Predictive Policing Using AI & ML for Domestic Law Enforcement: Critical Analysis & Framework Development in EU

August 2023

University of Glasgow: 2696514M
Dublin City University: 21108951
Charles University: 42030775

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

Word count: 23246
Supervisor: Dr. Anzhelika Solovyeva
Date of Submission: 21-08-2023

Abstract

This research provides a critical analysis of predictive policing systems within Germany and Netherlands, as part of the larger European Union (EU), where the focus is on biases and issues within processes that need to be addressed using a human security approach to help improve and redesign these systems to integrate them into a normative framework such as the EU AI Act. As case studies, location-based and person-based predictive policing practices in Germany and Netherlands were identified and chosen for analysis through a deductive analytical approach using a relevant theoretical framework. Qualitative analysis was employed on information collated from interviews with researchers and industry experts from the respective regions, as well as published journals in the field of predictive law enforcement in EU. The literature review has information on the data collected from journals as part of the analysis, which was later employed extensively within the empirical analysis section. Unlike most literature reviewed; this research explicitly applies human security as a theoretical approach to predictive policing practices as well as applying relevant theories to explain the probable solutions provided by experts. By employing a human security approach integrated with algorithmic bias theory, the results of the data analysed showed that although there are positives that need to be considered in these systems, with the probable development and introduction of the EU AI Act it is highly likely these systems will need to be revised and, in some cases, abolished, due to the inherently biased processes and datasets. It is possible to improve the systems though however that will require strong collaboration with external stakeholders, disaggregated collection and analysis of data to curb bias in datasets, and the introduction of a standardised, independent auditing body and frameworks. However, even with all these steps there can be issues with standardisation from individual states.

Acknowledgement

I would like to offer my sincerest gratitude to my core supervisor, Dr. Anzhelika Solovyeva of Charles University Prague for her guidance regarding the whole research process and providing useful remarks when I was writing the thesis. The feedbacks were helpful in enhancing the quality of my work. I would like to express my sincerest gratitude to Junior Prof. Dr. Lucia Sommerer, author of the book "Self-Imposed Algorithmic Thoughtlessness and the Automation of Crime Control" and Dr Lorella Viola, Author of "The Humanities in the Digital" for their valuable insights in this field. Special thanks to Dr Gwen van Eijk, Policy Officer Technology and Human Rights at Amnesty International for her detailed explanation on the research work that has been done by human rights organisations within this field.

List of Figures

Fig 1: 2^{nd} gen assessment (Itälunni, 2018)………………...…………......11

Fig 2: Source of bias in Simplified Policing Algorithms (FRA, 2022)…28

Fig 3: CAS interface (Oosterloo et al., 2018)……………………..…….49

Fig 4: Two-stage assessment (Itälunni, 2018)….……………………….53

Fig 5: Runaway feedback loops (FRA, 2022)…………………………..62

Fig 6: Bias Mitigation Techniques (FRA, 2022)……………………......75

List of Acronyms

AI- Artificial Intelligence

GDPR-General Data Protection Regulation

ML- Machine Learning

PBF- Person-Based Forecasting
STS- Socio-technical systems approach

SP- Structured Professional judgement

Contents

Abstract i

Acknowledgement ii

List of Figures iii

List of Acronyms iv

1. Introduction 1

1.1 Background 1

1.2 Objectives & Motivation 4

1.3 Research Structure 5

2.Conceptual Background 7

2.1 Types of Predictive Policing & incorporation of AI 8

2.2 Literature Review 12

2.3 Research Question 18

3.Theoretical Framework 19

3.1 Definitions 19

3.2 Framework 21

3.2.1 Human Security theory: 22

3.2.2 Algorithmic Bias theory: 25

3.3 Counter bias Frameworks 30

3.3.1 Algorithmic Fairness Theory: 30

3.3.2 Socio-Technical Systems (STS) Approach: 32

3.4 AI Act 36

4.Methodology 38

4.1 Research Design & Strategy 38

4.2 Case Selection 42

4.3 Participants 45

4.4 Limitations 46

5.Empirical Analysis 48

5.1 Use cases 48

5.2 Arguments for improvement instead of discontinuation 54

5.3 Analysis 56

5.3.1 Greater collaboration with non-state actors in education & NGOs is necessary ……………………………………………………………...56

5.3.2 Disaggregated collection and processing of data will reduce 'feedback loop' of biases ……………………………………………………..61

5.3.3 Independent Audit body to curb transparency & bias: 67

5.3.4 EU Standardisation could be attained through EU AI Act, but the roadblocks are more state-oriented: 72

5.4 Discussion..74

6.Conclusion...80

Bibliography...85

Appendix 97

Introduction

1.1 Background

The permeation of technology into different facets of society has allowed the implementation of technological solutions that go beyond resolving common problems. Complex systems such as law enforcement are incorporating technology to assist in identifying and catching criminals. However, technology has evolved to a point where it doesn't simply help solve complex issues for society but also defines social constructs. To find criminals is not enough anymore, to identify potential criminals pre-emptively, before a crime occurs is the current phase of technology. This led to the development of predictive policing technologies. Ferguson (2012) defined predictive policing as a strategy employed by law enforcement to develop and utilise information and analysis thereby stimulating proactive crime prevention. By amalgamating big data and predictive analytics, law enforcement and state agencies are able to deploy resources with high efficacy in areas that have a relatively high propensity for crime. This is possible with the analysis of historical crime data to predict patterns of future crime including the type, location and the most probable perpetrator. Predictive policing systems though are varied and can be designed and developed using different methodologies to fulfil different purposes. Moreover, as these systems have become more and more ingrained into law enforcement practices and by extension, society itself, they have the innate ability to conduct social influence or be influenced by biases pre-existing within certain societies.

Although the development of such systems by state agencies might have been with the intention of creating a "neutral" approach to proactive law enforcement by omitting implicit biases prevalent within human decision-making, questions have been asked regarding how individual profiling is done. Pre-emptive criminal profiling can not only affect the perceptions and attitude society has towards individuals, and by extension their group, but can also be catalysed by implicit biases already present within that society. These biases can permeate from the datasets chosen to the algorithms and processes used within these systems by heavily relying upon a set of assumptions. Biases can taint algorithms in the same way an individual's assumptions are reinforced, instead of resolving existing biases.

Some concerns were expressed by Richardson et al. (2019) regarding the accuracy of crime data that was being used for predictive policing namely flawed data that can be racially biased that could lead to inaccurate and systematically biased predictions towards minorities. Even when high-quality and historically accurate information is used within such systems, some processes are empirically biased and susceptible to feedback loops, wherein the same biased results are fed into the system as input. This would lead to the same neighbourhoods being targeted for extensive patrols and surveillance due to relatively lower socio-economic standards, and individuals within certain minorities would also be heavily surveilled. If predictions lead to the stigmatisation of individuals and specific communities, then potential harm can be afflicted from human security perspective through bias within algorithms. With the European Union Council adopting its common position regarding the Artificial Intelligence Act (AI Act), extending restrictions posed by the GDPR on decision making to a holistic regulation to govern AI systems within the EU is well on its way to become a reality.

El Atillah (2023) though stated that research conducted at Standford University's Center for Research on Foundation Models (CRFM) concluded that most AI models within EU, whether it be state-controlled or private, don't comply with the AI Act. From data resources to computational processes, the algorithms used as well as the method of implementation, compliance levels within leading predictive practices are high with some being due to the training data used, thus making it difficult for state bodies as well as private organisations to adhere to the regulations that could be set. As the research has led to the belief that there can be potentially massive resistance from both private as well as state bodies in adopting the proposed legalisation, some companies have expressed their desire to leave the EU thereby halting AI advancements. One of the concerns being that these regulations could curb the speed of innovative practices both on a state and private level. Since this resistance could extend to EU member states when state goals don't adhere to EU-wide goals, extensive collaboration among policymakers and tertiary organisations have been suggested to build a cohesive solution that adheres to the proposed AI Act while also ensuring state or private demands are met. This collaboration not only requires a proper framework but also identification of issues within current AI models, such as predictive policing, and how those can be resolved to develop better systems to help develop such framework and improve predictive practices, instead of discontinuing those.

With the probability of the EU AI Act becoming more of a reality, the need for understanding the social and technical problems that plagued predictive policing systems is essential now. Although most of the published research and interviews analysed focused solely on criticising predictive policing practices, some were identified were the positive impacts could also be analysed using similar theoretical framework and methodology. This would be beneficial as well in creating a case for improving predictive policing systems instead of discontinuing them. But what methodology can be used to identify the biases and issues that exist within current predictive policing systems that could help improve these processes? Given that AI biases can be inherently invisible due to how it is quantified across systems thereby making those difficult to resolve, a thorough theoretical framework is required for investigation of both socio-technical and social aspects related to such systems. By going beyond technical issues and gauging how design rationalities, namely data selection, process creation and theoretical assumptions are intrinsically linked with social impact, this report will attempt to investigate how these systems aren't entirely neutral and can cause negative social impact to certain communities, primarily minorities. Such detrimental social impact could be due to using information that contain pre-existing social biases, creating processes that pertain to those biases, isolating the system from tertiary and external stakeholders, or a combination of these factors. With a theoretical foundation built on human security with strong focus on the community security aspect to better understand the respective aspects affecting minorities, the social effects of and variables that affect predictive policing can be identified. While these theories can help analyse the social aspects of predictive policing systems, namely the stakeholders involved, and the role of state in creation of these systems, algorithm bias theory will analyse the technical aspects namely the datasets used, the data sources chosen, and the processes employed.

1.2 Objectives & Motivation

This section is devoted towards understanding the core objectives and motivations regarding pursuing this specific research topic. Although studies have been made criticising predictive practices within Europe, such as those in Germany and Netherlands, by identifying specific aspects of design and implementation, there has been a lack of research in explaining how those critiqued aspects could be detrimental towards human security. By explicitly applying human security theory, instead of implicit application that is prevalent in EU predictive policing research, to the data gathered from published journals, as well as interviews with researchers and industry experts within this field, the thesis attempts to explain how those practices could have adverse effects on different aspects of human security. To add to that, as a theoretical approach is applied to explain the issues within predictive policing, a counter-bias theoretical methodology can be applied to explain statements provided by interviewees and researchers on improving the system, thereby stimulating the development of an improvement framework in future research. This would be ideal for policymakers if the EU AI Act proposal does go through by allowing them to not only analyse and critique predictive policing systems through a human security approach but also develop strategies for improving these systems that are theoretically robust. It must be stressed that providing recommendations are beyond the scope of this research, and the inclusion of counter-bias methodologies are to explain the critique and statements provided in journals and by interviewees from a human security perspective. Furthermore, a lack of quantitative assessments made it difficult to critically assess predictive policing practices in Germany and Netherlands quantitatively. Thus, qualitative data regarding the practices have been analysed alongside applying relevant theoretical framework to understand where the practices could be detrimental towards human security.

1.3 Research Structure

The thesis has been divided into 6 chapter with sub-sections dedicated to each of those. First, in chapter 2 the conceptual framework will cover the state of arts by dividing that into types of predictive policing(2.1), a comprehensive literature review (2.2), and the core research question derived (2.3). Following that chapter 3 will focus on the theoretical framework namely the definitions used (3.1), the framework itself (3.2), the counter bias frameworks (3.3) and a short section dedicated to understanding the AI Act (3.4). Next, chapter 4 will encompass the research design and strategy used (4.1), the participants involved (4.2), followed by the limitations of the research (4.3). Following that, chapter

5 will provide the use cases selected (5.1), arguments for improvement of predictive policing practices (5.2), the empirical analysis of data collected (5.3), followed by the discussion section (5.4). Finally, chapter 6 will be the concluding chapter, followed by the bibliography and appendix.

Conceptual Background

This chapter will introduce the core concepts related to this research starting from the types of predictive policing mechanisms before moving into a comprehensive review of published literature within the realm of predictive policing and finally, leading to the inception of the research question. The chapter will be useful for not only developing conceptual understanding regarding predictive policing systems but also identifying gaps in published literature and research, thereby contributing to defining the scope of the study. With the incorporation of a diverse portfolio of tools namely facial recognition, machine learning, and big data, AI has already expanded and permeated into various social sectors including healthcare, marketing, cybersecurity and the military. Joh (2016) however mentioned that the implementation of these tools as a predictive policing process allows law enforcement to further expand their abilities in determining and arresting or detaining potential culprits. In determining culprits though law enforcement assert a degree of power that raises concerns primarily in its efficacy to distinguish probable culprits from innocent victims. The power to determine culprits rises through the integration of AI, both applied AI and generalised AI. Nadikattu (2016) defined applied AI as carrying out specific tasks through human thought process stimulation, alternatively generalised AI is utilised to create machine intelligence systems that can be applied as a real person for any responsibility. Given that states consistently have to prioritise the investment that is made into various sectors from AI to policing, Grace (2021) referred to predictive policing as a type of "smart policing" that allows more efficient utilisation of resources while ensuring better results.

2.1 Types of Predictive Policing & incorporation of AI

Big Data Predictive Analytics: Eckerson (2007) defined predictive analytics as a sub-genre of AI that predicts future events and behaviours through detecting patterns and relationships within large volumes of data. Data is analysed inductively in this process by using AI, Statistics, ML, Neural Computing and Computational Mathematics. As law enforcement has already started developing proactive crime prevention strategies, forecasting techniques have become imperative towards determining when and where crimes could occur, which has only been possible due to better data management within law enforcement agencies. Such proactive policing, as opposed to reactive policing techniques are aimed at identifying and predicting patterns, and causal elements correlated towards crime (Fitzpatrick et al., 2019). Crime forecasting using big data has allowed the police to identify crime hotspots for specific offences such as home burglary, thus providing direct support to law enforcement agencies. Gorr & Harries (2003) further added that by using key variables like economic condition, and trend levels of crime rates, allowed law enforcement agencies to efficiently deploy manpower across provinces while managing individual workloads thereby shifting to a more prevention-based approach instead of an enforcement-oriented one. DNA databases become one of the most recurring points of surfacing information through Big Data as it allows expansion in the usage of criminal DNA primarily to resolve both past and present criminal investigations (Neiva et al., 2022). By integrating genetic with non-genetic data, Big Data analyses data to potentially produce new information that might be vital for police investigations. For example, law enforcement agencies within the EU utilise Big Data to link information across diverse databases to derive information crucial for criminal investigations.

Although some organisations require Big Data mechanics, not all of them are completely aware of how this can be processed primarily due to the diversity, volume, and complexity of structures datasets can possess. Montasari (2023b) further reiterated that Big Data volume poses the most challenge in IT infrastructure while the velocity at which this is processed, stored and analysed could influence the next set of data is developed or collated. Regulating velocity thereby is vital to ensure that analytical systems like predictive policing processes are capable of identifying patterns within data and providing appropriate insights. To add to that, the diversity of data that can be obtained from various sources or originate within or outside an organisation requires standardisation before processing to ensure there is no consistent deviation. Furthermore, the veracity of data, which concerns the quality of data can be affected by the source of data, and how the data is processed (Gillis, 2021; Demchenko et al., 2013). Big data allows the analysis of large sets of data, which can then be used to find patterns within those datasets through data mining.

ML & Data mining: Data mining identifies patterns to extract information from Big Datasets through a combination of machine learning, and statistics. Prabakarn and Mitra (2018) identified Decision trees, Rule Induction, Genetic Algorithm, Nearest Neighbour Method and Artificial Neural Networks as core data mining techniques that are used within predictive policing. Methods such as the Hidden Markov Model (HMM) can be used for crime pattern identification primarily in cases of fraud and detection. The model comprises two random variables that use statistics to sequentially modify the state thus allowing users to analyse both monitored as well as hidden events. Through this model, processes are able to identify the likelihood of crime within areas while also learning reiteratively from the data provided. Big Data Analytics therefore has assisted in identifying central members and subgroups as well as organised crime through extracting hidden network structures, which has also stimulated criminal behaviour profiling and crime cluster detection. Oatley et al. (2006) defined another method, called Link Analysis that is more focused on affinity association, which identifies relationships amongst data and the corresponding association rule. The ability to correlate copious amounts of data makes this method ideal to interpret coordinated criminal behaviour such as home burglaries, by establishing relationships between criminal entities based on similar vehicles, addresses and bank accounts. As this method can be used to detect future threats, applying it to social media profiles and content could also prove to be beneficial in proactively identifying bombs or other forms of threats. Big data, ML and data mining allows the analysis of large sets of data through methods such as pattern identification, which is an essential component of multilateral approach towards analysis that is employed in the design, development and implementation of person-based forecasting predictive policing.

Person-based forecasting: Fitzpatrick et al. (2019) defined Person-Based Forecasting (PBF) as a prediction-based method to identify individuals who could be associated with previous crimes, gang connection or other risk indicators. PBF attempts to gauge an individual's likelihood of being involved in criminal activity either as a perpetrator or victim through analysing their previous patterns of criminal activity or interactions with the police. When compared to geospatial, statistical forecasting, person-based predictive policing practices within Europe relies upon second-generation assessments that are integrated with structured professional judgement as seen in the next page:

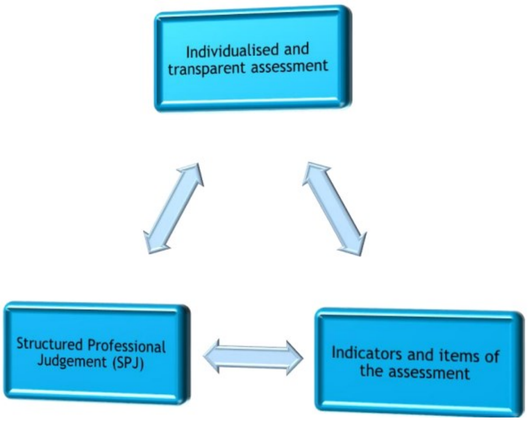

Fig 1: Second generation assessment structure (Itälunni, 2018)

With focus on more severe risks such as violence and sexual violence as opposed to burglaries, these provide individual risk assessments with a greater scope of interpretation (Rettenberger, 2016). Developed through the Violent Extremist Risk Assessment (VERA) mechanism used in Canada, the tool was initially developed to identify violent political extremism through structured professional judgement (SP) before being expanded to other types of crimes. Alongside SP, individual risk assessment identifies whether the risk is low, moderate or high when concerning an individual, as well as demographic factors such as gender and age. This approach was further evolved in the Netherlands by adding 34 indicators related to violent extremism such as attitudes and ideologies, social context and motivation to create a systematic and structured assessment tool (Pressman et al., 2016). As a semi-structured approach, this process is understood to be more transparent with greater number relevant variables that could determine violent acts as well as people who weren't initially flagged to the police. This approach does require the provision of information in a proper structure and documentation regarding one's biographical histories and allegiances to already known criminal or terrorist organisations, which would allow a probable perpetrator to be addressed on an appropriate scale of being a relative person.

The following section will extensively review all published journals and articles on predictive policing in EU, namely Germany and Netherlands, to identify gaps, namely if there are any journals where an explicit human security-oriented approach is used to identify issues within such systems as well as understanding the core complaints and solutions provided by experts in this field.

2.2 Literature Review

The literature review will focus on analysing published literature on predictive policing, not only pertaining to Europe, but beyond to develop a core understanding of what predictive policing systems entails, the critique and benefits associated with predictive policing systems, any relevant literature that has applied a human security approach and finally, the core gaps in previous literature that have been covered by this research.

Data-driven practices in policing can be traced back to the 1990s when CompStat model was implemented by the New York City Police Department (NYPD) to identify crime patterns and allocate resources efficiently through data-driven analysis of crime reports and other data sources. This eventually led to the development of the PredPol system in 2010 by the Los Angeles Police Department (LAPD) that used historical crime data to proactively identify areas where future crimes are probable and subsequently dispatch law enforcement in those areas. (Shapiro, 2017) Within the UK, which is the most developed country around EU in such predictive practices, data analysis has been integrated into domestic law enforcement practices since 2004, with crime-hotspot mapping having been implemented since 2013, the first being ProMap. (Montasari, 2023a) In comparison, European countries namely Germany and Netherlands are more nascent in terms of implementing predictive policing practices having started utilisation and investigation since 2015 to understand the social implications and how social biases might have shaped such systems.

As a data-driven approach for not only resource allocation but proactive crime prevention, concerns have risen regarding the potential implications these might have towards the society, minority communities and even individuals. The aim of the literature review is to examine critical research on predictive policing practices before narrowing down the focus to Germany and Netherlands specifically as well as understand how current research have not yet applied a human security approach.

Critique of Predictive Policing: Although there has been significant development within EU, some of the core underlying critiques have been consistent namely the reliance on correlations and patterns within datasets instead of analysing and identifying the root cause of crime that could help proactively reduce crime more efficiently. Thus, for regions like Germany and Netherlands, the predictive policing systems are currently limited to crime control accomplished by surveillance and short-term investigations, thereby restricting more sustainable strategies to combat crime that could involve resolving social issues that can catalyse criminal outcomes (Sommerer, 2022). Predictive policing systems have also been criticised for influencing the behaviour of domestic law enforcement towards individuals as well as communities. Strikwerda (2021) mentioned how such predictive systems essentially create a foregone conclusive result by directing police officers to individuals, as the officers will then link every act an individual performs to certain traits that can be mapped on a suspicious individual. This enables police officers to use actions such as searches and investigative detentions as they have a reason to interfere due to the system. These patterns of over-policing even extend to locations as certain predictive policing tools can identify locations where a crime is likely to occur, directing police officers there, and then map behavioural patterns of individuals to those living there. This increases the risks associated with false imprisonment and convictions, particularly if the areas are populated by economically impoverished or racial minorities. Ferguson (2015) argues that police officers have always attempted to identify crime-prone individuals and hotspots, even before the development of predictive policing systems, however their judgments have been marred with racial and class biases. Study by Vepřek et al. (2020) exhibited that users and developers of predictive systems believe it uses a neutral, quantitative approach could provide more credible results. Dr Lucia Sommerer though in her book "Self-Imposed Algorithmic Thoughtlessness and the Automation of Crime Control," have argued that such predictive systems aren't entirely neutral as the development of processes and necessary training and input datasets are reliant on human discretion allowing pre-existing biases to trickle down into these processes.

Sommerer (2022) added that as predictive policing systems within Germany are reliant heavily on data, as the processes are trained using specific datasets, if those datasets lack accuracy, then the system will be imbalanced. Such imbalance will replicate, reproduce and in many cases cumulate, any incomplete or inaccurate results produced by the system. For certain groups, namely racial minorities, this can have an adverse effect, making them heavily vulnerable. If certain communities are already prejudiced against by domestic law enforcement, and with a higher number of crimes within that community, then more crime data will be documented and used as input for the systems. This would ultimately enhance discriminatory effects through negative feedback loops, even when protected characteristics are replaced by proxies that circumstantially correlates to characteristics like religious or ethnic background. These effects can be further exacerbated due to greater police interference or over-policing that increases the chances of false positives. Sprenger & Brodowski (2023) added that the risks associated with discriminatory results due to bias programming and datasets can be traced back to an overreliance on the central belief that the information used is accurate, with very little room for improvement or criticism due to a general lack of transparency in design, development and implementation of such systems.

Although there has been extensive research in the field of predictive policing, most are in more established regions namely UK and USA, where either the regulations have also been catered to the application of predictive policing or such systems have been abolished in certain states or provinces. To add to that, research within the realm of person-based predictive policing practices is more difficult than location-based due to the inherently opaque nature of decision-making. Furthermore, a core gap identified through the literature review has been that most research within this field uses AI and ML as the focal issue, with predictive policing practices forming parts of it rather than the crux of the research, which it is in this research. Also, the previous research delineates the effects of such practices on different aspects of human security implicitly and partially, with strong focus provided on either individual or community security, a pitfall this research avoids by explicitly adopting a human security theoretical approach. Finally, as most research is highly critical of such practices, understanding and improving predictive policing practices is relatively nascent in Europe, and with the probability of the EU AI Act standardising, improving and in some cases, abolishing such practices, comprehending improvement strategies from industry research pioneers and experts is vital for future research.

Benefits of Predictive Policing: As mentioned earlier, most research is critical when it comes to predictive policing, however understanding the benefits of such practices could help create a case for improvement rather than abolishment. Some researches have been identified where the benefits formed the core focus. Vepřek et al. (2020) mentioned how given the constraints domestic law enforcement must work with in terms of resources, predictive policing mechanisms allow the analysis of greater amounts of data within considerably less time to allow the police to be more instantaneous in their reactions. Santos (2013) added that predictive policing practices can potentially allow domestic law enforcement to allocate resources more effectively through data-driven decision making. This is achieved by enabling targeted operations to be carried out in an optimised manner through efficient resource allocation, and deployment. Finally, by automating analytical tasks that could have otherwise required additional human resources, forecasting, and modelling potential criminals, victims and crime locations become faster and essentially data-oriented only as human biases are prevented. A case for the improvement of such practices has been made in the empirical analysis to exhibit why the research aims at understanding core issues and bridging the relevant solutions through a theoretical framework to improve predictive systems.

Human Security Approach towards analysis: Bennett Moses & Chan (2018) applied a human security approach implicitly and partially to predictive policing systems, not contained within Europe, to understand how biases and issues in these systems manifest. One of the core reasons identified was the assumption that crime datasets and data categories employed accurately reflected reality without factoring in lack of police reports from minorities due to pre-existing biases against them. By using tainted datasets as training and core input for these systems, predictions reinforce existing stereotypes against specific neighbourhoods or communities, thus increasing police attention towards those areas or groups. However, no attempts were made to understand how these predictive systems can influence, and shape social structures, which forms a central part of this research through application of community security aspect, among other aspects, of human security. Ferguson (2015) relays the same sentiments about crime reports as people of colour and poor people have more disproportionate contact with domestic law enforcement. Thus, if race and socio-economic class are variables that determine individual or group risk factors, predictive policing tools can reinforce or accumulate biases. Discrimination and increased surveillance were identified as common themes that affect human security issues in terms of community and political security respectively. To add to that, stakeholder assumptions that the relevant set of historical data or research that critique state-level analysis are irrelevant and doesn't need to be factored into the system, creates a lack of room for improvement (Strikwerda, 2021). States prevent collaboration as well as the optimisation of datasets and practices by inducing a lack of transparency in how these systems are developed and function as criticism is avoided. As aforementioned, several research have attempted to use the human security approach however applying it to AI and ML specifically instead of a particular practice within that realm i.e., predictive policing. Moreover, from researches that have been reviewed, there are no researches that have attempted to understand the issues and expert solutions pertaining to predictive policing systems using a human security approach coupled with anti-bias methodologies, as most have an implicit application of human security rather than an explicit one, which limited the theoretical approach unlike this research. No research was found where an in-depth analysis of predictive policing practices was conducted through the integration of a human security approach, as most either employed that implicitly or used predictive policing as a periphery of AI as exhibited above. Finally, as the EU AI Act is intrinsically based on human security by providing greater importance to the individual and communities within a state than the state itself, a normative framework that can further enhance the human security approach to identify, understand and solve issues within predictive policing thus exists that was unavailable in EU before.

Based on all the literature that have been reviewed, the core gap identified has been the application of human security approach towards understanding the issues within predictive policing systems. This research therefore adopts a human security approach coupled with algorithmic bias theory to expand and improve on such previous researches by providing a comprehensive look into how predictive systems work, identifying and explaining the issues within them that are related to human security, and how those can be resolved using anti-bias and socio-technical methodologies. By centralising human security approach in the analysis of predictive policing practices, this research covers a core limitation of previous attempts within this field, leading to the central research question.

2.3 Research Question

Integrating the conceptual background and after investigating the published literature extensively with the gaps in literature identified above, allowed the creation of the core research question:
"What are the problems and reasons for those in predictive policing practices within EU, and how can those be tackled?"

Relevant and adequate data will be collected and analysed during the research to answer the core question which has been used to construct the interview questionnaire for the interviewees. The core question has two aspects: first on identifying and understanding the problems through the theoretical framework, the latter part regarding tackling the problem would be by applying counter-bias frameworks to identify and understand solutions for those problems in the discussions section.

Theoretical Framework

The theoretical framework has been developed through a multi-layered approach to answer the two segments of the question. The first segment focuses on the problems with predictive policing systems, where the collated information from journals as well as interviews will be analysed using human security theory complemented by algorithmic bias theory as both are relevant to understand the continued, accumulated and reiterative adverse effects of predictive policing systems on different aspects of human security. The second segment of the research question focuses on the solutions and understanding those through a theoretical framework, which is done by analysing the collated information from interviews and journals through a combination of algorithmic fairness theory and socio-technical systems approach. Both sections are interconnected as the former identifies and explains the issues theoretically whereas the subsequent counter-bias framework explains solutions suggested theoretically that can be reduce the issues by integrating algorithmic fairness and socio-technical systems approach in the improvement and development of predictive policing systems. The following section explains the definitions of core concepts and keywords that have been used within the theoretical framework and beyond.

3.1 Definitions

For the theoretical framework specific definitions of key terms were used that have been delineated below:

AI (Artificial Intelligence)- Russell & Norvig (2021) defined AI as intelligent instruments or processes that are capable of receiving different variables within a social environment before performing actions upon the attained information. The AI Act defined AI systems as systems designed by humans for specific purposes that are developed using methods like machine learning, logic and statistical methods etc., which also allow these systems to interact with one another. (Shi, 2023)

ML (Machine Learning)- Russell & Norvig (2021) further identified ML as a subset of AI where a processor observes datasets to develop a model based on those, which then uses pattern recognition and analysis to build a hypothesis pertaining to the world that can be used to resolve certain problems, in this case, within the society. For predictive policing particularly, Jordan & Mitchell (2015) described it as the AI methodology used to build computers that automatically improve themselves through experience instead of explicit programming, machine learning is the intersection between Computer Science, and Statistics to help make informed decisions under uncertainty, which may have been unreadable otherwise. ML accomplishes tasks while learning and improving from the experience. Ray (2019) Predictive policing practices encompasses mainly the task-oriented and the cognitive simulation features of machine learning that focuses on developing systems that improve the performances across a set of predetermined tasks by investigating and simulating human learning processes. (Carbonell et al., 1983) In the area of Data Mining, which is relevant within predictive policing practices, ML becomes key towards utilising historical processes to improve decision making while stimulating processes to be adaptive for the users. ML algorithm types that are crucial within predictive policing are Supervised, Unsupervised and Reinforcement Learning.

Big Data- Ongsulee et al. (2018) identified Big Data as predictive and user behaviour analytics to extract value from data, structured, semi-structured or even unstructured datasets. As predictive policing processes involve high-volume, diverse information assets that require enhanced decision-making and process automation methods, business intelligence and analytics are used to process copious amounts of data.

Algorithms- Defined by Finn (2017) algorithms represent a set of rules defining a sequence of instructions to perform a specific task such as solving a problem.

Bias- Kaufmann et al. (2019) defined biases in the context of predictive policing systems as algorithmic outputs that would consistently put certain subgroups that are defined by race, gender, and other social categories, at a disadvantage.

Predictive Algorithms- Gandomi & Haider (2015) defined predictive algorithms as a form of analytical tool that employs a diverse array of techniques to predict future outcomes by relying upon historical and current data. This definition is further expanded by Nowotny (2021) as an extension of AI and ML technology to predict the future of different social constructs.

3.2 Framework

There are several theories that could be applied in gauging the presence of biases within predictive policing systems and how such biases could be detrimental towards the retention and attainment of human security within a state. For the benefit of the research human security theory coupled with algorithmic bias theory have been selected to analyse the journals and interviews. This section is divided into two subsections that goes on to explain both theories and how those are relevant to the analysis of predictive policing practices. Issues in predictive policing systems and their effects on human security are explained through human security theory where some of the core aspects would be utliised. Subsequently, algorithmic bias theory would be integrated to the former to comprehend how the adverse effects identified are propagated and accumulated across both individual and community security.

3.2.1 Human Security theory:

Human security theory moves away from archaic state-oriented National Security paradigms by deepening and stretching its definition. The United Nations Report (1994) described Human Security as a universal concern that consists of several interdependent aspects that are first and foremost people-centric and is heavily focused on prevention. This concept has been expanded to 7 components namely: Economic, Food, Health, Environmental, Personal, Community, and Political Security that are correlated to one another. The Canadian and Japanese interpretations of Human Security offer a narrower and broader conception with focus on freedom from fear or the survival of one's dignity and pride, respectively. The Japanese definition adheres closely to the traditional definition but also attempts to enforce physical, economic, health and political protection. They do so by addressing structural issues that may be detrimental to human security through policies and multilateral discussions with other countries. (Bosold & Werthes, 2005)

For Europe specifically, Albrecht et al. (2004) developed the Human Security Doctrine that outlined the concept of Human Security as defined by the European Union. The report insisted that the European Union (EU) has started creating security policies that are built on human security and not state security with a strong focus on individual freedom from basic insecurities namely the right to economy, health, community, personal and political security. This approach also highlighted the protection of every individual and not only those of the Union's borders. One of the core principles introduced within this doctrine would be the primacy and promotion of human rights above state interests. This approach categorises human rights as the primacy, so protection of civilians from any forms of harm becomes more imperative than defeating the enemy. This entails the protection of minority communities from potential discrimination at the hands of state authorities, which is probable in predictive policing practices. This goal can be achieved through the development of state authority that can uphold all aspects of human security, through collaboration with internal and external stakeholders. Furthermore, by employing a bottom-up approach where strategies and processes are designed for the community and not the state, which can be done by collaboratively working with international organisations, a multilateral method of execution can be adopted. Multilateralism allows policy and strategies to trickle down effectively to state organisations as a standardised and cooperative method of execution is employed to mechanisms such as predictive policing.

To analyse data from interviews and published journals and articles on predictive policing, the approach employed was the one that uses the 1994 UN definition of Human Security, namely the 7 aforementioned dimensions as a framework where AI and ML processes are investigated with focus on how these mechanisms can adversely impact or cause risk to a certain aspect of human security. For the purpose of this research, the following aspects would be focused upon to identify how biases within predictive policing algorithms could adversely affect human security, namely through economy, community and personal aspects. Individual state governments, namely the German and Dutch law enforcement agencies, were at the core of this research. With its intersectionality, holistic and people-centric approach towards prevention of risks, this framework can be employed to identify motives as well as lacking in predictive policing processes in Europe. Within the realm of human security, primarily for this research, community security is key to understanding how different racially or economically marginalised communities can be adversely affected by predictive policing systems.

Geller et al. (2017) believes that preexisting discrimination coupled with technological policing processes result in creating a bias towards racialized individuals from economically deprived communities. Supporters of predictive policing believe that it represents a cost-effective solution for controlling crime reliant on scientific data, however even when race isn't explicitly utilised as a variable within prediction algorithms, these processes still discriminately target racialized communities. Thus, it can be argued that the greater the discrimination lies within the laws, history and legal structures within a state, the higher the chances are that the bias would be reiterated and reinforced through predictive policing techniques. Delgado & Stefancic (2023) used "social construction" to explain that communities and racial minorities are a product of social thought and relations rather than being inherently fixed or objective. However, reinforced biases against specific communities can lead to different types of over-policing namely: *banishment* from one's homes, *containment* by restricting people from certain places, or *blight* that neglect certain communities to maintain class hierarchies. (DaViera et al., 2023)

DaViera et al. (2023) identified the types of state violence that is common when such reinforced-bias oriented policing lead to namely: "*banishment* (i.e., removal from communities and homes), *containment* (restricting people from certain places), *blight* (neglecting communities to maintain race and class hierarchies), *extraction* (taking wealth and resources, divesting from communities) and *elimination* (killing and incarcerating people)." By applying the different types of over-policing that can affect specific minority communities based on different aspects of human security, Benjamin (2023) further added some of the mechanisms that are used within predictive policing to exercise discrimination against minority communities namely engineered inequity, which describes biases that are inherent within the design of processes and technologies, and default discrimination, which highlights biases that are the result of indifference towards social and historical contexts when developing technologies. The subsequent section explains the manifestation of these biases within algorithms and processes, and how those are propagated and accumulated, enabling it's integration to human security theory that will be applied to the data collected to understand how these biases create a reinforcing loop of discrimination against minorities and individuals.

3.2.2 Algorithmic Bias theory:

Buolamwini & Gebru (2018) defined algorithmic bias as the systematic but discriminatory practices and processes that are the product of machine learning algorithms programmed with biased data, which results in biased outcomes towards a certain group of individuals. Algorithms are thus developed to ensure that decision-making is reduced to a number under the assumption that a neutral algorithmic approach is employed, however this neutrality is difficult to decipher due to the opaqueness of algorithms (Jackson, 2018). Biases would be less visible if they are deeply embedded within algorithms, which is when the algorithms reflect the very biases they were presumed to ignore. Yetim (2011) attributed such biases within algorithmic designs to the lack of diversity in the group of people responsible for designing the algorithm. When data privacy organisations and legislatures were introduced in America to obstruct organisations from deriving applicant's criminal convictions, employers started utilising race as a proxy for criminal convictions thereby leading to Black applicants receiving fewer callbacks. Agan & Starr (2018) attributed this to the racial gap in felony convictions that immigrants and people of colour are subjected to, which allowed predictive algorithms to extrapolate from singular entities to collective groups i.e., races. Such biases are difficult to identify as the reasonings are correlated to the very biases that are denied. Instead of removing the implicit biases that cloud human judgement, by relying upon data that are clustered with the very same biases, discrimination becomes an embedded trait within the algorithm.

Predictive Policing processes can be categorised as a "feedback loop," as it is reliant upon algorithms that are tainted and reinforce the same assumptions prevalent within and perpetuated by law enforcement data (Purves, 2022). By sending more patrol units to neighbourhoods that have been identified as hotspots for criminal activity, greater number of arrests would be made and fed into the system thereby reiterating the feedback loop. Such loops are a result of the "ratchet effect", which Harcourt (2006) defined as law enforcement resources dedicating resources to investigating, searching and subsequently arresting members from a certain proportionally high-offending group. This distribution of arrest creates a disproportionate representation of that group thus exemplifying the sampling distribution police rely on, where they profile frequent offenders and cluster them into members of a specific group. Feedback loops within a system can led to a "runaway," which describes an overestimation of results due to the data fed leading to disproportionate over policing of certain neighbourhoods or individuals. In the Netherlands, this becomes more evident when some predictive policing systems only target vehicles that have plates from East European countries.

Garzcarek & Steuer (2019) added that the very nature of applied algorithms, pattern recognition alongside clustering and categorisation, when applied to humans will create or reinforce certain prejudices. Statistically, this translates to the construction of pre-existing perceptions about an individual due to experience and information the algorithm gathers from others that are assigned to the same group, in this case ethnicity or race. As the algorithm is designed to assign new individuals based on the characteristics measured within a certain group, predictions on one's future behaviour are directly proportional to the behaviours observed previously within others of the same group. Predictive algorithms have a propensity to generate bias based on measured characteristics leading to an individual's assignment to a group that is correlated but not related causally to the variables they are judged upon. These biases are generated irrespective of the input database being an accurate representation for the larger population based on measured characteristics. In practice, this bias is similar to the prejudice humans construct through their own experiences and arbitrary assessment of similarities within certain communities where data is analysed using a certain method that is only suitable for correlations, to create judgments that require individualistic causal reasoning. Within predictive systems, this bias can be compounded and becomes a systematic bias towards certain groups, which is further bolstered due to the algorithm using protected characteristics as a decisive variable.

The FRA report (2022) further added that data quality could impact feedback loops as an analysis exhibited that varying rates of reporting by victims could lead to creation of biased feedback loops. By discounting low reporting rates, these systems are unable to perceive ‘true crime rate’ leading to incorrect predictions and policy decisions. The report also provided evidence suggesting low levels of reporting due to discrimination on the victims’ personal and socio-economic aspects, which is especially applicable for ethnic minority and immigrant groups. Furthermore, by putting excessive weight on training data, machine learning algorithms have a propensity to develop runaway feedback loop more swiftly as this prevents the control of how predictions overreact to small signals within data.

Bias in algorithms normally occurs due to the inclusion of proxies, neutral pieces of information that are strongly related to a protected characteristic. Reliance on protected characteristics leads to discrimination even when coded parameters, training data and input variables are used. These characteristics are usually easy to spot when integrated and help assist the assessment of the predictive system based on those characteristics. And although it may be simpler to prevent discrimination based on explicit protected characteristics (e.g., ethnicity), it is more difficult to prevent those connected to proxies (such as names as a proxy for ethnicity) as they are diverse and almost limitless. In predictive policing, proxies could lead to certain neighbourhoods being targeted more for policing activities due to the larger population of minorities.

The same report has also indicated that algorithmic processes can go beyond discriminating on pre-existing prejudice and discriminatory e.g., by combining demographic and behavioural variables for profiling (FRA, 2022). As these parameters serve as proxies for already protected characteristics, any new combinations are already factored within European non-discrimination law. However, if the algorithms discriminate against individuals based on a collection of parameters that have yet to be identified or that aren't correlated to existing protected characteristics then it would lead to the emergence of new disadvantaged groups or individuals such as online gamers who play excessive first-person shooter games. In these cases, prevention of discrimination based on legally protected characteristics might be insufficient, which is further compounded by the lack of transparency of AI and algorithmic systems.

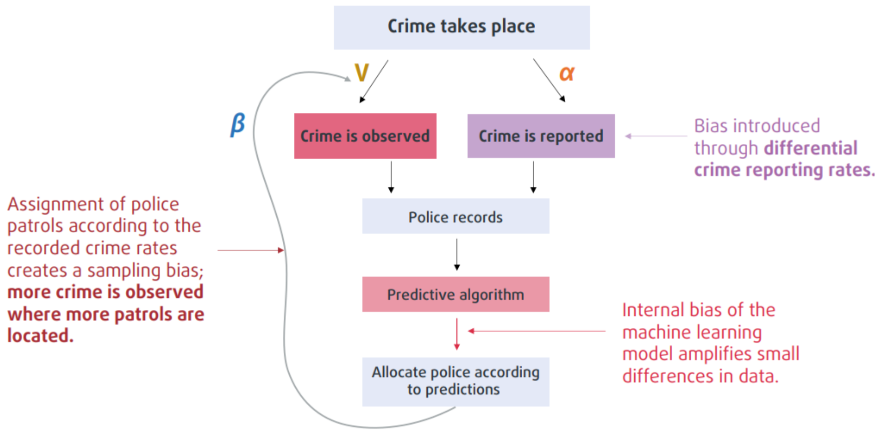

Fig 2: Source of bias in Simplified Policing Algorithms (FRA, 2022)

The FRA Report (2022) outlined Predictive Policing Algorithms alongside the biases that inhibit these systems in the above diagram. Crime Reporting rate (α), police distribution (β), observability of crime (V) i.e., the likelihood of detecting a certain crime and the assumed 'true' crime distribution (Ω) have been represented above. While α and V parameters take into account both all recorded reported and detected crimes, Ω accounts for assumed and nonrecorded crime. The aforementioned parameters remain constant throughout while is initially determined, and the only value that changes over time to identify feedback loop formation. The report concluded that although initial distribution of police force wasn't relevant in feedback loop formation, the internal bias of the algorithm models accelerates the formation even when no real differences could be amplified through the use of random variation. These variations create fictitious differences, continuously amplifying them. Erroneous allocation and surveillance can occur when crime observability or crime reporting varies significantly across neighbourhoods. True crime distribution though is close to uniform, but the potential misallocation of law enforcement resources can adversely affect that. This misallocation could occur due to excessive police being allocated to a neighbourhood with highest crime rate causing runaway feedback loops, or higher concentration of more observable crime within a low crime rate neighbourhood where a feedback loop form sending additional police patrols. Algorithmic bias theory would therefore help investigate the permeation of bias within different aspects of algorithms from data to process while human security theory will help understand the detrimental effects of these biases and other predictive system issues on different aspects of human security, not necessarily minority communities only. To understand the statements provided within journals and by interviewees on better collaboration, and a standardised auditing framework among others, the following section on anti-bias frameworks is required to explain how algorithmic biases and human security issues can be countered using a theoretical approach to stimulate a more social-centric technological development, which would require algorithmic fairness theory coupled with socio-technical systems approach theory.

3.3 Counter bias Frameworks

In the empirical analysis, counter-bias frameworks serve as instrumental tools to interpret expert perspectives from interviews and literature, aiming to enhance the impartiality of predictive policing processes in Europe. The discussion sub-section emphasizes this framework as a conduit for understanding expert solutions theoretically. Hanna et al. (2020) posit that rather than ignoring race, it should be "denaturalised" in predictive policing methodologies, viewing it as a multifaceted and relational social construct. This approach, supported by Benthall & Haynes (2019) and Sen & Wasow (2016), entails manipulating associated variables and treating race as a 'bundle of sticks'—a composite of elements, some modifiable. The overarching shift, as Hanna et al. (2020) underline, should be from examining race's effects to exploring the impact of racism. Adopting the system creation approach by Ford & Haware (2010) to predictive policing can effectively stratify race, correlating it with other system parameters. This chapter delves deeper into these counter-bias methodologies.

3.3.1 Algorithmic Fairness Theory:

Hanna et al. (2020) devised the methodology of employing Algorithmic Fairness Theory when examining Predictive Processes. This approach is viable as predictive procedures that extensively focus on social constructs tend to mitigate bias assume a simplistic conceptualisation of race as a singular variable instead of an ingrained concept that affects several variables. Thus, simplifying the social, economic and political complexity of racial categories. To add to that, any process that is built with race as a concept gravitate towards treating groups as interchangeable thereby ignoring the unique biases and oppressions that different groups could be facing. This approach can thereby severely limit the efficacy of algorithmic fairness as well as the ability to analyse or intervene thus reaffirming pre-existing racial biases. Furthermore, due to lack of research, the methodology that is used to conceptualise or operationalise race as a technique to ensure algorithmic fairness has received fairly low attention. Additionally, by discounting and providing a lack of focus into racial categories when those are adopted into processes, the algorithms developed require proper fairness assessment. Focusing on categories and measurement processes should be important and understanding that a disaggregate analysis alongside incorporating a greater nuanced understanding of measurement would allow more efficient analyses and formation of predictive policing procedures. For predictive policing the following ways were determined as methodologies to create a framework for predictive policing algorithms where biases are accurately accounted for:

- **Group Fairness in Algorithmic Bias:** Hanna et al. (2020) emphasises equalising statistics across dataset subgroups for fairness. One approach is ensuring equal positive predictions and similar rates of false positives and false negatives across groups. Kleinberg et al. (2016) suggests calibrations that are unbiased by protected attributes. Racial categories can be mathematically translated and integrated into counter-bias systems. However, the methodology is criticised for oversimplifying racial issues and not adequately representing unique oppressions of minority groups. A more inclusive approach acknowledges individual minority experiences.
- **Conceptualising and Operationalising Race:** Racial bias in algorithms can be tackled by critically understanding and defining race, as noted by Andrus & Gilbert (2019). Important considerations include re-evaluating historically prejudiced census categories, understanding race's evolving definition, and ensuring transparency in racial categorisations. Assessments should determine if race affects algorithm

outcomes. Collecting diverse racial metrics allows datasets to be better scrutinised and enhanced by research communities.

- **Disaggregated Analysis:** Effective algorithm audits need group definitions that consider cultural and demographic differences. This approach contrasts group fairness and focuses on essential racial aspects relevant to predictive policing. Data collection methods, racial categorisation, and diverse evaluation techniques help detect social inequalities in AI systems. Howell & Emerson (2017) and Penner & Saperstein (2015) offer distinct measures to understand racial inequities, but combining them provides deeper insight into racial bias in sociotechnical structures.
- **Centring Marginalised Groups' Perspectives:** Counter-bias frameworks should consider the perspectives of racially oppressed groups. As Anderson (2009) suggests, understanding injustices from the viewpoint of affected populations can help. Including marginalised voices in the early stages of system design fosters a participatory approach, ensuring their integral role in predictive policing system development.

Although algorithmic fairness theory will assist in exhibiting the recommendations provided by interviewees and journals on how algorithmic biases could be solved in predictive policing systems, the necessity to integrate residents and minorities as focal aspects of design and development alongside research organisations, and educational institutes, which can be found in abundance in both published journals and interview statements, can be explained by combining a socio-technical design methodology.

3.3.2 Socio-Technical Systems (STS) Approach:

In the area of big data analytics, there have been extensive studies that have applied information system theories such as algorithmic fairness theory, but theories that combine social and technical elements through an information systems theory approach have yet to be explored. (Gupta & George, 2016) To add to that, Günther et al. (2017) advised integrating both human and technical features when developing and executing information systems. As these features continue to develop due to the augmentation resulting from the performance of social and technological sectors and big data permeates into different aspects of society, shifting how decision-making processes function, social spheres become more affected due to real-time decision-making processes such as predictive policing processes.

Trist (1981) and Clegg (2000) developed a socio-technical framework to identify and manage human factors in technical environments, as well as a methodology to redesign work processes within an organisation that would stimulate more effective combination of technological and human resources, which in this case would be human security. Although some have described STS as an amalgamation of multiple interacting parts in a system namely the social and technical subsystem, others believe that STS is more concerned as a group of principles that shows how the organisation could attain harmony between technical and social subsystems. Carayon (2006) segregated components within an organisation into three components, as per the STS approach:

- The social subsystem in the case of predictive policing would comprise of the perceptions, beliefs and relationships, both within and between groups in this case namely minorities and domestic law enforcement agencies.
- The technical subsystem encompasses the tools, knowledge and techniques namely the algorithm, data and practices involved in the development and execution of predictive policing processes.
- Finally, the environmental subsystem balances the social and technical subsystems alongside diverse stakeholders, in this case, government, regulations, domestic law enforcement, and diverse race groups.

Carayon (2006) and Trist (1981) noted that while social and technical subsystems hail from human and natural sciences respectively, they work in tandem, transforming input into output. They possess mutual causality—being independent yet interdependent. The STS approach, with principles applicable to socio-technical processes, has seen various interpretations, like Cherns' (1976) nine principles for joint optimisation and Clegg's (2000) focus on design. However, only those relevant to this research will be elaborated, particularly Chern's STS principles tied to predictive policing, which are crucial for addressing biases.

- Compatibility- Cherns (1987) emphasized that system goals must align with its design process, involving affected parties and users. For predictive policing, surveys can gather diverse group inputs.
- Variance Control- Cherns (1987) underscored controlling unexpected deviations in standard processes. In predictive policing, this means rigorous quality control for process, data, and outcomes.
- Information Flow- Cherns (1987) advocated for stakeholders, including citizens, to monitor shared information and access feedback. This ensures informed stakeholders and an evolving process based on feedback.
- Power and Authority- According to Cherns (1987), stakeholder groups should control resources in line with their responsibilities. In predictive policing, minorities should influence data and processes, while government agencies remain accountable.
- Support Congruence- Both Closs et al. (2008) and Cherns (1987) emphasized systems supporting and reinforcing social structures. For predictive policing, this means training law enforcement against racial bias and maintaining consistent data processes.
- Incompletion- Cherns (1987) valued a system's continual improvement post-implementation. Law enforcement must recognize and address flaws in arrest data before integration.

In comparison to the aforementioned principles, the ones Clegg (2000) provides are a systematic view of design considerations alongside content-principles of the information system and the process-principles attached to the design of the process. The above factors are required to develop an efficient and less-biased predictive policing mechanism; however, the following represent the principles used to develop said-system:

- Design is Systematic- Clegg (2000) argued that the interplay between social and technical subsystems becomes evident during implementation. For predictive policing, citizens' security is paramount, and technology serves to enhance this security. Instead of a strict command-and-control method, predictive policing should be viewed as a system that both depends on and shapes societal constructs, namely human security.
- Design is an Extended Social Process- Predictive policing isn't just a one-off process but continues evolving post-implementation. While designers and officials play a role, citizens' perspectives shape the system. Feedback mechanisms, like surveys, and legal frameworks can influence its design. The allocation of tasks in such systems should be fluid, taking into account feasibility studies.
- System Components Should be Congruent- Predictive policing designs must prioritise human security. Clegg (2000) stressed the need for unbiased performance assessment and clear information dissemination. The importance of system transparency is underscored by Simmler et al. (2023) and Oosterloo et al. (2018) highlighting the challenges with location-oriented systems like PreCobs and CAS. Knoechelmann (2022) also noted the opacity in person-based forecasting systems like RADAR-iTE's. The FRA Report (2022) emphasised the benefits of transparency and the potential of explainable AI.

- Problems Should be Controlled at the Source- Issues should be addressed at their root, as highlighted by Clegg (2000). Predictive policing design should embrace STS principles, incorporating both social and technical facets. Regular evaluations should be performed, and multi-disciplinary knowledge is essential for effective system design. Resources, both in terms of expertise and investment, should be directed towards the intertwined social and technical aspects of the system. Moreover, a mechanism that filters perspectives from various stakeholders, while integrating feedback, is essential.

Therefore, by applying STS approach the need to integrate biases that can permeate into minority communities and subsequent necessary bias mitigation strategies developed from algorithmic fairness theory will be encapsulated. Alongside that adopting a socio-technical systems approach will also assist in explaining why educational and research organisations need to collaborate with the state in designing and developing predictive systems in a transparent method and the need for consistently improving such systems through a standardised audit framework, which can be developed in conjunction with the EU AI Act.

3.4 AI Act

The AI Act, developed and presented by the EU Commission, is the first proposal to regulate EU-wide AI systems through standardised rules. The AI Act is not a part of the core theoretical framework but rather a normative framework that is being used to gauge how predictive policing systems within EU can have issues and biases that need to be addressed, and the need to adopt a human security oriented approach in development and implementation of predictive systems within EU, which can be applied by following similar methodologies as the ones applied here.

The council first adopted a common position in defining AI more narrowly as a system that is designed to operate with some level of autonomy based on human and machine level inputs to infer a given set of objectives using either machine learning or logic-based approaches to generate predictions, recommendations, or decisions that influence the environment the system interacts with. (Pingen, 2023) Furthermore, by adopting a risk-based regulatory approach to address risk without hindering improvement or increasing costs, the AI Act assesses systems based on risk level, which exhibits the use and potential impact on the safety and fundamental rights of people. Risk situations are divided into four: minimal, limited, high-risk, and unacceptable. Unacceptable risk scenarios prohibit systems that manipulates people's behaviour subconsciously or exploits specific vulnerable groups through risk-based scoring systems or real-time biometrics, which entails some aspects of predictive policing systems. Some predictive policing practices have been classified as a high-risk scenario, namely the classification of natural people that fosters inequality within the community, which harms fundamental rights. For these scenarios, and oversight and transparency to explain and allow critique of decision-making processes is essential. (Shi, 2023) It must be noted though that currently high-risk AI systems are the primary focus of regulation, with concerns being expressed by research organisations on the extent to which this regulation can be passed and implemented across EU states as predictive policing practices can be diverse and thus categorised as either high risk or unacceptable. The European Commission plans to create a framework that will consultation and sharing of best practices across EU for AI systems. Dr Pingen (2023) mentioned how The Common Position extended banning behaviour manipulation based on age or disability to one's social or economic position, prohibiting social scoring by private as well as state bodies.

Methodology

4.1 Research Design & Strategy

This research adopts a unique approach to analysing predictive policing practices commencing with a comparative case study design by analysing predictive policing approaches within two countries namely Germany and Netherlands using a human security and algorithmic bias lens. This will help analyse common themes in terms of challenges predictive policing systems that can be explained through the theoretical approach, as well as stimulating the application of the theoretical framework to understand the solutions provided. To supplement and go beyond, this research will contain an in-depth analysis and discussion of the results of the case studies supplemented through a theoretical framework and interviews conducted to invite a new perspective of investigating predictive systems and insight prospective future research.

A qualitative case study approach is being used for this research due to the lack of sufficient quantitative data and focus on understanding core issues in predictive policing through analysing journals and interviews. Data collated from published journals and research alongside the interviews that will be conducted with educational industry experts and authors would be connected to the theories outline above in the theoretical framework using a deductive approach to develop conclusions. This approach was identified after being recommended by Yin (2014) as an ideal analytical approach that can be employed to case studies, and then expanded upon by Pearse (2019) for its inherent ability to use pattern matching and thematic analysis to find and analyse specific case studies relevant to the research. A deductive analytical approach is employed as theories and frameworks are applied to observations of predictive policing processes to answer specific questions leading to specific conclusions. Applying human security theories alongside Algorithmic Bias theory would help identify shortcomings within existing predictive policing practices. Subsequently, using Algorithmic Fairness Theory with Socio-Technical Systems approach to counteract the identified criticisms will help in formulating probable framework and methodologies for more unbiased system development and implementation.

The methodology to be applied upon the research is reliant on the information that would be attained, which needs to pertain to the following aspects:

1. Data and analysis of processes used to design and develop predictive policing systems in Netherlands and Germany
    a. Types of location and person-based predictive policing practices and the technology used in association with those.
    b. Mapping Development practices used to create these systems.
2. Biases and flaws within current processes from human security and human rights perspective
    a. Identifying flaws in terms of consistent and reiterative biases and the probably reasonings behind those
    b. Identifying flaws in standardised practices and using theoretical framework for explanation
3. Using anti-bias methodologies to address flaws within current predictive processes.

A multi-method research strategy is therefore applied through combining the case study research, literature review and expert interviews. To gather the above information, the following methods will also be applied:

- **Qualitative study of issues pertaining to human security within person-based and location-based predictive policing practices in Netherlands and Germany** regarding the data and processes used in design, development and implementation. This would focus on biases prevalent within the system, the stakeholders currently operating these processes as well as any further drawbacks of this system such as transparency, lack of auditing bodies and standardisation. Relevant literature, namely published journals, would be studied and accumulated for this section with further information obtained from interviews with researchers and experts within the field of predictive policing.

- **Analytical study of biases and structural issues** that permeate as part of predictive policing practices where the focus is on state security and not on human security. This will be conducted primarily through secondary research of relevant literature on state developed predictive practices and how those can essentially function as insulated, "black box" feedback loops. This area would be stimulated further through interviews with representatives from human rights and research organisations namely Amnesty International thus allowing identification of issues within current systems as well as formulation of counter-bias strategies to help create systems that are more accurate and unbiased.

The interviews would be used to analyse the perceptions of relevant stakeholders and experts within the field of research in predictive policing practices. The scope of the primary data for the research would be gathered from personal interviews with the following stakeholders:

- Researchers within the field of algorithmic bias in predictive mechanisms
- Experts in the field of digital humanities and socio-technical infrastructures
- Researchers in the field of human rights within Europe

The interview questionnaires would be developed reliant on the information that has already been derived from the journals, books and research of each researcher and aligning those to the extensive literary research conducted as part of state of art. These questions would focus on:

- The current condition of human security and human rights as a result of predictive policing practices in Netherlands and Germany and what factors stimulate bias against minorities.
- The steps that could be taken to ensure that these processes are designed and implemented to cater towards human security, and why this shift is required.
- The stakeholders that need to be involved to what extent within the development and employment of such predictive systems to uphold human security.
- Any roadblocks currently that could prevent the optimisation of predictive processes or collaboration across different stakeholders and the steps that can be taken to prevent these circumstances.

The interview guides were structured for each participant separately and have been provided in the appendices, semi-formal interviews were conducted with the reformulation of some questions during the interview relying upon the replies of the participants, to ensure the interview would steer towards finding the relevant answers. By developing the research strategy, the cases that were finally chosen for analysis were determined and the considerations to do so have been outlined in the following section.

4.2 Case Selection

Given the broad applications of predictive policing globally, a holistic study would require investigating a diverse group of practices, which weren't possible due to limited time frame and number of interviewees, who are experts in certain types of predictive policing practices only. Thus, the case selection of this research has been delineated below:

- Predictive policing practices in Germany and Netherlands are the core considerations of this research due to adopting both location and person-based predictive policing, the level of research already conducted in this field, and the interviewees available amongst the contacted ones. Netherlands being the first country in the world to deploy predictive policing on a national scale, was an ideal choice as it allowed the research to gauge how these practices are ingrained and thus can penetrate and shape different aspects of society. Alternatively, choosing German predictive practices was ideal due the current state of flux of these practices due to numerous instances of German government deprioritising certain practices for being unconstitutional, allowing the research to examine systems that were more critical. The overarching constant of the EU AI Act being standardised and implemented across both also played a role in the case selection. Furthermore, the possible implementation of the EU AI Act as compared to more developed regulations in the UK and USA, provides this research with the unique opportunity to investigate possible causes, roadblocks as well as stimulants that can assist in implementing a standardised regulation within EU.

- Although there are several types of predictive policing practices that exist, the research focuses on person-based forecasting and location-based predictive policing, as in both cases the effect it has on a select individual or minority community can be gauged using the different aspects of human security. For person-based forecasting, the research looks at VERA-2R in the Netherlands and RADAR-iTE in Germany. For location-based forecasting, CAS in the Netherlands and PreCobs in Germany are practices that were investigated. These predictive policing systems were specifically chosen due to the similarity in the design and implementation of these systems, as well as how all four have significant critique within research but have a lack of human security approach towards the analysis. In Germany, there are at least 5 different geospatial systems used for location-based prediction policing, however PreCobs was specifically chosen due to the longevity of its usage, but also due to the variance to which it used across to region as some areas have chosen to discontinue its usage. CAS was chosen for the Netherlands as it is the oldest and has been deployed on a national scale. VERA-2R was picked for the Netherlands due to the expansive way it is used within and outside Europe by judicial professionals, and how ingrained it has been in the Netherlands since 2016. Alternatively, RADAR-iTE has been in used within Germany since 2017, while being continuously improved factoring in ethical and legal aspects, thus this was identified as a suitable case due to the number of iterations of improvement it has gone through.

- These practices were also chosen due to the implicit application of human security theory that were identified in published researches and journals thus making it more convenient to apply the theoretical framework explicitly.
- Predictive policing systems have existed in the UK since as early as 2013, and relevant literature has been encountered several times during the literature review. However, to ensure the research has a uniform approach considering Brexit, and the EU AI Act, which is still in

development and will use a normative framework, predictive policing systems within the above EU countries were chosen to be analysed only.

- The law enforcement practices and regulations are varied between different countries, namely Netherlands and Germany, and might very well have affected the development and deployment of predictive policing systems in those countries. However, integrating different legal systems and laws would have extended the project beyond the socio-technical sphere it currently lies within. The research will thus abstain from conducting any investigation into legal frameworks of these countries instead the use cases will specifically focus on predictive policing systems within the aforementioned countries. A general view of European standards has been included in defining GDPR and the planned EU AI Act.

- Although the planned EU AI Act plays a role in the motivation for this research, it must be stressed that as the Act has not been passed yet, the research will not go into great details regarding it, only employing a peripheral understanding of what this may entail for predictive policing within the EU.

- Finally, the cases selected were chosen as there were some published research on the positive impact of these systems within the regions they were implemented, as well as how integrated they were to domestic law enforcement practices, which can create a strong case for improving the systems instead of abolishing them.

Once the cases were selected, the interviewees required for primary research were determined based on their expertise in the field as well as the region of their research, the comprehensive list has been provided in the following section.

4.3 Participants

The interviewees were selected based on their academic and research background in predictive policing, as well as their contribution in the field of human rights within predictive policing systems in Netherlands and Germany respectively. The interviewees for this research and their background which correlates to them being chosen to have been given below:

- Junior Prof. Dr. Lucia Sommerer (interview 1): Author of the book "Self-Imposed Algorithmic Thoughtlessness and the Automation of Crime Control: A Study of Person-Based Predictive Policing and the Algorithmic Turn," who specialises on automated crime control & algorithmic biases within Germany. Being a pioneer in the field of the intersection between criminology and emerging technology, her PhD thesis on the usage of algorithms to calculate criminal risks was awarded the Körber Prize as well as the Young Scholar Award of the German Criminological Society, which was later developed through extensive research into the aforementioned book.

- Dr Lorella Viola: Author of The Humanities in the Digital (interview 2): Beyond Critical Digital Humanities and Research Associate at the Centre for Contemporary and Digital History (C²DH), a pillar in developing Digital Humanities methodologies that quantitatively and qualitatively analyse social patterns integrated within technological processes primarily, whose research is primarily located within Netherlands.

- Dr Gwen van Eijk (interview 3): Policy Officer Technology and Human Rights at Amnesty International Netherlands, and former Assistant Professor of Criminology at Erasmus University Rotterdam, who has extensive research experience namely in Discriminatory risk profiling for minorities using race, primarily within the Netherlands.

Although the interviewees selected were adept for the purpose of this research, it must be stressed that there were certain limitations that needed to be factored in during the research that have been identified in the following section.

4.4 Limitations

Although a qualitative approach has been employed for the purpose of this study, time constraints prevented more comprehensive research that could have been stimulated further through integrating quantitative research as well. Investigating the number of false or overturned convictions due to initial results obtained from predictive policing practices could have been applicable, if that information was readily available. Furthermore, due to the nature of research, and the criticism directed towards such state-controlled practices by research organisations, a core limitation of this research was finding legal or state representatives who were responsible for creating and maintaining such predictive systems within the Netherlands and Germany respectively. To add to that, although the research is aimed at looking into EU countries, the number of countries were limited to two, namely Germany and Netherlands due to lack of time and budget constraints. Other countries such as Denmark and Belgium, could be investigated in future research papers. Also, as the EU AI Act is being applied as a normative framework, other developed regulations namely GDPR aren't utilised within this research as the effects of such practices on those regulations have been widely investigated as per the literature review and most predictive policing practices are public and developed in accordance with those regulations. Also, as most of the published journals adopted an extensively critical approach towards predictive policing systems within EU, it was highly difficult to find and extensively analyse published research that highlights the positive aspects of implementation. As mentioned beforehand, recommendations are beyond the scope of this research and the counter bias theoretical framework is developed to analyse why and how researchers and the interviewees advocate for standardised auditing, transparency, and a collaborative framework to improve predictive systems. It must also be stressed that research in predictive policing is still an emerging field, namely within the regions mentioned, therefore there was a considerable lack of available data, and the qualitative case-focused nature of this research will make transferability difficult when being applied to other regions. Lastly, for the benefit of the research, the interviewees were limited to educational and research topic experts within Netherlands and Germany to help explain the systems, the issues related to these systems within those specific countries and how those issues can be resolved by a combination of EU, educational and state-controlled and collaborative interventions.

Empirical Analysis

For the empirical analysis, the person-based and location-based predictive policing use cases for the Netherlands and Germany are described thoroughly including all the processes within those from inception to data accumulation to the use of data by stakeholders. After clearly outlining the processes within these predictive systems, information collated from journals regarding the systems is filtered through the theoretical framework to understand the problems within the systems as well as develop probable steps that can be utilised for resolving those issues.

5.1 Use cases

Netherlands- Oosterloo et al. (2018) mentioned how the integration of datafication and automation within different aspects of bureaucracy led to development of and adoption of Intelligence-Led Policing by the Dutch Police, which was stimulated by moving away from local teams to the formation of a national police force. This resulted in the development of the *Central National Crime Database* which led to the creation of the **Crime Anticipation System** (CAS). A data-driven mechanism to proactively prevent crimes through statistics, CAS relies heavily on the following data sources: BVI (*Central Crime Database*), GBA (*Municipal Administration*) and CBS (*Demographics from Statistics Netherlands*). Although CAS has access to and utilises various kinds of data, it can only use data back up till three years. Willems & Doeleman (2014) classified this into three types of information. The first type is socio-economic from the Central *Bureau of Statistics* (CBS), which focuses on the age, and economic position, namely incomes of an individual alongside the social benefits present within an area. The second is from *Basisvoorziening Informatie* (BVI) which consists of previous crimes, locations and known criminals, as collated by the police force. Finally, the third type is derived from the *Municipal Administration* (BAG), which consists of streets and addresses to create the map where predictions are made.

Fig 3: CAS interface picture colour coded for risk factors including red for burglary, green for street robbery, blue for youth disturbance etc. (Oosterloo et al., 2018)

Categorised as a spatiotemporal prediction mechanism, CAS focuses on not only hot spots but also timing within a city, instead of identifying high-risk individuals, thus constructing heat maps as demonstrated above. These maps illustrate locations that are at a higher risk of high-impact crimes, while leaving blocks for those that have lower risks, which can increase over time. (Mali et al., 2017) As CAS specifically identifies the top three percent of high-risk areas to correlate to the police force capacity, improving efficient distribution of resources would be its core function. After being constructed for the Amsterdam Police in 2013, testing was conducted in cities like Hoorn and Hoefkade, among others, and although the results were inconclusive, it was made available across all law enforcement teams in the Netherlands in May 2017. Some of the complaints came from users themselves i.e., police officers regarding the use of the system and the interface. Being deployed in 110 base teams out of 167 as of 2018, CAS plays a major role in determining police patrol locations. Calculating high-risk locations is conducted through directional coefficients that are determined by factors including timelines of crimes, more recent crime trends, as well as number of incidents over a specific period of time, among others.

To predict locations, CAS aggregates information of a specific zip code through utilising open demographic data before integrating that into the risk metrics. Variables such as the number of inhabitants within a specific household as well as an area, the composition within a household, property values, empty properties, number of people who receive income within a household, anyone who receives social benefits and total income within an area, among others. The system is therefore more suited to urban areas as rural areas might not possess statistically valid numbers of people and households, or diverse economic factors that are accounted for as part of the process.

In the Netherlands, a combination of VERA-2R and IR46 are used as well for person-based predictive policing. The country report by Orana and Perteshi (2022) identified the **Violent Extremism Risk Assessment 2 Revised** (VERA-2R) as a predictive policing instrument that uses fixed risk indicators categorised into belief, social context, history, commitment and motivation. Additional variables can be added for personalised assessment, but at core it allows domestic law enforcement officials to prioritise and structure relevant information to understand what needs to be investigated or what type of other information needs to be accumulated. 34 indicators are used in this system to categorise an individual as high, medium or low risk, alongside extraneous variables such as criminal history and mental disorder. Individuals aren't categorised on a simply numerical basis but also the weight of each variable and the information available to create reports that can be presented in the court or used to design in-person interventions such as surveillance or incarceration. The same report also described the **Islamic Radicalisation 46** (IR46) tool used by the Dutch Police, which functions as an early warning system to identify individuals within the Islamic community who have a relatively greater willingness to commit violence. To measure the level of radicalisation this system categorises radicalisation into four levels of severity namely preliminary phase, social estrangement, jihadization, and finally, jihad/extremism. This is done by providing 46 variables determining social context and ideology, although similar to VERA-2R police officers can add more information. As a tool for screening radicalisation within a community but given how unreliable and poor current understanding of the determinants of radicalisation are, the Dutch government has been working to improve risk assessment mechanisms for radicalised individuals.

Germany- Developed and defined by IfmPt (2018), a private institute based in Germany, the main goal of **Pre Crime Observation System (PreCobs)**is to reduce burglary numbers by forecasting risk-prone areas at an early stage. These predictions are heavily influenced by a "near repeat" approach where risky areas are identified after burglary occurs, as it correlates past events alongside a series of burglaries within these areas as an indicator of that being a target of future crime. Shapiro (2017) abbreviated this definition by describing it as a geospatial algorithm modelled to generate risk profiles for locations based on spatial proximity to an initial incident. Using a relatively low amount of data consisting primarily of police records of reporter burglaries, PreCobs extracts additional information namely time, object, modus operandi, damage and exact geographical location to generate predictions. Information obtained is compared with pre-identified triggers (e.g., a list of modi operandi) and anti-trigger criteria (e.g., the use of keys) as references to identify or oppose future near repeats respectively. (Simmler et al., 2023) Patrols and extraneous surveillance would then be deployed for areas that were identified risk prone. PreCobs goes beyond real-time analysis by verifying predictions through simulations such as evaluating past predictions in comparison with actual occurrences to authenticate accuracy before reincorporating successfully predicted near repeats consistently into the algorithm. The dataset thus is constantly expanded by relying upon past events ensuring that the system is consistently learning, however the system isn't entirely machine learning oriented as the predictions are reliant on entered 'if-then' decisions.

Vepřek et al. (2020) further stated how development of predictive policing systems within Germany can be either through commercial partnership such as PreCobs or in-house such as NRW, however cooperation with technology companies is vital in both cases beyond development especially in training police employees and data analysis units. To add to that, predictive policing mechanisms could utilise either decentralised or centralised interpretation of probabilities both in terms of development and analysis. For resource allocation, calculated probabilities are either transferred to specialised units to draw their own conclusions, which is the case of PreCobs. Alternatively, similar mechanisms like NorthRhine Westphalia (NRW), a central analysis department, is responsible for generating and interpreting the probabilities before disseminating abstract recommendations to domestic law enforcement units. For Germany, the core function of these systems would be improved information and knowledge management within law enforcement by cooperating and coordinating within traditional preventive service units (Präventionsdienst Stellen). It should be noted though that knowledge transfer between different state actors in Germany remain informal and sporadic due to organisational challenges and lack of standardisation.

Law enforcement in Germany have also implemented the **RADAR-iTE** (rule-based analysis of potentially destructive offenders to assess the acute risk of Islamist terrorism) as part of their proactive efforts to curb crime and in this case specifically terrorism. Developed in cooperation with the Forensic Psychology Department of the University of Konstanz by the Federal Criminal Police Office (BKA), to assess and identify Militant-Salafis through a multi-stage risk assessment approach. A multi-stage assessment grouping individuals into increased, moderate or high-risk categories, allows specific risk measure advice for recognised high-risk individuals to be formulated. Bautze (2018) believed that the system stepped away from religious biases by employing over 70 questions to assess the individual risk of violence based on the information from previous events that may have occurred in that person's life. Some of the questions that may be included would be evidence of Jihad motivated travelling, violent background, any mental illnesses, access to weapons as well as personal, community and social events and activity.

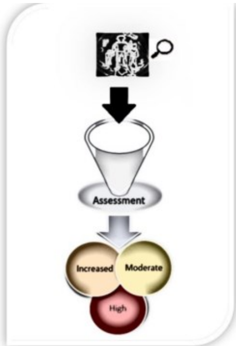

Fig 4: Two-stage assessment (Itälunni, 2018)

Knoechelmann (2022) further described that the process starts with a police officer in-charge of monitoring an individual, accumulating all relevant information regarding that individual. Afterwards the questionnaire is used to score that person and categorise him or her as into the aforementioned buckets of "high-risk" or "moderate risk," which is relayed back to the officer. The following section focuses on some of the benefits of these practices that should be considered when considering improvements and applying the EU AI Act to those.

5.2 Arguments for improvement instead of discontinuation

As the EU AI Act has developed greater inhibitions against predictive policing from the initial drafts, the research had to integrate arguments for improvement of these systems to ensure that the objective of the research stayed on analysis of biases and issues within predictive policing that can help improve the systems. While the efficacy of predictive policing systems can be debated, some statistics are helpful in corroborating the partial effectiveness of such systems. Study conducted by De Graauw (2014) found that in Netherlands, CAS predicted 15% home burglaries accurately alongside 36% with high probability. In the same study, 36% mugging were predicted accurately while 57% were predicted with high certainty. Dutch Law enforcement use ‘flex teams,’ adaptive police resourcing stimulated by predictive policing systems that allows them to effectively allocate resources to predefined areas for policing. (Hardyns & Rummens, 2018) Such adaptive form of policing is correlated to the swiftness with which predictive policing systems are able to identify patterns within large datasets and also enables 5-10% reduction in costs incurred when employing human resources. Law enforcement agencies are thus able to not only act more swiftly but also with higher efficacy in terms of resources, which may be deterred if such practices are abolished.

Furthermore, study conducted by Lum et al. (2011) exhibited that employing concentrated law enforcement with specific police strategies within small spatial has the highest potential to succeed. Predictive policing systems points a police officer in the right direction by either assigning them to high risk areas or including those in that in their usual patrol routes. (Hardyns & Rummens, 2018) By making patrol routes more controlled and concentrated the chances of actively preventing or responding to a crime increases, but alongside that it acts as a deterrent for crimes within those specific areas.

Most of the critique for predictive policing systems is directed towards the use of person-based practices where profiling based on specific patterns of behaviour can enable racial or socio-economic profiling. However, Susser (2021) noted that predictive systems don't necessarily identify individuals as prospective criminals or victims but rather assists the police in making those decisions of whether one ought to be considered a suspect or not. This can stimulate more efficient and swift profiling as it provides law enforcement with the necessary information regarding individuals to make adept decisions. This has a strong probability though to enable biases already present within law enforcement regarding minorities. Hardyns & Rummens (2018) added how such predictive systems allows law enforcement to shift their focus to core tasks alongside empowering them to make more calculated decisions thus making them the centre of the system. When applying human security approach to this set of information, personal security is key as it actively attempts to account for the security of every individual within a community, without any form of bias for or against any. However, being a pattern-recognition system, there are discriminatory issues within the datasets and processes employed that must be addressed. Finally, as part of community security, predictive systems was developed with the emphasis of protecting all those who are within a state as part of a larger community, irrespective of their racial or socioeconomic status. By discounting racial, ethnic or socioeconomic biases, relying entirely on pattern recognition within datasets, these systems aim to apply an unbiased analysis of data. However, it must be stressed that by not accounting for the biased and oppressive behaviour certain communities could face within a region, these systems can actively enable or reinforce those biases through reiterative decision-making couple with biased datasets. The next section analyses the information collected from research articles, journals as well as the interviews conducted by applying the theoretical framework to the collated information.

5.3 Analysis

Each section of the empirical analysis is divided into two-parts, one for location-based predictive policing systems and the other for person-based forecasting. This will help separate the two predictive processes and categorise each case study into their respective type of predictive policing, as well as helping the comprehension of the identify individual problems, while also ensuring that each analysis is specifically catered to the aforementioned cases.

5.3.1 Greater collaboration with non-state actors in education & NGOs is necessary for more accurate datasets and optimised processes:

Itälunni (2022) identified that enabling transparency and accountability within state organisations, namely law enforcement agencies, when constructing and deploying predictive policing mechanisms requires effective collaboration with NGOs and educational organisations. International NGOs such as Amnesty International are operating at a high scale responding to human security threats while also measuring the performances of political mechanisms such as predictive policing processes. As these NGOs not only commit themselves to supporting UN and state agencies but also civil societies, namely minorities, they are vital stakeholders when projects are designed and deployed to uphold human security. Within the EU, there may be issues in creating standardised relationships and collaborations between state organisations and NGO networks as mentioned by Prof Lucia Sommerer (interview 1) in her interview due to the different political and legal affiliations and perceptions each state has. Furthermore, separating an individual from his or her religious affiliations and focusing on their relationship to secularism can only be fostered through the help of humanitarian and educational organisations with the appropriate research to support these claims. Dr Viola (interview 2) added in her interview that as nobody within state bodies have the technical or social expertise to decipher complex algorithms within the space of predictive policing, interdisciplinary collaboration is necessary alongside holistic collaboration with the government and other entities. By tackling the problem from multiple angles, including academia, research and NGOs, holistic resolution of individual and community interests is probable for both location-based & person-based predictive forecasting.

Location-based: Oosterloo et al. (2018) identified that one of the main drawbacks of CAS would be its reliance on location-identification for predicting specific crimes related to a specific place, namely burglary, pickpocketing etc. This isolates the mechanism from focusing on a specific demographic within criminals thus leading to limitations in accuracy. Greater accuracy in predictions requires more nuanced data such as the precise times of an event, which is harder for some crimes such as burglary when compared to others such as mugging. This has a detrimental effect on predicted outcomes as the BVI operates with exact times only. The EFRA Report on Bias in Algorithms (2022) added that data is also influenced by how diverse the detection of different types of crime is as some crimes are easy to detect and record such as car thefts where people have incentive reports to use for insurance claims. Certain demographics are associated with these types of crimes as compared to ones that are harder to detect such as financial crimes and fraud. This causes further biased predictions as algorithms become focused solely on crimes where information is more readily available and recorded by the police, which can be compounded by increased surveillance leading to more observed crimes and thus more biased crime reports.

Person-based: For PreCobs, Gerstner (2018) identified that the limited evaluation period coupled with small trial size areas and the absence of experimental design could impact how effective the system is in combating or reducing crime. Evaluation of the system has shown inconsistent results in how it can reduce burglary within a specific area. Scanlan (2019) added that research showed reduction in crime due to predictive policing systems being conducted by software companies in collaboration with academics, thus adversely impacting the credibility of the research due to shared interests. Furthermore, if potential offenders comprehend the usage of greater patrolling and surveillance of hotspots, they could seek to shift their activities elsewhere thus leading to the displacement of crime in less surveilled areas. Orana and Perteshi (2022) outlined that although there is a strong perception that the VERA-2R is a comprehensive risk assessment tool, it comes with its own fair share of limitations firstly as it is resource-intensive requiring long processing time as well as practitioners to be adept at the use of the instrument. To add to that, as the tool requires cohesive collaboration as a unanimous consensus mechanism, external stakeholders namely within mental health and educational research are required to provide information for it to produce a reliable result. Thus, for VERA-2R to be beneficial to society, standardised and cohesive trust and cooperation within state and external stakeholders is necessary, which doesn't exist within the EU yet. Furthermore, as predictive policing systems rely upon pattern identification where historical crime data is the input to predict high risk crime individuals, the data does have limitations exhibited by decades of criminal research, namely that it doesn't constitute all criminal offences that can be utilised as a representative random sample. Some countries have been known to use victimisation surveys that randomly sample the population to understand their experiences of crime, ranging from burglary to online fraud as well as whether they report these to law enforcement authorities, which allows an estimation of crime that has not been factored into the statistics. Itälunni (2018) further stated that for mechanisms like RADAR-iTE identifying probable Islamic terrorists is difficult due to lack of accurate data as this is reliant on religious and political attitudes that can be hard to track especially when certain terrorist organisations are more decentralised. Individuals who are categorised as high-risk might be immune to adverse effects due to already being investigated; publicised numbers only indicate half of the suspects as low risk while 40% are classified as high-risk. Thus, two-step mechanisms can help reduce the number of people observed, namely as high risk. However, when applied to an entire population, only morality and law can prevent misuse and oversight. (Garzcarek & Steuer, 2019) One of the larger criticisms of such a system as outlined by Itälunni (2022) would be the predisposition of terrorism or an expectation of terrorism arising from Islamic culture thus creating more Islamophobia within the community particularly with the influx of refugees.

By applying Human Security theory to the aforementioned practices of predictive policing, it can be deciphered that there is a lack of integration of individuals as a core focus within these systems. As the development of these systems are reliant on the state in the centre, and the humans acting as a component of the system, biased results are produced due to codification of social categories. As Schoenherr et al. (2023) mentioned previously that the integration of different perspectives from various stakeholders would help make predictions more accurate. A state-focused design approach amalgamated with human-oriented development would require the collaboration of external stakeholders namely educational and human rights organisations however, it would ensure that the system isn't entirely reliant on insulated perceptions such as those of system architects. Furthermore, by applying analogical reasoning as explained by Hobeika et al. (2016) previously, the co-relationships between different stakeholders can be identified using the knowledge they are providing and can be used to identify and integrate new stakeholders. To add to that, the EU Human Security Doctrine (Albrecht et al., 2004) explained such increased focus on collaboration would allow multilateralism, which in turn would enable strategies and policy improvements to transfer effectively to state organisations from external experts and relevant stakeholders.
Das (2020) applied the learnings from human security through the subversive AI (SAI) framework to assist in the development of better predictive policing systems by separating design into, technical, social and ethical goals. The

technical goal would be to create filters that designers can apply to systems that ensures the system or process is interpreted uniformly, holistically across multiple stakeholders, which can be implemented by amalgamating black-box system development with guidelines for affected stakeholders on how the process is applied to their data. Although the technical goal is vital for system design, for the benefit of this research the social and ethical goals are considered key. The social goal empowers people, namely those who are disproportionately negatively affected by predictive policing systems such as religious minorities. The ethical goal is the reduction of power possessed by state agencies in the usage and application of machine learning even when the privacy and civil rights of individuals and specific minority groups are adversely affected. Haggart (2019) identified the core threat within algorithms that subversive AI tries to prevent and/or resolve, the processing and transformation of large datasets into sensitive outputs that can lead to adverse predictions towards minorities. The need for diverse stakeholder groups to be involved in process design as different stakeholders bring different expertise to the system development is further explained by the EU human security doctrine. By using a multilateral approach, research and educational organisations can employ mixed-method formative studies through semi-structured interviews of human rights activities, to map biases within sensitive data when designing systems, thereby helping in developing predictive systems in collaboration with state agencies. Following the identification of these biases, human rights organisations can enable co-design workshops to help design the desired solution as then the developed process would align with the core of predictive policing systems development: human security. Finally, the workshops would lead to the integration of a multilateral human security framework where the standardised and cooperative execution of black-box reiterative systems ensure designer biases are constrained. Datasets and developed algorithms have to be validated against the needs of human users and stakeholders who could be affected by its employment, which could be done by combining qualitative methodologies on derived data, with quantitative methodologies that is already used within state agencies for the development of technical systems that are moulded and can mould social structures.

5.3.2 Disaggregated collection and processing of data will reduce 'feedback loop' of biases that require mitigation to prevent over-policing of individuals and locations:

Instead of relying upon state entities for data and information, the scope of accumulation has to be expanded to both human rights as well as educational organisations. In the Netherlands, BVI, GBA and CBS collectively provide this information for CAS, whereas in Germany RADAR-iTE derives its input from BKA and LKA. Garzcarek & Steuer (2019) stated above how identical biases from multiple data sources can become a systematic bias towards minorities especially when there is a monopoly in data ownership and sources, which can then evolve into a universal norm. Shapiro (2017) previously mentioned how state agencies categorically exclude crime data available from secondary or tertiary sources namely human rights and educational institutions, to ensure the algorithm works as they intend it to. Dr Viola (interview 2) also stated how the data sets used for training the algorithms are categorised and assembled to adhere to state goals, instead of fulfilling societal goals associated with human security. For an algorithm or predictive process to operate efficiently, choosing the right data set is paramount. Waters et al. (2023) mentioned how study conducted by European Union's' Agency for Fundamental Rights (FRA) found that AI usage in resource allocation can further exacerbate feedback loops beyond a probabilistic model as exhibited below:

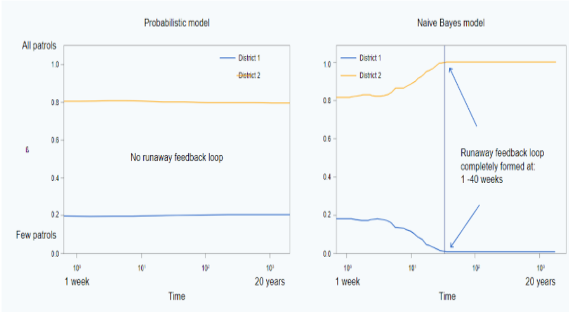

Fig 5: Runaway feedback loops in predictive policing (FRA, 2022)

The study above highlighted how runaway feedback loops become more ingrained when areas that are initially overly targeted through predictive policing have ethnic minorities. This allows discriminatory over-policing policies to be justified. The effect and cause behind such bias towards specific communities can be explained through the community security aspect of human security. The selection of data within predictive policing, for example, is only isolated to state-controlled entities currently, and it is apparent that a combination of default discrimination and coded exposure are prevalent within these datasets, namely in the presence of biases that don't integrate false accusations and racial stereotypes, respectively. To add to that, feedback mechanisms within the algorithm become active and aggravated due to tainted sources of data thereby reinforcing negative assumptions towards minorities, eventually becoming overestimations and leading to overreactions by law enforcement. These issues become more apparent when applying the human security framework to predictive policing practices as it becomes apparent that dimensions namely economic, political, community and personal security are adversely affected for minorities as current processes are still heavily focused on state security and not on upholding human security, as was developed by Albrecht et al. (2004) in the EU Human Security Doctrine. Even when investigating two-stage assessment, as the biases are not only preserved in historical data but also a part of the social processes, a "cumulative disadvantage" develops and reinforces the discrimination minorities are subjected to through methods like over-policing. Each type of over-policing towards specific communities contributes towards adversely affecting certain categories of human security. Banishment and containment would affect economic, food, personal and health dimensions of human security, blight would affect the community dimension while extraction and elimination will affect the economic, personal, community, food and health dimensions. Furthermore, the inherent biases that are present in the designers and developers of the system, as explained by Dr Viola (interview 2), at numerous points throughout the system compounds these biases further as the selection of data and design of the system is correlated to the beneficiaries i.e., state-controlled entities.

Location-based: The EFRA Report on Bias in Algorithms (2022) identified how the variables used create a predisposition towards the construction of a family and its economic factors i.e., usually criminals are from broken families. This correlates to individuals who are recipients of social benefits, which infers that poorer people are more likely to indulge in criminal behaviour, thus creating a bias against lower economic classes. Before 2017, CAS also consisted of certain variables that explicitly created biases against minority races including children who didn't have two Dutch nationals as parents or whether an individual is "Western" in comparison to Netherlands. Yanow & Van der Haar (2013) believed this created a racial determinism against South American, Africans and Asian countries as well as former Dutch colonies that were deemed as "non-Western." This is further exacerbated by the absence of variables that would indicate an individual as "Western," due to the apparent lack of predictive value it added. Although this can be interpreted as ethnic origins being unsuitable for a predictor of crime, the possibility this translates to implicitly woven aspects of ethnicity within the system might be greater. Research conducted by Driessen et al. (2014) found no correlation between the race or ethnic origins of a person and his or her criminal behaviour. But CAS still utilises ethnicity as an implicit factor through demographic indicators, even when evidence has shown them as non-causal variables.

Van Der Woude (2023) highlighted that the Dutch National Police's "Project Sensing" predictive policing tool, predominantly used in Roermond, exhibits bias towards Eastern Europeans. Originally designed to counteract shoplifting and pickpocketing, a study by Amnesty International (2020) found that while 60% of suspects were Dutch, the system disproportionately targeted Eastern European-origin individuals and vehicles for scrutiny. Using CAS extensions, this system collects and analyses vehicle data, such as brand and model, to predict potential criminal activities in Roermond's shopping areas. Cars deemed high-risk trigger a law enforcement alert, allowing patrol officers to intervene based on various car details. Under Dutch law, these officers can stop vehicles based on these alerts and potentially detain individuals. This approach, although aiming to prevent shoplifting and pickpocketing, primarily targets Eastern Europeans, thereby stretching beyond traffic safety considerations.

Although the Dutch National Police's internal reports deny a singular focus on "mobile banditry"—an organized crime type associated with Eastern European groups—Van Der Woude (2023) noted the reports' acknowledgment of its rise in the region. While the reports assert they don't rely on crime statistics exaggerating Central and Eastern European (CEE) national involvement, verifying these claims proves challenging. The police, while using "Project Sensing" to address mobile banditry, have cited data indicating the primary culprits hail from Eastern European countries like Poland and Romania. These reports suggest that CEE nationals in the Netherlands, even if legally present, are perceived as potential criminals. Thus, while race isn't directly a factor in the system's design, nationality and ethnicity are implicitly used, perpetuating biases and justifying discrimination against minorities.

Person-based: Çankaya (2015) posits that predictive policing variables are modern extensions of traditional police indicators. Studies show that officers often rely on a blend of cultural, biological, and physical traits to decide when to question someone, mirroring demographic data of "known offenders" from databases like CBS and BVI. By comparing physical traits with location and time, police inherently incorporate racial biases into both algorithmic and practical decision-making (Çankaya, 2012). Such biases, stemming from officials' subconscious racial profiling, create feedback loops, intensifying the racial discrimination faced by minorities. This "cumulative disadvantage" not only perpetuates bias but also augments societal inequality and jeopardizes minority groups' security (Wilson, 2011).

Efforts to understand radicalization, like the Dutch Government's European database of Terrorist Offenders (ETD), which collects comprehensive data on terrorists, may fall victim to these feedback loops. The FRA report on Algorithmic Bias (2022) indicates that without a diverse data contribution from entities beyond law enforcement, the biases entrenched in historical crime data persist. For predictive policing, which depends on identifying patterns from historical crime data, there's an inherent limitation. Such data doesn't encompass all crimes, making it non-representative. Some nations use victimization surveys to gauge unreported crimes, but each predictive system has its inherent flaws.

As algorithmic classification relies upon data from the past, a concern for such processes would be not accounting for other members within a community whose behaviour on an individual level could deviate from what is considered the norm over time. Another critique as discovered by Garczarek & Steuer (2019) would be the inherent human intervention during the development and scoring of variables for these processes, even when it is labelled as an automation, as this adds an additional layer of bias to the setting of parameters and variables. Shapiro (2017) further stated that some state agencies take liberty in voluntarily excluding certain data sources when concerning human rights and educational institutions, and crime data when training algorithms, which isn't apparent due to the lack of transparency. By disallowing public scrutiny alongside social and legal accountability through a lack of transparency, the lack of openness of predictive policing systems makes it resistant towards review by independent actors. Whether being developed by private actors or in-house, these algorithms are sheltered from criticism even when transparency has to be a priority due to the technical nature of the process. Panelli (2018) expressed concerns regarding the “black box” method most law enforcement agencies employ to withhold disclosing specific information pertaining to the algorithm and processes employed, which makes it considerably difficult for citizens as well as policy makers to understand. There are some software developers who do provide information on how the system works, the data used and any oversight mechanisms that are employed for data inaccuracy.

Dr Viola (interview 2) mentioned that the data has to be improved alongside the algorithm and how it functions as the selection and creation of these datasets involve specific decisions that correlate certain biases that are inherent in law enforcement. Dr Gwen (interview 3) mentioned how there is a lack of regulation as there is no legal basis for predictive policing, which allows biased data to permeate from different sources and be used in practice. The Project Sensing predictive policing process in Netherlands is an adept example of how biased data could be without oversight and lack of data sources. Dr Gwen (interview 3) mentioned how in this process, Dutch police only monitor license plates and country codes that are not of Netherlands through cameras, without any tertiary sources to help improve their understanding of crimes. The Bias in Algorithms report by EFRA (2022) supported these claims by adding that data quality, if left unimproved, would impact feedback loops, especially as historical minorities have low reporting rates making true crime rates difficult to perceive. This is more apparent in process like RADAR-iTE, where Itälunni (2018) explained how these mechanisms have difficulty in identifying probable terrorists due to lack of accurate data and the decentralisation of their activities. As pattern identification is central to the algorithm, the dataset will introduce limitations to its efficacy due to being incomplete and not including all criminal offences.

5.3.3 Independent Audit body to curb transparency & bias:

Shapiro (2017) and the EFRA Report on Bias in Algorithms (2022) stated how a lack of transparency coupled with lack of accountability within these processes curbs public scrutiny and accountability towards state agencies that develop and implement these systems. To add to that, the predisposition towards certain minorities is also prevalent within these systems such as the belief that terrorism is on the rise in Islamic culture, which led to the design and implementation of RADAR-iTE. Dr Viola (interview 2) stated such opaque, "black box" design comes from the implicit beliefs surrounding tech and namely techno-determinism and the infallibility of predictive processes. Dr Gwen (interview 3) added to that by saying that advocacy is more difficult than criticism due to the lack of details provided in how the system works beyond the hypothetical, which policymakers require to develop standardised policies.
Transparency is required regarding the collection, training as well as the categorisation of data to allow the study and improvement of these processes. As current processes don't document any forms of methods and practices, the predictive judgments made relies not only upon incomplete and unstructured datasets as mentioned above, but also the biases of the developers of the system as well as the internal factors of the companies and state bodies responsible which includes both economic and political interests. Shapiro (2017) added how some state agencies, for example, are prone to exclude certain data sources, namely those concerning human rights and educational or research organisations. Dr Viola (interview 2) attributes this to a lack of "location of liability," which is required by predictive processes to ensure that the consequences of the technology are fairly investigated and so is the accountability, which can only be possible through establishing an independent auditing organisation.

Location-based: For CAS, similar to other systems, the simplification of information is also detrimental. Data officers, who are at the core of deciphering the information provided to them through contextual knowledge and their own expertise, receive simplified information with some elements being omitted. While supporters of the system believe providing data officers with limited information enables them to conduct more extensive research on risk areas devoid of guidance from CAS, one could argue that this could lead to misinterpretations as crime is a complex problem to analyse. By keeping the system closed for users, the Dutch police have attempted to change the user from someone who visualises and reproduces results to someone who can enrich them through qualitative explanation. Oosterloo et al. (2018) though added that the extra layer of interpretations could enable probable biases from permeating into the system. As the production of data already consists of biases against minorities such as arresting specific groups, processed data allows humans to contextualise and frame results through their own knowledge, but the sheer volume of data makes it a time-consuming process to enhance data leading to predictions that keep users obscured from the process.

Oosterloo et al. (2018) therefore believed that the knowledge utilised alongside the methods, validity, justifications and scope i.e., the epistemology behind CAS is heavily influenced by preexisting culture of discrimination against certain groups within law enforcement. Even from an ontological perspective by determining individuals and places, it becomes apparent that CAS is not a neutral system but rather a social construction of reality, shaped by how Dutch law enforcement views physical characteristics, economic variables and other behaviour as being adversely correlated.

Person-based: For person-based predictive policing systems in Germany, the BKA and the State Criminal Police Offices (LKA) have kept considerable information regarding the entire design, development and different variables used in analysis for RADAR-iTE within a "black box," similar to other predictive policing mechanisms there is a lack of transparency. To add to that, Fernandez & de Lasala (2021) mentioned the variables used, accounting for personal life, social life, social media, travel history and criminal history, which can only be applied to small marginal groups and not holistically as generalisations aren't possible. The EFRA Report on Bias in Algorithms (2022) stated that alongside lack of transparency, the lack of accountability would be due to the complexity of the system. Due to a combination of intellectual property law and the complexity of the algorithms being used, law enforcement officials might not possess the appropriate information and training necessary to accurately use the system or understand what the process is doing. Finding errors or biases becomes difficult due to this and an over-reliance on algorithms can lead to biased results from this. Such forms of biased results due to AI represent a type of engineering inequity that become more prominent when variables such as previous arrests for any crime namely narcotics or being a victim of aggravated assault or shooting are used as predictors, when there is credible research that shows how minorities are more likely to be arrested within European communities. Furthermore, through the promotion of default discrimination within the crime data that is used for predictive policing, and numerous statistics showing that non-white individuals are more likely to be arrested. Racial classification embedded within state institutions would be reinforced racial stigmatisation in civil society in ways that are relevant to machine learning systems design. States conjure race as a factor through tertiary variables within predictive policing practices to ensure social, economic, and political inequalities that are embedded in state and civic institutional practices.

Ugwudike (2022) criticised the lack of focus given on AI systems developed by justice systems and the deployment of AI audits that place extensive focus on technical components only. Kazim et al. (2021) added to that by stating auditing organisations have to integrate the non-technical and social aspects such as design principles as current datasets prove disadvantageous for low-income and racialised groups that are historically vulnerable to domestic law enforcement.

Dr Viola (interview 2) believes that this can be accomplished through human oversight within audits to ensure fairness and reduce data inaccuracy, but this also means that human scholars, who are already aware how to critique digital processes, require complete understanding of the mechanisms and the rationale behind those used within the system. It must be stressed though that in certain states within the USA, such audits and testing have stimulated lawsuits and protests that have ultimately led to the discontinuation of such systems, results which Dr Viola (interview 2) feels could be replicated if applied to the EU.

Dr Sommerer (interview 1) added that an auditing body would not only ensure greater trust between law enforcement and citizens as well as human rights organisations and other stakeholders but also allow the systems to stop functioning as solely self-learning AI. As law enforcement and state bodies might not entirely comprehend biases that are present within the system, by creating an independent auditing body to monitor and supervise the algorithms at regular intervals these biases can be identified proactively and mitigated to improve the system. Current systems lack any form of openness and thus are resistant towards review, and subsequent evolution being sheltered from criticism. As mentioned by Panelli (2018) the prevention of disclosure of specific information concerning the algorithm and processes, debilitates policy makers as well as citizens from assisting in improving it. Audit Representatives, therefore, must include individuals who are adept at human security and civil rights, among those who are more knowledgeable in legal and technical areas. Das (2020) added that through conduction of standardised audit and evaluation of AI systems, design moves away from theory into human security-centred predictive policing systems. By removing human-based analysis or inputs when designing systems, the bias towards specific minorities can be restricted. Combining such design practices with more conventional methods of ensuring human privacy and security such as end-to-end encryption and protected networks could resolve issues in user-centred design practices for AI. Such methods of human-centred design process prevent AI from inferring sensitive biases through identified information. Bolkar (2023) mentioned that the AI Act does propose standardised auditing, regulation and monitoring practices based on the following aspects, however there is a lack of clarity regarding framework and practices that need to be addressed:

- Registration of algorithms and ML models within the EU database
- Disclosure of all content that is generated through the model.
- Ensure prevention of content that goes against human rights and discrimination.
- Training data has to be copyrighted and provided to the auditing body in summary.

By applying the identified theoretical framework, the purpose and development of an auditing process can be understood and subsequently stimulated. When applying human security theory to predictive policing systems, namely the Human Security Doctrine developed for EU (Albrecht et al., 2004), the primacy of human rights becomes key and how minority communities are discriminated against due to the state's adherence to designing uncontrolled systems for states and not the citizens that don't allow state accountability to be called into question. This approach in identifying problems and biases within the system and solving those at the source can be explained and developed through STS approach as transparency and congruence must be key components of the system to enable greater collaboration between diverse stakeholders. (Schoenherr et al., 2023)

An auditing process must be focused on the following intersecting aspects:

- structural and systematic evaluation based on race, gender and other biases that are harmful towards marginalised subgroups, this should include both the algorithm designed and the data used to develop the process, alongside the methodology used for data selection.

- legal critiques focused on human rights violations namely within the sphere of privacy concerns.
- deficiency in substantial transparency and accountability of the process that would be a part of both design and technical criticism.

5.3.4 EU Standardisation could be attained through EU AI Act, but the roadblocks are more state-oriented:

Montasari (2023a) elaborated that one of the core reasons why standardisation is a necessity would be the diversity of data that could be obtained from different sources or originate within a state-controlled organisation that doesn't take into account deviation of patterns. Standardisation of practices as well as auditing frameworks would not only ensure cooperation and revision of practices and algorithms but also foster greater trust between the community and state agencies.

Dr Sommerer (interview 1) mentioned how privacy and human dignity, for example, are fundamental rights as per the German Constitution and the EU Charter of Fundamental Rights, alongside the right to equal treatment. However, the application of predictive policing systems, both as PreCobs and two-stage assessment systems like RADAR-iTE, translates to state security taking precedence of aspects of human security, with equal treatment and privacy representing community and personal security, unlike what ought to be the focus as per the EU. EU-wide Standardisation could help curb some of the issues that are currently prevalent in predictive systems however problems persist in how legal practices are developed and implemented within individual states. Dr Sommerer (interview 1) added that although EU law attempts to implement similar laws in different states and harmonise practices, developments in criminal law and policing are entirely reliant on member states. The different standards of policing and legal practices across member states thereby creates a hindrance in the standardisation of practices, because even if all states were to come to an agreement, only the minimal standards would be possible to achieve. Dr Gwen (interview 3) further added that although AI Act might help in standardising some aspects of predictive policing practices, high chances are that governments will try to get around this one of the reasons would be the lack of proper definition of AI practices. Standardisations is also heavily reliant on how efficient auditing and oversight practices are.

Dr Sommerer (interview 1) further stated that the European Court of Justice has agreed on halting self-learning algorithms, namely neural networks, due to the tricky ethical and legal landscape states might find themselves in. However, the creation and implementation of this legislation is entirely reliant on what practices are defined as self-learning and how that definition incorporates scholarly research and academia focused on human security. These definitions thereby become the responsibility on cohesive collaboration across the different stakeholders involved namely law enforcement, educational and human rights institutes, and state agencies, among others.

The EU human security doctrine developed by Albrecht et al. (2004) clearly stated how standardisation is necessary for predictive processes and requires multilateralism to allow both policies and strategies to be revised and improvised then seamlessly integrated into state-controlled practices through cooperation with external stakeholders. To add to that, the different aspects of STS can also be used to identify the need to create a standardised system by shifting the core of the system from state to human security. Social category codification would be one of the ways that can be achieved as it assesses algorithm accuracy by taking into account the diversity of communities present and the conflict, they could present towards one another. By understanding and identifying the social inequalities that are apparent within flawed datasets, sources and points of entry within the process can be reviewed and improved through standardised frameworks, as well as creating standardised frameworks for such processes. (Schoenherr et al., 2023)

5.4 Discussion

This section is dedicated to understanding the solutions of the some of the core issues obtained above from the empirical analysis. These solutions and recommendations were provided by interviewees as well as published articles, however a theoretical approach developed by combining algorithmic fairness theory and socio-technical systems approach is applied to understand the theoretical reasoning behind the solutions to the problems identified.

In 5.3.1, to stimulate greater collaboration with non-state actors to optimise processes and develop better datasets, STS could be applied as it would separate design and development into technical, social and ethical goals. If each goal were to be connected and managed by relevant stakeholders then not only would this foster collaboration but also ensure accountability, and most importantly, transparency. To ensure social and ethical goals are met, marginalised communities have to be empowered and state power has to be reduced when considering privacy and civil rights, which can only be done by involving human rights organisations to play an active role in the design of the system. Organisations like Amnesty International can work with researchers and educational institutes to analyse data and anti-bias design strategies that can then be used within the system to increase accuracy of predictions.

The STS approach can also help enable disaggregated collection and processing of data to reduce ‘feedback loop’ of biases, which was identified in 5.3.2. Choosing and implementing data sources can be achieved through the use of variance control to identify and resolve errors thereby preventing faulty patterns as well as consistently improve the system through better improvement and completion of datasets. For state agencies and system developers, it's imperative to understand that the system is an extended social process, where different datasets and perceptions have to be integrated, through citizens' surveys and questionnaires, to ensure holistic development and implementation. Victimisation surveys, for example, could help curb some of the algorithmic bias present in the system as this allows an extended estimation of crime that has not been included in historical crime data. Clegg (2000) also mentioned how the STS approach emphasises on controlling the data deviations at the point of origin through continuous evaluation and filtration. Data can be further controlled by hen applying algorithmic fairness theory to the anti-bias framework, one of the methods of which would be to identify neutral proxy variables within datasets to ensure feedback loops can be mitigated through a combination of algorithm fairness processes such as downsampling as suggested by Ensign et al. (2018) and winsorization. An experimental method of using downsampling was employed in the FRA report (2022) below:

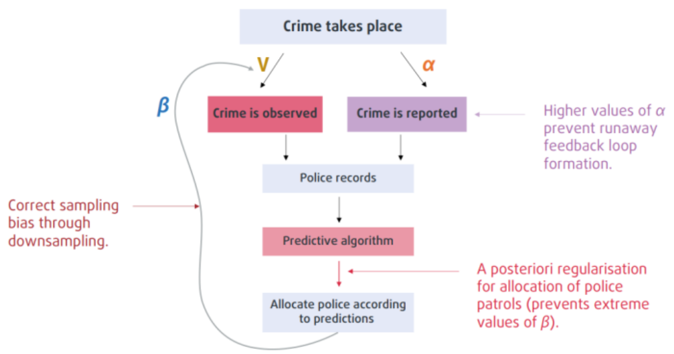

Fig 6: Bias Mitigation Techniques to reduce feedback loop formation (FRA, 2022)

To mitigate feedback loop formation within algorithms that are used for location-based predictive policing the following methods could also be employed that originated from education and research institutions:

- reporting rates can be increased to circumvent over-policing as one of the bias sources would be biased crime reporting rates. This can be achieved by increasing the volume of objective data used as initial inputs.
- Statistical and programming techniques namely winsorization can be used to limit extreme values or outliers to arbitrary values to reduce impact and ensure the preservation of the overall distribution of the variable.
- Furthermore, "overfitting", which is the tendency of machines to focus too much on patterns found within training data, has to be account for through regularisation. Regularisation prevents algorithms from predicting values that are too extreme through the inclusion of a mathematical restriction, thus constricting overfitting. This value requires criticism to ensure the predictions are useful while feedback loops are prevented.
- To add to that, downsampling, which was introduced by Ensign et al. (2018) can also be employed. Downsampling involves assigning specific probabilities of recording to crime events that counteract the overly strong predictions by choosing a subset of data points from the original dataset either randomly or over regular intervals through systematic sampling. This allows greater information to be derived from a shorter subset while also preventing the development of a feedback or norm.
- Finally, equal crime reporting rates across neighbourhoods are vital in mitigating formation of feedback loops, although this is reliant on true crime rates and how those are interpreted using multiple external stakeholders and data sources. The FRA study stated that community trust within law enforcement deters over time as people start reporting crimes such as burglaries in the first place due to the belief that the police wouldn't do anything.

Although predictive algorithms attempt to quantify individuals as such instead of categorising them as part of a larger community, minorities or otherwise, due to the large role datasets play within their development and implementation, measures have to be taken to disable this. Consequently, Hanna et al. (2020) proposed using “equalised mistreatment” to ensure there is an equal rate of positive prediction across diverse groups, as well as identifying the injustices inflicted on marginalised groups using and critically identifying racial perceptions that are bound by data sources and census categories. Scanlan (2019) also added that internally developed algorithms, data validation process must be conducted prior to input and any sets of crime data that doesn’t accurately reflect both reported and unreported crime reports must be adjusted accordingly or the organisations should be penalised. The most common theme among the literature investigated as well as the interviews conducted would be the disaggregated analysis of data through audit frameworks by an independent body, which has been analysed 5.3.3 and to understand the solutions and statements provided, the STS approach can be applied again.

Socio-technical systems (STS) approach could help develop the necessary framework by catalysing information flow as a core deliverable as all stakeholders must be made aware of the data used and how. Power and authority as well as would be require the level of control to be regulated amongst state and law enforcement agencies as well as external stakeholders, by ensuring state bodies are accountable for their actions through exercising limitations and transparency in their control. Furthermore, the framework must support and reinforce social structures through developing and auditing training programs that must be provided to law enforcement officials to curb inherent bias towards minorities. Finally, when designing auditing frameworks for and subsequently improving predictive policing systems the following design aspects must be followed:

- Making the design systematic by explicit identifying and integrating interdependent social factors namely human security as core deliverables for both auditing and core predictive policing system
- Integrating environmental factors such as government mandates, and legal frameworks, and allocating those across both human and technological resources. As the design of predictive policing systems is shaped socially, mechanisms such as surveys and structured questionnaires allow citizens to have their perceptions recognised and integrated into the development and improvement of the system. Pre-emptive impact assessments are one such methodology that evaluates social and design logics within AI systems with the purpose being the identification of negative social impact namely discrimination alongside other broader social consequences that might arise.
- Controlling biases and other problems at the source through proactive, multi-disciplinary, mitigation means
- Ensuring that the system goals are standardised and aligned with those to support social structures, in this case, human security.

It must be stressed though that the data-selection process requires separate solutions by investigating theoretical assumptions as well as going beyond categorising such AI systems as being neutral tools that produce outputs based on data patterns, without accounting for the human or social structures that could influence how they operate. This directly counters tech-determinism by identifying the affect societal norms could have on the development of technology thereby holding developers accountable for the processes they create. Benjamin (2023) defined techno-determinism as the faulty perception that society could be affected by but won't affect technological development, as AI systems aren't independent of developers or larger social systems. AI outputs thereby are results of the data used as well as the theories employed in design and implementation that need to be audited to identify source of algorithmic bias. Mugari & Obioha (2021) suggested a two-pronged approach for this involving an independent review board acting as an external mechanism to act as an oversight extending beyond predictive policing to general law enforcement practices. This would not only allow all civil society groups to act as oversight but also hold state agencies accountable. By applying algorithmic fairness theory this is further elaborated.

Hanna et al. (2020) implored the need for disaggregated groups, as part of algorithmic fairness theory, to standardise algorithm audit frameworks. By conducting analysis of differences between different communities and how predictive algorithms could have biases towards specific communities, by measuring variation in inequalities combined with racial categorisation could help standardise predictive processes. In such an algorithmic fairness theory-based approach, the data collection, sources as well as the processes used, all become a part of a holistic standardised framework. The next chapter will include the final thoughts and conclusions regarding the research.

Conclusion

In the ever-evolving landscape of law enforcement and technological integration, predictive policing stands at the forefront of innovation and controversy. Drawing on the observations of Mugari & Obioha (2021) the discussion around predictive policing systems and their efficacy in crime control and resource allocation for understaffed domestic law enforcement agencies, must not outweigh how useful the system can be. Very few research or study have called for predictive policing practices to be ceased, instead pointing out improvement areas primarily in the realm of data reliability, addressing biases, transparency, accountability and the adverse effects these practices can have on human rights. The possible development and implementation of the EU AI Act though could heavily restrict or dissolve predictive policing processes as some of the methods used currently are classified as unacceptable. While there's undeniable potential in these systems, it's crucial to weigh them against concerns like data reliability, transparency, potential biases, and the broader implications on human rights.

With the prevalent belief that most predictive practices would be categorised as high-risk, due to disconnect between strategies designed to protect the state and those designed to protect EU, states and policymakers have to identify what practices within predictive policing processes need to be altered and how. By taking a human security-oriented approach towards identifying and defining these issues, supported by other relevant theories, not only is it possible to address these issues but also develop the approach needed to solve them.

To ensure that the theories can be applied to all forms of predictive policing practices in EU namely big data, ML, and person-based forecasting, the analysis on location-based forecasting systems namely CAS and PreCobs, as well as person-based forecasting systems in VERA-2R and Radar-iTE within

Netherlands and Germany allowed all the core approaches to be considered. And although state agencies have employed multi-stage assessment strategies, namely within person-based forecasting to curb errors, literary research of the practices proved otherwise. Existence of biased variables namely social, economic and race within the predictive process, even when research has indicated minimal correlation of these variables with crime, stimulates a cumulative disadvantage for minorities through reinforcement of established, embedded biases in datasets. Furthermore, as these practices rely upon datasets derived from state-controlled agencies, consistently relying on predictive decisions without any form of scrutiny could lead to over-policing and create feedback loops that catalyse law enforcement agencies to subconsciously perform racial profiling. To add to that, by using a "black box" approach to keep processes and datasets used in predictive systems a secret, a lack of transparency also stimulates a lack of accountability and trust in the process and state organisations. The inherent human intervention in the categorisation and selection of datasets and data sources alongside how the process is used therefore required a multi-layered theoretical approach to understand and resolve.

To ensure both social and socio-technical aspects are covered, a multi-layered theoretical approach was utilised to help analyse the information accumulated. Firstly, by using human security theory, namely focusing on the core components of economy, community and personal, the adverse effects that predictive policing practices have upon specific aspects of human security were identified, gauged, analysed and understood. Algorithmic bias theory was also applied to understand the technical nature of how biases permeate from different state-controlled sources and datasets into the process and development of predictive systems, and how such reinforced biases conjure "feedback loops." These loops would stimulate the reinforcement of social biases that constitute the development and implementation of predictive systems, leading to overestimation of results i.e., runaways. To help exhibit that the issues within predictive policing systems identified through human security and algorithmic bias framework, can be countered through anti-bias theoretical framework, algorithmic fairness theory and socio-technical systems approach were subsequently applied. The former helped develop anti-bias strategies namely the need for disaggregated systems for analysis and critique to help optimise processes and using marginalised communities as the focus of the system for improvement or development. Subsequently, adopting a socio-technical systems approach towards resolving the issues within such predictive systems, core variables of system improvement and new system development came to light namely the need for a transparent development process, an independent auditing body with a standardised auditing framework and controlling dataset issues by involving external research and educational organisations. Bringing the theoretical framework together was the application of the EU AI act, which is yet to be implemented, as a normative framework to understand both limitations, and probable solutions that could help improve predictive processes instead of abolishing those. To help supplement the information collected from journals and published articles, three interviewees were selected who are educational and research experts in the field of predictive policing in Germany and Netherlands.

The empirical analysis of the information collated from research journals and interviews using the theoretical framework identified a need for greater collaboration in development and improvement of predictive systems alongside disaggregate collection and analysis of datasets, and an independent auditing body for standardisation of these processes. Human security, particularly personal and economic security, explained how the analysed predictive policing systems, both location-based and person-based, need to divert focus from state and integrate the humans affected, namely from minority communities into the system. As very few within state organisations might have the technical knowledge to develop and manage such predictive systems, and even fewer might possess the social knowledge to improve these systems, segregation of goals is required. A human-oriented development of systems through separation of technical, social and ethical goals, and having different external and internal stakeholders manage those would increase accountability and cooperation between state and non-state actors. Technical issues relating to algorithmic bias, namely feedback loops and runaways, due to extensive reliance on inaccurate or incomplete state-controlled datasets reinforces discrimination against minorities creating a cumulative disadvantage leading to overenforcement upon individuals and locations alike. To resolve such types of algorithmic bias that adversely affect minority community security, algorithmic fairness theory coupled with socio-technical approach was used. Controlling data deviations at the source alongside negating neutral proxy variables within algorithms, as well as statistical techniques like winsorization, overfitting and downsampling are efficient techniques that could resolve feedback loops. An independent auditing body that helps standardise practices through regulations such as the EU AI Act, while conducting structural and systematic audits through a two-pronged approach would allow external human rights, educational and civic stakeholders to participate in developing and optimising predictive systems alongside citizens. This would also ensure state-organisations are held accountable by enabling greater transparency to increase fairness and decrease data inaccuracy, thereby enhancing trust between law enforcement and minority communities. By using STS approach, current processes can be optimised to conduct structural and systematic evaluation of protected characteristics by keeping the system design explicit, including environmental factors, and ensuring system goals are standardised by controlling biases. Finally, although framework and strategy have been provided on standardisation through normative frameworks like the EU AI Act, state-oriented roadblocks are highly probable as state security policies lack explicitly human-oriented approaches, which will require restructuring and adopting state security policies.

Perhaps these improvements could be initiated by altering how predictive policing is viewed. Perry et al. (2013) identified that unlike current methods of implementation, predictive policing is not an isolated strategy to combat crime, but rather a comprehensive and correlated crime prevention strategy that requires social and subject-matter experts to become an efficient process. Prediction is part of the more expansive method of crime control, and requires transparency, counter-bias initiatives as well as a standardised auditing framework for reiterative improvement. Instead of relying on solely the digital landscape, the significant contribution that socio-technical and social humanities could offer must be integrated to state-level understanding of security and consequently predictive policing practices. By shifting away from an isolated definition and becoming an expansive one, states will stimulate greater collaboration with external stakeholders as well as communities to help develop and enhance such system while ensuring trust between state entities and communities. By ensuring practices are audited, biases are resolved and external stakeholders, namely the ones most affected i.e., citizens and minority communities, are actively involved, it is possible to save predictive policing systems by improving them, when the EU AI Act is implemented.

Bibliography

1. Agan, A. and Starr, S. (2018) ‘Ban the box, criminal records, and racial discrimination: A field experiment’, *The Quarterly Journal of Economics*, 133(1), pp. 191–235.
2. Albrecht, U. *et al.* (2004) ‘A human security doctrine for Europe: the Barcelona Report of the Study Group on Europe’s Security Capabilities’.
3. Anderson, E. (2009) ‘Toward a non-ideal, relational methodology for political philosophy: Comments on Schwartzman’s challenging liberalism’, *Hypatia*, 24(4), pp. 130–145.
4. Andrus, M. and Gilbert, T.K. (2019) ‘Towards a just theory of measurement: A principled social measurement assurance program for machine learning’, in. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 445–451.
5. Bautze, K. (2018) ‘Wie gefährlich sind „Gefährder “?’, *Kritische Justiz*, 51(2), pp. 205–212.
6. Benjamin, R. (2023) ‘Race after technology’, in *Social Theory Re-Wired*. Routledge, pp. 405–415.
7. Bennett Moses, L. and Chan, J. (2018) ‘Algorithmic prediction in policing: assumptions, evaluation, and accountability’, *Policing and society*, 28(7), pp. 806–822.
8. Benthall, S. and Haynes, B.D. (2019) ‘Racial categories in machine learning’, in. *Proceedings of the conference on fairness, accountability, and transparency*, pp. 289–298.
9. Bolkar, S. (2023) ‘EU AI act: The Regulatory Framework on the usage of machine learning in the European Union’, *InfoQ*. Available at: https://www.infoq.com/news/2023/07/eu-ai-act/ (Accessed: 19 July 2023).

10. Bosold, D. and Werthes, S. (2005) ‘Human security in practice: Canadian and Japanese experiences’, *Internationale Politik und Gesellschaft*, 1(2005), pp. 84–101.
11. Buolamwini, J. and Gebru, T. (2018) ‘Gender shades: Intersectional accuracy disparities in commercial gender classification’, in. *Conference on fairness, accountability and transparency*, PMLR, pp. 77–91.
12. Çankaya, S. (2012) ‘De controle van marsmannetjes en ander schorriemorrie’.
13. Çankaya, S. (2015) ‘De politiële surveillance van ras en etniciteit’, *LG Moor, J. Janssen, M. Easton en A. Verhage (red.), Cahiers Politiestudies: Etnisch profileren en interne diversiteit bij de politie*, pp. 13–33.
14. Carayon, P. (2006) ‘Human factors of complex sociotechnical systems’, *Applied Ergonomics*, 37(4), pp. 525–535. Available at: https://doi.org/10.1016/j.apergo.2006.04.011.
15. Carbonell, J.G., Michalski, R.S. and Mitchell, T.M. (1983) ‘An overview of machine learning’, *Machine learning*, pp. 3–23.
16. Cherns, A. (1976) ‘The principles of sociotechnical design’, *Human relations*, 29(8), pp. 783–792.
17. Clegg, C.W. (2000) ‘Sociotechnical principles for system design’, *Applied ergonomics*, 31(5), pp. 463–477.
18. Closs, D.J. *et al.* (2008) ‘Toward a theory of competencies for the management of product complexity: six case studies’, *Journal of Operations Management*, 26(5), pp. 590–610.
19. Das, S. (2020) ‘Subversive AI: Resisting automated algorithmic surveillance with human-centered adversarial machine learning’, in. *Resistance AI Workshop at NeurIPS*, p. 4.

20. DaViera, A.L. *et al.* (2023) ‘Risk, race, and predictive policing: A critical race theory analysis of the strategic subject list’, *American journal of community psychology* [Preprint].
21. De Graauw, J.S. (2014) ‘Tijdruimtelijk voorspellen van criminele incidenten’.
22. Delgado, R. and Stefancic, J. (2023) *Critical race theory: An introduction*. NyU press.
23. Demchenko, Y. *et al.* (2013) ‘Addressing big data issues in scientific data infrastructure’, in. *2013 International conference on collaboration technologies and systems (CTS)*, IEEE, pp. 48–55.
24. Driessen, F.M.H.M., Duursma, F. and Broekhuizen, J. (2014) *De ontwikkeling van de criminaliteit van Rotterdamse autochtone en allochtone jongeren van 12 tot 18 jaar. De rol van achterstanden, ouders, normen en vrienden*. Reed Business Information.
25. Eckerson, W.W. (2007) ‘What is Predictive Analytics?’, in *Predictive Analytics Extending the Value of Your Data Warehousing Investment TDWI Best Practices Report*, pp. 1–36.
26. EFRA. (2022) *Bias in algorithms: artificial intelligence and discrimination.* LU: Publications Office. Available at: https://data.europa.eu/doi/10.2811/25847 (Accessed: 23 August 2023).
27. El Atillah, I. (2023) ‘Most AI models don’t comply with the EU’s AI Act, study finds, AI models don’t comply with the EU’s AI Act, according to a Stanford study.’, *euronews*. Available at: https://www.euronews.com/next/2023/07/10/ai-models-dont-comply-with-the-eus-ai-act-according-to-a-stanford-study (Accessed: 21 July 2023).
28. Ensign, D. *et al.* (2018) ‘Runaway feedback loops in predictive policing’, in. *Conference on fairness, accountability and transparency*, PMLR, pp. 160–171.

29. Ferguson, A.G. (2012) ‘Predictive policing and reasonable suspicion’, *Emory LJ*, 62, p. 259.
30. Ferguson, A.G. (2015) ‘Big data and predictive reasonable suspicion’, *University of Pennsylvania law review*, pp. 327–410.
31. Fernandez, C. and de Lasala, F. (2021) ‘Risk Assessment in Prison’, *Brussels, Belgium: Radicalisation Awareness Network*, pp. 2021–04.
32. Finn, E. (2018) *What algorithms want: Imagination in the age of computing*. mit Press.
33. Fitzpatrick, D.J., Gorr, W.L. and Neill, D.B. (2019) ‘Keeping score: Predictive analytics in policing’, *Annual Review of Criminology*, 2, pp. 473–491.
34. Ford, C.L. and Harawa, N.T. (2010) ‘A new conceptualization of ethnicity for social epidemiologic and health equity research’, *Social science & medicine*, 71(2), pp. 251–258.
35. Gandomi, A. and Haider, M. (2015) ‘Beyond the hype: Big data concepts, methods, and analytics’, *International journal of information management*, 35(2), pp. 137–144.
36. Garzcarek, U. and Steuer, D. (2019) ‘Approaching ethical guidelines for data scientists’, *Applications in statistical computing: From music data analysis to industrial quality improvement*, pp. 151–169.
37. Geller, A., Fagan, J. and Tyler, T. (2017) ‘Police contact and mental health’, *Columbia public law research paper* [Preprint], (14–571).
38. Gerstner, D. (2018) ‘Predictive policing in the context of residential burglary: An empirical illustration on the basis of a pilot project in Baden-Württemberg, Germany’, *European Journal for Security Research*, 3(2), pp. 115–138.
39. Gillis, A.S. (2021) ‘The 5 V’s of big data, 5 Vs of Big Data’, *techtarget*. Available at: https://www.techtarget.com/searchdatamanagement/definition/5-Vs-of-big-data (Accessed: 23 July 2023).

40. Gorr, W. and Harries, R. (2003) 'Introduction to crime forecasting', *International Journal of Forecasting*, 19(4), pp. 551–555.
41. Grace, J. (2021) ''Algorithmic impropriety'in UK policing contexts: A developing narrative?', in *Predictive Policing and Artificial Intelligence*. Routledge, pp. 237–253.
42. Günther, W.A. *et al.* (2017) 'Debating big data: A literature review on realizing value from big data', *The Journal of Strategic Information Systems*, 26(3), pp. 191–209.
43. Gupta, M. and George, J.F. (2016) 'Toward the development of a big data analytics capability', *Information & Management*, 53(8), pp. 1049–1064.
44. Haggart, B. (2019) 'The age of surveillance capitalism: The fight for a human future at the new frontier of power, S. Zuboff (2018)', *journal of digital media & policy*, 10(2), pp. 229–243.
45. Hanna, A. *et al.* (2020) 'Towards a critical race methodology in algorithmic fairness', in. *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp. 501–512.
46. Harcourt, B.E. (2019) *Against prediction: Profiling, policing, and punishing in an actuarial age*. University of Chicago Press.
47. Hardyns, W. and Rummens, A. (2018) 'Predictive policing as a new tool for law enforcement? Recent developments and challenges', *European journal on criminal policy and research*, 24, pp. 201–218.
48. Hobeika, L. *et al.* (2016) 'General and specialized brain correlates for analogical reasoning: A meta-analysis of functional imaging studies', *Human brain mapping*, 37(5), pp. 1953–1969.
49. Howell, J. and Emerson, M.O. (2017) 'So what "should" we use? Evaluating the impact of five racial measures on markers of social inequality', *Sociology of Race and Ethnicity*, 3(1), pp. 14–30.

50. IfmPt (2018) *Predictive policing made in Germany*, *PRECOBS Predictive Policing*. Available at: http://ifmpt.de (Accessed: 16 June 2023).
51. Itälunni, J. (2018) 'Development of the RADAR-iTE instruction card for operational and educational purposes'.
52. Itälunni, J. (2022) 'Human Security And The European Economy: Review of the sociological and economic situation in the EU sphere'.
53. Jackson, J.R. (2018) 'Algorithmic bias', *Journal of Leadership, Accountability and Ethics*, 15(4), pp. 55–65.
54. Joh, E.E. (2016) 'The new surveillance discretion: Automated suspicion, big data, and policing', *Harv. L. & Pol'y Rev.*, 10, p. 15.
55. Jordan, M.I. and Mitchell, T.M. (2015) 'Machine learning: Trends, perspectives, and prospects', *Science*, 349(6245), pp. 255–260.
56. Kaufmann, M., Egbert, S. and Leese, M. (2019) 'Predictive policing and the politics of patterns', *The British journal of criminology*, 59(3), pp. 674–692.
57. Kazim, E. *et al.* (2021) 'Systematizing audit in algorithmic recruitment', *Journal of Intelligence*, 9(3), p. 46.
58. Kleinberg, J., Mullainathan, S. and Raghavan, M. (2016) 'Inherent trade-offs in the fair determination of risk scores', *arXiv preprint arXiv:1609.05807* [Preprint].
59. Knoechelmann, M. (2022) 'DSM-5 personality disorder terrorism risk assessment: A four-part validation study'.
60. Lum, C., Koper, C.S. and Telep, C.W. (2011) 'The Evidence-Based Policing Matrix', *Journal of Experimental Criminology*, 7(1), pp. 3–26. Available at: https://doi.org/10.1007/s11292-010-9108-2.
61. Mali, B., Bronkhorst-Giesen, C. and Den Hengst, M. (2017) *Predictive policing: lessen voor de toekomst: een evaluatie van de landelijke pilot*.

62. Montasari, R. (2023a) *Countering Cyberterrorism: The Confluence of Artificial Intelligence, Cyber Forensics and Digital Policing in US and UK National Cybersecurity*. Springer Nature.
63. Montasari, R. (2023b) 'The Application of Big Data Predictive Analytics and Surveillance Technologies in the Field of Policing', in *Countering Cyberterrorism: The Confluence of Artificial Intelligence, Cyber Forensics and Digital Policing in US and UK National Cybersecurity*. Springer, pp. 81–114.
64. Mugari, I. and Obioha, E.E. (2021) 'Predictive policing and crime control in the United States of America and Europe: trends in a decade of research and the future of predictive policing', *Social Sciences*, 10(6), p. 234.
65. Nadikattu, R.R. (2016) 'The emerging role of artificial intelligence in modern society', *International Journal of Creative Research Thoughts* [Preprint].
66. Neiva, L., Granja, R. and Machado, H. (2022) 'Big Data applied to criminal investigations: expectations of professionals of police cooperation in the European Union', *Policing and Society*, 32(10), pp. 1167–1179.
67. Nowotny, H. (2021) *IN AI WE TRUST: power, illusion and control of predictive algorithms*. John Wiley & Sons.
68. Oatley, G., Ewart, B. and Zeleznikow, J. (2007) 'Decision support systems for police: Lessons from the application of data mining techniques to "soft" forensic evidence', *Artificial Intelligence and Law*, 14(1–2), pp. 35–100. Available at: https://doi.org/10.1007/s10506-006-9023-z.
69. Ongsulee, P. *et al.* (2018b) 'Big Data, Predictive Analytics and Machine Learning', in *2018 16th International Conference on ICT and Knowledge Engineering (ICT&KE). 2018 16th International Conference on ICT and Knowledge Engineering (ICT&KE)*, Bangkok:

IEEE, pp. 1–6. Available at: https://doi.org/10.1109/ICTKE.2018.8612393.

70. Oosterloo, S. *et al.* (2018) 'The Politics and Biases of the "Crime Anticipation System" of the Dutch Police', *Proceedings of the International Workshop on Bias in Information, Algorithms, and Systems*, 2103, pp. 30–41.
71. Orana, A. and Perteshi, S. (2022) *Disengagement, rehabilitation and reintegration of foreign terrorist fighters*. DRIVE. Available at: https://qkss.org/en/publikimet/disengagement-rehabilitation-reintegration-foreign-terrorist-fiighters (Accessed: 19 July 2023).
72. Panelli, J. (2018) 'Ethics of Predictive Policing', *Viterbi Conversations in Ethics*, 2(2).
73. Pearse, N. (2019) 'An Illustration of a Deductive Pattern Matching Procedure in Qualitative Leadership Research', *Electronic Journal of Business Research Methods*, 17(3), p. 264. Available at: https://doi.org/10.34190/JBRM.17.3.004.
74. Penner, A.M. and Saperstein, A. (2015) 'Disentangling the Effects of Racial Self-identification and Classification by Others: The Case of Arrest', *Demography*, 52(3), pp. 1017–1024. Available at: https://doi.org/10.1007/s13524-015-0394-1.
75. Perry, W.L. (2013) *Predictive policing: the role of crime forecasting in law enforcement operations*. Santa Monica, CA: RAND.
76. Pingen (2023) *Council's Common Position on Artificial Intelligence Act*, *eucrim*. Available at: https://eucrim.eu/news/councils-common-position-on-artificial-intelligence-act/ (Accessed: 28 July 2023).
77. Pressman, E. *et al.* (2016) *VERA-2R violent extremism risk assessment version 2 revised: A structured professional judgement approach*, *europa.ca*. Available at: https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran/collection-inspiring-practices/ran-practices/violent-extremism-risk-

assessment-version-2-revised-vera-2r-pressman-rinne-duits-flockton-2016_en (Accessed: 3 August 2023).

78. Purves, D. (2022) 'Fairness in Algorithmic Policing', *Journal of the American Philosophical Association*, 8(4), pp. 741–761. Available at: https://doi.org/10.1017/apa.2021.39.
79. Ray, S. (2019) 'A Quick Review of Machine Learning Algorithms', in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Faridabad, India: IEEE, pp. 35–39. Available at: https://doi.org/10.1109/COMITCon.2019.8862451.
80. Rettenberger, M. (2016) 'Die Einschätzung der Gefährlichkeit bei extremistischer Gewalt und Terrorismus', *Kriminalistik*, 70.2016(8–9), pp. 532–537.
81. Richardson, R., Schultz, J.M. and Crawford, K. (2019) 'Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice.', *NYUL Review Online*, 94(2), pp. 192–233.
82. Russell, S.J. and Norvig, P. (2021) *Artificial intelligence: a modern approach*. Fourth edition. Hoboken: Pearson (Pearson series in artificial intelligence).
83. Santos, R.B. (2013) *Crime analysis with crime mapping*. 3rd ed. Thousand Oaks, Calif: SAGE.
84. Scanlan, J. (2019) 'Auditing Predictive Policing', *Brigham Young University Prelaw Review*, 33(1), p. 4.
85. Schoenherr, J.R. *et al.* (2023) 'Designing AI Using a Human-Centered Approach: Explainability and Accuracy Toward Trustworthiness', *IEEE Transactions on Technology and Society*, 4(1), pp. 9–23. Available at: https://doi.org/10.1109/TTS.2023.3257627.

86. Sen, M. and Wasow, O. (2016) ‘Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics’, *Annual Review of Political Science*, 19(1), pp. 499–522. Available at: https://doi.org/10.1146/annurev-polisci-032015-010015.
87. Shapiro, A. (2017) ‘Reform predictive policing’, *Nature*, 541(7638), pp. 458–460. Available at: https://doi.org/10.1038/541458a.
88. Simmler, M. *et al.* (2023) ‘Smart criminal justice: exploring the use of algorithms in the Swiss criminal justice system’, *Artificial Intelligence and Law*, 31(2), pp. 213–237. Available at: https://doi.org/10.1007/s10506-022-09310-1.
89. Sommerer, L.M. (2022) *Self-imposed Algorithmic Thoughtlessness and the Automation of Crime Control: A study of person-based predictive policing and the algorithmic turn*. Nomos Verlagsgesellschaft mbH & Co. KG. Available at: https://doi.org/10.5771/9783748929550.
90. Sprenger, J. and Brodowski, D. (2023) ‘“Predictive policing”, “Predictive Justice”, and the use of “Artificial Intelligence” in the Administration of Criminal Justice in Germany’. Available at: https://doi.org/10.22028/D291-39980.
91. Strikwerda, L. (2021) ‘Predictive policing: The risks associated with risk assessment’, *The Police Journal: Theory, Practice and Principles*, 94(3), pp. 422–436. Available at: https://doi.org/10.1177/0032258X20947749.
92. Trist, E.L. (1981) *The evolution of socio-technical systems*. Toronto: Ontario Quality of Working Life Centre.
93. Ugwudike, P. (2022) ‘AI audits for assessing design logics and building ethical systems: the case of predictive policing algorithms’, *AI and Ethics*, 2(1), pp. 199–208. Available at: https://doi.org/10.1007/s43681-021-00117-5.
94. UNDP (United Nations Development Programme) (1994) ‘Human Development Report 1994’, *UNDP (United Nations Development*

Programme) [Preprint]. Available at: https://hdr.undp.org/content/human-development-report-1994 (Accessed: 12 August 2023).

95. Van Der Woude, M. (2023) 'Securitizing mobility: Profiling "non-core" Europeans', *Security Dialogue*, p. 096701062211250. Available at: https://doi.org/10.1177/09670106221125041.
96. Vepřek, L.H. *et al.* (2020) 'Beyond Effectiveness : Legitimising Predictive Policing in Germany', *Kriminologie - Das Online-Journal | Criminology - The Online Journal*, pp. 423-443 Seiten. Available at: https://doi.org/10.18716/OJS/KRIMOJ/2020.3.3.
97. Waters, P., Veal, C. and Raper, M. (2023) *The perils of feedback loops in machine learning: predictive policing*, *Lexology*. Available at: https://www.lexology.com/library/detail.aspx?g=c8fff116-2112-48dd-841c-f9d1688d722b (Accessed: 23 August 2023).
98. *We Sense Trouble* (2020). Amnesty International. Available at: https://www.amnesty.org/en/documents/eur35/2971/2020/en/ (Accessed: 21 July 2023).
99. Willems, D. and Doeleman, R. (no date) 'Predictive Policing–wens of werkelijkheid.', *Het tijdschrift voor de Politie*, 76(4), p. 5.
100. Wilson, G. (2011) 'Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage', *Contemporary Sociology: A Journal of Reviews*, 40(3), pp. 306–307. Available at: https://doi.org/10.1177/0094306110404515m.
101. Yanow, D. and Van Der Haar, M. (2013) 'People out of place: allochthony and autochthony in the Netherlands' identity discourse — metaphors and categories in action', *Journal of International Relations and Development*, 16(2), pp. 227–261. Available at: https://doi.org/10.1057/jird.2012.13.

102. Yetim, F. (2011) ‘A Set Of Critical Heuristics For Value Sensitive Designers And Users Of Persuasive Systems’, *ECIS 2011 Proceedings*, p. 185.

103. Yin, R.K. (2014) *Case study research: design and methods*. Fifth edition. Los Angeles: SAGE.

Appendix

Appendix A- Interview questionnaires

Interview with Dr Lucia Sommerer

1. In your book "Self-imposed Algorithmic Thoughtlessness and the Automation of Crime Control," you discuss the concept of "person-based predictive policing." How would you define this concept?

2. With the introduction of facial recognition systems and processing larger volumes of data across multiple surfaces including social media platforms, could you explain its potential implications for civil liberties namely human security?

3. The EU Parliament are working on developing the AI Act. This act is focused heavily on AI Safety Privacy with certain exemptions being given to open source AI as well as some forms of predictive policing. However there has been significant discussion about how this can be put into practice. How do you think law enforcement agencies ensure that predictive policing models are used in a manner that is compatible with human rights and anti-discrimination laws? (European Parliament finally passes its AI Act • The Register)

4. When investigating German, UK or Dutch predictive policing systems, were you able to identify any particular biases that feel were appropriately accounted for when designing these systems? Are there any you felt were NOT accounted for in the process or practice?

5. Do you believe that domestic law enforcement agencies and state actors in the EU have taken steps or have a roadmap to effectively address these biases now or eventually?

6. For design and development software organisations form the epicenter even when the system is developed in-house. However not much importance is given to other stakeholders such as educational or human rights organisations (e.g. Amnesty International) even when they contribute extensively through research to the improvement of such sociotechnical systems. Do you feel that effective collaboration between state agencies and external stakeholders is vital for such systems to succeed?

7. What legal frameworks exist in the EU to regulate the use of predictive policing, and do you see any gaps or areas for improvement in these frameworks?

8. With some having already identified challenges in practice when it comes to legislations such as the AI Act, can you speak of any challenges in developing a shared EU-wide approach to the use of predictive policing? What are some of the opportunities standardisation of predictive policing systems might present to states?

9. How do you think predictive policing models should be audited and evaluated to ensure that they are fair, accurate, and effective?

10. What kind of counter bias measures would you think would help evolve predictive policing practices?

11. In your opinion, how can law enforcement agencies improve public trust when using predictive policing methods?

12. In the coming years, how do you see the legal and ethical landscape of predictive policing evolving in the EU, and what role can academics and researchers play in shaping this evolution?

13. Any drawbacks on standardisation? State level or EU level collaboration hindrance?

Interview with Dr Lorella Viola

1. In the context of predictive policing, how do you see the integration of digital technologies and methodologies from the field of humanities playing a role in enhancing law enforcement practices?

2. How do you think critical perspectives and methodologies from the digital humanities field address potential biases and ethical concerns in the development and deployment of predictive policing algorithms?

3. Can you provide insights into how digital technologies have transformed research practices in the humanities field, and what are the key implications?

4. Can you discuss any specific examples where humanistic insights have been utilized to critically evaluate and improve the effectiveness of predictive policing strategies?

5. How can data interpretation and contextual understanding, drawing from the humanities, enhance the transparency and accountability of predictive policing algorithms to the communities they serve?

6. What are some of the key challenges and ethical considerations that researchers and practitioners need to address when working with digital technologies such as predictive policing in the humanities?

7. How can incorporating digital humanities methodologies in predictive policing research and practice help address the cultural and social implications of these technologies in Europe?

8. When considering biases within big data that permeates into social constructs like predictive policing systems, how much can you attribute that to the lack of diversity in information sources? What would be other concerns beyond bias and how can one address that through effective collaboration?(anserewd)

9. You mentioned how audits have caused predictive policing systems to be discontinued in the USA, do you feel within Europe standardized practices can be developed or do you feel eventually when these systems are audited they would be discontinued as well?**

10. You talked about how one of the concerns for systems like predictive policing would be the limits of the system, even when AI and ML are concerned. Do you believe these limitations can be attributed more to the current state of socio technical development, the design and development process, or the selection of data?

11. In your opinion, what are the key considerations or challenges when incorporating humanistic approaches into the design and evaluation of practices like predictive policing algorithms?

Interview with Dr Gwen Van Eijk

1. What would you say are some of the key findings of Amnesty regarding predictive policing practices in the EU namely the Netherlands? Have you identified any specific human rights concerns?

2. What are some of the challenges that you have faced in advocating for the protection of human rights in the context of predictive policing? How has Amnesty sought to overcome these challenges?

3. How important would you say standardization on a state and EU-level for predictive policing practices would be for resolving these human rights violations? Do you think EU level standardization is possible?

4. What role do you think technological organizations can play in ensuring and upholding human rights standards in the context of predictive policing?

5. When looking at the work Amnesty International and even educational institutions are doing in this field, and the lack of state-level approach towards research, how imperative would you say collaboration with external actors is to ensure and uphold human rights standards? How can they collaborate with law enforcement and state agencies to address potential risks?

6. Would you be able to discuss any recent cases or examples where Amnesty has documented instances of human rights violations or abuses related to predictive policing practices?

7. Do you think it’s possible for EU-wide collaboration on this issue? What would you say are some of the roadblocks?

8. What are some of the ongoing initiatives conducted by Amnesty International that aim to raise awareness and promote accountability in relation to predictive policing?

9. Finally, do you see a future where predictive policing practices are beneficial without the current biases or drawbacks? What would you say are the core challenges that need to be overcome for such a future?

Appendix B- Interview transcripts

Interview transcript for Dr Lucia Sommerer

Prof Lucia: [00:00:00] Is it good like this?

Fariha: This is perfect. Actually. Can you hear me? Yeah.

Prof Lucia: Yeah. And just like, let me know if the sound quality gets, uh, worse during our talk or anything. Just let me know.

Fariha: All right. I will. Thank you so much. Uh, should I start the interview?

Prof Lucia: Yes, perfect. You can start the recording or whatever you need.

Fariha: Yes, uh, this is a recorded interview and I have started recording just now. Uh, thank you so much for agreeing to the interview. Uh, I'll start with a short introduction. My name is Praeha Mansoor. Uh, and I'm doing an Erasmus program called, uh, uh, International Master in Security, Intelligence, and Strategic Studies.

My thesis is focused on socio technical relations between Uh, predictive policing and human security, um, and your extensive research and expertise in the field of algorithmic systems and crime control make you an ideal guest, make [00:01:00] you an ideal guest for this discussion. As an esteemed professor, author, uh, and a thought leader, you have made a significant contributions to our understanding of ethical and societal implications of automated crime control systems.

So my first question would be. In your book, Self Imposed Algorithmic Thoughtlessness and Automation of Crime Control, you discuss the concept of person based predictive policing. How would you define this concept?

Prof Lucia: How I define person based predictive policing? Yes. Um, yes. It was important for me because generally, when I was writing the book, the public discussion about predictive policing was like all over the place, like everyone was talking about the negative effects of predictive policing in general.

And, uh, many people did not make the distinction [00:02:00] between person based and location based predictive policing. But it's really important to make that distinction because, um, there are different legal consequences and real world consequences attached to it. Um, you know, location based predictive policing, um, is, is just predicting the locations of future crimes, primarily burglaries.

And then the police squad is sent there more frequently to deter criminals that might come along. But, you know, it's not affecting an individual, um, an individual directly. It may affect them indirectly, but not directly. Person based predictive policing, however, is directly affecting an individual. It is when a computer, um, creates a risk score for an individual, for me and you, for our future likelihood [00:03:00] of committing, um, a crime.

Fariha: Um, this is something that I wanted to mention. Initially, I was only looking into location based predictive policing in Germany, but from your book, I, uh, got, uh, introduced to the concept of, uh, radar it, e di di, uh, r i a s, uh, and flight data pattern recognition. So I'm looking into those as well. Thank you so much.

Mm-hmm. , this brings us to our, my next question. Which would be with the introduction of a facial recognition system and, uh, processing larger volumes of data across multiple surfaces, including social media platforms. Uh, what do you think the potential implications for civil liberties would be, namely human security in this regard?

Prof Lucia: Um, well, privacy as a fundamental right [00:04:00] enshrined in German, uh, in the German constitution and in the EU, um, Charter of Fundamental Rights, privacy, um, would definitely be affected. Um, the right to equal treatment, um. Would be affected potentially if any of these technologies operate in a in a discriminatory way discriminating against some groups in the society For example making more mistakes with certain subgroups of society I'm sure you're very aware of Um, that there are studies how algorithms, um, especially facial recognition algorithms are, um, not as well trained on dark skin.

So the facial recognition, um, algorithm may [00:05:00] make a self positive more frequently, um, on a person with dark skin compared to a person with lighter skin. Just one example of many ways how an algorithm can. Uh, discriminatory. Um, so I expect impact on privacy on the right to equal treatment. Mm-hmm. , um, those are the main ones.

There are others. Um, as you've read in my book, I, I think it's in my book, you can also think about. Whether even aspects of human dignity, the rights to human dignity, which is enshrined in the German constitution, would be affected. You know, this doesn't immediately have like, um, an equal right on the EU level.

It's very special to German law. But you could consider if a person, a human, you and me, if we are reduced just to data points. Yes. To be fed into a [00:06:00] machine and then a prediction. Mm-hmm. , if and if no human is looking at, um, at us really, but just the computer. Then if our human dignity would be affected in, um, some way, um, that is one, one, uh, claim you could make under German law.

Fariha: Yes. If there is no human oversight, there might be. a greater chance of error, and we might be reduced to only numbers. Um, this brings me to my next question, which is, uh, the EU Parliament passed an AI Act four days ago. This Act focuses heavily on AI safety privacy with certain exemptions being given to open source AI, as well as some forms of predictive policing.

However, there, uh, has been significant discussion about how this can be put into practice. How do you think, uh, law enforcement agencies ensure that predictive policing models are used in a [00:07:00] manner that is compatible with human rights and anti discrimination laws?

Prof Lucia: Yeah. Um. Sorry, I wasn't sure. You said the EU Act was passed, right?

Yes. Um, um, are you sure it was passed or that it was just a preliminary, um, decision making on the Act? I think it

Fariha: was a preliminary decision making.

Prof Lucia: Yeah, yeah, yeah, right, because it, it, the leak, if you said it was passed, you know, it would already, like, be enforced. Yes. Yeah, it, it, we have to be... Um, careful with the right legal, um, language, um, and yeah, sorry, I'm, I'm, I'm sorry I was distracted.

Could you try to phrase the question again, just a little shorter, maybe?

Fariha: Um, the gist would be how do you think the law enforcement [00:08:00] agencies ensure that predictive policing models are used in a way that is compatible with human rights and anti discrimination laws?

Prof Lucia: Yeah. Um, you know, there has, uh, like recently been this judgment by the European Court of Justice on predictive policing.

Are you aware of that judgment? Uh, no, I am not. Yeah. Um, I'm just trying to, to look it up so I can send it to you. Okay. Um, there's some, there's some information in there that the European court of justice said, like, um, we in Europe were not allowed in law enforcement to use, uh, self learning AI, for example.

Um, so that would be interesting for you. And that would be one element of answering your question. Um, one way the police could ensure that it's in line with the [00:09:00] laws is to exclude self learning, um, self learning A. I. Um, and the big question being, what really is self learning? A. I. Is it any kind of machine learning or any or a special kind of supervised unsupervised machine learning.

That's the big question when you look at that judgment, but that's one thing. The other thing is I mean, I could tell you what they should do. Um, they're not doing that much. Like, I could tell you that they should set up an independent body to over, to, to check the algorithm and like regular intervals to make sure everything is going according To plan and to have this independent body to do that.

Um, that is not within the police organization itself, but like an independent, uh, body. [00:10:00] Like, I think that would be necessary, but we don't have that right now in Germany or anywhere else.

Fariha: Okay. Um, you actually answered my next question as well. Uh, so another question I have is when investigating German, UK or Dutch predictive

policing systems, uh, were you able to identify any particular biases that feel were appropriately accounted for when designing these systems?
Uh, are there any, uh, you felt were not accounted for in the process or practice?
Prof Lucia: I felt in general, what I've seen in my research back then, you know, that was, uh, 2018, 19 there wasn't a big awareness in within the police for potential biases. I more felt that the police thought [00:11:00] it, well, it is not a bias. And for example, foreigners. come up more frequently in an algorithm that is designed to predict terrorist crimes or Islamic, uh, extremism.
Um, I think the thinking of the police was more like, well, it makes, it makes sense. These are the people that commit these kind of crimes, so why should it be biased? It's, it's not discriminating, it's only reflecting reality. So I think that was more the thinking in the police at the, at the time, and maybe not very sensitized to the issue.
Fariha: Okay. Um, also, do you feel that effective collaboration between state agencies and external stakeholders, uh, such as, uh, human rights organizations like Amnesty International, uh, is, is it vital for such systems to succeed?
Prof Lucia: Um, I think it's [00:12:00] not happening right now. Um, there would be a benefit, I think, to have civil rights, um, representatives in an independent review body of these special police algorithms.
Um, that would be a benefit. But, you know, you would first have to have this independent body to do the review, and no one has set it up yet. Yes.
Fariha: Um, also with, uh, some having already identified, uh, challenges in practice when it comes to legalizations such as the AI Act, can you speak of any challenges in developing a shared EU wide approach to the use of predictive
Prof Lucia: policing?
Challenges? Um, well, the most obvious one is, is a legal one because, [00:13:00] you know, a lot of EU law is harmonized. In many areas, the EU tries to, um, have the same law, similar laws in every country. However, there's one big exception, and that is criminal law and policing. In this area, the member states in Europe are...

Very much. Um, it's very, the policing and the criminal law is very important to every member state. So this is kind of the last area where member states say and where the EU law says, you know, you don't have to make everything equal and the same. You know, you don't have to have the same police and criminal law in each and every member state.

Yes. So, um, that means that there are different standards for. Police law, criminal law in every member state. So that is a hindrance to [00:14:00] having a uniform approach to anything related to police in Europe. Um, you were talking specifically with respect to the EU AI Act. Yes. Um, yeah, it's, well.

Fariha: All right. Um, you, uh, you talked about having a human oversight, uh, organization or a mechanism that would, uh, sort of, uh, audit, uh, what's happening, uh, in the, in, in predictive policing and how the AI is learning. Other than that, do you think, uh, is, uh, What can the law enforcement agencies do to improve public trust when, uh, using predictive policing methods because there is a lot of skepticism, uh, surrounding it?[00:15:00]

Prof Lucia: Hmm. Yeah. I think that would be the most important one. Which is, you know, which plays into one notion of transparency. If you have an independent body, you are more transparent. Um, I think, uh, trust is related to transparency in a way, and, and that, that independent body is already one element in transparency, and that's, that's the only way you could try to persuade them, but it, it will be, it will be hard because you know there are studies who generally say that, um, um, You know, humans don't like it when a computer does something so far reaching and impactful as policing.

Fariha: Yes. And judging by, uh, the current studies and the literature that's already there, um, the current models or the [00:16:00] way predictive policing is working, it is based on a lot of historical data as well. And as you have mentioned, there is no... Body that would kind of audit the entire process. So this is absolutely very important.

Um, this brings me to my last question, which would be in the coming years. How do you see the legal and ethical landscape of predictive policing evolving in the EU? And what role can the academics and researchers play in shaping this evolution?

Prof Lucia: The movement will probably go maybe into neural networks more and more and more and more complex forms of

AI. However, that development is hindered by that judgment by the European Court of Justice that said [00:17:00] no self learning algorithms are to be used. But now it depends on how the legal scholarship and practice will interpret. The meaning of self learning system, you know, that depends on that very much. Okay.

And I think academia has a really important role to play to... You know, not just in the abstract, write about these issues, but to develop hands on recommendations with, so we can have a dialogue with academia and the police and the practice and the politicians. And our job is to make our research understandable and to communicate our research.

So our task is not just to do the research and write a really big... Book that no one is reading, but our task is also to go out and talk to people and [00:18:00] politicians about this.

Fariha: All right. Um, my questions were over, but I was wondering another thing, uh, which is if there are, let's say, standardized laws on predictive policing in EU, do you think there would be any drawbacks of that?

Prof Lucia: Um, not necessarily. I think that would be a good thing if there would be like standards for this. Um, but the fear could be, of course, that if so many countries must agree, you know, they will focus on the minimum standard, on the lowest standards possible and agreeable. So it could be that the standard is somehow too low, you know?

Yes.

Fariha: All right. Um, I think I'm all done with the questions. Um, thank you so much for taking the time to answer them so thoroughly. Um,

Prof Lucia: I will be... Thank you [00:19:00] for searching this cool

Fariha: topic. Thank you so much. I will be sending you a copy of the result after I'm done with my thesis and I'm done defending it.
But thanks once again.
Prof Lucia: Yeah, but I'm sure it will be fine. Um, the, the transcript and everything. And I'm just, let me just look up. Are you interested in knowing about the European Court of Justice judgment? I, I was just about to look it up. Yes, I am. I was wondering where I have it. Just a second. Yeah, sure.
Um,
there it is.[00:20:00]
It's in, it's in English. Okay. Okay. Like, okay, how can I, how can I send it to you? Like, oh, I can, I can just dictate. Okay?
Fariha: Uh, you can email it to me, uh, if
Prof Lucia: that's easier. I, no, I, wait, what was your email again? Let me look it up. Uh,
Fariha: it was, uh, I can email you from my email again. Would that be any easier? Um. I will send you an email right away.
Uh, so that it's, you can see which address I reached you out for.
Prof Lucia: Went to University of...
Fariha: It's University of Glasgow. I've already sent you an email.
Prof Lucia: Oh, Glasgow. I think my sister went to Glasgow as well. Oh, that's great.
Fariha: [00:21:00] And... What did she study? Uh,
Prof Lucia: I think Anthropology. Okay,
Fariha: that's great. Uh, I have sent you an email from 269 6514 M student. gla. ac. uk. Yeah,
Prof Lucia: there it is. Yeah.
So, I will just send you the, um, article. No, no, I'll just send you the name of the court decision. Okay. S E R... And the exact page where the information is on. Alright. And then I think you will be able to find it.
Fariha: Yes, I should be able to do so. If you just send me the name. That

Prof Lucia: would be fine. It's in the...

Yeah, and then just search the [00:22:00] document for like the... Signal words like self learning and it's the European Court of Justice has, it has the judgments in all kinds of languages. So it does have it in German, but also in English and French, like in every language, I think. Okay. That sounds perfect.

Cool. Yeah, and I'm, I'm happy to, you know. Um, see the finished product of your, of your research in the end.

Fariha: Oh, absolutely. I will be sending you a copy. Cool. Yeah,

Prof Lucia: I just, I just

Fariha: sent you. Thank you. I think I received it. Wait a second.

Prof Lucia: Um, I haven't... Yeah, it will be there. Yeah. In a second. My email is on the floor.

Fariha: Yes, I, I, I received [00:23:00] it. Thank you so much again. Okay. Thank you. Bye bye. Bye bye. Good night.

Interview transcript for Dr Lorella Viola

Fariha Mansur: Hello. yeah.

Lorella Viola (she/her): Hi, Video.

Fariha Mansur: Hi, can you hear me?

Lorella Viola (she/her): Yes, I can hear you. Okay, sure. Sure. Sure. Then I will stop mine. If that's okay, then

Lorella Viola (she/her): no problem

Fariha Mansur: So I'll start with a little bit of background about myself also. First of all, thank you so much for agreeing to the interview.

Fariha Mansur: my name is, and I'm doing an Erasmus program called International Masters and Security Intelligence and Strategic Studies. My thesis is focused on the social technical relations between predictive policing and human security as discussed. This is a recorded interview. But you can ask me to stop the recording anytime.

Lorella Viola (she/her): Okay.

Fariha Mansur: coming to. Why, I'm interviewing you your book, the humanities and the digital. It has been an absolute game changer for me in terms of understanding how technology, ethics and the humanities intersect. especially when it comes to practices like predictive policing.

Fariha Mansur: I decided to reach out to you because your expertise and insights fit perfectly with the teams I'm exploring in my thesis. Your work has given me some seriously thought provoking perspective on how digital technologies can transform different fields, including humanities, since predictive policing relies on advanced algorithms and data analytics. I believe your insights would be a very valuable for my research.

Fariha Mansur: so I'll just start with my first question. If that's all right.

Lorella Viola (she/her): Sure, please.

Fariha Mansur: in the context of predictive policing, how do you see the integration of digital technologies and methodologies from the field of humanities being a role in enhancing law enforcement practices.

Lorella Viola (she/her): Okay?

Lorella Viola (she/her): So okay, so obviously, the predictive policing method like other methods used as a technological methods used for let's say, bigger governmental decisions over integrated in the backing banking sector, for example, on the health security he to management, or a credit management, and so on, and so forth. are always in a way ambiguous in terms of how these methods are. In fact, developed, designed the algorithms shared. so let's say it's always a bit to take how these technologies are developed and almost nobody in a in the government, for example, as they as I think the expertise to understand and unpack these these algorithms, which are extremely, extremely complex.

Lorella Viola (she/her): So clearly, the way also the data sets for training. These algorithms are, in a way chosen, but also categorized and assembled and put together is always a bit to make and certainly controversial. So there are many, many problems with the method in general and I'm sure you are. You are well aware that already in a C in in the United States that had piloted this method, there have been many problems and some cities stop to the program altogether. Years of policy, because of course, of all the problems that that turned out to have to call but also because they externally audited this codes. And of course it was in a wake, as again controversial. And he, he ended up over policing neighborhoods. They were already over at least.

Lorella Viola (she/her): So the the coming back to your question about how the humanities could if I understood the correctly, the question how the humanities could contribute to the implementation of these methods. for for bigger governmental decisions. So clearly they' the the humanities have much to recommend, of course. however, I wouldn't really in a way, position that as a one either, or sort of the conversation. So either the sciences or the humanities.

Lorella Viola (she/her): I would recommend a much more holistic approach, and, in fact, training more humanity scholars, as we are already doing digital skills. But, on the other hand, also training more the scientists for humanities. also. in a way skills also. And mindset, and of course, critical skills to principally I'm not saying that hard side to not have this case, of course, I'm just saying the fact that we always we have trained to always critique and unpack. And really, in a way, try to understand really the the the mechanisms of the the rationale behind everything. whereas, of course, in computer science, the the goal is always to automate everything. Rather so My idea would be, or my not my my recommendation. But I I think, in my opinion.

Lorella Viola (she/her): what could be a game changer is really to introduce more, more transparency if it practically full transparency into into the dangers of blindly trusting these supposed to rigor, right? Because the technology is pitched as being absolutely so perfect and never infallible. But of course this is not true. And the developers of the of the technology know that very well. Obviously But my! So first of all, already already starting from a place of transparency into how the data sets were collected into, how they were trained into, how they were categorized. Also, would already open up at least the floor for discussion and for being, for for this technology, for these algorithms to be audited and in a way also studied, but also improved. Obviously, so this would be, this would be, of course, a a a big improvement, I think. already.

Fariha Mansur: I think that's a related question I to relate to the question you just answered. You mentioned how audits have cost predictive policing systems to be discontinued in the Us. Do you feel within the Europe? If there were standardized practices that could be developed in terms of predictive policing. Do you feel eventually, when these systems would be audited, they they would be discontinued as well.

Lorella Viola: What? When this? What did you say? Sorry. So when these systems will be audited? Yes, if there was a standardized practice within Europe in terms of predictive policing. naturally, there would be an auditing body eventually. Do you think they would be discontinued if the practices were audited? Well, obviously, I I I do not know, obviously. But looking at what has happened. the United States. perhaps. this that. There are several examples so predicting policing programs being discontinued right following not just all, but also lawsuits, and they despite that in in in the United States there are several cities that still adopt predictive policy. But again, in Europe. It's still it in my At the best of my knowledge, in in Europe. the cities that have started adopting predictive policing. We are still at the test or pilot stage. so

Lorella Viola (she/her): I don't know. I don't know whether it will ever go beyond the the pilots. The pilot state to be honest and and I don't know whether you I'm sure you know about this. But there isn't too much talk about this right? I mean, it's it's quite also opaque the whole.

Fariha Mansur: The fact that even some Ctc. As some some countries in in Europe already testing this. I don't think many people know about this. So I am doing my research on Germany and Netherlands. Specifically. So, I'm looking into location base and person base, predictive policing at the moment. coming to my next question.

You talked about how one of the concerns for systems like predictive policing would be the limits of the system even when A AI and Ml. Are concerned. Do you believe these limitations can be attributed more to the current state of socio-technical development or the design and the development process, or the selection of data?

Lorella Viola (she/her): Well, it's the 2 sides of the same coin, isn't it? So you can't really. So in order for you to understand how an algorithm works, you have to look at how the data must be prepared in order for the algorithm to work properly. And when I say properly, I mean automatically, so clearly, you can't really.

In my opinion, you can't really say, Oh, it's the algorithm's fault, or it's it's already you. We need to improve the data. But we can leave the algorithm as it is, or we need to improve the algorithm. We can leave the data as it is, what things need to be go go together right? So clearly, every time a data set is assembled and created. Whoever created that data set made specific decisions and specific choices at so many points during the collection and the categorization of the data set that. And I think these choices are hardly ever documented.

And it's very in any case, even if they are documented, that kind of documentation is not really shared. Perhaps it's left, is it? It stays internal. But it's not really clear how the researcher or the designer of the data set categorize the specific categories. and every time a, we put something in a category that obviously expresses a clear world view and that clear judgment that we are making about the specific category. And because we are talking about social categories. Here, we're talking about people. Basically. So we are, whoever creates that particular data set is making decisions on about who that you know that that specific person, that the specific human category should fit, and where should not. So so clearly? Again, I think transparency an open disclosure on how the datasets, together with how obviously the algorithm is developed and works should already be a big step forward. I would say.

Fariha Mansur: absolutely. Yeah. My next question would be, what are some of the key challenges and ethical considerations that researchers and practitioners need to address when they're working with digital technology, such as predictive policing?
Lorella Viola (she/her): so every so so the the what I talk about in my book is that there is a little still awareness, of the fact that these whatever we create a digital object, this digital object becomes something in a way alive. And it's all right. and in this me that he has consequences. So what I call in the book is digital consequences. So this awareness, often is missing in whoever deals with digital objects embedded to the object. I mean everything right. It can be a a digital data set, and it can be. I don't know a tweet. It can be anything, of course, to here. We're talking about prediction policing. So in this particular case, the consequences are a normal for real humans, right? but what what I think is, still needs. I don't know. Stressing and introducing in the in the main discourse is that these technologies are far from being unbiased, and of course, They embed all the past decisions that people have in a way entrenched in the creation except for the algorithm. And this includes, even not just the biases of the person who who the other people who developed the the technology, but in fact, also, perhaps internal factors and dynamics in in the company that you know hired the people who develop to these algorithms right? There are also so many factors and agenda, so that certainly dictated the creation of this specific technology. And of course, on on top of everything, there is always economic interest rather than the good of of of our society. but also a political interest. So these are never are never really far away. And this technology is never developed, you know, within the government, right to the government. it a obviously outsource. This is to to companies external companies, you know who's business model is, of course, to make it money. So so clearly this, this is something that we always have to they are in mind very clearly that the the main goal or at least the rational behind which this this technology was developed is

to make money and every every time we use this technology, we are, in fact introducing some consequences. and that is something that I talk about in the book is that Well done. A boy the calls this location of liability. which is the fact that because this automated system feels like you, you have no control over it, right? If it's like all the computer said it. So it must be true. It's not my fault. So this is evidence-based. The decision. It's accurate, and it's UN biased. So because of these predominant, this course and the consequence, the consequences of using this technology are. in a way, given, you know, without any remorse to to to to this system, that that that create the made of the decision. So and this is what what they are, Dian boy, the call this location of liability. So so, for which decision makers are distanced from the humanity that is affected by this automated procedures. and this is something that obviously we need to. We need to to to talk about. And we need to introduce in the main in the main discourse, this could be our contribution.
Fariha Mansur: So when you talk about con concerns beyond bias in systems like predictive policing. how do you think in real life we can address that? And do you think effective collaboration would be of any benefit in terms of that?
Lorella Viola (she/her): So certainly. Well, I don't know what you mean by collaboration, whether you mean interdisciplinary collaboration, or with the government, or what kind of collaboration do you mean?
Lorella Viola (she/her): Hmm, yes, of course, absolutely. It's Again. A holistic approach also means tackling the the the issue from many angles, right from different point of view also from from from different areas, also society. So yes, Academia. Of course, the research government Ngos, I think it's this need needs addressing. Because, again. we know, in fact, very little about this technology still, and how it evolves, and because it evolves very quickly and very unpredictably. Also, we should be extremely careful on how, when these systems are used to to make societal decisions. So absolutely, there should be more more again, transparency, because you know it. You know the for collaboration to occur. You have to be transparent on how they use and

implement systems. So absolutely. this, this? yeah, in absolutely. This could be also hoped for. Right? Yeah.

Fariha Mansur: can you discuss any specific examples where humanistic insights have been utilized to particularly evaluate and improve the effectiveness of predictive policing strategies.

Lorella Viola (she/her): no, I don't think so. I not nothing to jump to to to mind, I think mostly is there? There have been many. Of course, the lawsuits, but there have been a people, of course, for testing, and this come from a several bodies, including Ngos, but also just people getting together at that because they were directly affected by the user predictive policing and because of the protests. then, you know there was a official auditing of these of these methods, but I don't think the humanity from the humanities, at least the field the as as far as I know, I know that there are groups. that, for instance, in in philosophy, did they? are involved with the ethics ethical issues in technology and in the use of particularly artificial intelligence for decision making processes. but I cannot. Yeah, no, I don't think so? No, but at least they're related specifically to predictive.

Fariha Mansur: you pretty much answered all the questions I asked and the ones I didn't ask. But to the last one, I will ask. in your opinion. what do you think are the key? Considerations or challenges are when incorporating humanistic approaches into the design and evaluation of practices like predictive policing algorithms.

Lorella Viola (she/her): Right? So so again, one I think, significant the contribution that the humanities has to offer to this? overweight, overwhelmingly digital landscape. Is that The constant reminder you know this in a way, content balance is this positivistic, this course, that still equals the removal of the human with the, you know, the promise of objectivity and fairness.

So really, now is this, Illusory? Right? Because it's just an illusion that we can always get rid of biases for as long as technology again, is is created by by humans. And we always be biases in the technology. so so absolutely, the our contribution as humanity scholars is the Again the the, the, the counterbalancing, a counterbalancing narrative that, reminds everyone that really, ironically, these systems are in desperate need of humans to to be unbiased, you know, or at least to to to reduce, to reduce the biases. So so yes, this would be one, absolutely 1 one bigger contribution. and again, this. it's not really like, I say, my book, this, you know a a sexist, racist and homophobic, a digital society. It's not so much a reflection. Human subjectivity in data and algorithms. The proof of it's that's what I say in the book, right to that it's we just pretend to that. If we use the technology, then we are, we're we're good to go. But it's it's quite the opposite. In fact. So this is this would be our our contribution. I think that makes absolute sense, because, it's ironic that we want an AI or AI system to do it. but the thing is, there would be a need of human oversight. Anyways, in this. Absolutely. Yeah, you can you? It's a again, at least at the moment. Right? We don't know. We don't know in the future the technology evolves up like blistering pace, really. But again, is it's a it's it's this this. The systems are really in desperate need of of few months to be in fact, supervised and constantly critiqued and checked and adjusted, and and so on, and so forth. So is again it, whatever I see some article talking about technology in the news. I I always see how it's just in in the collective consciousness that, you know, technology is a perfect to the computers to never get anything wrong. The everything is, you know. As for as long as a computer says it, then you know, we cannot argue against So I don't know. I think this this really needs it needs changing. And work like yours. for example, are already a clear indication that there is an awareness that this is not true. and the hype around the technology needs recycling. So so so this is good, I think. slowly but surely we get there hopefully.

Fariha Mansur: Absolutely I would like to thank you again for taking time off your busy schedule for this and I will be sending you a copy of my thesis after it's done.

Lorella Viola (she/her): Oh, thank you.

Fariha Mansur: But yeah, thanks again.

Fariha Mansur: I hope you have a lovely day.

Lorella Viola (she/her): Oh, thank you, you, too, and the good luck with your studies that you

Lorella Viola (she/her): bye, bye, bye, bye, bye, bye.

Interview transcript for Dr Gwen Van Eijk

Fariha Mansur: Hello! Hello! I Hello!

Hi, Gwen, can you hear me?

Gwen van Eijk: Hi. yeah. Good holiday.

Fariha Mansur: Hope you're doing well.

Gwen van Eijk: Yes, how are you? I'm doing good, too. Thank you for agreeing to the interview.

Fariha Mansur: My name is, and I'm doing an Erasmus program called International Master and Security intelligence and strategic studies. My thesis is focused on the socio-technical relations between predictive policing and human security for a Netherlands. I'm focusing on Cass. there are 2 are, and I are 46. So people, base and person location based predictive policing as discuss. This is a recorded interview. But you can ask me to start the recording any time.

Gwen van Eijk: yes, I would prefer if you wouldn't record actually, although, yeah, I realized that I usually allow recording of voice. So if you want, I can turn the video off. Is that all right? No, that that that's fine. I yeah, I sort of. But then I, yeah, I realized that. yeah, for interview. I it's very inconvenient for you. You would have survived. it's okay.

Fariha Mansur: But if if if that helps. This is only for my thesis, and after the thesis is done every like recording would be destroyed, and I would be send you up you a copy of my thesis as well.

Okay, that that would be great, anyway. Okay, no problem. So I read a mystery report, the xenophobic machines. And we sense trouble, and I'm really honored to have the opportunity to discuss these important issues with you and gain insights from your expertise. I believe our conversation will take greatly contribute to my research. So shall we begin?

Gwen van Eijk: Sure I have. I I I can help. So let me know what you want to know.

Fariha Mansur: Okay, so my first question would be, what would you say? Are some of the key findings of, unless you regarding predictive policing practices in EU, namely, the Netherlands, and have you identified any specific human rights concerns?

Gwen van Eijk: Yes, I would say generally what we found is that First of all, there is no legal basis to do predicted policing. So it's based on very general policing. the police act in an evidence which does not allow for any impact on on that nice price on the fundamentalized So that's where we're the problem starts and then we see also in practice that there is no regulation because there's no legal basis there. So recommendation, and that, you know, oversight. so that's that's the sort of framework and then, as it comes to the tools that are, use themselves we find that they are discriminatory and In some cases also they? so that's a human rights issue of for discrimination. obviously. And then in some cases, I also tend to process personal data in a way that it didn't. So it's a privacy of of people. That's what we found in the recent trouble report, which was a predictive policing experiment by the Dutch police, which was focused on as of the European citizens. The tracking license plates and the country code of the license place and also focused on just by design on people who are not designing. And it's because it's used. The cameras at the Amdr cameras, the automated fight number. So the cameras about the highways. it recorded a a vehicle. coming into the city, and okay.

Fariha Mansur: hmm. For my second question, I would say, what are some of the challenges that you have faced in advocating for the protection of human rights in the context of predictive policing. And how has Msc. Sought to overcome these challenges?

Gwen van Eijk: well, first of all, I a huge challenge in in addressing this issue is that we often don't have the details of how it works. Kind of data are used. are there any evaluations done?

And so I think, yeah, we were lucky for the predicted policing experiments and and models. and also for the xenophobic machine to have to have the data. But often we see the other countries as well, and also these other models that you just mentioned that you're looking into. We have quite some trouble. yeah, I'm covering the relevance details. So that's a problem. When you want to advocate for a human rights based user predicted policing. Or Perhaps you want. We want to. we have call for the then on the and the context of a new and it's very difficult to do advocacy when you don't have details, because when you want to explain what's the problem with these instruments, you want to go beyond hypothetical? fares and problems. it may be violating privacy. that doesn't scare policymakers enough. Because I believe that for efficiency reasons and for security users the So that's yeah. That's a struggle that we that we have. And in in advocating, that's policy members tend to think. it's yeah. It has value in terms of efficiency and security that really helps assigning police where they are needed most, that it helps in preventing crime. And then, when we point out the the problems and the dangers of predicted that we have trouble that really giving out the details. Information in general, we have no evidence to make those claims.

Fariha Mansur: So would you say, it's the lack of transparency that makes it most difficult?

Gwen van Eijk: Yeah, definitely. Yeah. And it doesn't mean that that our arguments are not that substantiated enough, but they might be based on what we know from the Us. Saying or that that are not to be used anymore. that goes for the instrument step to touch reports. About they're not no longer in use. So you're always talking about historical examples or examples elsewhere. yes, I mean, I would say.

Fariha Mansur: even though I'm just doing a research on predictive policing. It's very difficult to get relevant research done on EU. I have to revert to a lot of research done like US-based research on predictive policing.

Gwen van Eijk: Yeah.

Fariha Mansur: yeah.

Fariha Mansur: actually, I'm coming to Jose the topic. So that's that's a challenge, it is. But I mean, this is why I chose the top topic, because there was a gap where I could contribute

Gwen van Eijk: sure, yes.

Fariha Mansur: yeah, coming to the third question, how important would you say standardization on a stage and EU level for predictive policing practices would be for resolving these human rights violations. and do you think that EU level standardization is even possible.

Gwen van Eijk: I think the there the the chances of of the prohibition are individual based. it looks very realistic that that's going to happen, my worry is that maybe not all recognized by. They might be having a they might have an algorithm. They might be automated. But when it's something gospel or find else artificial intelligence that depends on how you get the definition. So in a way, AI act will help to standardize But the government will try to find a way around this. That's my expectation by saying it's not. It's just a single algorithm. it doesn't have. So that's the word, And but but then, yeah, everything. The success of the standardization will also depend a lot on the oversight oversight body is capable of actually investigate you. And do we? Is there enough transparency for us to know what's happening and for us, and join us, for example, to investigate because we know that our very important, that you said how the policymakers even define AI like what they consider to be AI. In the first place. Yeah, it's because of the the 3 instruments. And the evidence that you mentioned to us is definitely AI and and the Dutch police will and also classify. It's a self, Marian algorithm. They they've already written. Now that it's machine learning. So they cannot really back out of that definition. but the other 2 instruments? I think they're they're definitely not self learning the very simple algorithms that I think they might be the previous. I don't know if that would classify. That's AI. So that that.

Gwen van Eijk: But there are some regulations for algorithms in on the Dutch Level National as well. no, I am actually

Fariha Mansur: I I would actually look into more like for the Dutch the country specific what you mentioned the regulations might be, because every every country would have their own specific regulations in terms of that as well.

Gwen van Eijk: yes. Although the AI Act pro overall substandard regulations. As it comes to AI But then in the Netherlands due to the child benefit scandal. That due to that schedule and new regulations have been called for by department for algorithms. So that's a bit of a lower standard which would yeah, include all the some. So there's a a new oversight body. It's up and running. Not yes, I mean, it's been installed that doesn't have with asking. Yes. but there, yeah, it will. They're planning to give it full capacity and full hours in the future. So a new upside body. an algorithm, registry? that would have to include yeah, all the the algorithms by in the public sector. But there may be exceptions for the because of the security reasons. because of their that if people know the risk and there will be a mandatory human rights impact assessment for all algorithms. And we still haven't seen a plan to actually implement this but that's what the party wants at for that. So it's going to happen. We think they are very important. what we think is still missing in the regulation on the Dutch level and also which will also not be clear in the is that to prevent the discrimination? it should not be enough to use a race or nationality or equivalent factors as indicators in this model. and also it shouldn't be allowed to be self learning algorithms to my. if they are self learning in the where they? So that sort of should be a prohibition on the got to the things that happened I've described in the.

Fariha Mansur: you know. Thank you. my next question would be when looking at the work that am nesty, international and even educational institutions are doing in this field. In terms of research, we can see that there is a lack in state level approach towards research and predictive policing. How imperative would you say collaboration with external actors is to ensure and uphold human rights standards

Gwen van Eijk: for the for the government to yes, we we think that sorry. important. It's. I think it's important to to consults in fact, it's people and communities. when you develop any policy specifically. specifically when it comes to has had a problem of racial programming and still still has. So that's still the problem of in general.

Gwen van Eijk: so I think that would be. We always have expertise on human rights. They tend to look at these instruments. Maybe at the best, from an ethics viewpoint. this is. But yeah, we're talking about rights that are actually in the International. I have to abide to the European Convention. so I think that we? Yeah, we could. We could, really be up there. although on the other end, I think government should have that actually inhabs. I should actually 10 more.

Fariha Mansur: Hmm, what role do you think technological organizations can play in ensuring and upholding human rights standards in the context of predictive policing?

Gwen van Eijk: yeah, because we see private. Yes, private actors. actors, and we see. So I think my general assessment would be that it's the same thing at at best they look at it from a viewpoint of of ethics.

Fariha Mansur: I know

Gwen van Eijk: it's not. It's not something that. and all discrimination in my design. I still. I think that that's not the first issue. Sorry it's still was our goal. And then, oh, yeah, we want to. Do, you know. Yeah. So I think that's that's something to be included in. included more in education.

Fariha Mansur: okay, so we already talked about EU wide collaboration in terms of predictive policing. What would you say? Are some of the major roadblocks in terms of that?

Fariha Mansur: Let me rephrase that. Do you think it's possible for an EU wide collaboration in terms of predictive policing. And if that happens, what would you say the major challenges might be.

Gwen van Eijk: do you mean in terms of policing or the National Police Department, or maybe. I don't know. This is more. I don't think much about this question, because I don't know much about it. I know that the Dutch police works, together with several other police departments in your. They do this they call perfectly. And then to yeah, we do this couple of days where they look for outstanding. I to find drugs bye and mobile. So they do work together, and they do look at each other. I think, yeah, it costs tomorrow, as you know.

Fariha Mansur: Okay, thank you. Next question would be, what are some of the ongoing initiatives conducted by amnesty international that aim to raise awareness and promote accountability in relation to a project of policing.

Gwen van Eijk: So we Our work consists of a couple of things we do research. we do login the political level. And we, you come here. Useful technological issues are some of the challenge to do public campaign, because it's a difficult topic for people to understand. but at the EU level there is a advocates for A human price based. AI act and we've called for prohibition. on all forms of just today there was a call published on the website. And just to make sure that I have the screen.

Fariha Mansur: Yeah, sure, no problem.

Gwen van Eijk: Okay? yeah. So that's that's our most important.

Fariha Mansur: Okay? Yeah. coming to my final question. I can sense your stand on predictive policing. But still, do you see a future where predictive policing practices can be actually beneficial without biases or drawbacks. And if that future ever comes. what would you say? The core challenges would be like some of the road blocks on the way.

Gwen van Eijk: yes, I don't realize that. That's yeah. My, my, my, my position may be this is a bit intense, but I'll make it for and which is also my mind with with the viewpoint as it has progressed since the but we haven't made public statements in that much detail at all. Not but I mean, what? What will remain a problem. Whatever the problem is that these models are using historical police data. So even if they are not discriminatory by design, for example, targeting or targeting, as it's the case with the matrix. you've come across that reports, are you from? even if it's not discriminatory design using, it's it's going to be very problematic because of yeah, I'm sorry.

Gwen van Eijk: Disadvantaged neighborhoods. I don't see really how you would overcome that. But also imagine that there are a very definite demographic where? this conference wouldn't be so. that where you wouldn't have these problems. So it depends. I mean, as long as it's it's focused on what we call calling criminology or street. I think that I could. Yeah, with the half of our time and And then, when I was another issue of basically, if I don't, if you don't, you don't you? You are yourself, and you're not, you know more than just like others in terms of certain characteristics. That is what we're. It's lumps together. People who have similar characteristics. which is yeah, just not in my. but I will always be a problem. so I don't. I don't really see how you would. Yeah. Fundamental. Call them to It might be a little bit different when you're looking at nice focused and it's a pleasing. But then, again, it depends a lot on what you're trying to predict.

Fariha Mansur: Okay.

Gwen van Eijk: yeah, I. I'm just yeah.

Fariha Mansur: no, that makes complete sense. I'll some of the things that you've said today. It's it's making me think. rethink some of the concepts that I was initially working on. But thank you. Thank you for everything you said To keep the interview strictly within 30 min. So that's what I'm gonna do. I'm done with all my questions, and thank you so much once again for for joining with me today. Yes, I appreciate that. You kept our time to the limits. And I'm sorry I don't have time to to talk. I wanted to.

Gwen van Eijk: and you have a deadline yourself, so I hope it's been helpful. And should you come across any, you know, maybe have a follow up question Do email me? I will be away for a couple of weeks. And alright. Okay, good luck. Thanks again. Bye, bye.

Gwen van Eijk: bye, bye.