' (2021) *Legal Studies*, pp. 1-19.

migrant's country of origin,[41] by processing data carried out by autonomous systems based on algorithms.

5. Which regulatory design? Old safeguards in innovative tools

All issues that have been already raised lead to a conclusive – and pivotal – question, which refers to the most appropriate regulatory framework, able to guarantee a trustworthy use of such tools. In this perspective, it is necessary to question the ability of existing rules to intercept the functionalities of these innovative tools, or whether, on the contrary, it is necessary to consider ad hoc regulatory regimes, which are able to detect and govern the specificity – in positive and negative terms – that can be associated with such devices and at the same time are capable of adapting to the evolution of the latter from a technological point of view. In this sense, the need to subordinate the use of autonomous systems based on artificial intelligence at least to a 'human rights impact assessment' may be an essential requirement, in order to verify whether their use in the management of migration flows does not end with prejudice to the rights of migrants and applicants for international protection.[42]

Within the European Union, the Proposal for a Regulation on artificial intelligence presented by the Commission explicitly mentions the use of artificial intelligence systems in the management of migration, asylum and border control, underlining the point that they 'affect people who are often in particularly vulnerable positions and who are dependent on the outcome of the actions of the competent public authorities'.[43] In light of these considerations, the Proposal affirms the particular importance of guaranteeing in these areas 'the accuracy, non-discriminatory nature and transparency' of AI systems, in order to guarantee 'respect for the fundamental rights of the persons concerned, in particular their rights to free movement, non-discrimination, protection of privacy and personal data, international protection and

---

[41] N. Kinchin, 'Technology, Displaced? The Risks and Potential of Artificial Intelligence for Fair, Effective, and Efficient Refugee Status Determination', cit., p. 15.

[42] See A. Beduschi, 'International migration management in the age of artificial intelligence', cit., p. 8.

[43] Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) (2021), Recital 39.

good administration'.[44] On this ground, in the Proposal the artificial intelligence systems used in these areas are classified as 'high risk' if they consist of '(...) polygraphs and similar tools to detect the emotional state of a natural person; for assessing certain risks posed by natural persons entering the territory of a Member State or applying for visa or asylum; for verifying the authenticity of the relevant documents of natural persons; for assisting competent public authorities for the examination of applications for asylum, visa and residence permits and associated complaints with regard to the objective to establish the eligibility of the natural persons applying for a status'.[45]

Therefore, the use by national and European authorities will be conditioned, once the proposal of Regulation will be enacted, to the compliance with predetermined requirements and safeguards, to be verified through a conformity verification procedure. Among such requirements and safeguards, it is useful to recall the duty to establish a 'risk management system', which is intended as 'a continuous iterative process run throughout the entire lifecycle' of the AI system and must comprise – among other elements – the analysis of 'known and foreseeable risks' associated with each high-risk AI system and, within the most appropriate risk management measures, it recalls the 'elimination or reduction of risks as far as possible through adequate design and development' (Article 9). With regard to data governance (Article 10), the Proposal refers to the need to examine, during the training, validation and testing data sets, the existence of 'possible biases' and of 'any possible data gaps or shortcomings, and how those gaps and shortcomings can be addressed'. In light of their implementation in the immigration governance context, the requirement that 'training, validation and testing data sets shall be relevant, representative, free of errors and complete' and that they must have the appropriate statistical properties, 'including, where applicable, as regards the persons or group of persons on which the high-risk AI system is intended to be used' (Article 10). Also 'the characteristics or elements that are particular to the specific geographical, behavioural or functional setting within which the high-risk AI system is intended to be used' must be taken into account during the training and validation data sets. These two latter requirements may have special relevance when applied to AI systems in the management of migration, asylum and border control.

In light of this potential legal framework, it would be useful to verify the compatibility of AI systems already experimented with at the EU borders (for example, the

---

[44] Ibidem.

[45] Ibidem.

aforementioned iBorderCtrl project) with what is envisaged for 'high-risk' systems by the Regulation Proposal, as well as the compatibility with the latter of the legislative acts already adopted in this context at the European level.[46]

---

[46] For a systematic analysis of these tools, European Parliamentary Research Service, 'Artificial Intelligence at EU borders. Overview of applications and key issues' (2021), available at https://www.europarl.europa.eu/thinktank/en/document/EPRS_IDA(2021)690706.