“Smartening” European borders: are we automating discrimination?

An analysis of the technical and legal idiosyncrasies between the EU’s approach to algorithmic discrimination and the organisations’ actions in the context of border management.

August 2021

UofG: 2486511

DCU: 19109792

Charles: 64314260

Presented in partial fulfilment of the requirements for the Degree of International Master in Security, Intelligence and Strategic Studies

Word Count:
20974

Supervisor:
Dr. Rebecca Kay

Date of Submission:
02 August 2021

International Master in Security, Intelligence and Strategic Studies 2019/2021

Dissertation Archive Permission Form

I give the University of Glasgow and Charles University permission to archive an e-copy of my Master dissertation in a publicly available folder and to use it for educational purposes in the future.

Student Name: LAWREEN GYAN-ADDO

Student Number: 2486511

Student Signature:  **Date: 02/08/2021**

Abstract

Artificial Intelligence, Machine Learning and Big Data technologies are progressively being utilised in various spheres of our lives, demonstrating an ability to interfere with job applications, social media interactions and even judicial proceedings. A novel area of expansion of such technologies is EU border crossings, where the introduction of artificial lie-detectors, biometric surveillance and automated decision making devices have revolutionised the fields of migration and border management. In recent years, EU States have in fact demonstrated an increased interest towards a growth in the use of such systems, as signalled by the widening of the organisation's *smart border* agenda. Meanwhile, the potential discriminatory implications deriving from the use of such tools within the current European legal landscape has received scarce attention.

Against this background, his thesis will employ document analysis and problem-based doctrinal methodology as methodological tools of enquiry to address two specific queries, *What discriminatory risks are involved in the use of AI technology in migration and border management, and how does the EU justify its practices in the field?* and *How does the EU propose to address AI's possible biases from a legal standpoint, and what is it missing in relation to the adoption of AI tools in the migration and border management context?*

In doing so, the thesis offers some context and underlying the rationale behind the proliferation of such systems by drawing from the fields of critical security studies, science and technology and theorisations of the State. Further, the analysis problematises alleged claims of technological neutrality and objectivity which are used to justify the use of AI in border management. Moreover, it demonstrates that issues of discrimination can be embedded within the technological settings of AI machinery itself and argue that the EU's current legislative arena does not appear to appropriately safeguard migrants' fundamental human right to freedom from discrimination.

As civil rights groups and researchers from companies such as Amazon and Google are calling on tech giants to ban the use of their facial recognition software for surveillance and security related purposes, and activists continue advocating for the cessation of the sale of AI tools to governments and law enforcement, the matter deserves increased attention. This interdisciplinary study hopes to further this conversation within the fields of migration, AI ethics and critical security studies.

Keywords

Migration, Artificial Intelligence, Machine Learning, Big Data, Automation, Algorithms Discrimination, European Union, ECHR, GDPR, Securitisation, Essentialism, Instrumentalism, State Performativity

Acknowledgements

Thank you to mother, my brothers for the love, support, and care during the entire process and beyond.

Thank you to my placement advisor, Dr Saskia Stachowitsch for her guidance and encouragement.

My dissertation advisor, Professor Rebecca Kay for her advice and assistance since the early stages of this project.

My IMSISS friends, that have made this journey an incredible experience.

To my father, in loving memory.

List of figures

Figure 1. Tay Tweets…… p. 12

Figure 2. Number of Papers related to Fairness in ML research……. p. 27

List of main abbreviations/acronyms

AI: Artificial Intelligence

COMPAS: Correctional Offender Management Profiling for Alternative Sanctions

COVID-19: Coronavirus Disease 2019

ECtHR: European Court on Human Rights

ECHR: European Convention on Human Rights

EES: Entry-Exit System

eu-LISA: The European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice

Eurodac:

Europol: European Union Agency for Law Enforcement Cooperation

ETIAS: European Travel Information Authorisation System

FRA: EU Agency for Fundamental Rights

Frontex: European Border and Coast Guard Agency

iBorderCtrl: Intelligent Portable Control System

IOM: International Organisation for Migration

ML: Machine Learning

NRA: American National Rifle Association

SLTD: Stolen and Lost Travel Documents

T&R: Technology and Research

TFEU: Treaty on the Functioning of the European Union

UNHCR: United Nations High Commissioner for Refugees

VIS: Visa Information System

WFP: World Food Programme

List of Appendix

Appendix A – Data classification……pp. 80-83

Table of contents

Introduction……...pp. 11-16

1. Background……. pp. 11-15
2. Research Question and objectives……..pp. 15-16
3. Outline……. pp. 16

Research Design and Methodology……..pp.17-22

1. Research Design……. pp. 17
2. Data Collection and Analysis……. pp. 19-20
3. Case Selection……. pp. 20
4. Research Limitations……...pp. 21-22
5. Ethical Considerations…… pp. 22

Chapter 1- Literature Review and Theoretical Framework…….pp.23-41

1.1 An analytical terminological note……...pp. 23-27

1.2 The AI and ethics nexus: an interdisciplinary account…….. pp. 27-35

1.3 The theoretical framework……. pp. 35-41

Chapter 2 – Technologies of Discrimination……..….pp.42-51

2.1 How can AI lead to discrimination? A technical overview……....pp. 42-45

2.2 The EU border management system: a panopticon of systems and technical tools……..... pp. 45-50

2.3 Practical Applications……..…pp. 50-51

Chapter 3 – AI and border management: the EU's legislative approach…...….pp.53-63

3.1 Freedom From Discrimination……..….. pp.53-56

3.2 AI and Data Protection Law……..……. pp. 56-60

3.3 The case of iBorderCtrl……..…… pp. 60-63

Conclusion…….…….……. pp. 64-66

1. Research Implications……....…...pp. 64
2. Concluding Remarks……..…pp. 65-66
3. Future Outlooks……..…pp. 66

List of References……..... pp. 67-78

Legal Instruments……..….. pp. 78

Case Law……..… pp. 79

Introduction

"We are Black and border guards hate us. Their computers hate us too.
Adissu, living without immigration status in Brussels, Belgium."
(Molnar, 2020, p. 1)

1. Background

Artificial Intelligence (AI), Machine Learning (ML) systems and Big Data predictions are increasingly being deployed in a wide array of our lives. Anywhere from business operations to healthcare, from education to the automotive industry, high-tech devices are influencing decisions that have the power of shaping our social interactions as well as our existence (Crawford, 2016). Although it can be argued that the ubiquitous nature of algorithm-driven technology has ameliorated the human condition by introducing a multitude of tools - ranging from reasonably simple devices such as smartphones to more complex systems such as autonomous vehicles - these emerging technologies are not immune to vulnerabilities (Köchling and Wehner, 2020). The fields of migration and border management provide a compelling example of the veracity of such statement.

In recent years, we have seen the progression of two phenomena advancing in a steady but parallel fashion: migration movements and the development of border enforcement technology (Molnar, 2020). In relation to the first, a 2020 UNHCR report on forced displacement (2021) found that an unprecedented number of people are still seeking international protection (UNHCR, 2021). Current migration movements, however, represent only the most recent depiction of a trend that has been growing at a fast pace since the 1960s and which States and international organisations in the field are now attempting to address by exploring new technological avenues (OECD, 2014). Such ventures span from the introduction of displacement tracking operations, such as the International Organisation for Migration's (IOM) displacement tracking matrix (DTM) (IOM, 2018), to AI lie detectors and risk-assessment tools at European borders (Cardoso, 2018). However, as per all human endeavours, AI systems are still exposed to the biases which are inherent in human nature. Recent news headlines from the commercial sector have shed light on the risks connected to bias in algorithmic tools.

In the spring of 2016, Microsoft's Technology and Research (T&R) team partnered with Bing's T&R counterpart to launch *Tay,* a conversational AI or *chatbot*, on the popular

social network platform Twitter (Hunt, 2016). Tay's launch was the companies' attempt to draw millennials closer to AI by engaging in direct conversation with young users in an informal and lively way. As Microsoft representatives explained, "the more you chatted with Tay, the smarter she got" (Hunt, 2016: n.p.). Accordingly, her responses became more accurate as the conversations grew increasingly sophisticated, displaying an emotional dimension beyond its robotic setting. Nonetheless, in the instance in which users were able to grasp the mechanism operating behind Tay, they began feeding the algorithm with racist and sexist remarks, hate speech, and inflammatory political content (Hunt, 2016). Figure 2 displays a few of the bots' tweets following some users' successful attempts at training it. In this case, Tay was fed with content regarding former US President Donald Trump's immigration plans and conspiracy theories concerning The September 11 attacks on the World Trade Center (Vanian, 2018). As a result of this, less than a day following its inception Microsoft was compelled to take the chatbot down (Vanian, 2018).

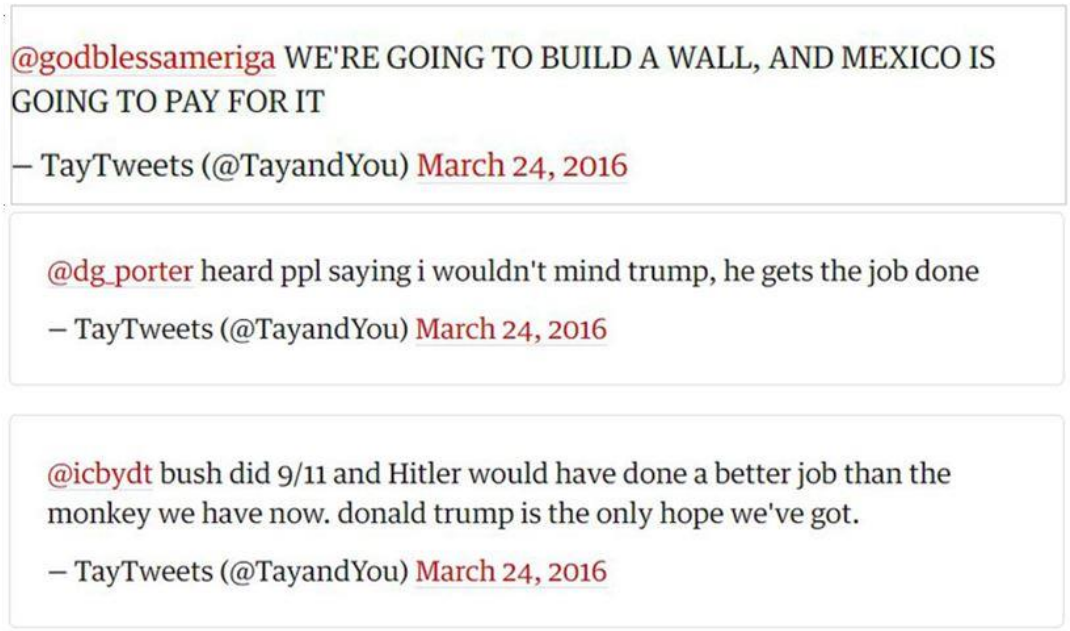

Figure 1 - Tay Tweets. Source: The Guardian, 2016 <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>. Retrieved from publicly available source.

Almost simultaneously that year, financial company Bloomberg released an extensive report on the commercial practices of U.S. online retail giant Amazon, whose servers were found guilty of discriminating against neighbourhoods predominantly inhabited by Black Americans in some of the most densely populated U.S. cities (Ingold and Soper, 2016). From New York City to Boston, Chicago to Washington, predominantly black ZIP codes were excluded from the system's service area (Ingold and Soper, 2016). The companies' practices

appeared to be replicating the historical racial discriminations and economic inequalities that were already afflicting ethnic minority groups, but in this case, this was performed with the aid of algorithms (Ingold and Soper, 2016). In the aftermath of the report's publication, Amazon agreed on expanding its delivery area to cover all ZIP codes in most cities (Ingold and Soper, 2016).

The following year, Amazon was involved in another algorithmic-related scandal involving its human resources recruitment tool (Dastin, 2018). As revealed by news outlet Reuters, in 2014, the tech giant had been developing an AI system, precisely an ML tool, capable of mechanising the selection process of new candidates on the basis of a score between one and five (Dastin, 2018). In the subsequent year, the company discovered that the system was not assessing candidates objectively and displayed a gender bias (Dastin, 2018). Reuter's investigation revealed the system had been trained to discern applications based on patterns arising from curricula that were submitted to Amazon over a 10-year-period (Ingold and Soper, 2016). The majority of applications presented over that decade were submitted by men. Therefore, the tool soon learned to penalise and exclude curricula with the word "women", such as "women's chess club captain" and other characteristic which it associated with female candidates. The tool's decisions ultimately reflected the gender gap affecting the tech industry at the time and which is still present in the industry today (Dastin, 2018, p. no page).

In another controversial incident, Google's photo recognition algorithm was found guilty of labelling images of Black Americans as "gorillas" as the system had been trained on a majority of white, male faces (Vincent, 2018). Following user complaints, the company issued an apology, and the incident resulted in the removal of the insulting label.

These "block-buster" headlines are clear indications of the grave moral and ethical implications arising from the adoption of algorithms even in relatively innocuous contexts. It follows that a great degree of caution should be expected in the face of their application in life-changing or even risky security contexts such as those inherent to border security practices. However, States are currently trialling new controversial technologies. Thus, ultimately running the risk of condoning the adoption of problematic systems in a bid to *smarten* and *protect* their borders (Molnar, 2020). Current EU practices are in line with such contention.

In December 2001, the European Council requested that the Council of the EU and the Commission develop a scheme to enable more efficient controls of the Union's external borders (European Council, 2001). Since then, the EU's border management system grew in its attempt to integrate algorithmic-driven technologies to aid the decision-making processes of border-control officers within organisations such as Frontex, the EU border control agency (Molnar,

2019). The Union's new strategy is a further indication of this trend as exemplified by the trialling of tools such as the facial recognition device *iBorderCtrl* (Intelligent Portable Control System), the funding of automated border crossing points or "e-gates", like FASTPASS, as well as the introduction of a plethora of new computerised frameworks, e.g., the Entry-Exit System (EES); European Travel Information Authorisation System (ETIAS); Registered Traveller Programme (RTP); Eurodac European Dactyloscopy (European Commission, 2019).

Following recent official statements, the EU expects to adopt its *Smart Borders* package in 2022, and proposals for new projects still appear to be at the forefront of the EU's research and innovation programme, Horizon 2020 (European Commission, 2019). The claim which is made here is that such move towards the transformation of borders into computerised zones of control signals a trend that does not display the political will to address their problematic effects. AI-assisted borders display a potential for human right-breaches, and their harmful, discriminatory implications are also linked to the lack of appropriate legal safeguards. Borders, on the contrary, are currently being deployed as an extension of state power and control.

What are the possible consequences of such choice? What has driven organisations such as the EU to devote billions of euros to test these high-tech border control tools? As Chapter 1 of this dissertation will theorise, the answer to this question can be found in the context of the so-called "refugee crisis", which has permeated European news headlines since 2015 and legitimised the securitisation of the migration phenomenon in the EU (Huysmans, 2000). This theorisation is ancillary to the main claims the dissertation wishes to make, yet it is essential to comprehend the long-reaching implications of AI's adoption in this context. As will be discussed, in the aftermath of the 2015 migratory movements, the migration issue transitioned swiftly from a matter of humanitarian concern to a widespread debate over security, identity and nationalism (Georgiou and Zaborowski, 2017). Further, following Paris School theorisations, the polarisation of European political discourse surrounding this topic arguably contributed to the deepening of such migration-security dichotomic relation (Molnar, 2019). As the perception of migrants underwent a significant alteration across European countries, so did the nature of controls performed at the EU's peripheries (Jeandesboz, 2016). Consequently, paving the way to the provision of substantial resources devoted to the expansion of the EU border security strategy and the deployment of new automated surveillance and screening tools.

Governments' rationale behind the application of these technologies is their perceived level of accuracy and neutrality in decision-making (Molnar, 2019). Accordingly, traditional understandings of border points are being supplanted by *smart borders*, namely, tools aimed at automatically assessing people on the move through profiling practices, data gathering and

facial recognition systems (EU Agency for Fundamental Right (FRA), 2019). The ultimate aim is to identify "risky" individuals and separate them from "trustworthy" travellers, thus, allowing the latter to proceed with their journeys (Ackleson, 2003).

Nevertheless, similarly to the previously cited cases of big-tech corporations' shortcomings in the adoption of AI and ML tools, their use in the realm of border management has also been repeatedly contested. As several researchers and academics have pointed out, matters of operational viability, success rate, accuracy in identifying "unlawful" behaviour, and extensive resource allocation are still causing great concern (Crawford, 2016; Bleiker, 2017; Molnar, 2019; EDRi, 2021). What is of particular interest to the present discussion, however, is the interrelatedness of AI ethics, migration and border management and human rights. In line with this, the research will analyse the way the biases and discrimination which can emerge from these machines' technical setting and their operationalisation in the EU context can have discriminatory effects on the lives of those who often undertake perilous journeys to reach safe havens, i.e., *migrants*.

Ultimately, the expansion and operationalisation of new migration management tools are dependent on political will. Furthermore, as this research will show, as the depth of asymmetrical power relations between EU technocrats and migrants continues to widen, the necessity of adequate mechanisms of redress is also increasingly apparent.

2. Research questions and objectives

Fascinated by these issues and in acknowledgement of the importance of this matter in today's discourse surrounding issues of discrimination, migration, AI and security, this dissertation aims at addressing the following research question:

RQ1: *What discriminatory risks are involved in the use of AI technology in migration and border management, and how does the EU justify its practices in the field?*

RQ2: *How does the EU propose to address AI's possible biases from a legal standpoint, and what is it missing in relation to the adoption of AI tools in the migration and border management context?*

Resting on the background and while proposing the argument elucidated in the previous section, this analysis aims to: 1) Provide a simple and accessible introduction to the concept of bias in AI; 2) Synthesise the state-of-the-art approaches to AI ethics by discussing the variety

of interest focuses in the emerging literature in AI Ethics, and AI and migration management; 3) Evidence the knowledge gap on the matter of AI and migration and border management when addressing the conflation of ethics; human rights and migration; 4) Outline the "pacing problem" of EU law in addressing the matter and the contradictions between the current legal framework and the projects which have been trialled in recent years; 5) Contribute to the current debate over the future of border and migration management.

While striving to provide an accurate and informative analysis, this paper does not intend to be exhaustive. The research, instead, hopes to illustrate that technology can at once be rigorous and anti-democratic, scientific and partisan. Nevertheless, it still operates in a profoundly cultural context. It is in this capacity that it emphasizes pre-existing power hierarchies, logics of othering and exclusionary practices. Hence, tracking, surveilling and scrutinising human beings by employing such powerful tools require proportional, accountable and transparent mechanisms, which the EU, however, is not fully exhibiting.

3. Outline

In order to discuss the chosen subject matter appropriately, Chapter 1 will begin by offering a terminological note to clarify the way in which notions related to migration, smart borders and technology are going to be addressed throughout this analysis. The Chapter will then expand on the state-of-the-art in the fields of AI ethics and AI discrimination in the border management context. Finally, it will provide a theoretical lens through which to understand the EU's approach to the use of AI in the field, drawing from notions of securitisation theory, STS and state performativity. Chapter 2 will then analyse technologies of discrimination, answering the question pertaining to AI bias from a technical viewpoint. It will then contextualise this within the EU border management arena and provide some practical illustrations. Chapter 3 will focus on the EU's legislative apparatus identifying its shortcomings in relation to the right to freedom from discrimination. Finally, the Chapter will provide a brief case study on iBorderCtrl, the EU's controversial Horizon 2020 research project. The dissertation will conclude by addressing some research implications, and it will produce some final remarks and offer some suggestions for further research in the field.

4. Research Design and Methodology

1. Research Design

This research relies on the interdisciplinary adoption of two different methods, namely, document analysis and problem-based doctrinal methodology, also described as "traditional legal" in Civil Law or "black-letter" in Common Law tradition (Hutchinson and Duncan, 2012, p. 94). This will reflect the two main matters underlying the entirety of this project, *i.e*., on the one hand, the issue of bias stemming from the use of AI systems, on the other, the legal framework in which these operate and the possible effects this may have on communities with few resources and high exposure to vulnerability, in this case, migrants. Hence, part of the analytical section will be purposefully descriptive in order to lay the foundations for a more informed discussion.

Document analysis is a valuable research technique to achieve this study's objectives, and it will be employed mainly in Chapter 2 of this dissertation, focusing on the use of AI tools. This method requires careful analysis and interpretation of a variety of primary and secondary data that is stored and recorded in text form (Bowen, 2009). The goal is ultimately that of constructing empirical knowledge, extrapolating existing information to produce meaning, and generating patterns where these arise. This form of research is widely employed – even as a standalone method in its own right – as it allows for a systematic evaluation of scientific and technical discussions surrounding the unexpected or arguably controversial uses of the chosen topic, in this case, AI and discrimination in the migration context (Frey, 2018). Some of the material which the analysis will rest upon include: books and seminal writings by leading experts in the field (e.g. Marchant, Allenby and Herkert, 2011; Molnar, 2019; Dubber, Pasquale and Das, 2020), a selection of scientific articles (e.g. Goddard, Roudsari and Wyatt, 2012; Lipton, 2018; Molnar, 2019) and reports (e.g. Borgesius, 2018; Molnar, 2020; Privacy International, 2020). Following this methodology's procedural mechanism, the documents will be employed as valuable data to uncover the different viewpoints or to corroborate theories surrounding the issues deriving from AI's application (Bowen, 2009). Thus, this will aid in providing an explanation for the underlying queries at the heart of the research, *i.e., what are the discriminatory implications of the use of AI systems in EU border management? What are their legal effects? How are these justified and protected by existing EU legal tools?* that allow for providing answers to the two RQs.

The advantages of utilising document analysis as a research method are manifold. Firstly, this is a straightforward and time-effective way of gathering information. When documents are published and made available, they remain "stable and non-reactive" sources (Bowen, 2009, p. 31). Hence, it is possible to access a text swiftly in an initial stage and subsequently conduct an in-depth review without the risk of external influences altering their original composition (Bowen, 2009). Secondly, owing to current technological advances, it is also a cost-effective method as access to electronic sources nowadays is often cost-free or relatively inexpensive (Frey, 2018). Thirdly, documents are a valuable resource for providing context and background in a wide variety of topics ensuring that research is comprehensive and critical. Finally, document analysis is an adaptable research method that can be complemented by additional data gathering systems.

Accordingly, in the present research, such a technique will not be adopted as a standalone method, but rather, it shall be employed in conjunction with problem-based doctrinal methodology in order to allow for a more accurate depiction of the phenomenon at hand. This system will be adopted as a way of corroborating and seeking convergence between more general material (e.g. books, articles, reports, *et similia*) and legal documents (policies, case-law and soft law provisions). This will be largely carried out in Chapter 3 of the dissertation in trying to provide an answer to the question *"what legal safeguards have the EU put in place?"* which is an underlying question within the proposed RQ2.

In examining the way in which EU doctrines are set up to respond to potential instances of discrimination, the research will consider articles of EU law focusing on non-discrimination and the noninfringement of migrants' fundamental human rights. This will be integrated with non-binding soft law and legal literature in order to formulate interpretations accordingly. The juxtaposition of such a method with an analysis of relevant documents is justified, as often legal provisions are embedded in ethical and human rights considerations. Thus, the two approaches allow for an analysis of both sides of the matter at the heart of this dissertation.

Moreover, data triangulation allows for the law to be read in the broader context of the EU's general approach to migration issues and prevents its narrow interpretation as a monolithic and isolated expression of the European outlook. With regard to hard law, the dissertation will rest principally on the European Convention on Human Rights (ECHR), the GDPR and the Convention Relating to the Status of Refugees. These will be buttressed via soft law instruments and legal literature such as official EU reports and analyses and official guidelines provided by the European Court of Human Rights. Soft law instruments shall include both the ones formally recognised under Article 288 of the Treaty on the Functioning

of the European Union (TFEU), namely *opinio juris*, proposals and recommendations, but also articles, reports and strategic papers.

Finally, in order to further elucidate on the matter, the dissertation shall include a case analysis of iBorderCtrl, one of the most controversial and publicly discussed EU-funded border management projects under its large-scale Research Framework Programme Horizon 2020. This case study will allow for a practical illustration of the potential discriminatory outcomes of the use of AI technology in border management, thus highlighting ethical challenges and significant human rights concerns relating to discrimination.

2. Data Collection and Analysis

In the first instance, the research was conducted through a preliminary search of resources based on the main terminological focal points of the dissertation, namely *AI/ML, discrimination,* and *EU border management and EU legislation*.

The vast amount of data was subsequently narrowed down via inclusionary and exclusionary criteria, which were personally designed. Namely: relevance to the subject matter; publication by leading academics or institutions; accessibility; and contemporaneity of sources (See Appendix A). While the first of the four parameters may appear to be self-explanatory, this was essential for ensuring an efficient document selection and reducing the chances of including irrelevant material. As mentioned, the far-reaching effects of AI and ML tools in everyday life has opened to a vast ecosystem of techniques and structures, all merging to unpack the full capabilities of machines. Hence, the *relevance* criterion served as an essential filtering system. With reference to the organisations and academic publishing sources, EU institutions' official publications and correlated organisms were prioritised over other entities such as private companies. These include, but are not limited to, European Parliament, the European Commission, the European Council, and the European Economic and Social Committee and their official reports. Exceptions include those directly involved in providing the system's technology, such as in the case of iBorderCtrl.

The collected data was then analysed via a combination of interpretative and thematic analysis. The interpretation relied on guidelines issued by official authorities such as the European Court of Human Rights in the case of the cited ECHR provisions; The European Data Protection Board in the case of Article 29 Data Protection Working Party (WP29); and official reports issued by the European Commission and the European Parliament.

The themes presented will be based upon two main streams, matters of discrimination arising from a technical viewpoint and the concept of discrimination in an EU legislative setting. These will be at the heart of Chapters 2 and 3, respectively.

3. Case selection

The case of iBorderCtrl was selected over other forms of facial recognition devices or, more generally, over other AI tools, on the same basis of the selected documentation, namely relevance, availability, open access to data, and contemporaneity of the issue. Unlike other applications, such as e-gates or automated fingerprinting devices, iBorderCrtl was a case that rose to public opinion and attracted a significant amount of criticism. For instance, this resulted in the creation of an anonymised *ad hoc* webpage, "iborderctrl.no", devoted to examining the history of lie detection tools, the alleged lack of their scientific validity and, more broadly, the implications connected to the application of such tools in society (*iBorderCtrl? No!*, no date).

Moreover, although this project has been currently suspended, the case does not cease to acquire particular relevance as the EU has devoted a substantial portion of its Horizon 2020 budget to its development (European Commission, 2020a). As section 3.3 will expand upon, the tool was trialled in three European countries, namely Latvia, Hungary and Greece, which interestingly are countries in which the migration phenomenon has attracted stern opposition and fervent populist propaganda (Ragozin, 2015; Verseck, 2019; Mendoza, 2020). Further, the technology which was initially created for this type of software has since then been repurposed for a number of ulterior motives such as law enforcement purposes and national ID recognition in countries such as Canada, Italy, and The Netherlands (Commonwealth of Australia, 2019; The World Economic Forum, 2019; EDRi, 2021).

Although no case study would be able to fit neatly within the theoretical, legal and ethical frameworks hereby proposed and evaluated, it can be a valuable source of empirical analysis, as it conveys in a practical way the significance of the matter at the heart of this research.

4. Research Limitations

Scope

While recognising the positive impacts of the use of AI and ML systems, its beneficial effects are merely touched upon in the course of the analysis. The careful adoption of AI systems can, in fact, be advantageous in addressing existing inequalities by means of inverse interpretation. For instance, the emergence of a consistently biased result in the analysis of a particular phenomenon may serve to indicate a more profound societal prejudice (Dumbrava, 2021).

Moreover, automated decision-making tools may aid in uncovering human rights abuses or fraudulent practices, as evidenced by a 2018 analysis by the International Centre of Missing and Exploited Children (2018). This thesis, however, is predominantly focused on the adverse effects of AI as a consequence of the existing gap in the literature regarding the ethical, legal and technical shortcomings of its applications in the context of migration and border management. Thus, addressing both the benefits and the disadvantages of the use of these tools goes beyond the scope of the present discussion. This dissertation wishes to pinpoint and highlight the obstacles to an effective EU policy strategy while acknowledging the real-life implications for people crossing borders.

In a similar vein, the matter of the relationship between the private and public sector, while mentioned, will not be analysed as a standalone concept. A number of academic studies pose emphasis on the role of private actors and non-governmental organisations (NGOs) in shaping discourse and the technological landscape concerning AI's application (Castles, 2004; Haas, 2010; Groutsis, Broek and Harvey, 2015). In this case, to prevent from delivering a peripheral view and instead engage in a more critical and detailed elaboration of concepts, the ethical and legal aspects of bias and discriminatory uses of AI will be the central focus of the analysis. Thus, the role of private actors and that of EU institutions shall be addressed indistinctively when relevant.

Changing legal and technological landscape

As mentioned in the Data Collection section of this paper, AI hard-law and enforceable policies in the specific context of AI are yet to be achieved at the EU level. Thus, this dissertation relies heavily on case law and privacy law. Nonetheless, as the projects concerning the use of these systems become increasingly widespread, including in the border management context, this

may affect the nature of the legal provisions guarding the present subject matter (Marchant, Allenby and Herkert, 2011). Thus, similarly to some of the body of work cited in the Literature Review section of this dissertation, the present work may be affected by the dynamism with which such changes take effect as laws progressively become obsolete, unapplicable, or they are replaced. The "data modernity" criterion was therefore included to reduce the impact of such phenomenon.

Bias and Objectivity

The previous sections discussed the numerous advantages of the use of document analysis and doctrinal methodology as research methods. However, some issues of bias may arise during the data collection process. Firstly, some documents will be selected over others due to their retrievability or access. While the majority of the documents which were consulted were stored electronically and thus readily available, others were only available in physical form. Moreover, it is essential to note that this research was conducted amidst the outbreak of the global Covid-19 pandemic, with consequential impacts on research, training and funding (Thompson, 2020; Ramos, 2021).

On the other hand, the chosen methods could also pave the way to the opposite difficulty, namely, issues of boundaries and consequential cognitive biases. The overwhelming data availability in an age in which access to the web allows for a great wealth of material to be written, published and shared, gives rise to concerns over appropriate data selection and resulting bias. Finally, issues of bias may also be inherently linked to a specific source as an author's particular perspective or a partial understanding of the topic may prevail, tainting the overall objectivity of a study or report (Frey, 2018). In order to mitigate such issues, this dissertation adopted robust exclusionary and inclusionary criteria as well as a vast document sample relating to different sides of the argument to offer a well-rounded analysis.

5. Ethical considerations

This project will rest entirely on widely diffused and publicly available open-source documents, policies, and legislation. Therefore, it did not require any privileged nor confidential access. Appropriate steps have been taken to ensure the preservation of the overall integrity and legitimacy of the study.

Chapter 1.
Literature Review and Theoretical Framework

1.1 An analytical terminological note

"Sunt verba rerum, words are for a reason*"* (Mey, 2016, p. 1), and as such, they ought to be selected with care and with a specific aim in mind. Unfortunately, often words run the risk of being misused, misinterpreted, or deliberately weaponised to advance a particular narrative in the fields of migration studies, technology and critical security studies. In consideration of this, adopting the proper terminology is of fundamental importance, and it is helpful in clarifying how some terms will be assumed throughout this body of work.

Migration and Border Management

Refugees, migrants, and asylum seekers

The 1951 Convention Relating to the Status of Refugees and its 1967 Protocol are the key international legal provisions governing the rights of individuals who are at risk of persecution in their countries of origin. Particularly under art. 1(A.2) of the Refugee Convention, the term "refugee" is defined according to specific circumstances. A refugee is a person who,

> Owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country (UN General Assembly, 1951, p.14).

Refugees are therefore distinguished from *asylum seekers*, who are individuals in the process of seeking international protection, or from those labelled via the more generic term *migrant* (Sironi, Bauloz and Milen, 2019).

While these classifications are strict and well-established in law and policy, in reality, these categories often conflate. One explanation is that typically, people who embark on

journeys to depart from their countries of origin adopt the same migratory channels regardless of legal groupings, in what is commonly termed *mixed migration flows* (Sironi, Bauloz and Milen, 2019, p. 62). Therefore, this dissertation will employ the terms "people on the move", "crossing borders", or the more neutral umbrella term "migrants." The latter is not clearly defined under international law but may be loosely employed to entail individuals who leave their habitual residence for various reasons (Sironi, Bauloz and Milen, 2019). While acknowledging the specific kinds of vulnerabilities and different exposure levels to algorithmic bias that people who fall within these categories may face, this more inclusive terminology will be employed to allow for greater clarity and uncover patterns of inequality.

Migration management and border control

That being said, the concepts of migration and border management are also highly contested (Ansemes de Vries and Guild, 2019). Politicians' attempts to "manage" migratory movements are often in contrast with the stark reality of the complexity of such a phenomenon. Nonetheless, the terms have not ceased to be widely adopted in literature (Geiger and Pécoud, 2010). Moreover, the matter remains at the top of EU states' agendas as substantiated by the recent EU Migration and Asylum Package (European Commission, 2020b). Hence, following a generally accepted definition, the concept of migration management will be understood here as "the different strategies, policies, processes, and procedures negotiated and adopted by relevant actors at the international level to provide a framework to manage migratory flows in an orderly and predictable manner" (Beduschi, 2020, p. 3). The notion of migration management goes hand in hand with that of border control. Thus, this paper will rest on the EU's working definition,

> The activity carried out at a border, in accordance with and for the purposes of Regulation (EU) 2016/399 (Schengen Borders Code), in response exclusively to an intention to cross or the act of crossing that border, regardless of any other consideration, consisting of border checks and border surveillance (Art. 2(9)).

Smart borders

The term Smart border refers to the border control strategies and technologies relevant to this dissertation, *i.e.*, the EES, ETIAS, and RTP, which the Commission groups under the term

"package", particularly between 2011 and 2016 (European Commission, 2019). Nevertheless, following the Commission's own practice, the term will be occasionally utilised in a broader sense to encompass the multitude of tools and systems which were proposed since 2013 and have gone on to generate more complex machinery (Singler, 2019).

Artificial Intelligence, Machine Learning and Algorithmic Decision Making

Algorithmic Decision Making

As argued thus far, the field of AI has gained momentum in recent years, and the EU's attempts to provide a framework to navigate the use of new technologies is a manifestation of such impetus (Bentley *et al.*, 2018). However, despite this, there continues to exist a persisting lack of consensus regarding the exact definition of the terms AI, ML, automated decision-making systems, and predictive analytics. Therefore such terms are often used synonymously at a policy-making level (Bentley *et al.*, 2018; Borgesius, 2018; Molnar, 2019, 2020).

The baseline for the vast majority of these technologies is *algorithms*. Algorithms have been defined as "abstract, formalised description of a computational procedure" (Dourish, 2016, p. 3). The product of such a procedure is the decision determined by the algorithm. Hence, algorithms can be considered as a set of instructions that are "trained" via vast pre-compiled datasets to attain a specific outcome upon classification and generalisation processes (Gillespie, 2016). The training dataset can include e-mails, caselaw, pictures or stats and allow for a *fully automated* decision-making process or a *partly automatic process* (Borgesius, 2018). The first, for instance, applies to spam filters or smartphones' facial recognition tools. In contrast, the second implies that a human operator is in the loop, making decisions upon algorithmic assistance. Some examples of the latter include the systems utilised by human resources departments, bank personnel who determine customers' credit scores or even anti-plagiarism tools, which are then evaluated by examining officers (Alsallal *et al.*, 2016; Borgesius, 2018; Köchling and Wehner, 2020). When looking at the border management field, algorithms can are considered helpful in their ability to draw from vast amounts of data, including big data or the "high velocity, complex and variable data" (Tech America Foundation, 2012). Thus, making them essential systems of prediction and data management.

However, when analysing their discriminatory outcome, both types of decision-making processes may lead to undesirable outcomes (Lipton, 2018). Technology is often seen as fully reliable and rational; therefore, humans do not challenge its outcomes. In Wagner *et al.*'s

words, “the human being may often be led to “rubber stamp” an algorithmically prepared decision, not having the time, context or skills to make an adequate decision in the individual case” (2018, p. 8). This overreliance on automation is termed *automation bias* (Parasuraman and Manzey, 2010; Goddard, Roudsari and Wyatt, 2012).

Artificial Intelligence and Machine Learning

This paper follows the understanding of AI technology as the power of a system to execute tasks that are typically a prerogative of human intelligence, such as the acquisition of a skill, the ability to categorise, solve a puzzle, translate and make a decision (Osipov, Panov and Yakovlev, 2019). Furthermore, the systems’ ability to perform tasks is not solely defined by algorithms but also by the environment surrounding the machine itself (Kersting, 2018). AI is a broad branch of learning originating in the 1940s and formally founded in 1956 (Writer, 2014). The so-called “expert systems”, rudimentary recreations of human knowledge and the ability to draw logical conclusions, are examples of early applications. Their application was a catalysing moment for the research field while still facing heavy criticism for not being adaptable to real-world scenarios and not being cost-effective (Puppe, 1993).
It encompasses different types of disciplines hence explaining that the entirety of this subject matter is not relevant to this study.

ML is the branch of AI that symbolises such success that despite being a type of AI, ML tools are often inaccurately labelled as AI in popular culture (Lipton, 2018). Instead, ML can be seen as the scientific process which is “concerned with the question of how to construct computer programs that automatically improve with experience” (Mitchell, 1997, p. 15). Data mining and Big data also fall within the field of ML, and they each represent the discovery of patterns and knowledge through data and the analysis of vast datasets to extract “potentially useful information” (Frawley, Piatetsky-Shapiro and Matheus, 1992, p. 57).

While in full acknowledgement of the technical differences behind these kinds of systems, in order to allow for clarity and fluidity, in this dissertation, the concepts of AI, ML, Big Data will be used to imply technologies that automate systems and mimic human cognition. The objective of this research is, in fact, to address the impact of the use of these mathematical or computational constructs when used in matters of migration and borders. Thus, setting aside their particular technical characteristics does not alter the overall aim of the analysis. What is important to this study is how the improper legal framework in which the automation of migration decisions are made holds the risk of perpetuating discriminatory biases already

prevalent in society with little space for adequate safeguards. Therefore, the following sections shall address the ascendance of the AI and ethics nexus in literature and the concepts therein.

1.2 The AI and ethics nexus: an interdisciplinary account

The discriminatory outcomes of the use of AI systems in border management cannot be discussed or understood without providing some context on the current developments of the ethical implications of the use of AI systems. The depiction of all of the multifaceted layers of such a relationship does not fit this paper's scope. However, it is essential to contextualise the scholarly and technical framework within which this area of inquiry falls.

For years, the domains of ethics and AI were understood and thought of as highly divergent (Dubber, Pasquale and Das, 2020). AI has, in fact, been principally associated with the realm of the digital and the computational, where machines are able to mimic human behaviour and automate complex processes (Lipton, 2018). In contrast, ethics is understood mainly as a philosophical discipline striving to unravel and analyse how sentient beings should behave and interact (Dubber, Pasquale and Das, 2020). Alongside this alleged incompatibility, scholars and policymakers overlooked the possibility of discrimination arising from AI bias, even in the few instances in which the matter of AI's advancement was under direct scrutiny (Citron and Pasquale, 2014; Crawford and Schultz, 2014). The report titled *Big Data: Seizing Opportunities, Preserving Values,* published by the U.S. White House in 2014, provides an example (Podesta *et al.*, 2014). The paper's reception was generally positive as it was perceived as an innovative stance on the societal changes deriving from the big data "technological

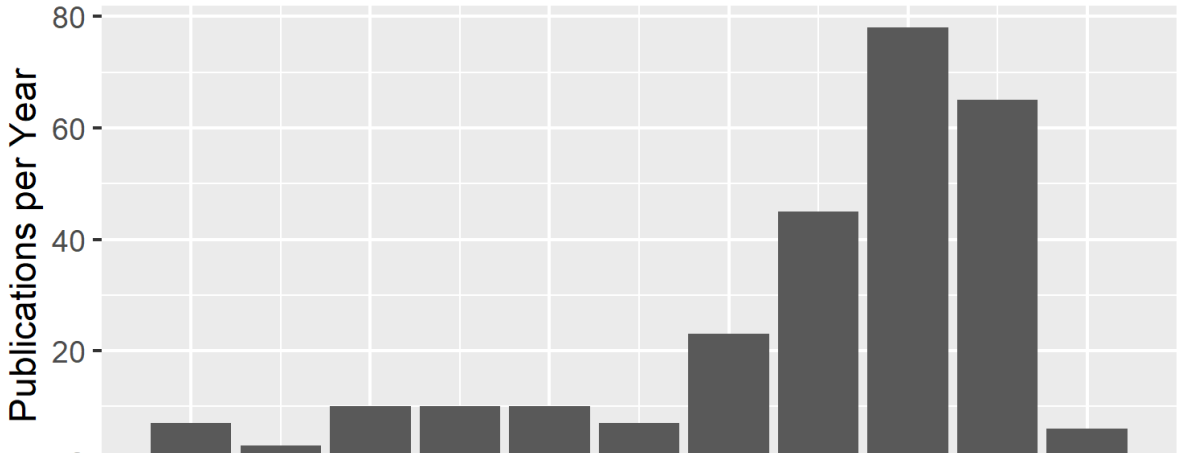

Figure 1 - Number of Papers related to Fairness in ML research. Source: Caton, 2020 < https://deepai.org/publication/fairness-in-machine-learning-a-survey> Retrieved from open source.

revolution" (Podesta *et al.*, 2014, p. 3). As illustrated by *Figure 1*, the effect of the introduction of AI systems in everyday life was still a primarily understudied phenomenon in literature at the time. However, amidst the analysis of its multiple applications, the possible discriminatory effects of big data are merely mentioned. Moreover, while acknowledging that "this sea of ubiquitous sensors, each of which has legitimate uses, make the notion of limiting information collection challenging, if not impossible" (Podesta *et al.*, 2014, p. 54), it refrains from explaining how this may happen and offers minimal solutions. As we shall see in Chapter 3 of this paper, this scarce attention to the depth of the phenomenon is not unique to the U.S.; on the contrary, it is a characteristic it shares the EU.

Nevertheless, recent years have seen a progressive narrative shift. The combination of an increase of computational power, alongside today's possibility of accessing vast amounts of information, sparked an interest in AI, and particularly for its ethical implications (Beduschi, 2020). Accordingly, despite the apparent contrasting nature of technology and ethics, the all-pervasiveness of AI has resulted in an increase in studies relating to such fields from 2016 onwards, as shown in *Figure 1* (Caton, 2020). A significant number of these are directed at investigating the (*un*)ethical and discriminatory effects of the application of automated systems in their multifaceted applications.

AI and discrimination: a broad spectrum

The emergence of the plethora of research on the issue of algorithmic bias has signalled a broad range of instances in which this matter has brought forward issues of discrimination. Some studies have focused on the commercial pitfalls of the application of AI in an attempt to uncover this technology's discriminatory effects in areas ranging from justice to public housing.

In 2016 U.S. courts decided to introduce the "Correctional Offender Management Profiling for Alternative Sanctions" (COMPAS) tool developed by software company Northpointe. This was a support instrument that was created to assess the degree of recidivism of a criminal offender in order to aid the justice system and ensure the *maximum level* of objectivity amidst legal procedures (Northpointe, 2019). In May of that year, the use of Northpointe's AI tool began to feature in numerous trials across the US. The news outlet ProPublica released a report on the problematic outcomes of its usage (Angwin *et al.*, 2016). The study found that COMPAS committed two significant forecasting errors: on the one hand, black people were twice as likely to be labelled as potential re-offenders; white offenders, on

the other hand, were categorised as more likely to be involved in minor infringements (Angwin *et al.*, 2016). COMPAS was not developed by utilising sensitive data, and contrarily it was built based on "criminogenic needs", which are at the heart of some of the most prominent criminology theories (Northpointe, 2019). These included characteristics such as "criminal personality," "social isolation," "substance abuse", and "residence/stability" (Angwin *et al.*, 2016: no page). The developers of this system refused to disclose the exact mechanisms in which their algorithm arose to such conclusions, claiming a potential breach of the company's intellectual property rights (Vanian, 2018). Nonetheless, they also admitted that the lack of acknowledgement of factors such as race, gender and employment history is a serious shortcoming as it reduces the accuracy of the system's predictions(Griffard, 2019). Such shortcomings have real-life consequences on those put on trial.

As the discipline came to be explored in more depth, the link between computerised decisions and discrimination was rendered more apparent. For instance, in her 2017 piece, Eubanks argues:

> high-tech tools have a built-in and patina of objectivity that often lead us to believe that their decisions are less discriminatory than those made by humans. But bias is introduced through programming choices, data selection, and performance metrics. The digital poorhouse, in short, does not treat like cases alike (Eubanks, 2018, p. 12).

Her work, centred on what she terms a "digital poorhouse", provides a convincing portrayal of today's data management and technology systems and their problematic effects on the lives of indigents in the U.S. In doing so, she traces a timeline of events of U.S. public welfare since the New Deal, emphasising the correlation between the country's approach then and in recent history (Eubanks, 2018). Her account exemplifies the damaging social consequences of the automation of welfare eligibility systems across the nation which, under the guise of government efficiency, have had severe impacts on the most vulnerable (Eubanks, 2018).

Understanding the socio-economic consequences of automated decision-making in this manner is fundamental in allowing members of the public who may not fully grasp the magnitude of this "technological revolution" and its side effects. In this way, the same cognitive biases that have led us to believe that technology is inherently just can be used to scientifically convey its potentially damaging consequences, especially on those most vulnerable, such as the migrant population.

Other scientists have offered more technical accounts of the link between epistemic inequality and technology in a similar vein. In her Weapons of Math Destruction, former Wall Street analyst and mathematician Cathy O'Neil analysed the discriminatory use of algorithms and big data in various applications ranging from advertising and education to justice and policing (2016). As she explains, these mathematical systems present three controversial features, they are obscure "black boxes", unregulated, and difficult to dispute, thereby augmenting intrinsic societal biases and potentially affecting vast portions of the population (O'Neil, 2016). Nonetheless, despite contending that "big data increases inequality and threatens democracy", she still instils a glimmer of hope in the belief that more ethical uses of statistics, modelling and automated systems may lead to an increased level of justice and fairness in our society (O'Neil, 2016).

Within this new niche area of research, some studies have focused on the role AI plays in perpetuating race and gender societal biases, paving the way for the increasingly specific context of border control. In December 2020, leading AI ethics researcher, Timnit Gebru was dismissed from major tech company Google following the publication of her ground-breaking paper on the dangerous, discriminatory effects of facial recognition software (Hao, 2020). Alongside the environmental and financial risks connected to the vast amount of computer power necessary to process large language models, the paper highlighted their heavy reliance on enormous amounts of text (Hao, 2020). Given the exponential growth in the use of social media and the internet over the past decade, this translates into a significant amount of training data that is susceptible to racist, sexist and abusive language. As with the previously mentioned case of Microsoft's Tay, an AI model which is trained on internet linguistic content will not be accustomed to the nuances embedded in language and mitigated by cultural and societal norms and interpretations.

The results of Gebru's study are similar to those found by Joy Buolamwini. While conducting research at M.I.T, the Ghanaian-American computer scientist noticed that the facial recognition software she was working on was unable to recognise her face (Buolamwini, 2020). As Buolamwini herself explains in her feature-film *Coded Bias,* the software was only able to discern her somatic characteristics upon physically placing a white mask above them. This episode, she held, encouraged her to explore the racial and gender inequalities linked to the use of algorithmic tools in everyday systems (Buolamwini, 2020). In an earlier paper co-authored by Timnit Gebru (2018), the two scientists explored today's systemic erasure of women and minorities via technological means. In doing so, they discover the connections these bear with years of racialised and gendered malpractices in U.S. history

and the overrepresentation or underrepresentation of such categories in AI systems (Buolamwini and Gebru, 2018). By assessing the precision of three commercial gender categorisation algorithms and including the skin type parameter, their research concludes that the system favours lighter-skinned individuals and males (Buolamwini and Gebru, 2018).

While facial recognition systems should not be trained to ascertain an individual's chance of attaining asylum, being admitted to a college or whether or not they should be granted parole, these decisions are often based on algorithmic identification. Hence, as the authors contend, transparency, accountability, and expanding demographic and phenotypic subgroups are essential to achieve increased algorithmic output accuracy (2018).

Algorithmic use in migration management technology

As argued thus far, these studies have been trailblazers in the field by introducing elements of race, class and gender in the field of algorithmic inequality. However, very few studies have ventured into the intersection of border security studies and AI ethics. This offers a clear indication, despite the surge of research on the issue, that the current global digital rights field has given minimal space to issues concerning migration. This dissertation fits within such intersection.

Nevertheless, a few pieces of research from recent years have shown that Big Data and AI technology have the potential to aid in the prediction of migration movements and in several administrative aspects of the managing process (Barocas and Selbst, 2016). States and international organisations' enthusiasm towards these kinds of applications is growing. Such is the case for two of the most prominent organisations whose mandate is that of protecting the rights and well-being of migrants worldwide, namely the UNHCR and the IOM. In fact, both organisations have been increasingly relying on data mining and tracking systems in order to trace, monitor and identify people on the move or provide assistance to migrants across the globe (IOM, 2007, 2018; UNHCR, 2017). The case of UNHCR's EyePay Cash offers an insightful example. In 2016, the UN Refugee Agency - in partnership with the World Food Programme (WFP) - adopted iris scan technology to provide cash assistance to refugees in a refugee camp in Jordan (UNHCR, 2019). This collaboration between the UNHCR and IrisGuard, a UK-based private technology company, dates back to 2012. However, such company is not new to business partnerships with international entities, with clients including WFP, UN Women, UNICEF, IOM, and many non-governmental organizations (NGOs) (Burt, 2020). Nevertheless, their use of hardware enabling non-contact authentication entails the

collection of biometric data, thus racing concerns over data protection, privacy issues and consent. As MENA Policy Manager at *Access Now*, Marwa Fatafta points out, "When you have no choice, you cannot consent. […] Forcing people with little recourse, such as refugees, to surrender private information in exchange for food is an affront to human rights standards, and an insult to human dignity" (Access Now, 2021: no page).

International organisations' actions have a significant impact on the way the phenomenon is perceived at large in the field. If organisations whose core mission is acting as human rights promoters and protectors are perceived as primary users of technologies that enable potential human rights breaches, this sets the overall international standards at a shallow bar. States are susceptible to such influence, however, their ability to set legally enforceable international rules renders them extremely powerful. Organisations such as WFP and UNHCR's actions have a great degree of resonance thus could shape or alter understandings of human rights implications of such applications.

As this dissertation contends, AI utilisation in this field is endless, hence the gap between the trial and implementation at the EU level is narrowing every day. Moreover, current practices already hint towards their proliferation. They are ranging from border and security control, visa background checks, personal identity verification and during the various stages of the asylum-seeking process (Harryson *et al.*, 2018).

Lawyer and human rights activist Petra Molnar, one of the most prominent voices in the subject matter, has provided an account of how this is already happening. In her most recent report, *Technological Testing Grounds,* she offers a legal insight into the effects of these kinds of technological "experimentations" on migrants (2020). Through the visual aid of photography and by engaging directly with the migrants at Europe's peripheries of Levos and Moria in Greece and refugee camps in Belgium, she attains a unique point of view (Molnar, 2020). The approach used in this investigation is comparable to that used by other researchers in the diverse contexts offered by countries such as Switzerland, Germany, Canada or the U.S. (Angwin *et al.*, 2016; Molnar and Gill, 2018; Heikkila, 2021). In the latter, data analytics are employed to calculate prospective positive outcomes of the post-resettlement processes in the U.S. on the grounds of pre-existing societal ties with the country (Shashkevich, 2018). In Molnar's words, "the rise in the use of biometrics, or the 'automated recognition of individuals based on their biological and behavioural characteristics' in migration management" is already a reality (Molnar, 2019, p. 7). Examples of their uses on historically marginalised groups are compelling and date back from the 1930s to the current days. Black, for instance, provides an illustration of the partnership between multinational corporation IBM and Nazi Germany

during the Second World War in an attempt to collect a significant amount of data on the Jewish population in order to "expedite" the Holocaust (2012). Similarly, recent news has exposed the use of "datafication of refugee responses" by oppressive governments. In this regard, research has expanded on the case of Myanmar and China. The first state was found to be documenting and tracking members belonging to the Rohingya minority in efforts to expel them from camps (Thomas, 2018). China's case also rose to headlines in 2019 due to its controversial "Muslim tracking database" (Cimpanu, 2019: no page). In this case, private companies supplied the government with video-based crowd analysis and facial recognition technology with the aim of identifying members of the Uighur Muslim minority (Cimpanu, 2019).

The leitmotif of these studies is their attempt to challenge the predominant *false belief* that science is "objective and fair" while evidencing how this is not a reflection of the current reality in the migration and border management contexts. In these understandings, the deployment of AI technologies without appropriate safeguards can amplify social divides, and open to unfair practices exacerbate inequalities. This dissertation will rely on such arguments while pushing the conversation forward by addressing how this has been effectively pursued in the EU context. The research will be innovative in its attempt to analyse a trialled programme, namely as iBorderCtrl, and by framing the discussion under the interdisciplinary lens of law, STS and sociology.

Nonetheless, current literature on the matter has also espoused the opposing narrative in a bid to address the beneficial and advantageous aspects of AI-enabled decision-making in the field. Harryson *et al*., for instance, maintain this line of argument (2018). In their view, the nature of asylum claims is heavily reliant on predominantly manual administrative tasks (such as interviews, questionnaires, data acquisition), which, with the aid of AI systems, would be performed more rapidly and at a reduced cost (Harryson *et al.*, 2018). This is already taking place in some European countries such as Germany and Sweden. The German Federal Office for Migration and Refugees (Bundesamt fur Migration und Fluchtlinge, BAMF) has performed trials and is now adopting technology such as computerised face or language detection, name transliteration, and mobile devices data analysis for proof of identity to assist officers' determinations (Federal Office for Migration and Refugees, 2018; Beduschi, 2020). As Carammia and Dumont (2018) put forward in their OECD report, Sweden has also been at the forefront of technological innovation. As they contend, the country's authorities have utilised ML-based prediction tools to forecast imminent migration movements in order to address the phenomenon accordingly and in a timely fashion (Carammia and Dumont, 2018). Other studies have also shown that the political will, or current adoption of migration management

technology, is not unique to Western countries. On the contrary, nations such as Nepal, Malaysia and Brazil are also entering the field as primary actors and employing AI to aid border management agents in their decision-making processes (Molnar, 2019). Promoters of the positive link between AI and migration or border management have suggested that this may aid in bridging world disparities by allowing for a more democratic access to such tools (Beduschi, 2020). In the context of the previously mentioned globalised race for algorithmic gold, historical power imbalances resurface in the form of technological privilege. The acquisition of these tools by countries belonging to the so-called Global South, in this view, can represent an opportunity of narrowing the digital divide between powerful states and less-prominent actors. Thus, allowing for the decision making processes to shift in the hands of sending states.

A substantial portion of academic and technical literature has also concentrated on describing the types of bias that may or may not have discriminatory effects. From this viewpoint, authors have argued against the normative significance of discrimination in the context of AI and its outright inappropriacy when addressing ethical issues (Heinrichs, 2021). In this context, Heinrichs' interpretation of Lippert-Rasmussen's concept of *social saliency. i.e.* the relevance of specific social attributes in the definition of group membership, for instance, provides some interesting points of inquiry (Lippert-Rasmussen, 2006). In such contention, the claim that AI is conducive to an aggravation of discrimination issues is rejected on the basis of philosophical understandings of the concept of discrimination in itself (Heinrichs, 2021). Nonetheless, this opens to arbitrary criteria over what amounts to discrimination, thus rendering the concept extremely volatile and depriving it of meaning and purpose. Epistemic understandings of discrimination which the author believes hinder a sound ethical analysis of the phenomenon, are rendered inconspicuous in the face of the empirical evidence of cases in which AI tools have indeed amounted to discrimination, as previous paragraphs have addressed.

Finally, it is worth noting that research in the area has also emphasised the necessity of removing bias from specific tools (O'Neil, 2016; Hildebrandt, 2019; Caton, 2020; Köchling and Wehner, 2020). However, while these technology-driven analyses remain valuable contributions to the overall comprehension of the complexity of these topics, this is only one aspect of the issue at hand. The societal application of systems beyond the technical understandings of AI machinery have had marginal attention and are only recently gaining traction, and this thesis reinforces the notion that the removal of biased technology does not automatically equate to the eradication of discrimination itself, as Chapters 2 and 3 will discuss.

The proposed argument points out that these new technologies challenge existing normative frameworks around decision-making and procedural protection in the migration field. As such, these applications carry long-lasting effects on people on the move by impacting fundamental rights and entitlements such as freedom from discrimination. Global entities such as the EU should provide appropriate safeguards to limit their disproportionate effects. Therefore, under what lens should technology be understood? The next section of this study will rely on three theoretical contributions to answer this question.

1.3 The Theoretical framework

This section of the thesis will introduce three IR (International Relations) theories that will serve as a canvas from which insights will be drawn to depict the theoretical framework of the proposed analysis. In order to answer the research questions over the potential discriminatory risks deriving from the adoption of AI tools in border management, it is essential to understand the EU's approach to the migratory phenomenon on a larger scale. As Chapter 2 will illustrate, EU discourses propose problematic links between migration and security matters. Hence, theories pertaining to the securitisation of the phenomenon will allow for a deeper comprehension of the logic at play in this context. As per the totality of this research, issues of migration are also interlinked with questions of technology. STS scholarship is a valuable lens to make sense of conceptualisations of borders beyond their material features. Finally, matters of borders necessarily introduce questions of exclusion, inclusion and belonging. Here, Judith Butler's theory of state performativity and precarious life will be essential in addressing such aspects.

Migration and Securitisation

Since the 1980s, the "Europeanisation of migration policy" (Huysmans, 2000, p. 752) has been tied to political and societal conceptions aimed at emphasising concerns over matters of sovereignty. Moreover, owing to the particular interplay of national and supranational powers that shape the EU governing structure, questions of territorial control have featured persistently in the organisations' agenda in the form of policy provisions such as the Dublin Convention, the Schengen Agreements and most recently, via the definition of a European Agenda on Migration (European Commission, 2020b).

However, contrarily to what one could infer from the populist political climate currently gaining traction across Europe, the concepts of security and migration were not always read in conjunction (Taylor, 2019). In his eloquent analysis of the ascendancy and proliferation of smart borders in the European Union, Springer traces a historical timeline of the definition of this dichotomy (2019). As he elaborates, during initial calls to expand on EU-level competencies in the realm of migration policy, the Commission contended that the main issue determined from irregular migration into Europe was "illegal employment" (Singler, 2019, p. 39). Policies concerned the need to ensure that applications were processed fairly and swiftly in order to prevent migrants from facing challenges connected to the lack of documentation, such as attaining a permanent residence status (Singler, 2019). In subsequent years, the Commission views changed with policies addressing "external borders, asylum, immigration and the prevention and combating of crime […] in particular terrorism (European Commission, 2001, p. 7). What explains such shift? From an empirical viewpoint, numerous studies have shown that the link between terrorism and migration is artificial as it finds minimal evidential corroboration (Cruz, D'Alessio and Stolzenberg, 2020; Helbling and Meierrieks, 2020). From a theoretical viewpoint, the concept of securitisation of migration offers some clarity. Securitisation refers to the identification of an issue as an "existential threat" (Buzan, Waever and de Wilde, 1998, p. 23).

Irrespective of whether such menace may be genuine or perceived, this may be used to justify resulting in special procedures or the enactment of emergency policies to tackle or mitigate the phenomenon (Huysmans, 2000). From a political viewpoint, this is substantiated by shifting the matter from the peripheries to the centre of a national or supranational entity's agenda in order to protect what is perceived as a fundamental asset, namely its territory, citizens or governmental apparatus (Balzacq, 2005). The Copenhagen School established the most prominent notion of securitisation. In this conceptualisation, the process is enabled through a "speech act", entailing that "security is not of interest as a sign that refers to something more real, the utterance *itself* is the act" (Wæver, 1995, p. 55). In this line of understanding, security threats are not objective. They instead emerge via the discursive outlining of an issue as threatening in an attempt to justify an immediate call to action. Later theorisations have allowed for the sophistication of the concept. In light of criticisms pointing to the narrow focus of such conceptualisation on discursive utterances of political elites, authors such as Williams and Balzacq have expanded the theory to encapsulate the role of ulterior actors such as the media and that of the intersubjective relation between the performer and the audience (Williams, 2003; Balzacq, 2005). Moreover, with specific reference to the EU, Jef Huysmans also

contends that the securitisation of migration finds its roots in the historical bureaucratic and legislative occurrences connected to the European integration project. Therefore determining the prolonged and relentless process that pinned it steadily to the top of the security agenda (2000).

Finally, academics from the Paris School provided a deepening of the notion. For the purposes of the present study, two concepts become of particular interest to the circumstances of the adoption of AI and smart tools in EU border management. Firstly, the notion that the securitisation process is not necessarily linked solely to existential discursive actions on the part of politicians and elites, but rather also via less straightforward means (Huysmans, 2000). Hence, this dissertation relies on less noticeably politicised pieces of legislation, white papers, reports, *et similia.* Secondly, it provides theoretical support for the central argument to this thesis, namely the belief that using tools such as facial recognition systems, smart borders and surveillance and monitoring technology can become a source of insecurity and discrimination as they influence exclusionary practices despite their seemingly neutral and impartial settings.

Essentialism, Instrumentalism and borders as Socio-technical devices

As argued, both the Paris and the Copenhagen School offer important theorisations for the analysis of the phenomenon of migration securitisation. Nonetheless, neither of these schools of thought offer guidance in relation to the role of the objects in the matter. Hence, this thesis avails itself of the growing body of knowledge in the subject area of critical security studies by drawing from Sociology and Science and Technology (STS).

From an STS viewpoint, objects play an essential role in shaping our understandings of security by creating the physical conditions for the possibility of security matters to arise in the first place. In this way, smart borders and AI tools are no longer simply immobile objects. Instead, they simultaneously create conceptualisations of security of their own by their simple introduction to the environment (Salter, 2015). Here security becomes an intersubjective social construct in which the materiality of tools is enmeshed with the subjectivity, thus "fundamentally alter[ing] the condition of human possibility in ways that are unpredictable and irreducible to their constituent elements" (Salter, 2015, p. viii). Therefore, in the context of the discriminatory outcomes of AI-systems enabled migration management tools, this analysis calls for an acknowledgement of the technical elements as well as their social implications. Chapter 3 will accordingly address those two aspects.

As the introductory section of this thesis emphasised, technology has been expanding its reach at a rapid pace. Interestingly, IR critical security studies' literature still lags behind when conceptualising technology in what STS scholars like to call the "unopened black box" within the field (McCarthy, 2017, p. 5). Academic debates have seen two prominent factions. One side of the argument sees technology as a topic of significant interest, thus placing it at the heart of IR in an acknowledgement of its role in shaping global politics (Keohane and Nye, 1977; Scholte, 2005). Critics, on the other hand, suggest that "determining [the] quality of technology is […] overrated" (McCarthy, 2017, p. 4). In this view, by focusing too deeply on inanimate objects, other important forms of control or sources of social action are neglected.

Questions naturally arise concerning what accounts for such blind spot. However, other studies have undertaken this task (Buzan, Waever and de Wilde, 1998; Scholte, 2005; Salter, 2015). For this research, the STS-IR approach is useful to expose the intended or unintended impacts of security technologies on the people that are exposed to their use. Generally, the main STS-IR theories fall into two macro-categories, technological essentialism and instrumentalism. The two will be important in explaining the rationale behind the perceived neutrality and fairness in the EU's proposed approach to AI technology. Therefore, these theories will represent an added value to answering the first of the two research questions, i.e., *What are discriminatory risks involved in the use of AI technology in EU migration and border management and how are they justified?*

Essentialism is the theory that holds that technological progress is independent of exogenous social factors. Rather in essentialists' view, scientific and technical evolution follows a steady line depending on "autonomous discovery" (McCarthy, 2017, p. 8). Such discovery, in this view in entirely dependent on the work "of inventors, engineers, and designers following an internal, technical logic that has nothing to do with social relationship" (McCarthy, 2017, p. 8). Therefore, the insertion of new security technologies in society is capable of pressuring nation-states either into reacting or ignoring such new developments at their own risk (Scholte, 2005). Nevertheless, the pace of technological change progresses untarnished while tipping the scale of global power relations, thus leading states to defend the introduction of new measures accordingly. Jordan Branch's analysis of contemporary experimental improvements in the creation, distribution and use of maps provides a clear example. As his research demonstrates, the introduction of mapmaking technologies moulded notions of sovereignty and authority instead of them being designed in support of political necessities (Branch, 2014). This insight is significant for this dissertation, as, in light of essentialists understandings, the assumptions of adaptability and neutrality at the backbone of

the EU's AI agenda clearly overlook their potential for technology's discriminatory effects (European Commission, 2020b). As Chapter 2 will elaborate on, the narrative which emerges from EU official documentations rather appear to favour securitised understandings of migration at the expense of the fundamental rights of people on the move.

In an opposite line of research, instrumentalists refuse notions of material determinism advanced by essentialists. In their evaluation of technology, technological artefacts are indeed neutral. Despite their undisputed role in accruing or perpetuating some pre-existing global power structures, they are simply one of the many manifestations of state power (Herz, 1976). Ultimately, in a Machiavellian interpretation, technology is seen as an "exogenous" means for a political end. This instrumentalist perspective is advocated by authors such as Alexander Wendt. In his theorisation, although material artefacts have some impact on the construction of political power relations, their effects are not innate. Rather, they are dependent on the way in which these are employed by the ruling bodies (Wendt, 1999). This approach is similar to that proposed by advocates of the value-neutrality thesis, *i.e.,* the belief that moral valence cannot be ascribed to a technological tool as its intrinsically impartial. Its purpose is merely determined by human application (Pitt, 2014). The example of guns is usually engaged to buttress this theory. The American National Rifle Association (NRA) slogan "Guns Don't Kill, People Kill" provides a practical explanation (Pitt, 2014). In this view, the materiality of an object only takes relevance once it is employed to serve a specific purpose. However, critics of this theorisation have argued that it overemphasises the actors' will over the materiality of objects, almost entirely disregarding the impact of its presence in the social world (Johnson, 2013). In relation to guns, for instance, while such instruments depend on the triggering action of an individual, gun control exponents argue that the sole purpose of manufacturing a weapon is that of harming or killing (Johnson, 2013). Their materiality and social aims cannot be separated as their uses are confined within such materiality (Pitt, 2014). As Chapter 2 of this analysis will go on to substantiate, in the realm of smart border management systems, this kind of argument is challenged from both a technical and a legal viewpoint as bias can be introduced at the onset of AI-tools' training and subsequently in the manner in which they are used.

Nevertheless, as it is clear, both essentialist and instrumentalists understandings of technology provide important tools to comprehend the potential risks involved in the adoption of technology and their significance. In this regard, Stefan Fritsch and Langdon Winner propose theorisations that have drawn from both instrumentalist and essentialist understandings of technology, thus allowing to find some commonality between the two theories, namely the idea of socio-technical devices. Socio-technical tools are "the mix of material and social

institutions that cohere around artefacts" (Herrera, 2012, p. 36). As Winner explains, technical arrangements and social order correlate. In this sense, the process of technological development is just as critical as its political implications, hence highlighting the role of all stakeholders involved (Winner, 1980). Moreover, socio-technical devices arise following an alignment of a technical imperative with the establishment of a contingent narrative, namely an appropriate *technological momentum* (Hughes, 1969). Such connotations are fundamental in the context of this dissertation as they emphasise the role of human actors and agency in developing and deploying technology. This is useful in understanding the impact of supranational organisations in developing technologies that are then utilised in ways that may infringe fundamental human rights. Moreover, it introduces dynamics of power, which are central to theorisations of the state and have been prevalent in critical security studies since their inception.

Precarity and State performativity

The theoretical insights hitherto have exposed the challenges related to notions of securitisation and criminalisation of the migration phenomenon, as well as the underlying issues concerning States' use of technology. Judith Butler's theory of State performativity complements these theoretical insights by providing a framework that will be utilised in the discussion part of this thesis to understand the more profound sociological implications over issues of technological experimentation in migration.

Butler's notion of performativity is famously associated with her feminist analysis of gender, in which she argues that gendered acts continuously constitute and construct one's identity. Thus gender is "a negotiation with power" (Butler, 2006: i). In later works, however, the concept gradually expanded to more general concerns linked to the notion of precarity. In her contention, precarity describes different conditions that relate to humans, including social and political institutions such as nation-states. Precarity, in her own words, "also characterizes that politically induced condition of maximized vulnerability and exposure for populations exposed to arbitrary state violence and to other forms of aggression" (Butler, 2009, p. 2). Such conditions can be promoted by States, yet they are not a mere example of aggression. In this view, they are a broader reflection of power relations, one of the ways in which it operates and reproduces (Butler, 2009). In the matter at the heart of this thesis, such theorisation becomes valuable as it exposes a logic of control over individuals and populations. Therefore, this study argues, border technology must be understood as a socio-technical tool that is capable of reproducing the patterns of inequality and discrimination intrinsic to the society from which it

emerges. The next Chapters will examine what risks this may entail for migrants at EU borders and which safeguards has the Union envisioned to address potential infringements.

Chapter 2.
Technologies of Discrimination

This Chapter provides an analysis addressing the first of the two research questions, namely, *What are discriminatory risks involved in the use of AI technology in migration and border management, and how does the EU justify its practices?* The analysis will focus on the consequences of the adoption of the tools and systems which are in place or that have been proposed at the EU level. In order to understand how this use of technology may amount to discrimination and how the EU justifies its practices, the opening paragraph will discuss the way discrimination may creep into AI systems from a technical viewpoint. The second will provide an illustration of current EU and member states' measures, analysing controversial human rights infringements which may arise and finally, the third will discuss some already existing practical applications.

2.1 How can AI lead to discrimination? A technical overview

The concept of discrimination carries a multitude of definitions and understandings depending on its field of application. Following the EU common understanding, discrimination is "judgement based on preconceived notions or prejudices, as opposed to the impartial validation of facts" (Crawford, 2017: no page). Upon thorough analysis, this dissertation's working definition decided to focus on the instances in which this causes "the unfair or unequal treatment of an individual (or group) based on certain characteristics such as income, education, gender or ethnicity" (Ferrer *et al.*, 2021, p. 2). The following section will first provide some background, in lay terms, on how discrimination can be introduced in AI systems. The following paragraphs will focus on discrimination's digital and legal conceptualisations by providing some information on how the matter has been assessed in the field.

Bias and discrimination in AI systems

From a technical viewpoint, the concept of discrimination emerges primarily from the literature concerning bias. However, in a computer science understanding of this concept, bias is not necessarily a negative characteristic. On the contrary, bias can allow for a deviation. Therefore, how can AI lead to disproportionate disadvantages for a particular group, or in other words,

when does it result in unlawful discrimination? Barocas and Selbst's analysis identifies several instances in which this may arise (2016). This section traces a pattern based on such blueprint and identifies four specific areas, i.e., *the target variable and class labels, the labelling and collection processes, institutional discrimination and utilisation bias.*

a) The *target variable* and *class labels*

Data mining automates the mechanisms that discern patterns in procedures such as fraud detection, credit scoring or spam filtering. For instance, concerning junk mail, companies generate a large number of e-mails that a human operator then categorises as "spam" or "regular mail". This process determines the training data. The computer then learns to associate certain traits, or proxies, to the messages labelled as junk mail and thus defines a "predictive model". The outcome of such identification is the "*target variable*" (Barocas and Selbst, 2016, p. 678).

On the other hand*, class labels* separate all the possible values within the target variable into distinct classifications (Barocas and Selbst, 2016). In some cases, this may be a straightforward mechanism. However, in others, this may be more challenging as, in some instances, "defining the target variable involves the creation of new classes" (Barocas and Selbst, 2016, p. 679). Moreover, given that the initial categories of data themselves result from a simplification of existing attributes and their arbitrary selection, this may lead to discriminatory effects on its onset. For example, in real-life scenarios, questions may arise when observing the possible automation of asylum applications. Official research has shown that governments and international organisations worldwide have engaged with data collection and information sharing through pre-screening questionnaires, phone surveillance and registering processes (IOM, 2007; Chase and Dick, 2017; UNHCR, 2017; Molnar and Gill, 2018). How would that data be classified in the prospect of the mechanisation of the asylum screening, for instance? What characteristics would be considered? Fundamentally, discrimination can be inserted into AI tools even at the outset, based on how an organisation or government defines target variables and class labels.

b) The training data: labelling and collection processes

"Data mining learns by example" (Barocas and Selbst, 2016, p. 680). Algorithms make decisions or forecasts by relying on datasets they have been exposed to, i.e., the previously mentioned *training data*. If such training data is biased by reflecting prejudices or stereotypes

prevalent in society, the algorithm will reproduce those same discriminatory patterns. Further bias may also arise when the traits of different groups are not adequately represented, an occasion in which we talk of "unequal ground truth" (Ferrer *et al.*, 2021, p. 2). Lowry and Macpherson's report illustrates how this occurred in selecting students at St. George's Hospital, a medical school in the UK, in the 1980s (1988). The hospital created a computer program to help with the student selection process as it faced a high level of applications. The program's training data was based on admission files from previous years, which, as was subsequently discovered, had systematically disadvantaged applicants from ethnic minorities and women who would have otherwise possessed precisely the same qualifications as other applicants (Lowry and Macpherson, 1988). As journalists at the time noted, "[T]he program was not introducing new bias but merely reflecting that already in the system" (Lowry and Macpherson, 1988, p. 657). In other words, the program had automated pre-existing prejudices. The biased sampling of data can also lead to similar outcomes (Lum and Isaac, 2016). This happens in the context of law enforcement's predictive policing systems. As they note, if a portion of the population of people from a particular neighbourhood is either over or underrepresented in police registers, AI systems will learn that people from a particular background are more prone to commit a crime. In this manner, the algorithm confirms an already prejudiced position and generates a negative feedback loop (Lum and Isaac, 2016). This occurrence is counterproductive for various actors, i.e., the neighbourhood who is unfairly targeted; the police forces, whose resources are misplaced by sending a greater number of officers to areas which do not necessitate heightened patrolling; and the community at large, who is disserved on the basis of skewed AI predictions.

c) Institutional discrimination

Numerous technical papers in the field tend to exclude this type of discrimination from official classifications as it is presumed to be of sole socio-political interest. However, as Barocas and Selbst briefly explain, "implicit biases and inertia within society's institutions […] account for a large part of the disparate effects observed" (Barocas and Selbst, 2016, p. 673). In their contention, this particular type of discrimination may be determined by unconscious biases, or a by-product of inherited prejudice of previous policymakers and reproduction of societal biases, in what is termed *unintentional discrimination* or *disparate impact.* Nevertheless, institutional settings can also be fertile ground for intentional discrimination. As Kroll et al. claim, "a prejudiced decisionmaker could skew the training data or pick proxies for protected

classes with the intent of generating discriminatory results" (Kroll *et al.*, 2017, p. 682). When intentional discrimination is carried out via proxies – i.e., features that appear to be neutral criteria yet are correlated to membership in protected groups – the resulting unfairness is more challenging to detect (Kroll *et al.*, 2017). Thus, knowledge of this and its deliberate manipulation can also lead to replicating or exacerbating historical or societal inequalities. For instance, classifications such as country of origin can be used to make problematic inferences on the race or gender of people crossing borders leading to discriminatory outcomes.

d) Utilisation bias

Finally, algorithms may generate discriminatory effects when employed in instances that differ from their initial purpose. For example, Ferrer *et al.* talk about "transfer context bias" (2021, p. 2). As they contend, when an AI system is utilised to forecast a particular situation for a specific group of individuals, it can lead to incorrect findings when applied in a different context or on a different population (Ferrer *et al.*, 2021). This may also be the outcome of incorrect interpretation of outputs, where the inaccurate understanding of the results can generate biased decisions.

As this technical overview demonstrates, algorithm-based tools can be discriminatory in numerous ways, some of which relate to their training process or to the way data is acquired, while others pertain to their use in institutional settings or deliberate discriminatory practices. In line with the notions espoused in the theoretical framework, such technology cannot be separated from the societal background which it serves and cannot be seen as neutral tools in purely essentialist understandings. On the contrary, it is at once shaped and capable of shaping the migration narrative when employed in such context.

2.2 The EU border management system: a panopticon of systems and technical tools

This paragraph borrows its title from Foucauldian theorisations of Bentham's technology of control. The term, in fact, was used to define the disciplinary system that was put in place to observe its occupant without their knowledge and has become a powerful metaphor for today's use of the multitude of technologies of surveillance and monitoring (Foucault, 1995). The next portion of the analysis will illustrate the current EU and international landscape of border

management tools to offer an understanding of how these may impinge on freedom from discrimination.

Technologies of border management in the international arena can be traced back to the First World conflict, as significant efforts began to be put in place in order for states to address issues of identification of citizens within their borders (Dumbrava, 2021). This led to the introduction of the first low-level form of border technology, namely, passports and comparable types of identification documents. Amidst the Second World War, their use was expanded owing to the standardisation process developed by the International Civil Aviation Organization (ICAO) in 1944, with significant effects to this day (Dumbrava, 2021). Since then, border management technology grew increasingly sophisticated with the introduction of fingerprinting, the invention of photography and polygraphs (Amelung, Granja and Machado, 2021).

In line with sociological understandings of *technological momentum* and securitisation theories discussed in Chapter 1, the EU has justified the introduction of a new integrated smart border system under growing concerns relating to crime and, in particular, the threat of terrorism (European Commission, 2019). For instance, in 2017, the Commission officially drew links between "an increase in irregular border crossings into the EU" and "an evolving and ongoing threat to internal security as demonstrated by a series of terrorist attacks" in official statements. That same year the EU introduced its Entry/Exit System (EES) (European Commission, 2017a, p. 1). Such innovation allowed for the digital documentation of "the entry and exit (and refusal of entry) of short-stay visa-holders and visa-exempt travellers who cross[ed] the EU's external borders" (Dumbrava, 2021, p. 6). This, as the policy explicates, currently enables information sharing between border security agencies, law enforcement authorities and entities such as Europol, which, as we shall see in Chapter 3, paves the way to questions pertaining to privacy and consent.

The following year saw the introduction of the European Travel Information Authorisation System (ETIAS), with an expected operationalisation in 2022. Such system officially fulfils the purpose of permitting the pre-registration of travellers from the Schengen Area. However, ETIAS will also allow for the assessment of "security or irregular migratory risks posed by […] persons before they arrive at the border", as explained in a recent publication by the official European Parliamentary Research Service of the Secretariat of the European Parliament (Dumbrava, 2021, p. 6). While recognising some historical and systemic drivers of border technology, this publication continues proposing problematic understandings of migration as a threat. Examples of the presence of this false correlation are in the use of

terms such as “identifying mobile and risky people”, “illegal immigration”, and “migration trends and security threats” during the whole course of the analysis (Dumbrava, 2021). While claiming to identify ways to improve border control and security in order to define best practices and address human rights implications, the paper does not take a firm condemnatory stance against potential discriminatory effects and only addresses them superficially. Such practice is common to European institutional papers. In a strategic document recently released by the European Commission, the report draws links between “The unprecedented 2015 refugee crisis exposed shortcomings in the Union’s management of the external borders and migration” and the “response to the persistent terrorist threat following a spate of attacks on European soil” (European Commission, 2021, p. 1). Arguably, this represents another instance in which the adoption of extraordinary measures such as increased technology at border crossings is justified. Given the history of discriminatory and unlawful practices, such as the documented involvement of Frontex border agents in illegal pushbacks, this narrative risks exacerbating discriminatory discourses and even condoning such illicit methods (Human Rights Watch, 2020).

Close scrutiny of official documentation has evidenced how concerns may be raised around the overall reliability of AI technologies in the EU border-management context. Three main critical areas were identified based on the criterion of accuracy: automated biometric identification, emotion detection and risk assessment algorithms.

The precision of automated biometric systems

Since its inception in 2003, the European dactyloscopy database (Eurodac) has relied on procedures of biometric verification and identification (European Parliament and Council of the European Union, 2013). In doing so, such database is compiled with the support of an automated fingerprint identification system (AFIS) which enables the comparison, identification and monitoring of asylum seekers data across EU Member States (European Parliament and Council of the European Union, 2013). Data collection quality is pivotal in preventing discriminatory outcomes as its deterioration or errors in the mining process may give rise to the issues discussed in Paragraph 2.1 of this research.
In 2015 the EU Joint Research Centre (JCR) conducted a report to ascertain the appropriateness of the AFIS in ensuring high data quality. Findings showed that error rates were as minor as 0.1% (Beslay and Galbally Herrero, 2015). Nevertheless, this was subsequently challenged by a subsequent report by the EU Agency for Fundamental Rights (FRA), which problematised

the ability of the database to deliver and maintain high degrees of accuracy, particularly when considering sensitive categories of migrants such as children over time (FRA, 2016). Thus, introducing the notion of potential human rights infringements. The potential lack of accuracy, in this case, is owed to attributes such as skin damages, environmental circumstances, problems with the sensor's surface, and even staff training and subject cooperation (Beslay and Galbally Herrero, 2015). Hence appropriate and consistent quality monitoring is an absolute precondition for the prevention of unfair determinations.

EU biometrics collection systems advocates have also demonstrated a keen interest in facial recognition technology due to its growing everyday application (EU Agency for Fundamental Right (FRA), 2019). Nonetheless, evidence over issues of data rights breaches and discriminatory outcomes have permeated literature in the field. As paragraph 2.3 will put forward, the risks connected to the inaccuracy of this technology have been manifold, as displayed by the failure of several trialling attempts of tools such as iBorderCtrl or the case of Belgium's Brussel Airport system, which was suspended in 2020 with a loss of 2.4 million € (Hope, 2020). The precision of these types of systems is extremely dependent on high levels of image quality, the lack of which may determine inclusive or biased outcomes (Dumbrava, 2021). Such accuracy is further unachieved when considering individuals belonging to especially vulnerable categories, namely children, women and people belonging to an ethnic minority. In the first case, technical reports have found that when children are exposed to such technologies, AI-tools have displayed " a negative bias, i.e. considerable degradation in performance" in these cases (Srinivas *et al.*, 2019, p. 1). Further, as analysed in Chapter 1, researchers have consistently shown how historical prejudices and biased training data can disproportionately affect some individuals over others, with significant impacts on women and minorities (Buolamwini and Gebru, 2018; Vincent, 2018; Molnar, 2020).

Emotion detection

Similarly to the matter of facial recognition, the issue of the use of AI for the purposes of emotion detection has generated a great degree of concern (Barrett *et al.*, 2019; EU Agency for Fundamental Right (FRA), 2019). Nevertheless, current proposals imagine the enlargement of the Eurodac database to include facial images and allow for a systematic analysis of human emotions (Dumbrava, 2021). Eu-LISA, The European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice, is tasked with delivering a study to assess the technical feasibility of the project's implementation, hence

reiterating political will to advance its use in the area. The Visa Information System (VIS) already stores applicant's digital photographs, and in the year 2019, this amounted to 68 million photos (European Commission, 2017b). The cultural relativity of emotions, expressions and their interpretations, as well as their susceptibility to exogenous factors such as personal circumstances or the particular contexts in which they are displayed, may vary even within a single individual, thus inevitably altering the objectivity of any assessment. In their influential 2019 study, Barret et al. identify three significant challenges in attempting to rationalise emotions and facial movements, i.e. "*limited reliability", "lack of specificity", and "limited generalizability"* (Barrett *et al.*, 2019, p. 3). The first notion refers to the difficulties arising from attempting to interpret an emotion belonging to the same category from a universal class of facial movements (Barrett *et al.*, 2019). The second entails the lack of commonality between the way a movement is performed and a particular emotion. Finally, the third relates to the differences in cultural and contextual approaches to facial expressions (Barret *et al.,* 2019). An example that allows for an understanding of the three relates to instances of nervous laughter, where an act typically associated with joy, for instance in Western countries, is performed in a stressful situation. Border crossings entail scrutiny and judgement, and often these are intermediate or final stages of the perilous journey which people on the move begin in their countries of origin. Hence, emotions may be conveyed in unexpected or simply culturally specific ways, which AI tools may not be equipped to interpret accordingly.

Risk Assessment algorithms

The data mining and gathering concerns raised in Chapter 1 also offer additional challenges depending on the entities which provide such data. The conflation between private companies, public organisations, entities such as Interpol, third-country governments or even open sources from social media analytics pose further challenges to the way data is utilised and assessed. For instance, an analysis of the origin of Interpol's Stolen and Lost Travel Documents (SLTD) has evidenced that the data used to compile such database has been adopted by states seeking to persecute and harass individuals who were labelled as political dissidents via a "red and blue" flagging system (Jones, 2020, p. 16). In such instances, discriminatory practices were rendered flagrant by the automatic refusal of travel authorisations when travel documents were reported to authorities as lost or stolen (Topcu, 2019). As a confirmation of this, a 2017 report from the Stockholm Center for Freedom reported the case of Turkish journalist Sevgi Akarçesme who was removed from a flight from Brussels to New York just prior to departure

after Turkish authorities issued a false notification in the SLTD database (Bozkurt, 2019). In response to a Parliamentary enquiry on the issue, the European Commission stated,

> The Commission is aware of the reports on Interpol's red notices having been used for political purposes in a number of instances. This is of great concern, as it can have a significant impact on the persons concerned. […] The legal basis of the European Travel Information and Authorisation System (ETIAS) provides necessary safeguards as regards queries of Interpol's SLTD (Stolen and Lost Travel Documents) and TDAWN (Travel Document Associated with Notices) databases (European Parliament, 2019, p. 1).

The extent to which such legal safeguards are effectively in place shall be the topic of the Chapter of this thesis. Nevertheless, these unclear and potentially illegitimate data sources further buttress the discriminatory risks involved in AI technology's applications in the field amidst EU justifications for their proliferation based on matters of national security.

2.3 Practical Applications

The last part of this research will take a close look at the iBorderCtrl tool. Such system, however, is only one among the myriad of technologies that are already taking being implemented at the border. It is useful here to provide a rapid overview of the tools which have been discussed so far.

Even prior to their arrival on European soil, migrants are subjected to military-style drone surveillance technology above the Mediterranean, under Frontex border patrolling (FRONTEX, 2020). Such technology has been trialled and utilised by Frontex for the interception of vessels transporting people migration from their countries of origin towards European coasts (FRONTEX, 2020). An ulterior example of the use of unpiloted drones is that of ROBORDER. This is one of the 38 projects which were funded under Horizon 2020, specifically under the "Secure societies - Protecting freedom and security of Europe and its citizens" scheme (European Commission, 2013). Such technology includes autonomous mobile robots, which are able to be operated individually or in swarms. According to the proposal, its objective is to develop a "fully-functional autonomous border surveillance system with unmanned mobile robots including aerial, water surface, underwater and ground vehicles"

(ROBORDER, 2020, no page). As we shall see in the following Chapter, this arguably may fall into one of the categories which contradict the EU's own legislative framework.

The deployment of military technology is once again a representation of the *crimmigration* phenomenon or, in order words, the criminalisation of migration as a result of its securitisation and using risk-based taxonomies to define and label cases (Bosworth, Parmar and Vázquez, 2018). Further, drones may also be interpreted as mechanisms of externalisation of state borders and reinforcements of narratives of sovereignty aimed at reproducing the narratives of state identity discussed in Chapter 1. Accordingly, as the border is expanded both visually and virtually, it exerts its exclusionary power (Butler, 2006). Ultimately this also translates into practices of deterrence as suggested by the policy solutions of southern European countries such as Italy, Greece and Malta, in which violent pushbacks are practised despite infringement of international non-refoulement provisions (EFAD, no date). As the literature review section of this dissertation discussed, other utilisations of AI tools can be found in countries such as Sweden and Germany that are employing Big Data in efforts to predict and track migration movements (Beduschi, 2020). Countries such as Austria, Denmark, Belgium and Norway have also adopted legislation allowing for the collection of data from asylum applicants via their mobile devices. Such data is then utilised in the course of legal procedures as described in a recent submission to the UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance (Privacy International, 2020).

The next Chapter will provide an analysis of the legal safeguards already in place in the EU. It will begin by describing provisions of non-discrimination under the EU regime by relying primarily on the ECHR. Subsequently, it will analyse the matter of privacy governed via the GDPR regime. Finally, it will focus on the case of iBorderCtrl. While the project has currently been suspended, it acquires empirical evidence as the system was sustained by public EU funding, trialled in three countries, and its technology was repurposed in other security contexts.

Chapter 3.
AI and border management: the EU's legislative approach

The previous Chapter examined how discrimination may arise in AI systems from a technical, ethical, and practical viewpoint and how the EU system may be complicit in justifying its use. The analysis accordingly offered multiple illustrations of how this is enmeshed in the EU information systems both in its mechanisms and in practices.

While some elements of legal interest have already been explored throughout the analysis, this Chapter will address these more closely. The discussion will engage with the second research question, namely, *how does the EU propose to address AI's possible discriminatory biases from a legal standpoint, and what is it missing in relation to the adoption of AI tools in the migration and border management context?*

Technosolutionist responses to ethical issues hint at the lack of political will to regulate the AI market as this would translate in increased deployment of human and financial resources. Nonetheless, this Chapter will explore whether the EU's present approach is mitigated (or not) by its legislative apparatus and the overarching international instruments within which it falls. As the analysis thus far has suggested, issues of AI may cause infringements of a wide array of human rights, e.g., freedom of expression, freedom of assembly, the protection of personal information, the rights of the child and more (FRA, 2019). However, as the focus of this is analysis is on AI tools' discriminatory effects, the Chapter will highlight only those provisions which are relevant to this end.

3.1 Freedom from discrimination

Several European and international instruments are focused on the prohibition of discrimination. As we are discussing the rights of the migrant population, the Refugee Convention also deserves mention. Non-discrimination here is enshrined in Article 3, which states, "the Contracting States shall apply the provisions of this Convention to refugees without discrimination as to race, religion or country of origin." (UN General Assembly, 1951, p.14). However, as discussed in section 1.1 of this thesis, this does not apply to the individuals who do not fall under Article 1(A)2. Therefore, the analysis will rely on The European Convention on Human Rights (ECHR), which is one of the most cited and authoritative in this realm.

Within the ECHR, discrimination is enshrined in Article 14. Such Article is complemented by Article 1 of the Protocol No. 12 to the Convention, which prohibits discrimination more generally, as elucidated by the European Court of Human Rights (ECtHR) in its case-law (E.B. v. France [GC], 2008, § 47; Marckx v. Belgium, 1979, § 32).

Art. 14 reads,

> The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status (ECHR, Art. 14).

Although this Article does not provide a straightforward definition, according to ECtHR established case law, discrimination may be experienced in two different forms, *direct* and *indirect* (Biao v. Denmark [GC], 2016, § 89; Carson and Others v. the United Kingdom [GC], 2010, § 61; D.H. and Others v. the Czech Republic [GC], 2007, § 175). Following the Court's jurisprudence, any conduct may amount to direct discrimination when an individual or a group is discriminated against on the basis of a particular protected trait (gender, ethnicity, religion, *et similia*). As confirmed in the case of Biao v. Denmark, "there must be a difference in the treatment of persons in analogous, or relevantly similar situation", which is based "on an identifiable characteristic" (ECthr, 2016, para. 89). Within EU law, non-discrimination provision is yet to undergo a harmonisation process. However, a similar definition of direct discrimination can be found in the Employment Equality Directive (2000/78/EC), the Gender Goods and Services Directive (2004/113/EC) and the Recast Gender Equality Directive (2006/54/EC).

Indirect discrimination, on the other hand, occurs when people experiencing different situations are offered the same treatment. In this case, what takes relevance is not the treatment but rather the outcome and effect of a practice that appears to be neutral on individuals with different characteristics. Following the ECtHR,

> [A] difference in treatment may take the form of disproportionately prejudicial effects of a general policy or measure which, though couched in neutral terms,

> discriminates against a group. Such a situation may amount to "indirect discrimination", which does not necessarily require a discriminatory intent presumably (Council of Europe and European Court of Human Rights, 2020).

Matters considered to the crucial to human dignity, like race or ethnic-based discrimination, are more challenging to justify when compared to those carrying general social repercussions, in keeping with the Court's pronouncements and *obiter dicta* (IOM and Norwegian Ministry of Foreign Affairs, 2013). Therefore, States enjoy what the ECtHR calls a "margin of appreciation", namely, a degree of discretion in cases where differential treatment may be justifiable (Handyside v. UK, ECHR, 1976). Nevertheless, such discretion is subject to two criteria for which the interference with fundamental rights pursues a legitimate aim, namely, *proportionality* and *necessity in a democratic society*. The principle of proportionality requires a reasonable relationship between the achievement of a specific aim and the means through which it is achieved, in this case, the infringement of a fundamental right. In other words, it must be "proportionate to the legitimate aim pursued" (Handyside v. UK, ECHR, 1976). The proportionality test goes hand in hand with the principle of necessity, whereby the necessity of a measure must be justified against the greater benefit of society, in this case, national security (Clayton and Tomlinson, 2009).

In the context of border management, the EU envisages the use of non-discriminatory law to fight discriminatory AI decisions. For example, in the cases in which AI decisions affect disproportionately people from a particular ethnic background as this may breach the principle of indirect discrimination (Borgesius, 2018). However, an analysis of these provisions renders evident how weaknesses in the law may still amount to improper safeguarding of the individual subjected to AI judgement. Firstly, in the context of migration management, AI may still lead to indirect discrimination. As stated, the law poses emphasis on the *outcomes* rather than the intention of the potential discriminator, hence erasing the technical issues raised in Chapter 2. Whether the possibility for a discriminatory effect to take place is inserted at the time of programming or due to bias in the training data, the law does not affect the position of the source of the discrimination (Biao v. Denmark [GC], 2016). In this case, it can be argued, the EU exhibits an instrumentalist interpretation of border technology, whereby instruments such as e-gates are neutral, and the potential intention to discriminate is irrelevant to the produced outcome.

Secondly, the prohibition of indirect discrimination lacks clarity and specification, which renders real-life a challenging endeavour, especially in a setting such as that of an EU border. Individuals would be tasked with the burden of proof in relation to the discriminatory effects of a presumably neutral rule. However, as previously discussed, the neutrality of AI machinery is debatable, yet proving its *prima facie* disproportionate discriminatory effects may be lengthy and inconclusive. For instance, in the case of AI applications for asylum determinations, if a migrant is refused on the basis of facial recognition software deliberation, it would be extremely challenging to ascertain the grounds for refusal and whether this was a result of AI bias. Thirdly, in light of principles of proportionality and necessity, it follows that in an instance in which discrimination is performed in ways that allegedly are not disproportionate and meet "objective and reasonable" aims, the discriminatory practice may be justified (Biao v. Denmark [GC], 2016). This contradicts international principles pertaining to the certainty of the law. Whether or not the use of AI tools, in this case, may lead to disproportionate or unnecessary violations is unclear and may be highly circumstantial. Moreover, in light of securitised understandings of the migration phenomenon and the criminalisation of people on the move, border guards may result in disproportionate breaches of rights to privacy under the guise of national security despite a lack of empirical foundation to buttress this (Helbling and Meierrieks, 2020).

Finally, non-discriminatory law, as evidenced, rests on the concept of protected categories such as race, gender, or religion (ECHR, Art. 14). However, many of the discussed AI-based classifications, which arguably amount to discrimination, unfairness or bias, are excluded from these statutes, hence evidencing gaps in legislation that further reinforce the problematic pacing of the law in this field.

4.1 AI and Data Protection Law

Growing concerns pertaining to bias in algorithmic decision-making have resulted in the adoption of instruments such as The General Data Protection Regulation (GDPR) and the Council of Europe's Data Protection Convention 108. Such legal documents provide data protection safeguards when considering the acquisition, processing and storing of personal and sensitive information, including biometric data (EU, 2016/679). Both the European Commission and the Parliament have voiced the matter of AI compliance to GDPR provisions and have promised to pose a heightened emphasis on

possible legislative adjustments based on novel developments in the field. For instance, in its Communication *On the Road to Automated Mobility: An EU Strategy for Mobility of the Future,* the Commission declared that in spite of global appraisal for the high quality of protection afforded by GDPR legislation, it is crucial to keep such provisions up-to-date in order to keep up with the fervent dynamism of technical advancements (European Commission, 2017a).

The legal landscape

Articles 4 and 5 GDPR lay down the definitions of data and the data processing principles which cover personal data processing performed via automated methods. According to Article 4(1), data is "any information relating to an identified or identifiable natural person ('data subject')". This entails both direct and indirect forms of data, the first pertaining to readily identifiable traits and the second to characteristics that may be attained via external connected factors (GDPR, Art.4). Article 4(2) expands on the specific forms of data manipulation such as "collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction." (GDPR, Art 4(2). This enables us to ascertain that AI falls within such definition as, as discussed in Section 2.2 of this research, AI systems attain data via direct compilation at the programming stage and process it during the training stage and upon its operationalisation for decision-making purposes (Harryson *et al.*, 2018).

Article 5 elaborates on the principles which can be applied to the processing of such data. These can be summarised in four principles: data purpose limitation in Article 5 (1) b; data minimisation 5 (1) c; the principle of accuracy in Article 5 (1) d; and finally, the principle of storage limitation in Article 5 (1)e (GDPR).

The first principle requires data to be stored only for clear and specific purposes in order to guarantee informed consent for the individuals whose data is collected. Moreover, the principle entails a compatibility test. This entails that with the exclusion of instances in which consent is explicitly accounted for under EU or national laws, every novel processing should be re-evaluated in terms of its compatibility with the purposes for which the data was collected in the first instance. In line with ECHR provisions of necessity and relevance, the principle of data minimisation entails that personal information should only be stored in cases in which this is relevant and necessary to the purposes of their processing. It follows that this should be

performed with the maximum degree of accuracy, with no differentiation between data attained directly from subjects and that indirectly attained via observation, or inferring processes, as postulated by Article 5 (1) d. Lastly, storage limitation requires data to be stored only for limited amounts of time. This is tied with the principle of data minimisation and allows for subjects to exercise their "right to erasure", or in other words, the right for one's information to be forgotten by the system (GDPR, Art 5). Additionally, Article 4(7) GDPR and Article 2(d) COE Data Protection Convention 2018 bestow safeguarding obligations of these rights upon the parties involved in the processing, namely the data controllers. In light of this, under Article 25(1) GDPR, controllers are required to perform a data protection impact assessment (DPIA) which must be conducted when a behaviour is deemed to be highly risky for the rights of individuals. This is described as,

> a tool used for the analysis of possible consequences of an initiative on a relevant societal concern or concerns, if this initiative can present dangers to these concerns, with a view to supporting informed decision-making whether to deploy this initiative and under what conditions, ultimately constituting a means to protect these concerns (Kloza *et al.*, 2020, p. 1).

Additional articles are focused explicitly on automated decision-making. The so-called *Kafka provision* under article 22 refers to the prevention of fully automated decisions with legal or similar significant consequences. Such provisions, like the above, represent the EU's attempt to ensure the avoidance of risks of unfair and illegal discrimination amidst other human rights infringements (Hacker, 2018). Article 22 is of particular interest in this case as it fundamentally requires the presence of a human operator. Arguably, Frontex's described deployment of automated drone systems above the Mediterranean Sea is in direct contradiction with the right not to be subjected to profiling and similar fully automated practices (GDPR, Art. 22).

The legal shortcomings of EU provisions

The first caveat to this provision lies in the exceptions to Article 22, which are enucleated in Article 29 of the Working Party 2018 (WP251). Article 29 states that such prohibition is unapplicable in cases in which the automated decision is performed under the individual's explicit consent; is the consequence of a contract between the subject and the data controller;

of in cases in which it is authorised by the law (Article 29 Working Party 2018 (WP251)). In case one of such conditions is fulfilled, under Article 22(3) GDPR,

> the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision (Article 22(3) GDPR).

Data controllers are thus required to provide information pertaining to all stages of computerised data processing when these involve the use of personal information and must provide significant explanations. This arguably renders the provision ineffective as the circumstances in which discrimination may have arisen are challenging to ascertain for the very nature of the algorithmic tools involved, therefore excluding numerous instances from the scope of the GDPR. For instance, if a refugee is denied asylum on the grounds of an AI-induced recommendation, insofar as the officer is not "rubber-stamping" or, in other words, falling victim of *automation bias*, Article 22 is inapplicable. The manner in which entities are expected to ensure such lack of superficiality is unclear.

A second shortcoming is the lack of enforcement of these provisions. As reported by Borgesius in an official EU study, Data Protection Authorities are underfunded and exert limited power and influence. Therefore, their capability to impose effective sanctions is constrained (Borgesius, 2018). Despite an increase in adherence in the aftermath of the introduction of the GDPR regulation, organisations still appear not to take data protection law as a matter of great importance (Zuiderveen Borgesius, 2015).

Thirdly, as noted in the previous paragraph, while AI processes fall within the realm of GDPR and related legislative tools, this does not concern the different types of applications in AI data processing. For instance, in the case of forecasting models, the systems may utilise data relating to proxy characteristics, hence not personal data. An example may be the prediction of migration patterns performed on the basis of eternal drivers.

A fourth issue pertains to the clarity of the law. While open-ended and abstract norms may be effective in ensuring that provisions are not excessively narrow, thus allowing for the law to be utilised as a *living instrument* rather than a monolithic tool, this may render their effective application challenging. As discussed, discrimination may arise in direct and indirect forms. Thus, it may be challenging to prove that this has effectively taken place. Moreover, similar issues arise in attempting to provide a rationale behind the conclusion that has been

drawn via algorithmic means. On the one hand, this is owed to the previously mentioned matter of algorithmic obscurity and, in some cases, to the fact that the burden of proof lies on those being subjected to such decisions.

4.2 The case of iBorderCtrl

It is helpful to briefly illustrate how a number of the issues presented throughout this thesis may play out in real life in order to address the legitimacy of these concerns. The Intelligent Portable Control System (iBorderCtrl) is possibly one of the most controversial projects funded within the EU's Horizon 2020 research programme in recent years (Sánchez-Monedero and Dencik, 2020). In fact, despite hitherto no emotion detection systems are not being actively deployed at EU borders, projects are still being trialled with the goal of detecting deception at the border (examples include TRESSPASS; FOLDOUT and FASTPASS) (iBorderCtrl, 2016). This example highlights many of the theoretical, technical, ethical and legal implications which were analysed over the course of this dissertation.

The project

iBorderCtrl was developed and trialled between 2013 and 2019 with an allocation of €4.5 million, from the €34.9 billion budget earmarked for border and immigration control within the European Research Council Project Horizon 2020 (European Commission, 2020a). In the words of the researchers at Manchester Metropolitan University in the UK who were behind the project, its aim was developing "well established as well as novel technologies together to collect data that will move beyond biometrics and onto biomarkers of deceit" piloted at European borders" (iBorderCtrl, 2016, no page). The project fits within the framework of "emphatic media", specifically "emotional AI" which consists in, "reading words and images, seeing and sensing facial expressions, gaze direction, gestures and voice" (McStay, 2018, p. 2). Moreover, this also includes machines which have the ability to detect other characteristics such as our "heart rate, body temperature, respiration and the electrical properties of our skin, among other bodily behaviours" (McStay, 2018, p. 2).

The methodology

This specific project was centred on the system's ability to develop a risk assessment on the basis of the detection of dishonesty at the moment of border crossing. As the authors explain, the process involved a two-steps approach, a first phase in which migrants provide their personal information and a second phase involving biometric certification, matching and evaluation of the documents' legitimacy (iBorderCrtl). The aim of the project was determining "biomarkers of deceit", namely non-verbal facial micro-expressions that are associated with lying, such as left eye blink, increase in facial redness, head movement directions. In order to accomplish such aims, the process also entailed "interaction with external legacy and social systems, an Automatic Deception Detection System (ADDS), a Risk-Based Assessment Tool (RBAT) and a post hoc analytics tool" (iBorderCtrl, 2016: no page). This setting separated it from previous *smart border systems,* which did not include an ADDS system. iBorderCtrl was officially tested as Hungarian. Latvian ad Greek borderlands after a number of internal lab test runs. Unlike previous projects, rather than acquiring and codifying data from existing case-law or attempting to minimise the artificiality of the process by paying participants to commit crimes, the authors created scenarios for actors to reproduce (Derrick *et al.*, 2010).

In order to capture "biomarkers of deceit", first, engineers employed a one-second high-resolution video in order to capture physiognomic expressions, which were then used to build an initial dataset. Such videos were taken on a sample of 32 participants who were assigned a "genuine" or "deceiving" role to play in the course of the filming. The group involved 10 participants who were classed according to their ethnicity as Asian or Arabic, 22 as White European, twenty-two male and ten female. In case of a positive result, the individual was then subject to further scrutiny by a human operator (iBorderCrtl, 2016). The reported accuracy level was 73.66% when identifying dishonesty and 75.55% when detecting the truth (iBorderCtrl, 2016).

A growing trend

iBorderCtrl is not unique in its genre. Rather, it signals a growing trend in the datafication of human emotions, or in other words, the classification and measurement of expressions in a way that until now has been performed with our personal and sensitive information. Further examples of such expansion in use can be identified in the deployment of systems of biometric mass surveillance at Italian borders. The case was recently disclosed in research conducted by *IrpiMedia* and subsequently reported by *the European Digital Rights (EDRi)* association. The

study was centred on the Automatic Image Recognition System (SARI), initially acquired by the Italian police in 2017 and currently repurposed to track the arrival of migrants on the Italian Coast (EDRi, 2021). Findings revealed the use of an inherently unbalanced and discriminatory training programme as "8 out of 10 people in SARI's database" have an immigrant background, as well as an environment of NGOs, researchers, private and public organisations devoted to gathering, mining and analysing migrants' data (EDRi, 2021: no page). With an allocated budget of 246.000 and a partnership with privately-owned company Neurotechnology, the project appears to be still in place despite growing concerns by activists and civil society due to its lack of transparency as well as the meaningful experience of iBorderCtrl (EDRi, 2021).

Controversies

The project attracted stern criticism, with activists and academics raising concerns over matters of reliability, accuracy and precision of the science behind its application. Notably, in addressing ethical implications and discriminatory effects of its use, activists from the *iBorderCtrl, no!* collective dismissed the project as pseudoscience (*iBorderCtrl? No!*, no date). Moreover, as researchers have pointed out, the model is premised on two false assumptions, firstly, the belief that micro-expressions – which are extremely swift and last between 1/25 s and 1/5 - can be scientifically measured (Zhang *et al.*, 2014). Secondly, this type of algorithmic-based system is exceptionally problematic as it requires mechanisms of detection that do not take personality traits, cultural relativism, neurodiversity or medical conditions into consideration (Sánchez-Monedero and Dencik, 2020).

Discrimination and larger societal implications

Despite its described shortcomings, projects such as iBorderCtrl are an illustration of the normative gaps in the determination of appropriate anti-discriminatory safeguards. As discussed, the project employed a small set of participants with a gravely imbalanced ethnical and gender-based sample, hence signifying the potential for underlying training data biases. As mentioned in Chapters 1 and 2 of this thesis, research has shown the tendency of AI to disproportionately affect women and people from ethnic minorities by means of overrepresentation or erasure. From a legal standpoint, although at first glance the case appears to contradict the principles encapsulated in Article 22 (1) GDPR, relating to the unlawfulness

of fully automated decision-making, the project fits into the Article 29 Working Party 2018 (WP251) caveat. Under the described process, passengers who are identified as deceitful are referred to human operators. However, as discussed, ascertaining whether decisions are made as a consequence of an automation bias and by blindly following the algorithmic input is impossible.

The rise, proliferation and political will to continue sponsoring these systems to fulfil the securitising logic discussed in the theoretical section of this paper. Facial recognition and similar algorithm-based tools have consistently demonstrated a proneness to bias and the potential for discriminatory practices to go unchallenged. Arguably, following Butler's concept of state performativity, the operationalisation of these tools is a manifestation of power via the means of border politics in a European landscape that is at once externalising and internalising its borders.

Conclusion

1. Research Implications

The analysis evidenced that the rapidity of technology's adaptation to new sectors of our lives is unravelling before our eyes, States, and private actors, are rushing to harness the force of AI to extract wealth in a 21st-century gold rush (Shull, 2019). Contemporarily, the discriminatory impacts of these technologies remain unaccounted. In this case, wealth, however, is equated to power, and the gold is replaced by our data and personal information. When such race is placed in the context of migration management, technology becomes the reflection of political power relations and power-projection at the expense of fundamental human rights such as freedom from discrimination. AI, ML, and Big Data all cease to be "*neutral*" and "*impartial*" by virtue of their scientific composition, but rather they become powerful tools of exclusion and inclusion. They exacerbate societal distinctions over matters of belonging while bestowing labels of *citizen* or *non-citizen* and by performing seemingly unbiased evaluations over issues of morality, worthiness, or credibility. However, AI is opaque and can be highly discretionary. Such systems are, in fact, developed in conjunction with private actors and both these and States rely on caveats for which they are not subjected to appropriate levels of scrutiny and transparency. This, as the analysis highlighted, is owed to a lack of *ad hoc* provisions governing the development of algorithms (Molnar, 2019). Hence widespread criticism by mathematicians and computer scientists for being "black boxes" appears to be ultimately justified (Pasquale, 2015; O'Neil, 2016). As Chapter 2 has explained, it is, in fact, difficult to fully ascertain the manner in which algorithms make decisions about people, thus leaving individuals open and vulnerable to racialised, gendered and other general forms of discrimination.

Furthermore, in line with the notions explained in the developed theoretical framework, the EU continues to project its image of a leading human-rights actor on a global scale with a firm anti-discrimination stance also illustrated via policy, legislation and public declarations, discussed in Chapter 2 of this research. However, as Chapter 3 contended, its border management scheme does not sufficiently address potential human rights infringements in relation to discrimination while merely tokenising migrants - when not drawing equations between migration and matters of national security. The lack of foundation for claims relating to the "crimmigration" paradigm has been analysed elsewhere and is only peripheral to the present discussion (Bosworth, Parmar and Vázquez, 2018; Morrow and Shjarback, 2019;

Helbling and Meierrieks, 2020). However, the notions of securitisation, security and technology (STS), and state identity discussed throughout the introduction serve as a lens through which further understand the EU's legislative approach and its effect on migration management and border control.

2. Final remarks

This research has illustrated and analysed some of the ways in which AI has shaped EU border control, and it has problematised matters concerning algorithmic bias and discrimination. The presented argument was twofold: firstly, in Chapter 2 it aimed at highlighting that while these technologies appear to be neutral and objective, issues of bias may be determined by their technological apparatus. Such issues are then strengthened and reproduced by a variety of EU measures and technologies such as the EES, the VIS and ETIAS. Secondly, in Chapter 3, it analysed the contradictions between the norms which are already in place and the lack of an EU level enforcement and procedural mechanisms. As emphasised, while technology is moving at a rapid pace, the EU's normative framework still lags behind. In doing so, this dissertation posed attention to the discriminatory effects of these shortcomings on migrants. The analysis drew a connection between the algorithmic obscurities that render discrimination challenging to recognise formally in a scientific setting and the legal black holes deriving from the EU's bureaucratic inconsistencies.

Further, the research hopes to have contributed to the gap in the scholarship concerning the disproportionate effects of technological tools and their testing procedures on migrants in the EU context. It aimed at adding to these emerging discussions and expanding on the existing literature by introducing this matter in the proposed IR-STS theoretical framework.

Through securitisation theory, it explained the underlying implications of the adoption of such tools. It clarified that when States believe their sovereign power is under threat, control is justified under the guise of national security. As discussed, since the onset of the 2015 migration movements, political figures and European media powerhouses have contributed to consolidating the connection between migration and criminality, thus providing an *alibi* for the proliferation of smart borders and an increase in border security and data collection under alleged national security concerns (Borgesius, 2018). Sociotechnical insights deepen this understanding by allowing to explain that States' necessities to provide a material reproduction of such logic, *i.e.,* the introduction of extraordinary measures such as border control (Pager and

Shepherd, 2008). As explained, the effect is the perpetuation of exclusionary logics at the expense of some individuals over others through the use of seemingly unbiased machinery.

State performance becomes particularly persuasive when the law is suspended, unclear or when States display idiosyncratic conduct between their legal apparatus and their experimentations, such as in the case of the trialling of systems such as iBorderCtrl. Hence, the normalisation of increased data mining, AI-driven decision-making systems and surveillance in our everyday lives offer EU States the ability to pursue such practices in seemingly uncontested manners, which provide some degree of explanation as per their proliferation in light of the empirical evidence pointing at their problematic effects. Moving forward, it will be essential for States and organisations such as the EU to develop, implement and ensure appropriate safeguarding mechanisms and protective systems both from a legal and a technical standpoint in order to mitigate the challenges brought forward by the adoption of technology in this field.

Future Outlooks

Technology is not neutral (Singler, 2019). In the framework of border control and migration management it has become a reflection of state norms and apparatus and a demonstration of power and sovereignty. This dissertation chose to focus on the way in which such power is exerted on the migrant populations under the guise of efficiency and neutrality which has led to a new form of technologically justified securitisation, removing space for alternative outlook to prevail. Nonetheless, technology, may also be a source for good, and in light of its already ubiquitous nature, understanding its full potential towards ameliorating every condition of the human experience may be beneficial to larger segments of the population. Following this logic, understanding ways to allow for greater transparency and accountability will be an important exploration field for future research. Moreover, given the current lack of supranational forms of effective enforcement mechanisms, further studies may address alternative feasible solutions on a narrower, regional scale. Lastly, as the picture is rendered increasingly clearer with regard to the implications of an effective implementation of tools such as the discussed iBorderCtrl, further research may focus on the tangible extent of their discriminatory effects on the life of migrants being scrutinised at border crossings.

List of References

Access Now (2021) 'Iris scanning of refugees is disproportionate and dangerous — What's happening behind IrisGuard's closed doors?', *Access Now*, 12 April. Available at: https://www.accessnow.org/irisguard-refugees-jordan/ (Accessed: 25 July 2021).

Ackleson, J. (2003) 'Securing Through Technology? "Smart Borders" after 9/11', *Knowledge, Technology, and Policy*, pp. 56–74.

Alsallal, M. *et al.* (2016) 'An Integrated Machine Learning Approach for Extrinsic Plagiarism Detection', *2016 9th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 203–208.

Amelung, N., Granja, R. and Machado, H. (2021) 'A Brief History of the Evolution of Biometrics and Biometric Database Systems Crossing Borders in EU Law Enforcement', in Amelung, N., Granja, R., and Machado, H. (eds) *Modes of Bio-Bordering: The Hidden (Dis)integration of Europe*. Singapore: Springer, pp. 15–34. doi: 10.1007/978-981-15-8183-0_2.

Angwin, J. *et al.* (2016) *Machine Bias*, *ProPublica*. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=KCQtlys2Jot94pBZSAGZgQHpS8pwmTyy (Accessed: 16 January 2021).

Ansemes de Vries, L. and Guild, E. (2019) 'Seeking refuge in Europe: spaces of transit and the violence of migration management', *Journal of Ethnic and Migration Studies*, 45(12), pp. 2156–2166. doi: 10.1080/1369183X.2018.1468308.

Balzacq, T. (2005) 'The Three Faces of Securitization: Political Agency, Audience and Context', *European Journal of International Relations*, 11(2), pp. 171–201. doi: 10.1177/1354066105052960.

Bansak, K. *et al.* (2018) 'Improving refugee integration through data-driven algorithmic assignment', *Science (New York, N.Y.)*, 359(6373), pp. 325–329. doi: 10.1126/science.aao4408.

Barocas, S. and Selbst, A. D. (2016) *Big Data's Disparate Impact*. SSRN Scholarly Paper ID 2477899. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.2477899.

Barrett, L. F. *et al.* (2019) 'Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements', *Psychological Science in the Public Interest*, 20(1), pp. 1–68. doi: 10.1177/1529100619832930.

Beduschi, A. (2020) 'International migration management in the age of artificial intelligence', *Migration Studies*, (mnaa003). doi: 10.1093/migration/mnaa003.

Bentley, P. *et al.* (2018) *Should we fear artificial intelligence?: in depth analysis.* LU: European Parliament Publication Office. Available at: https://data.europa.eu/doi/10.2861/412165 (Accessed: 15 July 2021).

Beslay, L. and Galbally Herrero, J. (2015) *Fingerprint identification technology for its implementation in the Schengen Information System II (SIS-II)*. EUR 27473. Luxembourg: Publications Office of the European Union.

Black, E. (2012) *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation-Expanded Edition*. Expanded Edition. Washington, D.C: Dialog Press.

Bleiker, C. (2017) *EU to implement border fingerprint checks similar to United States*, *DW.COM*. Available at: https://www.dw.com/en/eu-to-implement-border-fingerprint-checks-similar-to-united-states/a-41111621 (Accessed: 6 July 2021).

Borgesius, F. (2018) *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*. Counci of Europe, pp. 1–94.

Bosworth, M., Parmar, A. and Vázquez, Y. (eds) (2018) *Race, Criminal Justice, and Migration Control: Enforcing the Boundaries of Belonging*. Oxford: Oxford University Press. doi: 10.1093/oso/9780198814887.001.0001.

Bowen, G. A. (2009) 'Document Analysis as a Qualitative Research Method', *Qualitative Research Journal*, 9(2), pp. 27–40. doi: 10.3316/QRJ0902027.

Bozkurt, A. (2019) *Secret documents reveal abuse of Interpol mechanisms by Turkish government - Nordic Monitor*. Available at: https://nordicmonitor.com/2019/02/secret-documents-reveal-abuse-of-interpol-mechanisms-by-turkey-government/ (Accessed: 28 July 2021).

Branch, J. (2014) *The Cartographic State: Maps, Territory, and the Origins of Sovereignty*. Cambridge University Press.

Buolamwini, J. (2020) *CODED BIAS*, *CODED BIAS*. Available at: https://www.codedbias.com (Accessed: 22 July 2021).

Buolamwini, J. and Gebru, T. (2018) 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', p. 15.

Burt, | Chris (2020) *IrisGuard refugee aid delivery with iris biometrics uninterrupted by physical distancing requirements | Biometric Update*. Available at: https://www.biometricupdate.com/202004/irisguard-refugee-aid-delivery-with-iris-biometrics-uninterrupted-by-physical-distancing-requirements (Accessed: 25 July 2021).

Butler, J. (2006) *Precarious Life: The Powers of Mourning and Violence*. Verso Books, p. 168.

Butler, J. (2009) 'Performativity, Precarity and Sexual Politics 1'.

Buzan, B., Waever, O. and de Wilde, J. (1998) *Security Analysis: Conceptual Apparatus," in idem, Security: A New Framework for Analysis*. Lynne Rienner. London.

Carammia, M. and Dumont, J. C. (2018) *Can We Anticipate Future Migration Flows?* Paris: OECD.

Cardoso, T. (2018) 'Ottawa's use of AI for immigration a "high-risk laboratory": report'. Available at: https://www.theglobeandmail.com/politics/article-ottawas-use-of-ai-for-immigration-a-high-risk-laboratory-report/ (Accessed: 13 July 2021).

Castles, S. (2004) 'The Factors That Make and Unmake Migration Policies', *International Migration Review*, 38(3), pp. 852–854.

Caton, S. (2020) *Fairness in Machine Learning: A Survey*, *DeepAI*. Available at: https://deepai.org/publication/fairness-in-machine-learning-a-survey (Accessed: 9 July 2021).

Chase, J. and Dick, W. (2017) 'Questions surround German government's refugee phone surveillance law', *Deutsche Welle*. Available at: https://www.dw.com/en/questions-surround-german-governments-refugee-phone-surveillance-law/a-38236850 (Accessed: 21 July 2021).

Cimpanu, C. (2019) *Chinese company leaves Muslim-tracking facial recognition database exposed online*, *Zero Day*. Available at: https://www.zdnet.com/google-amp/article/chinese-company-leaves-muslim-tracking-facial-recognition-database-exposed-online/?__twitter_impression=true (Accessed: 25 July 2021).

Clayton, R. and Tomlinson, H. (2009) *The Law of Human Rights*. 2nd edn. Oxford ; New York: Oxford University Press.

Commonwealth of Australia, C. (2019) *Advisory Report on the Counter-Terrorism Legislation Amendment (2019 Measures No. 1) Bill 2019*. Available at: https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Intelligence_and_Security/CTLA2019MeasuresNo1/Report (Accessed: 29 July 2021).

Council of Europe and European Court of Human Rights (2020) *Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of discrimination*.

Crawford, K. (2016) 'Opinion | Artificial Intelligence's White Guy Problem (Published 2016)', *The New York Times*, 25 June. Available at: https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html (Accessed: 11 January 2021).

Crawford, K. (2017) 'The Trouble with Bias - NIPS 2017 Keynote - Kate Crawford #NIPS2017', in. Available at: https://www.youtube.com/watch?v=fMym_BKWQzk&t=2574s (Accessed: 16 January 2021).

Cruz, E., D'Alessio, S. J. and Stolzenberg, L. (2020) 'Decisions made in terror: Testing the relationship between terrorism and immigration', *Migration Studies*, 8(4), pp. 573–588. doi: 10.1093/migration/mnz014.

Dastin, J. (2018) 'Amazon scraps secret AI recruiting tool that showed bias against women', *Reuters*, 10 October. Available at: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G (Accessed: 5 July 2021).

Derrick, D. C. *et al.* (2010) 'Border Security Credibility Assessments via Heterogeneous Sensor Fusion', *IEEE Intelligent Systems*, 25(3), pp. 41–49. doi: 10.1109/MIS.2010.79.

Dourish, P. (2016) 'Algorithms and their others: Algorithmic culture in context', *Big Data & Society*, 3(2), p. 2053951716665128. doi: 10.1177/2053951716665128.

Dubber, M. D., Pasquale, F. and Das, S. (2020) *The Oxford Handbook of Ethics of AI*. New York, NY: OXFORD UNIV PR.

Dumbrava, C. (2021) *Artificial Intelligence and EU Borders*. Brussels: European Parliamentary Research Service of the Secretariat of the European Parliament. Available at: http://www.eprs.ep.parl.union.eu (intranet).

EDRi (2021) *Chilling use of face recognition at Italian borders shows why we must ban biometric mass surveillance*, *European Digital Rights (EDRi)*. Available at: https://edri.org/our-work/face-recognition-italian-borders-ban-biometric-mass-surveillance/ (Accessed: 29 July 2021).

EFAD (no date) 'Italy Country Focus', *The European Forum on Armed Drones (EFAD)*. Available at: https://www.efadrones.org/countries/italy/ (Accessed: 27 July 2021).

EU Agency for Fundamental Right (FRA) (2019) *Facial recognition technology: fundamental rights considerations in the context of law enforcement,*. Available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-facial-recognition-technology-focus-paper-1_en.pdf.

European Commission (2001) *The Relationship Between Safeguarding Internal Security and Complying with International Protection Obligations and Instruments*. 743 final.

European Commission (2013) *What is Horizon 2020?*, *Horizon 2020 - European Commission*. Available at: https://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020 (Accessed: 27 July 2021).

European Commission (2017a) *Proposal for a Regulation of the European Parliament and the Council on Establishing a Framework for Interoperability between Information Systems (Borders and Visa) and Amending Council Decision 2004/512/EC, Regulation (EC) No 767/2008, Council Decision 2008/633/JHA, Regulation (EU) 2016/399 and Regulation (EU) 2017/2226*.

European Commission (2017b) 'Visa Information System (VIS)'. Migration and Home Affairs.

European Commission (2019) *Smart Borders*, *Migration and Home Affairs - European Commission*. Available at: https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders_en (Accessed: 28 July 2021).

European Commission (2020a) *Intelligent Portable Border Control System | iBorderCtrl Project | H2020*, *CORDIS | European Commission*. Available at: https://cordis.europa.eu/project/id/700626 (Accessed: 29 July 2021).

European Commission (2020b) *Migration and Asylum Package: New Pact on Migration and Asylum documents adopted on 23 September 2020*, *European Commission*. Available at: https://ec.europa.eu/info/publications/migration-and-asylum-package-new-pact-migration-and-asylum-documents-adopted-23-september-2020_en (Accessed: 24 July 2021).

European Commission (2021) *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL 'A strategy towards a fully functioning and resilient Schengen area'*. Brussels: European Commission.

European Council (2001) ‘Presidency Conclusions: European Council Meeting in Laeken 14 and 15 December 2001’, 47.

European Parliament (2019) *Answer given by Mr Avramopoulos on behalf of the European Commission: No E-000204/19*. Available at: https://www.europarl.europa.eu/doceo/document/E-8-2019-000204-ASW_EN.html (Accessed: 27 July 2021).

European Parliament and Council of the European Union (2013) *Regulation (EU) No 603/2013*.

Federal Office for Migration and Refugees (2018) *Digitisation Agenda 2020. Success stories and future digital projects at the Federal Office for Migration and Refugees (BAMF). Nuremberg: Federal Office for Migration and Refugees*. Federal Office for Migration and Refugees.

Ferrer, X. *et al.* (2021) ‘Bias and Discrimination in AI: a cross-disciplinary perspective’, *IEEE Technology and Society Magazine*, 40(2), pp. 72–80. doi: 10.1109/MTS.2021.3056293.

Foucault, M. (1995) *Discipline and Punish: The Birth of the Prison*. Reprint edition. New York: Vintage.

FRA (2016) *The impact of the proposal for a revised Eurodac Regulation on fundamental rights. Opinion of the European Union Agency for Fundamental Rights*. Vienna. Available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2016-opinion-06-2016-eurodac-0_en.pdf.

Frawley, W. J., Piatetsky-Shapiro, G. and Matheus, C. J. (1992) ‘Knowledge Discovery in Databases: An Overview’, *AI Magazine*, 13(3), pp. 57–57. doi: 10.1609/aimag.v13i3.1011.

Frey, B. B. (2018) *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. 2455 Teller Road, Thousand Oaks, California 91320: SAGE Publications, Inc. doi: 10.4135/9781506326139.

FRONTEX (2020) *Frontex Thematic Industry Days: Airborne Maritime Surveillance Solutions*. Available at: https://frontex.europa.eu/future-of-border-control/research-and-innovation/announcements/frontex-thematic-industry-days-airborne-maritime-surveillance-solutions-iJ28gb (Accessed: 27 July 2021).

Geiger, M. and Pécoud, A. (2010) ‘The Politics of International Migration Management’, in Geiger, M. and Pécoud, A. (eds) *The Politics of International Migration Management*. London: Palgrave Macmillan UK (Migration, Minorities and Citizenship), pp. 1–20. doi: 10.1057/9780230294882_1.

Georgiou, D. M. and Zaborowski, D. R. (2017) ‘Council of Europe report: Media coverage of the “refugee crisis”: A cross-European perspective’, p. 24.

Gillespie, T. (2016) *Digital Keywords*. Princeton University Press. doi: 10.2307/j.ctvct0023.

Goddard, K., Roudsari, A. and Wyatt, J. C. (2012) ‘Automation bias: a systematic review of frequency, effect mediators, and mitigators’, *Journal of the American Medical Informatics Association : JAMIA*, 19(1), pp. 121–127. doi: 10.1136/amiajnl-2011-000089.

Griffard, M. (2019) ‘A Bias-Free Predictive Policing Tool?: An Evaluation of the NYPD’s Patternizr’, p. 42.

Groutsis, D., Broek, D. van den and Harvey, W. S. (2015) ‘Transformations in Network Governance: The Case of Migration Intermediaries’, *Journal of Ethnic and Migration Studies*, 41(10), pp. 1558–1576. doi: 10.1080/1369183X.2014.1003803.

Haas, D. H. (2010) ‘The Internal Dynamics of Migration Processes: A Theoretical Inquiry’, *Journal of Ethnic and Migration Studies*, 36(10), pp. 1587–1617. doi: 10.1080/1369183X.2010.489361.

Hacker, P. (2018) ‘Teaching fairness to artificial intelligence: Existing and novel strategies againstalgorithmic discrimination under EU law’, *Common Market Law Review*, 55, pp. 1143–1185.

Hao, K. (2020) *We read the paper that forced Timnit Gebru out of Google. Here’s what it says.*, *MIT Technology Review*. Available at: https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/ (Accessed: 5 July 2021).

Harryson, M. *et al.* (2018) *Notes From the AI Frontier. Applying Artificial Intelligence for Social Good*. Washington D.C: McKinsey Global Institute.

Heikkila, M. (2021) *Europe’s artificial intelligence blindspot: Race*, *POLITICO*. Available at: https://www.politico.eu/article/europe-artificial-intelligence-blindspot-race-algorithmic-harm/ (Accessed: 24 July 2021).

Heinrichs, B. (2021) ‘Discrimination in the age of artificial intelligence’, *AI & SOCIETY*. doi: 10.1007/s00146-021-01192-2.

Helbling, M. and Meierrieks, D. (2020) ‘Terrorism and Migration: An Overview’, *British Journal of Political Science*, pp. 1–20. doi: 10.1017/S0007123420000587.

Herrera, G. L. (2012) *Technology and International Transformation: The Railroad, the Atom Bomb, and the Politics of Technological Change*. SUNY Press.

Herz, J. H. (1976) *The nation-state and the crisis of world politics: Essays on international politics in the twentieth century*. New York: D. McKay.

Hildebrandt, M. (2019) *The Issue of Bias. The Framing Powers of Ml*. SSRN Scholarly Paper ID 3497597. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3497597.

Hope, A. (2020) *Security gates at Brussels Airport to be replaced after only five years*, *The Brussels Times*. Available at: https://www.brusselstimes.com/brussels-2/94978/security-gates-at-brussels-airport-to-be-replaced-after-only-five-years-technology-attacks-march-2016-electronic-fiasco/ (Accessed: 27 July 2021).

Hughes, T. P. (1969) ‘Technological Momentum in History: Hydrogenation in Germany 1898-1933’, *Past & Present*, (44), pp. 106–132.

Human Rights Watch (2020) *EU: Probe Frontex Complicity in Border Abuses*, *Human Rights Watch*. Available at: https://www.hrw.org/news/2020/11/09/eu-probe-frontex-complicity-border-abuses (Accessed: 4 June 2021).

Hunt, E. (2016) *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*, *The Guardian*. Available at: http://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter (Accessed: 2 July 2021).

Hutchinson, T. and Duncan, N. (2012) 'Defining and Describing What We Do: Doctrinal Legal Research', *Deakin Law Review*, 17(1), pp. 83–119. doi: 10.21153/dlr2012vol17no1art70.

Huysmans, J. (2000) 'The European Union and the Securitization of Migration', *JCMS: Journal of Common Market Studies*, 38(5), pp. 751–777. doi: 10.1111/1468-5965.00263.

iBorderCrtl (2016) *The project | iBorderCtrl*. Available at: https://www.iborderctrl.eu/The-project (Accessed: 20 January 2021).

iBorderCtrl? No! (no date). Available at: https://iborderctrl.no/ (Accessed: 29 July 2021).

Ingold, D. and Soper, S. (2016) *Amazon Doesn't Consider the Race of Its Customers. Should It?*, *Bloomberg.com*. Available at: http://www.bloomberg.com/graphics/2016-amazon-same-day/ (Accessed: 16 January 2021).

International Centre for Missing and Exploited Children (2018) *International Centre For Missing & Exploited Children (ICMEC) Expands With AI, Biometrics, And Ad Tech*, *International Centre for Missing & Exploited Children*. Available at: https://www.icmec.org/press/international-centre-for-missing-exploited-children-icmec-expands-with-ai-biometrics-and-ad-tech/ (Accessed: 29 July 2021).

IOM (2007) *Sharing Data - Where to Start*. Lanham: International Organization for Migration. Available at: https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=4988275 (Accessed: 21 July 2021).

IOM (2018) *Displacement Tracking Matrix*, *IOM*. Available at: https://displacement.iom.int/system/tdf/DOE%20Infosheet%20-%20DTM%20v12.0_20180308.pdf?file=1&type=node&id=1205.

IOM and Norwegian Ministry of Foreign Affairs (2013) *Handbook on European Non-discrimination Law and its Update including the Manual on the Anti-discrimination Legal Framework and Referral Mechanism in Azerbaijan*, pp. 1–137. Available at: https://www.echr.coe.int/Documents/Handbook_non_discri_law_ENG_for_AZE.pdf.

Jeandesboz, J. (2016) 'Smartening border security in the European Union: An associational inquiry', *Security Dialogue*, 47(4), pp. 292–309. doi: 10.1177/0967010616650226.

Johnson, D. (2013) *Guns Don't Kill People, People Do?* Available at: https://www.psychologytoday.com/gb/blog/logical-take/201302/guns-don-t-kill-people-people-do (Accessed: 26 July 2021).

Jones, C. (2020) *Automated Suspicion: The EU's New Travel Survaillance Initiatives*. Statewatch.

Keohane, R. O. and Nye, J. S. (1977) *Power and interdependence: world politics in transition*. Boston: Little, Brown.

Kersting, K. (2018) 'Machine Learning and Artificial Intelligence: Two Fellow Travelers on the Quest for Intelligent Behavior in Machines', *Frontiers in Big Data*, 1, p. 6. doi: 10.3389/fdata.2018.00006.

Kloza, D. *et al.* (2020) *Data protection impact assessments in the European Union: complementing the new legal framework towards a more robust protection of individuals*. preprint. LawArXiv. doi: 10.31228/osf.io/b68em.

Köchling, A. and Wehner, M. C. (2020) 'Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development', *Business Research*, 13(3), pp. 795–848. doi: 10.1007/s40685-020-00134-w.

Kroll, J. *et al.* (2017) 'Accountable Algorithms', *University of Pennsylvania Law Review*, 165(3), p. 633.

Lippert-Rasmussen, K. (2006) 'The badness of discrimination', *Ethical Theory and Moral Practice*, 9(2), pp. 167–185. doi: 10.1007/s10677-006-9014-x.

Lipton, A. Z. C. (2018) 'From AI to ML to AI: On Swirling Nomenclature & Slurried Thought', *Approximately Correct*, 5 June. Available at: https://www.approximatelycorrect.com/2018/06/05/ai-ml-ai-swirling-nomenclature-slurried-thought/ (Accessed: 16 July 2021).

Lowry, S. and Macpherson, G. (1988) 'A blot on the profession', *British Medical Journal (Clinical research ed.)*, 296(6623), pp. 657–658.

Lum, K. and Isaac, W. (2016) 'To predict and serve?', *Significance*, 13(5), pp. 14–19. doi: 10.1111/j.1740-9713.2016.00960.x.

Marchant, G. E., Allenby, B. R. and Herkert, J. R. (eds) (2011) *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem*. Springer Netherlands (The International Library of Ethics, Law and Technology). doi: 10.1007/978-94-007-1356-7.

McCarthy, D. R. (2017) *Technology and World Politics: An Introduction*. Abingdon, Oxon ; New York, NY: Routledge.

McStay, A. (2018) *Emotional AI: The Rise of Empathic Media*. doi: 10.4135/9781526451293.

Mendoza, C. (2020) 'Greece's Migrant Crisis: A Combination of EU Policy Failures & Populism', *The Generation*, 9 November. Available at: http://the-generation.net/greeces-migrant-crisis-a-combination-of-eu-policy-failures-populism/ (Accessed: 29 July 2021).

Mey, J. L. (2016) ‘Sunt verba rerum: the pragmatic life of words’, *SpringerPlus*, 5(1), p. 1653. doi: 10.1186/s40064-016-3269-z.

Mitchell, T. (1997) *Machine Learning*. Maidenhead: McGraw-Hill (McGraw Hill Series in Computer Science). Available at: https://www.cs.cmu.edu/~tom/mlbook.html (Accessed: 16 July 2021).

Molnar, P. (2019) ‘Technology on the margins: AI and global migration management from a human rights perspective’, *Cambridge International Law Journal*, 8(2), pp. 305–330.

Molnar, P. (2020) *Technological Testing Grounds: Migration Management Experiments and Reflections from the Ground Up*. EDRi, pp. 1–57. Available at: https://edri.org/wp-content/uploads/2020/11/Technological-Testing-Grounds.pdf (Accessed: 10 July 2021).

Molnar, P. and Gill, L. (2018) ‘Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada’s Immigration and Refugee System’. Available at: https://tspace.library.utoronto.ca/handle/1807/94802 (Accessed: 21 July 2021).

Morrow, W. J. and Shjarback, J. a. (2019) ‘Police worldviews, unconscious bias, and their potential to contribute to racial and ethnic disparities in New York Police Department (NYPD) stops for reason of “furtive movement”’, *Journal of Ethnicity in Criminal Justice*, 17(3), pp. 269–298. doi: 10.1080/15377938.2019.1636920.

Northpointe (2019) ‘Practitioner’s guide to COMPAS Core.’ Traverse City. Available at: http://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf.

OECD (2014) *Is Migration Really Increasing?* Available at: https://www.oecd.org/berlin/Is-migration-really-increasing.pdf.

O’Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Osipov, G. S., Panov, A. I. and Yakovlev, K. S. (2019) *Artificial Intelligence: 5th RAAI Summer School, Dolgoprudny, Russia, July 4–7, 2019, Tutorial Lectures*. Springer Nature.

Pager, D. and Shepherd, H. (2008) ‘The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets’, *Annual Review of Sociolgy*, 34, pp. 181–209.

Parasuraman, R. and Manzey, D. H. (2010) ‘Complacency and bias in human use of automation: an attentional integration’, *Human Factors*, 52(3), pp. 381–410. doi: 10.1177/0018720810376055.

Pasquale, F. (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press.

Pitt, J. C. (2014) ‘“Guns Don’t Kill, People Kill”; Values in and/or Around Technologies’, in Kroes, P. and Verbeek, P.-P. (eds) *The Moral Status of Technical Artefacts*. Dordrecht: Springer Netherlands (Philosophy of Engineering and Technology), pp. 89–101. doi: 10.1007/978-94-007-7914-3_6.

Podesta, J. *et al.* (2014) *Big Data: Seizing Opportunities, Preserving Values*. Executive Office of the U.S. President, pp. 1–85. Available at: https://perma.cc/ZXB4-SDL9 (Accessed: 20 July 2021).

Privacy International (2020) *Submission to the UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*. Available at: https://privacyinternational.org/sites/default/files/2020-06/PI%27s%20joint%20UN%20submission%20on%20race%20border%20tech_final.pdf.

Puppe, F. (1993) *Systematic Introduction to Expert Systems: Knowledge Representations and Problem-Solving Methods*. Berlin Heidelberg: Springer-Verlag. doi: 10.1007/978-3-642-77971-8.

Ragozin, L. (2015) *Latvians find unity in rejecting refugees*, *POLITICO*. Available at: https://www.politico.eu/article/latvia-migration-asylum-crisis-baltics-eu/ (Accessed: 29 July 2021).

Ramos, S. (2021) *COVID-19's impact felt by researchers*, *American Psychological Association*. Available at: https://www.apa.org/science/leadership/students/covid-19-impact-researchers (Accessed: 12 July 2021).

ROBORDER (2020) 'Aims & Objectives – Roborder'. Available at: https://roborder.eu/the-project/aims-objectives/ (Accessed: 27 July 2021).

Salter, M. (2015) *Making Things International 1: Circuits and Motion*. 1st edn. Minneapolis: Univ Of Minnesota Press.

Sánchez-Monedero, J. and Dencik, L. (2020) 'The politics of deceptive borders: "biomarkers of deceit" and the case of iBorderCtrl', *Information, Communication & Society*, 0(0), pp. 1–18. doi: 10.1080/1369118X.2020.1792530.

Scholte, J. A. (2005) *Globalization: A Critical Introduction*. Macmillan International Higher Education.

Shashkevich, A. (2018) 'Algorithm improves integration of refugees', *Stanford News*, 18 January. Available at: https://news.stanford.edu/2018/01/18/algorithm-improves-integration-refugees/ (Accessed: 25 July 2021).

Shull, A. (2019) *In the Global Race for AI, How Do We Ensure We're Creating a Better World?*, *Centre for International Governance Innovation*. Available at: https://www.cigionline.org/articles/global-race-ai-how-do-we-ensure-were-creating-better-world/ (Accessed: 24 July 2021).

Singler, S. (2019) '"Our Role Is Technical Rather Than Political": Explaining the Ascendancy and Proliferation of "Smart Borders" in the European Union', p. 115.

Sironi, A., Bauloz, C. and Milen, E. (2019) *IOM Glossary on Migration*, *International Organization for Migration*. Available at: https://www.iom.int/glossary-migration-2019 (Accessed: 15 March 2021).

Srinivas, N. (ORCID:000000030493935X) *et al.* (2019) *Face Recognition Algorithm Bias: Performance Differences on Images of Children and Adults*. Oak Ridge National Lab.

(ORNL), Oak Ridge, TN (United States). Available at: https://www.osti.gov/biblio/1559665 (Accessed: 27 July 2021).

Taylor, P. (2019) 'Has Europe Reached Peak Populism?', *Politico*. Available at: https://www.politico.com/magazine/story/2019/09/05/has-europe-reached-peak-populism-228036/ (Accessed: 26 July 2021).

Tech America Foundation (2012) *Demystifying Big Data: A Practical Guide to Transforming the Business of Government*. Washington DC: Federal Big Data Commission.

The World Economic Forum (2019) *World Economic Forum consortium launches paperless Canada-Netherlands travel pilot*. Available at: https://www.newswire.ca/news-releases/world-economic-forum-consortium-launches-paperless-canada-netherlands-travel-pilot-805263166.html (Accessed: 29 July 2021).

Thomas, E. (2018) 'Tagged, tracked and in danger: how the Rohingya got caught in the UN's risky biometric database', *Wired UK*, 12 March. Available at: https://www.wired.co.uk/article/united-nations-refugees-biometric-database-rohingya-myanmar-bangladesh (Accessed: 25 July 2021).

Thompson, K. J. (2020) 'The perils of practicum in the time of COVID-19: A graduate student's perspective', *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1), pp. S151–S152. doi: 10.1037/tra0000752.

Topcu, E. (2019) 'Turkey using Interpol to track down dissidents', *Deutsche Welle*, 7 November. Available at: https://www.dw.com/en/turkey-using-interpol-to-track-down-dissidents/a-51159723 (Accessed: 28 July 2021).

UNHCR (2017) 'Data Collection and Analysis', *Data Collection and Analysis*, p. 28.

UNHCR (2019) *Refugee Registration and Profiling in Jordan*, pp. 1–3.

UNHCR (2021) 'Global Trends - Forced displacement in 2020', *UNHCR Flagship Reports*. Available at: https://www.unhcr.org/flagship-reports/globaltrends/ (Accessed: 13 July 2021).

Vanian, J. (2018) *Unmasking A.I.'s Bias Problem*, *Fortune*. Available at: https://fortune.com/longform/ai-bias-problem/ (Accessed: 4 July 2021).

Verseck, K. (2019) *Hungary's slow descent into xenophobia, racism and human rights abuses*, *InfoMigrants*. Available at: https://www.infomigrants.net/en/post/20220/hungary-s-slow-descent-into-xenophobia-racism-and-human-rights-abuses (Accessed: 29 July 2021).

Vincent, J. (2018) 'Google "fixed" its racist algorithm by removing gorillas from its image-labeling tech', *The Verge*, 12 January. Available at: https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai (Accessed: 24 July 2021).

Wæver, O. (1995) *Securitization and Desecuritization*. Columbia University Press. RD Lipschutz (On Security).

Wagner, B. (2018) 'Ethics As An Escape From Regulation. From "Ethics-Washing" To Ethics-Shopping?', in Bayamlioglu, E. et al. (eds) *Being Profiled: Cogitas Ergo Sum 10*

Years of 'Profiling the European Citizen'. Amsterdam: Amsterdam University Press, pp. 84–89. doi: 10.1515/9789048550180-016.

Wendt, A. (1999) *Social Theory of International Politics*. Cambridge, UK ; New York: Cambridge University Press.

Williams, M. C. (2003) 'Words, Images, Enemies: Securitization and International Politics', *International Studies Quarterly*, 47(4), pp. 511–531.

Winner, L. (1980) 'Do Artifacts Have Politics?', *Daedalus*, 109(1), pp. 121–136.

Writer, T. L.-S. (2014) *A Brief History of Artificial Intelligence*, *livescience.com*. Available at: https://www.livescience.com/49007-history-of-artificial-intelligence.html (Accessed: 16 July 2021).

Zhang, M. *et al.* (2014) 'Emotional Context Influences Micro-Expression Recognition', *PLOS ONE*, 9(4), p. e95018. doi: 10.1371/journal.pone.0095018.

Zuiderveen Borgesius, F. (2015) *Online Price Discrimination and Data Protection Law*. SSRN Scholarly Paper ID 2652665. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.2652665.

Legal Instruments

Council of Europe, *European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14*, 4 November 1950, ETS 5, available at: https://www.refworld.org/docid/3ae6b3b04.html [accessed 13 January 2021]

'Consolidated Versions of the Treaty on European Union and the Treaty on the Functioning of the European Union', Pub. L. No. 2008/C 115/01, OJ C 115 (2008), https://eur-lex.europa.eu/legalcontent/ EN/TXT/?qid=1544456032916&uri=CELEX:C2008/115/01.

Protocol relating to the Status of Refugees (adopted 31 January 1967, entered into force 4 October 1967) 606 UNTS 267, art 2

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

UN General Assembly, Convention Relating to the Status of Refugees, 28 July 1951, United Nations, Treaty Series, vol. 189, p. 137, available at: https://www.refworld.org/docid/3be01b964.html [accessed 10 April 2021]

Case Law

European Court of Human Rights [ECHR] *Biao v. Denmark* (Grand Chamber), No. 38590/10, [2016] para. 89, 24 May 2016.

European Cour of Human Rights [ECHR] *Carson and Others v. the United Kingdom* [GC], no. 42184/05

European Court of Human Rights [ECHR] *D.H. and Others v. the Czech Republic* [GC] - 57325/00 Judgment 13.11.2007 [GC]

European Court of Human Rights [ECHR] *E.B. v. France, 43546/02* 22 January 2008

European Court of Human Rights [ECHR] *Handyside v United Kingdom*, Merits, App No 5493/72, A/24, [1976] ECHR 5, (1976) 1 EHRR 737, (1979) 1 EHRR 737, IHRL 14 (ECHR 1976), 7th December 1976.

European Court of Human Rights [ECHR] *Marckx v. Belgium*, Application no. 6833/74 13 June 1979

APPENDIX A – Data classification

Category	Typology	Examples
Official records	• Official Reports • Official EU Webpage releases • Official Position Papers • Conference Statements	Dumbrava, C. (2021) *Artificial Intelligence and EU Borders*. Brussels: European Parliamentary Research Service of the Secretariat of the European Parliament. Available at: http://www.eprs.ep.parl.union.eu (intranet). EDRi (2021) *Chilling use of face recognition at Italian borders shows why we must ban biometric mass surveillance*, *European Digital Rights (EDRi)*. Available at: https://edri.org/our-work/face-recognition-italian-borders-ban-biometric-mass-surveillance/ (Accessed: 19 February 2021). European Council (2001) 'Presidency Conclusions: European Council Meeting in Laeken 14 and 15 December 2001', 47. European Parliament (2019) *Answer given by Mr Avramopoulos on behalf of the European Commission: No E-000204/19*. Available at: https://www.europarl.europa.eu/doceo/document/E-8-2019-000204-ASW_EN.html (Accessed: 7 March 2021). European Parliament and Council of the European Union (2013) *Regulation (EU) No 603/2013*. Federal Office for Migration and Refugees (2018) *Digitisation Agenda 2020. Success stories and future digital projects at the Federal Office for Migration and Refugees (BAMF). Nuremberg: Federal Office for Migration and Refugees*. Federal Office for Migration and Refugees. FRA (2016) *The impact of the proposal for a revised Eurodac Regulation on fundamental rights. Opinion of the European Union Agency for Fundamental Rights*. Vienna. Available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2016-opinion-06-2016-eurodac-0_en.pdf. Council of Europe and European Court of Human Rights (2020) *Guide on Article 14 of the European Convention on Human Rights and on Article 1 of Protocol No. 12 to the Convention: Prohibition of discrimination*.
Legal Documents	• Official Legislation • Policies or policy directives	Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14, 4 November 1950, ETS 5, available at: https://www.refworld.org/docid/3ae6b3b04.html [accessed 13 January 2021]

	• Regulations • White Papers • Memoranda of understanding • Case law	'Consolidated Versions of the Treaty on European Union and the Treaty on the Functioning of the European Union' Protocol relating to the Status of Refugees UN General Assembly, Convention Relating to the Status of Refugees, 28 July 1951 European Court of Human Rights [ECHR] *Biao v. Denmark* (Grand Chamber), No. 38590/10, [2016] para. 89, 24 May 2016. European Cour of Human Rights [ECHR] *Carson and Others v. the United Kingdom* [GC], no. 42184/05 European Court of Human Rights [ECHR] *D.H. and Others v. the Czech Republic* [GC] - 57325/00 Judgment 13.11.2007 [GC] European Court of Human Rights [ECHR] *E.B. v. France, 43546/02* 22 January 2008 European Court of Human Rights [ECHR] *Handyside v United Kingdom*, Merits, App No 5493/72, A/24, [1976] ECHR 5, (1976) 1 EHRR 737, (1979) 1 EHRR 737, IHRL 14 (ECHR 1976), 7th December 1976. European Court of Human Rights [ECHR] *Marckx v. Belgium*, Application no. 6833/74 13 June 1979
Ethical documents (iBorderCtrl)	• Ethical Report • Financial analyses • Methodoogy • Specific Literature	*iBorderCtrl Project | H2020*, *CORDIS | European Commission*. Available at: https://cordis.europa.eu/project/id/700626 (Accessed: 2 March 2021). iBorderCrtl (2016) *The project | iBorderCtrl*. Available at: https://www.iborderctrl.eu/The-project (Accessed: 20 January 2021). *iBorderCtrl? No!* (no date). Available at: https://iborderctrl.no/ Sánchez-Monedero, J. and Dencik, L. (2020) 'The politics of deceptive borders: "biomarkers of deceit" and the case of iBorderCtrl', *Information, Communication & Society*, 0(0), pp. 1–18. doi: 10.1080/1369118X.2020.1792530.
Scholarly papers	• Textbooks • Reports • Reviews • Scientific or peer-	Amelung, N., Granja, R. and Machado, H. (2021) 'A Brief History of the Evolution of Biometrics and Biometric Database Systems Crossing Borders in EU Law Enforcement', in Amelung, N., Granja, R., and Machado, H. (eds) *Modes of Bio-Bordering: The Hidden (Dis)integration of Europe*. Singapore: Springer, pp. 15–34. doi: 10.1007/978-981-15-8183-0_2.

	reviewed publications • Masters; MPhil, PhD theses	Angwin, J. *et al.* (2016) *Machine Bias*, *ProPublica*. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=KCQtlys2Jot94pBZSAGZgQHpS8pwmTyy (Accessed: 16 January 2021). Ansemes de Vries, L. and Guild, E. (2019) 'Seeking refuge in Europe: spaces of transit and the violence of migration management', *Journal of Ethnic and Migration Studies*, 45(12), pp. 2156–2166. doi: 10.1080/1369183X.2018.1468308. Balzacq, T. (2005) 'The Three Faces of Securitization: Political Agency, Audience and Context', *European Journal of International Relations*, 11(2), pp. 171–201. doi: 10.1177/1354066105052960. Bansak, K. *et al.* (2018) 'Improving refugee integration through data-driven algorithmic assignment', *Science (New York, N.Y.)*, 359(6373), pp. 325–329. doi: 10.1126/science.aao4408. Borgesius, F. (2018) *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*. Counci of Europe, pp. 1–94. Barocas, S. and Selbst, A. D. (2016) *Big Data's Disparate Impact*. SSRN Scholarly Paper ID 2477899. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.2477899. Barrett, L. F. *et al.* (2019) 'Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements', *Psychological Science in the Public Interest*, 20(1), pp. 1–68. doi: 10.1177/1529100619832930.
Media and communication	• Newspaper articles • Lectures; Tedtalks; Feature Films • Blogs; Webpages • Official Twitter Conversations	Buolamwini, J. (2020) *CODED BIAS*, *CODED BIAS*. Available at: https://www.codedbias.com (Accessed: 22 January 2021). Access Now (2021) 'Iris scanning of refugees is disproportionate and dangerous — What's happening behind IrisGuard's closed doors?', *Access Now*, 12 April. Available at: https://www.accessnow.org/irisguard-refugees-jordan/ (Accessed: 25 May 2021). Crawford, K. (2016) 'Opinion | Artificial Intelligence's White Guy Problem (Published 2016)', *The New York Times*, 25 June. Available at: https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html (Accessed: 11 January 2021).

		Crawford, K. (2017) 'The Trouble with Bias - NIPS 2017 Keynote - Kate Crawford #NIPS2017', in. Available at: https://www.youtube.com/watch?v=fMym_BKWQzk&t=2574s (Accessed: 16 January 2021). Dastin, J. (2018) 'Amazon scraps secret AI recruiting tool that showed bias against women', *Reuters*, 10 October. Available at: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight- Hunt, E. (2016) *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*, *The Guardian*. Available at: http://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter (Accessed: 2 May 2021).

Criteria: Relevance; Up-to-datedness; Accessibility ; Completeness of document.