Artificial Intelligence in Law Enforcement Surveillance: Citizen Perspectives, Resistance and Counterstrategies

Yasmine Ezzeddine

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

November 2024

Candidate Declaration

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.
2. None of the material contained in the thesis has been used in any other submission for an academic award.
3. I am aware of and understand the University's policy on plagiarism and certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged. Data used in *Paper 1, Paper 2* and *Paper 3* stems from joint collaborations as part of the multinational research project AIDA (funded under EU Horizon 2020 program, grant number 883596). I played a major role in the data collection, and the analysis, and write-up of the publications are entirely my own work, supported by my DoS. Contributions from researchers of partner countries are explicitly stated and acknowledged in the publications.
4. The work undertaken towards the thesis has been conducted in accordance with the SHU Principles of Integrity in Research and the SHU Research Ethics Policy.
5. The word count of the thesis is 25,736 exc. Appendices.

Name	Yasmine Ezzeddine
Date	November 2024
Award	Article-based PhD in Computing
Research Institute	CENTRIC
Director of Studies	Petra Saskia Bayerl

Foreword

”بسم الله الرحمن الرحيم وَمَا تَوْفِيقِي إِلَّا بِاللَّهِ عَلَيْهِ تَوَكَّلْتُ “

As I write these words, my heart is heavy with sorrow for my beloved Lebanon, a beautiful country that is now engulfed in war. The sounds of conflict echo in my thoughts, awakening the similar atrocities that I witnessed as a child. My prayers are with my family, my people, and my homeland. This thesis, written during such turbulent times, is a testament to the strength that comes from the love and hope I hold as a Lebanese person, for a better tomorrow.

To my dearest mother, you sacrificed everything so that I could stand here today. You raised me with unwavering dedication, guiding me through every hardship, supporting my dreams with a strength I will never be able to repay. You gave me everything when you had nothing, and it is your love and sacrifices that have shaped the person I have become.متشکرم مامان بدون تو ممکن نبود.

To my soulmate, my husband, Walid—this journey would not have been possible without you. You sold the business, left your home, and followed me so that I could chase my dreams. Your endless encouragement, patience, and belief in me gave me the courage to keep going. I am forever grateful for your sacrifices, for you are the reason I was able to pursue this PhD. Thank you for standing by me through every step of this journey. Always and forever.

And to my beautiful baby boy, Amir—this work is dedicated to you. At just six months old, you've already shown us the meaning of true love and hope, especially in these dark days. As you grow, I want you to know that everything is possible if you believe in yourself and work hard. And this thesis is a testament to the power of dreams, perseverance, and the love that surrounds you.

I am forever grateful to the guidance and support of my PhD supervisors, especially my DoS Saskia, my PhD colleagues, and to the wonderful CENTRIC family with whom I have shared the good, the bad and the sleeplessness!

May this work inspire hope for a future where peace, knowledge, and justice prevail.

Abstract

Artificial Intelligence (AI) has rapidly become a key component in the realm of security and law enforcement, offering unprecedented capabilities for enhancing safety, operational efficiency and crime prevention. However, the integration of AI into these fields has raised significant concerns among citizens, particularly around privacy, ethical oversight, and the potential for misuse. As AI-driven surveillance becomes more pervasive, it is crucial to understand the varied perspectives of citizens and the strategies they employ to navigate and resist these technologies.

Citizens' voices are essential in shaping ethical and responsible AI policies, as they reflect the concerns, values, and lived experiences of those directly affected by surveillance practices. Including these perspectives ensures that AI deployment aligns with public expectations, enhances transparency, and fosters greater trust between law enforcement agencies and the communities they serve.

This thesis explores citizens' reactions to the use of AI in law enforcement, focusing on their responses, resistance, and counterstrategies. This research achieves its aims and objectives through three studies, independently designed by embedding innovative mixed methods ranging from in-depth interviews to online experiments and privacy walks, this research provides new empirical insights by exploring the nuanced ways in which citizens engage with AI surveillance, articulate their concerns, and develop strategies to navigate or resist its impact.

The findings reveal a nuanced landscape of citizen perspectives: while some participants recognize the potential benefits of AI in enhancing security, many express deep concerns about the implications for privacy, data ownership, and the broader social impact of AI surveillance. The study identifies a range of viewpoints, from cautious acceptance to active resistance, highlighting the specific counterstrategies that citizens adopt to protect themselves from what they perceive as intrusive surveillance practices.

Briefly said, this thesis offers both theoretical insights and practical recommendations aimed at bridging the gap between citizens, law enforcement, and technology developers. It emphasizes the importance of transparency, the need for robust ethical frameworks, and the value of ongoing dialogue to ensure that AI technologies are deployed in ways that respect public trust and protect civil liberties, while still enhancing security.

Table of Contents

Key Terms and Abbreviations

- **AI (Artificial Intelligence):** Simulation of human intelligence processes by machines, especially computer systems, including learning, reasoning, and self-correction.
- **LEAs (Law Enforcement Agencies):** Organizations responsible for enforcing laws, maintaining public order, and protecting citizens.
- **Predictive Policing:** AI-driven methodology used to anticipate potential criminal activity based on data analysis.
- **Surveillance:** Observing or monitoring individuals, groups, or systems for security, management, or control purposes.
- **Privacy Walks:** Research method where participants engage with real-world surveillance environments to provide contextualized feedback.
- **Algorithmic Bias:** Systematic and repeatable errors in AI systems that lead to unfair outcomes for certain groups.
- **Facial Recognition Technology:** AI-based system capable of identifying or verifying individuals based on facial features.

- **Computer Vision:** Subfield of AI focused on enabling machines to interpret and make decisions based on visual data.
- **LLMs (Large Language Models):** Advanced AI models designed to process and generate human-like text based on extensive datasets.
- **Offline Surveillance:** Monitoring activities occurring in physical spaces through technologies like CCTV, biometrics, and drones.
- **Online Surveillance:** Monitoring digital activity, such as browsing habits and social media usage, typically facilitated by AIdriven algorithms.

Introduction

Background and Context

The integration of Artificial Intelligence (AI) into law enforcement and security sectors is fundamentally reshaping traditional surveillance and data practices. AI, defined as the simulation of human intelligence processes by machines, particularly computer systems, encompasses learning, reasoning, problem-solving, and perception (Russell & Norvig, 2016). This technological evolution is not only enhancing the capabilities of LEAs but is also fundamentally altering the landscape of public safety. Technologies such as facial recognition, biometric scanning, and automated number plate recognition (ANPR) are now embedded in the everyday operations of policing, providing officers with the ability to process vast quantities of data at unprecedented speeds (Ferguson, 2017; Norris & Armstrong, 1999). However, these advancements also give rise to a multitude of ethical, privacy, and civil liberty concerns, challenging traditional understandings of security and surveillance (Lyon, 2007; Macnish, 2021).

This research specifically investigates AI-driven surveillance technologies used in law enforcement, focusing on machine learning algorithms, computer vision systems, and predictive analytics. These systems are designed to identify patterns, detect anomalies, and provide actionable insights from vast datasets, often in real-time. A critical social factor influencing the adoption of AI in real-world scenarios is the public's perception of trust in these technologies. Trust plays a pivotal role in shaping citizen acceptance, as it reflects concerns about data security, accountability, and the perceived fairness of AI decisions.

Understanding and addressing these perceptions is fundamental to leveraging AI effectively in law enforcement contexts.

In fact, the debate over balancing security with the safeguarding of individual privacy has intensified as these technologies have become

more pervasive (DiVaio et al., 2022). Surveillance, methodically defined as "focused, systematic and routine attention to personal details for purposes of influence, management, protection or detection" (Lyon, 2007, p.14), now extends beyond simple observation to encompass predictive capabilities that pose new challenges to privacy and personal freedom. The shift from reactive to predictive policing models raises concerns about potential biases encoded within AI algorithms, leading to disproportionate targeting of certain demographic groups (Richardson et al., 2019).

The ubiquity of AI-powered surveillance technologies means that every action and reaction can potentially be monitored and recorded, turning private citizens into involuntary data contributors (Gates, 1996; Ball, 2002). This widespread data collection is pivotal for security purposes but also raises significant concerns about data ownership, user consent, and the ethical deployment of AI tools (Andrejevic, 2007; Petersen & Taylor, 2012). The narrative often oscillates between the enhanced capabilities provided by AI, such as predictive policing and crowd surveillance, and the chilling effects these technologies have on individual behaviour and societal norms (Stoycheff et al., 2020) leading to individuals altering their behaviour due to the fear of being watched. Moreover, public apprehension about who controls these AI technologies and the potential for their misuse is increasingly evident. Discussions around ownership and control reflect diverse perspectives on how surveillance tools should be deployed within law enforcement contexts, highlighting the need for stringent safeguards and transparent governance (Chohan & Hu, 2020; Schuilenburg & Peeters, 2020). The debate over ownership is crucial, as it determines who has access to the data collected and how it is used. The perceived legitimacy of these surveillance tools is closely tied to public trust in the institutions that manage them. This is particularly significant in discussions about the role of private companies in providing AI technologies to law enforcement, where concerns about profit motives and accountability arise (Mann et al., 2003).

Not to forget that the convergence of online and offline surveillance practices, amplified by the capabilities of AI, represents a significant shift in how both public and private spaces are monitored and controlled. Online surveillance, facilitated by AI algorithms, captures vast amounts of personal data through social media platforms, search engines, and other digital services, often without explicit user consent (Zuboff, 2019). This form of surveillance is deeply intertwined with economic interests, where personal data becomes a commodity, leading to what Zuboff (2019) describes as "surveillance capitalism." In contrast, offline surveillance encompasses the physical tracking of

individuals through AI technologies such as facial recognition and ANPR, often justified under the guise of enhancing public safety (Norris & Armstrong, 1999; Macnish, 2021). The intersection of these two realms creates a pervasive surveillance ecosystem that blurs the boundaries between public and private life. This dual surveillance raises profound ethical questions about autonomy, consent, and the power dynamics inherent in AI technologies (Lyon, 2007). Critically, it underscores the importance of developing robust legal and ethical frameworks that address the specificities of both online and offline surveillance, ensuring that AI's potential is harnessed in ways that respect individual rights and societal values.

AI in Policing: Purpose, Extent and Uses

The integration of Artificial Intelligence (AI) into policing has transformed both routine operations and strategic decision-making processes for law enforcement agencies (LEAs). AI technologies are now embedded in numerous facets of law enforcement activities, designed to improve efficiency, enhance predictive capabilities, and reduce human error in both investigative and preventative measures. The primary purpose of AI in policing is to optimize the management of data-driven tasks such as crime prediction, suspect identification, and threat assessment, enabling more informed, real-time decision-making (Babuta & Oswald, 2020).

The extent of AI adoption in policing varies globally, with some regions advancing rapidly in the deployment of these technologies, while others maintain a more cautious approach due to concerns over ethics, civil liberties, and the lack of clear regulatory frameworks (Završnik, 2020). In nations like the United States, the United Kingdom, and China, AI technologies such as facial recognition, biometric scanning, and predictive policing models have become routine tools in both public surveillance and criminal investigations (Ferguson, 2017; Smith, 2020). Conversely, many European countries, although technologically advanced, have been more restrained in adopting these tools, particularly due to public concern over privacy rights and the potential for misuse by LEAs (Oswald & Grace, 2021).

The types of AI technologies used by police forces span several domains. Facial recognition technology, for instance, is widely implemented to identify suspects in crowds or compare images against databases of known offenders (Buolamwini & Gebru, 2018). Predictive policing models, which rely on algorithms to identify crime hotspots and allocate resources accordingly, are another prevalent application, although these tools have sparked significant debate due to their potential to perpetuate biases present in the data (Brayne, 2021).

Additionally, AI-driven systems are being utilized for real-time data analysis, automating tasks such as video surveillance monitoring, license plate recognition, and anomaly detection in traffic patterns (Babuta & Oswald, 2020). Another growing use of AI is in cybersecurity, where LEAs deploy AI algorithms to detect and prevent cybercrimes, from fraud detection to monitoring online extremism (Završnik, 2020). This digital extension of AI's role in law enforcement illustrates its versatility, as it moves beyond physical crime scenes to address virtual threats that are increasingly prevalent in the modern technological landscape.

As AI becomes more ingrained in policing, it is essential to examine the implications of its use, especially concerning public trust, transparency, and accountability. While AI has the potential to make policing more effective and efficient, its deployment must be carefully managed to ensure that the technology is applied in a way that upholds civil liberties and respects the rights of citizens. The following sections will explore these concerns in greater depth, particularly focusing on the ethical implications of AI-driven surveillance, the public's reactions to its use, and the potential for counterstrategies and safeguards to mitigate its risks.

Surveillance and AI in Policing

The use of AI in policing encapsulates both the potential for enhancing operational efficiencies and the risks related to privacy violations and discrimination. AI-enhanced systems like CCTV and facial recognition offer significant improvements in public safety, yet they require rigorous oversight to prevent abuses and ensure they do not lead to discriminatory practices (Babuta & Oswald, 2020; Trottier, 2017). The necessity for ongoing evaluation is paramount to ensuring these technologies serve the public interest while respecting privacy and civil liberties (EHRC, 2020; Davis, 2021). For instance, facial recognition technology, while useful for identifying suspects, has been criticized for its higher error rates among minority groups, which could lead to unjust profiling and wrongful arrests (Buolamwini & Gebru, 2018). However, the accuracy of facial recognition algorithms have demonstrated improvements with error rates as low as 1.5% in controlled environments, though performance dropped significantly in diverse settings due to variations in lighting and demographics. These findings are consistent with recent benchmarks in AI, which report accuracy rates of over 99% for facial recognition in ideal conditions but a noticeable decline when addressing real-world variability (e.g., NIST FRVT 2022).

Expanding further, the integration of AI technologies in law enforcement also demands an analysis of their impact on procedural justice. The ability of AI systems to process and analyze vast amounts of data can lead to more informed and nuanced policing strategies. However, concerns about algorithmic bias and the potential for disproportionate targeting of marginalized communities necessitate transparent and accountable policing practices (Richardson, et al., 2019). Algorithmic transparency is crucial in ensuring that the data used in these systems are representative and that the outcomes do not perpetuate existing societal inequalities. Research has shown that transparent communication regarding the use and scope of AI tools in policing can enhance public trust and cooperation with law enforcement efforts (Smith, 2020).

Adding to these considerations, recent discussions in academic and policy-making spheres emphasize the need for establishing clear and enforceable safeguards that dictate the terms of use, data storage, and access to information derived through AI systems. These discussions, informed by research on citizen perspectives and empirical data, suggest that well-defined safeguards can prevent misuse and enhance the legitimacy and acceptance of surveillance practices among the public (Ezzeddine et al., 2022). Safeguards could include measures such as independent audits of AI systems, public transparency reports, and strict data minimization practices to ensure that only necessary data are collected and retained. Moreover, the concept of 'responsible AI' has emerged as a crucial theme in the discourse surrounding AI in policing. Indeed, ‘responsible AI’ refers to the development and deployment of AI systems that adhere to ethical principles, such as fairness, accountability, and transparency (Constantinescu et al., 2021). In the context of law enforcement, this involves ensuring that AI technologies do not exacerbate existing social inequalities or infringe upon individual rights. The challenge lies in balancing the potential benefits of AI, such as improved crime detection, with the need to protect freedoms and civil liberties.

Public Attitudes towards AI Use by LEAs

The deployment of AI technologies in law enforcement is profoundly influenced by public opinion, shaped by perceptions of fairness, efficacy, and transparency. Research indicates a conditional acceptance among the public, contingent upon clear demonstrations of the benefits outweighing the risks and robust mechanisms to protect against misuse (Saura et al., 2022). Engaging with the public to discuss the design and implementation of these technologies is crucial for aligning law enforcement practices with societal values and ethical standards. Public

involvement in the decision-making process can also lead to more effective and widely accepted AI policies (Mann et al., 2003). Further exploring public attitudes, recent studies have shown a nuanced view where public support varies significantly based on the type of AI application in law enforcement. For instance, there is generally higher acceptance of AI for investigatory purposes compared to real-time surveillance, which many perceive as more intrusive (Jones, 2021). Moreover, the extent to which these technologies are embraced by the public also hinges on the perceived efficacy of AI in enhancing safety without compromising personal privacy (Lee & Lee, 2019). Public trust in AI technologies is also influenced by the transparency of their deployment and the perceived impartiality of the algorithms used (Andrejevic, 2007).

Additionally, recent findings further underscore the varied views held by the public regarding ownership of AI surveillance tools. The legitimacy of these tools, as perceived by the public, is closely tied to who owns and operates them. Clear, transparent ownership and operational guidelines are perceived as potential approaches to alleviate concerns and positively enhance community cooperation with law enforcement initiatives, thus fostering a more security-conscious, yet privacy-respecting public atmosphere (Ezzeddine & Bayerl, 2024). The question of ownership is particularly pertinent in the context of public-private partnerships, where private companies often provide the technological infrastructure for AI surveillance. Ensuring that these partnerships are governed by clear and transparent agreements is crucial for maintaining public trust.

Moreover, the public's willingness to accept AI in law enforcement is also shaped by broader societal attitudes towards surveillance: while some individuals may view surveillance as a necessary tool for ensuring security, others may perceive it as a threat to their personal freedoms (Macnish, 2021). This dichotomy reflects the complex and often contradictory nature of public attitudes towards AI surveillance. As such, any attempt to integrate AI into law enforcement must be accompanied by pre-emptive efforts to educate the public about the technology's benefits and risks, as well as to address their concerns about privacy and autonomy.

Conclusion

This thesis explores the complex interplay between AI, surveillance, and public perception, aiming to provide insights into how these technologies are received and resisted by communities. By examining the empowering and inhibitory effects of surveillance, this research seeks to foster a dialogue among technology developers, policymakers,

and the public, guiding the ethical integration of AI in law enforcement to enhance security without compromising fundamental rights. This comprehensive analysis aims to contribute to a balanced understanding of the benefits and challenges posed by AI in policing, underscoring the importance of ethical considerations in the deployment of surveillance technologies.

Aim, Objectives, and Research Questions

Aim:

The primary aim of this thesis is to investigate the public's perception and response to the use of AI in law enforcement, focusing on understanding the nuanced views of citizens regarding AI-driven surveillance and the counterstrategies they employ to mitigate perceived threats to privacy and freedoms.

Objectives:

1. To examine the extent and context of AI application in law enforcement and its implications for privacy and civil liberties.
2. To analyse public attitudes towards AI-driven law enforcement tools, identifying factors that influence acceptance or resistance.
3. To explore the variety of counterstrategies employed by citizens in response to AI surveillance, assessing their effectiveness and implications for policy.
4. To provide recommendations for policymakers, LEAs and AI digital designers that can ensure a balance between technological advances and the protection of civil liberties, based on empirical findings and public engagement.

Research Questions:

1. What are the primary concerns of the public regarding the use of AI in law enforcement surveillance?
2. How do citizens perceive the benefits and risks associated with AIdriven surveillance by LEAs?
3. What counterstrategies do citizens employ to resist or cope with AI surveillance, and how effective are these strategies?
4. What policy measures can be implemented to ensure the ethical use of AI in law enforcement while maintaining public trust and safeguarding privacy and freedoms?

Thesis Structure

Chapter 1: Introduction

This chapter sets the stage for the thesis, providing an overview of the integration of AI in law enforcement and highlighting the associated

public concerns. It outlines the research aim, objectives, and questions that guide the study, establishing the foundation for a detailed examination of AI in law enforcement.

Chapter 2: Literature Review – Surveillance, Resistance, Counterstrategies, Safeguards and Ownership

- **Public Reactions to Surveillance**: Examines how different citizens perceive and react to surveillance, online or offline, influenced by factors like trust in institutions, perceived security benefits, and privacy concerns.
- **Resistance and Counterstrategies**: Explores both individual and collective forms of resistance to surveillance. This includes digital countermeasures such as encryption and VPNs, as well as offline tactics such as avoiding surveillance-heavy areas.
- **Safeguards in Surveillance**: This section reviews the ethical, legal, and technical safeguards that are essential for responsible use of surveillance technologies. It discusses how transparency, regulatory frameworks, and oversight mechanisms are implemented to protect privacy and civil liberties. Additionally, this section will discuss the distinction between online and offline AI surveillance, examining how each influences behavioural changes in individuals. It will critically analyse how the pervasive presence of surveillance technologies, whether digital or physical, alters individual behaviour, induces chilling effects, and prompts resistance strategies.
- **Ownership of Surveillance Tools**: Evaluation of citizen perception on who holds the power to deploy surveillance tools, control the collected data, and the implications of these ownership structures on privacy and power dynamics within society.

Chapter 3: Methodology Details the qualitative and quantitative research methods used to collect data, including the design, sampling, and analysis techniques. This chapter explains the choice of methodologies and their appropriateness for addressing the research questions, emphasizing the validity and reliability of the approaches.

Chapter 4: Publications, Candidate Contributions followed by Refection and Integration of Findings This chapter presents a detailed analysis of the data collected through various research methods. It integrates insights from the published articles and assesses how they align with or contrast against the primary data gathered, exploring themes such as trust, privacy, and the effectiveness of resistance strategies.

Chapter 5: Discussion Interprets the findings from the analysis, linking back to the literature reviewed in Chapter 2 and the insights from the published articles. Discusses the broader implications of the findings for policy makers, law enforcement agencies, and technology developers, particularly focusing on how to balance technological advances with ethical considerations and public expectations.

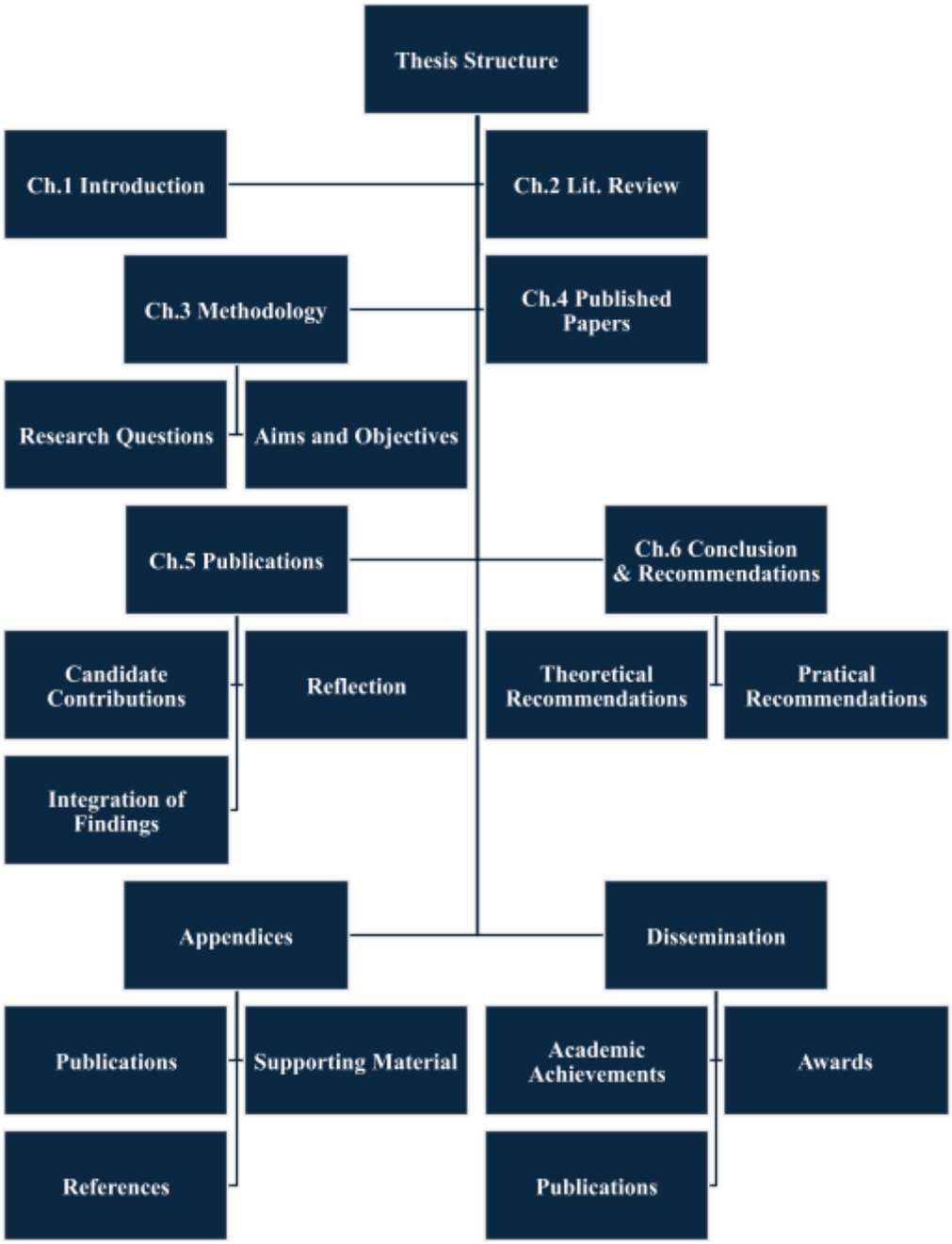
Chapter 6: Conclusions and Recommendations

Concludes the thesis by summarizing the key insights and findings. It highlights the theoretical implications of this research and provides practical recommendations for law enforcement agencies and policy makers on implementing AI technologies in a manner that respects privacy and enhances public safety without infringing on civil liberties.

Appendices and References Includes the published papers, supporting materials and full bibliographic references to provide comprehensive support for the research findings and conclusions drawn.

Dissemination and Publications from the PhD Outlines how the thesis findings have been shared through peer-reviewed publications, conference presentations, and collaboration with law enforcement and AI stakeholders. It also highlights the obtained awards and the research's impact in opportunities of international police training.

Figure 1. Thesis Structure Outline



Literature Review

Transformative Impact of AI on Law Enforcement and Surveillance Practices

The integration of AI within law enforcement and security sectors has ushered in a significant transformation in surveillance practices. AI is not merely enhancing the operational capacities of LEAs but reshaping the entire conceptual framework of surveillance. These technological innovations have led to more automated, predictive, and pervasive forms of oversight. Beyond traditional mechanisms, AI now enables far more sophisticated methods of data analysis and threat detection (Ceyhan, 2012; Završnik, 2020). The result is a paradigm shift in how public safety is managed, characterized by predictive policing, real-time

behavioural analysis, and biometric tracking technologies such as facial recognition and gait analysis (Brayne, 2020).

This shift toward AI-driven law enforcement has prompted both praise for its efficiency and criticism regarding its implications for civil liberties. Proponents highlight AI's ability to process massive amounts of data at speeds and accuracies unattainable by human operators (Babuta, 2020). Technologies like ANPR and predictive policing tools can help LEAs allocate resources more effectively and have shown cost efficiency, reducing investigative time by an estimated 35%, leading to preventing crime by identifying hotspots and at-risk individuals (Ferguson, 2017). However, these benefits were tempered by runtime delays in urban settings with complex surveillance networks, highlighting the need for robust computational infrastructure." Similarly, biometric surveillance systems, including facial recognition and iris scans, promise improved identification processes, enabling law enforcement to act swiftly in identifying suspects (Magalhães & Sánchez, 2021).

Nevertheless, such advancements have raised substantial concerns about ethics, privacy, and civil rights. Moreover, while AI enables the aggregation and analysis of vast amounts of data, its deployment often lacks the necessary transparency and accountability measures required to safeguard public trust (Yeung, 2018).

Moreover, the pervasive nature of AI surveillance has led to a reconceptualization of the public's relationship with surveillance technologies. No longer limited to traditional surveillance methods, AIdriven systems now operate with a level of autonomy and efficiency that far exceeds human capabilities. This has resulted in a form of surveillance that is not only omnipresent but also increasingly invisible, embedded within the very fabric of digital (online) and physical (offline) infrastructures (Andrejevic, 2007; Zuboff, 2019). The implications of this shift are profound, as they challenge the very notion of privacy in the digital age and necessitate a re-evaluation of the legal and ethical frameworks governing surveillance practices. Furthermore, the integration of AI into surveillance specifically has amplified concerns surrounding data ownership and consent. The vast amounts of data collected by these systems are often stored and processed without the explicit consent of those being surveilled, raising critical issues regarding the commodification of personal information and the potential for misuse (Andrejevic, 2007; Petersen & Taylor, 2012). This is particularly pertinent in the context of AI, where the data used to train algorithms can have far reaching consequences, influencing everything from law enforcement decisions to societal norms and behaviours. A critical element in the debate is the question of accountability and

transparency in AI deployment. Unlike traditional surveillance systems that rely on human judgment, AI systems operate based on algorithms that may not always be transparent or understandable to the public or even to the operators using them (Pasquale, 2015). The lack of clarity regarding how AI systems make decisions presents significant challenges to maintaining public trust in law enforcement practices. As Binns (2018) points out, the "black box" nature of many AI systems means that errors or biases in decision-making are difficult to detect and correct. This opacity threatens the foundational principles of due process and legal accountability, raising concerns about the ethical use of AI in law enforcement.

As AI continues to evolve, the challenges it presents to the principles of justice, equity, and democracy become more pronounced. The debate surrounding AI surveillance is not merely a technological one; it is deeply rooted in social, political, and ethical considerations that demand careful scrutiny and regulation. The need for robust safeguards, transparent governance, and public accountability is more urgent than ever as societies grapple with the implications of AI in policing and surveillance (Macnish, 2021; Lyon, 2014). The evolving landscape of AI-driven surveillance underscores the importance of a multidisciplinary approach to understanding and addressing these issues, one that is informed by legal, ethical, and technological perspectives alike.

AI in Policing and Security

Historical Development and Current State

The integration of Artificial Intelligence (AI) into law enforcement is part of a larger and ongoing trend toward the digitalization and automation of security practices. This transition, which began with the early use of computerized systems in policing during the late 20th century, has evolved significantly as technological advancements have accelerated.

Today, AI technologies are deeply embedded in a range of law enforcement activities, transforming not only operational capacities but also the broader paradigms of surveillance and public safety management.

Evolution of AI in Policing

AI's journey into law enforcement began with rudimentary data systems that helped officers manage records, perform basic criminal analyses, and generate insights based on historical data (Brayne, 2020). Over time, these systems expanded to incorporate advanced algorithms

capable of making predictions based on large-scale data sets, laying the foundation for contemporary predictive policing. Predictive policing represents one of the most prominent applications of AI in law enforcement, leveraging machine learning models to forecast crime hotspots, enabling more strategic resource allocation (Ferguson, 2017). Through algorithmic analysis of past crime data, LEAs can predict areas likely to experience criminal activity, theoretically allowing for pre-emptive interventions (Babuta & Oswald, 2020). However, this method is not without its controversies. Critics have noted the potential for AI to perpetuate existing societal biases, particularly when historical crime data is used to train the algorithms, which can lead to over-policing of marginalized communities (Richardson et al., 2019).

Technological Mechanisms and Tools

AI in law enforcement is not limited to predictive policing; it encompasses a range of tools that utilize sophisticated algorithms for various tasks. One such tool is facial recognition technology, which enables LEAs to identify individuals in public spaces or at critical security checkpoints (Magalhães & Sánchez, 2021). This application has proven effective in high-stakes environments like airports, stadiums, and large public gatherings, where swift identification is critical. Biometric scanning is another AI-powered technology that facilitates the identification and tracking of individuals based on unique physical characteristics such as fingerprints, iris patterns, or voice recognition (O'Neil, 2016). These AI-driven technologies offer unprecedented accuracy and speed, enabling faster responses to potential threats. However, the accuracy of facial recognition systems has been called into question, particularly in terms of racial bias. Studies have demonstrated that these systems are less accurate when identifying individuals with darker skin tones, leading to concerns about their application in diverse populations (Buolamwini & Gebru, 2018). This has raised ethical questions surrounding fairness and accountability in AI surveillance, as misidentification can result in wrongful detentions.

Global Adoption and Implementation

AI-powered surveillance tools have seen varying levels of adoption worldwide, with certain regions at the forefront of integrating these technologies into law enforcement. In the United States, major cities such as Chicago and Los Angeles have implemented AI-enhanced surveillance systems to monitor crime, particularly using predictive policing (Brayne, 2020). The United Kingdom has taken a similar approach, relying on AI technologies like ANPR and facial recognition for public surveillance (Magalhães & Sánchez, 2021). In China, the

government has deployed AI to unprecedented levels in its surveillance network, using facial recognition and biometric data to monitor public spaces in real time (Creemers, 2018). While these technologies aim to enhance public safety, their use has ignited debates over privacy, with critics arguing that they represent a significant intrusion into civil liberties.

Future Developments: AI's Trajectory in Policing

The future of AI in law enforcement is poised to expand as AI technologies continue to evolve. Machine learning algorithms are expected to become even more refined, with improved predictive capabilities and less bias in their datasets. AI could potentially be used in autonomous surveillance systems, including drone policing and smart city initiatives, which rely on AI to monitor urban environments 24/7 (Kitchin, 2014). Smart cities are increasingly integrating AI into their infrastructure, providing constant real-time surveillance through interconnected systems of cameras, sensors, and biometric recognition tools (Parviainen & Ridell, 2021). These AI-powered ecosystems are designed to improve public safety but require careful consideration regarding transparency, accountability, and citizen consent.

Recent advancements in AI highlight its expanding role in law enforcement. For example, generative AI technologies are now being piloted for real-time scenario simulation, enabling better strategic planning. Additionally, transformer-based models like GPT-4 are being explored for natural language analysis to improve intelligence gathering. These developments underscore the shift toward more adaptable and scalable AI solutions for modern policing (e.g., Agarwal et al., 2023; OpenAI Research, 2023).

There is also the potential for AI-assisted behavioural analysis, which could monitor public behaviour in real-time, flagging abnormal or suspicious activity. While this may prove useful for preventing crime, it raises significant concerns about privacy, particularly as public spaces become sites of constant surveillance (Zuboff, 2019). This brings into focus the ongoing ethical debate surrounding the balance between the benefits of AI in enhancing security and the risks of overreach and abuse in personal data collection (Lyon, 2014).

Digital versus Physical AI-surveillance

The dichotomy between digital (online) and physical (offline) surveillance represents a pivotal aspect in comprehending the complex implications of AI-driven technologies within law enforcement. Online surveillance, often characterized by its pervasive and largely invisible

nature, captures user behaviour, preferences, and personal data across digital platforms, creating a vast repository of information that is frequently collected without explicit consent (Andrejevic, 2007; Lyon, 2007). Conversely, offline surveillance is more overt, manifested through physical monitoring tools such as CCTV cameras, facial recognition systems, and biometric scanners, which directly interact with individuals in public and private spaces (Zhang & Qiu, 2022).

Understanding public reactions to these differing forms of surveillance is crucial, especially as the integration of AI amplifies concerns surrounding privacy, consent, and freedoms. In fact, evaluating the intersection of these surveillance modalities sheds light on how individuals perceive and navigate their environments under the constant observation of AI technologies, whether in digital or physical contexts. Such an examination is essential for developing robust ethical frameworks that address the specific challenges posed by both online and offline surveillance, thereby enriching the broader discourse on responsible AI usage in law enforcement and public safety (Zuboff, 2019; Macnish, 2021).

Public Reactions to Surveillance

Public reactions to surveillance are influenced by various factors, including cultural context, the perceived trade-off between security and privacy, and individual sensitivities to civil liberties issues (Haggerty & Samatas, 2010). Moreover, Lyon (2014) notes that public acceptance of surveillance often hinges on the transparency of its operations and the justifications provided by governing bodies. However, as Zuboff (2019) critically analyses, there is an increasing sense of resignation among the public, a phenomenon she describes as "surveillance capitalism," where data privacy concerns are overshadowed by economic interests.

Despite this, a significant portion of the public actively resists surveillance, driven by concerns about privacy, autonomy, and the potential for abuse. This resistance is often articulated through activism, legal challenges, and the adoption of privacy-enhancing technologies (Monahan, 2006). Public reactions to surveillance are thus not monolithic but rather reflect a spectrum of responses ranging from acceptance and resignation to active resistance. Understanding these diverse reactions is crucial for developing surveillance practices that align with public values and respect individual rights.

Surveillance-triggered Behavioural Changes

Psychological and Social Impacts

The widespread adoption of AI-driven surveillance systems has had significant implications for individual and societal behaviour. Namely,

"chilling effect" where individuals alter their behaviour when they are aware of being watched, leading to self-censorship and reduced freedom of expression (Gilliom, 2001). This effect has been documented in various contexts, from online activities to public spaces and interactions with law enforcement (Stoycheff, 2016). Additionally, AI surveillance exerts a profound psychological impact on individuals, manifesting in increased anxiety, stress, and a perceived loss of autonomy. The constant awareness of being monitored can lead to self-censorship, where individuals refrain from engaging in activities that might attract scrutiny (Dattatray Deulkar & Gupta, 2020). This self-censorship can have farreaching implications for democratic processes, as it discourages free expression and participation in public discourse (Rainie & Madden, 2015).

Socially, the presence of AI surveillance technologies in public spaces alters how people interact with one another and their environment. For example, the deployment of facial recognition cameras in urban areas has influenced how individuals navigate these spaces, often leading them to avoid areas where they might be captured on camera (Ball, 2002). This can create a pervasive atmosphere of unease and mistrust, particularly in communities disproportionately subject to surveillance (Gilliom, 2001). The social fabric is thus reshaped by the omnipresence of surveillance, imposing new forms of behaviour and interaction.

Coping Mechanisms and Behavioural Adaptations

In response to the pervasive nature of AI surveillance, individuals and communities have developed various coping mechanisms and strategies to adapt to this new reality. These strategies range from subtle behavioural changes to more overt forms of resistance, such as using technology to evade detection (Mann et al., 2003). For instance, people might avoid specific areas known for heavy surveillance or employ tactics such as wearing makeup that confounds facial recognition software (Tzovieli & Elovici, 2021).

The development of these coping mechanisms underscores the broader societal impact of AI surveillance: as individuals become accustomed to being constantly monitored, there is a risk of desensitization to surveillance or the internalization of the norms it imposes (Foucault, 1977). This normalization of surveillance blurs the boundaries between public and private spaces, leading to a gradual erosion of privacy and autonomy. Moreover, the social implications of these adaptations reflect broader trends in how surveillance shapes power dynamics and social structures (Andrejevic, 2007).

Another noteworthy coping strategy emerges in the form of braiding hair as a resistance tool against surveillance technologies, particularly

facial recognition systems. In their exploration of surveillance cultures, Klauser and Albrechtslund (2014) emphasize how simple, everyday practices, such as altering one's appearance, serve as a form of resistance to AI surveillance. In fact, braiding hair has emerged as a subtle yet significant act of resistance, as certain styles of hair braiding can interfere with facial recognition algorithms (Isaac, 2021). This practice highlights the creative ways in which individuals contest the power of surveillance technologies in everyday life. Much like the use of makeup or accessories to disrupt algorithmic precision, hair braiding reflects the agency of individuals in reclaiming control over their bodies in a heavily surveilled environment. These adaptations underscore a broader societal resistance to surveillance, revealing how personal identity and cultural expressions become intertwined with tactical evasion. Such behavioural modifications not only reflect a form of symbolic resistance but also raise critical questions about the growing practice of surveillance on individual autonomy and cultural practices.

Thus, by leveraging the physical manipulation of appearance, citizens can navigate spaces under heavy AI surveillance without compromising their autonomy. This approach highlights the tension between technological systems designed to monitor and the human ability to disrupt them through creativity. Moreover, such behavioural adaptations signify a deeper critique of the proliferation of surveillance technologies, where the resistance to surveillance becomes a cultural phenomenon rooted in individual identity and self-expression (Parviainen & Ridell, 2021). The discourse on coping mechanisms underscores the need for a nuanced understanding of the social dynamics at play when AI technologies intersect with daily life. As AI surveillance continues to evolve, it is likely that these forms of resistance will become more diverse and sophisticated, mirroring the advancing technologies that they seek to subvert (Isaac, 2021).

Resistance and Counterstrategies

Historical and Modern Forms of Resistance

Resistance to AI surveillance has manifested in various forms, from individual acts of defiance to organized campaigns that challenge the deployment of surveillance technologies. Counter-surveillance strategies, often referred to as "sousveillance," involve using technology to monitor those in power, effectively reversing the traditional dynamics of surveillance (Mann & Ferenbok, 2013). This concept represents a grassroots challenge to the top-down control of data by governments and corporations, empowering individuals to engage in their own forms of oversight.

Historically, resistance to surveillance predates AI and digital technologies, emerging from fundamental concerns over privacy, autonomy, and the exercise of control by authorities. However, the digital era, particularly with the rise of AI, has amplified these practices. The advent of the internet and AI technologies has led to more sophisticated forms of resistance, offering individuals and organizations new tools to protect their privacy and anonymity (Monahan, 2006). One prominent modern method is the use of encryption tools and anonymizing technologies, such as the Tor browser, which allows individuals to conceal their identities and activities online (Saura et al., 2022). These technologies are crucial in countering the pervasive nature of AI-driven digital surveillance, which monitors users' data, online behaviour, and personal interactions across platforms.

One form of resistance is neutralization, where individuals seek to undermine the efficacy of surveillance systems through various technical means. In the context of AI surveillance, this often involves using tactics that obscure or confuse AI algorithms. For example, adversarial attacks on facial recognition systems—where specific patterns or makeup are applied to confound the algorithm's ability to identify individuals— represent a growing form of technical resistance. Research from BenGurion University, for example, shows how simple techniques such as modifying makeup can prevent AI systems from accurately identifying individuals (Ben-Gurion University, 2020). These methods challenge the reliability of AI technologies by exploiting their limitations, particularly their reliance on data patterns that can be intentionally disrupted.

Obfuscation is another tactic that has gained traction in response to AI surveillance. This strategy involves deliberately providing misleading or excessive data to surveillance systems, effectively "flooding" the algorithm with irrelevant information. This dilutes the utility of surveillance and reduces the likelihood of individuals being identified or profiled accurately (Brunton & Nissenbaum, 2015). Obfuscation has been especially relevant in digital spaces, where users can create false data trails or use multiple identities to confuse tracking algorithms.

Alternatively, the use of physical forms of resistance, such as clothing and makeup, is also prevalent in offline environments. Beyond digital obfuscation, clothing designed to confuse AI surveillance systems, such as anti-surveillance garments, is an emerging form of resistance. These garments are designed with patterns that disrupt facial recognition algorithms, making it difficult for AI systems to track or identify

wearers in public spaces (Harvey, 2020). Similarly, makeup techniques can be used to distort the facial features that AI systems rely on, effectively rendering the wearer unrecognizable to surveillance cameras (Saura et al., 2022). These methods highlight the intersection of technology, fashion, and resistance, where individuals adapt their appearance to challenge the pervasive monitoring of AI systems.

Anti-Surveillance Activism and Legal Challenges

Modern resistance to AI surveillance also includes organized efforts by civil society organizations and legal actions aimed at restricting or banning AI technologies in law enforcement. A notable example is the Reclaim Your Face campaign, a Europe-wide initiative that seeks to ban biometric mass surveillance technologies, particularly facial recognition. This campaign, led by the European Digital Rights (EDRi) network, mobilizes citizens to challenge the growing use of AI-powered surveillance by governments and corporations (EDRi, 2021). The campaign advocates for stronger regulatory frameworks that protect individuals' privacy and limit the scope of AI surveillance in public spaces. Similarly, in the U.S., several cities, including San Francisco and Boston, have implemented legal bans on facial recognition technology used by law enforcement, following grassroots campaigns and public outcry over concerns of discrimination and privacy violations (Benjamin, 2020).

In fact, legal challenges have proven to be an effective form of resistance, particularly in contexts where citizens and advocacy groups have contested the unchecked use of AI surveillance. For instance, lawsuits filed against the use of facial recognition by police departments have led to significant legal victories, including rulings that restrict or prohibit the use of these technologies in public policing (Hill, 2020). These legal challenges underscore the role of the judiciary in moderating the balance between security measures and individual freedoms, ensuring that surveillance practices are subjected to scrutiny and oversight.

Organized Digital Resistance: Civil Society Platforms

Civil society organizations have also played a pivotal role in resisting AI surveillance by developing platforms and tools to expose and challenge these technologies. For example, the Patrick J. McGovern Foundation provides AI tools to organizations that advocate for transparency and inclusivity in AI governance (World Economic Forum, 2020). By empowering marginalized communities to participate in discussions about AI governance, these platforms serve as critical

spaces for public resistance against the encroachment of AI technologies on civil liberties.

Impact and Effectiveness of Resistance

The impact of resistance efforts against AI surveillance varies across different contexts and strategies. In some cases, these efforts have led to meaningful policy reforms, such as the introduction of stricter regulations on AI use by law enforcement or outright bans on certain technologies (Hill, 2020). However, resistance is often met with substantial opposition from governments and corporations, which argue that AI surveillance is essential for public safety and security.

Nevertheless, the ongoing resistance to AI surveillance is crucial for maintaining democratic oversight over these technologies. By challenging the status quo and raising awareness of the risks associated with unchecked surveillance, these efforts help ensure that AI technologies are subject to public scrutiny and remain aligned with societal values (Monahan, 2011). Additionally, the proliferation of both digital and physical resistance strategies signifies a growing recognition of the power imbalance between citizens and surveillance agencies and the necessity of empowering individuals to reclaim control over their personal data and identities.

Legal and Ethical Safeguards

As the use of AI in law enforcement expands, there is an increasing need for legal and ethical safeguards to govern the deployment of these technologies. These safeguards are essential for protecting individual rights, ensuring accountability, and maintaining public trust in the use of AI by law enforcement agencies. The necessity for implementing safeguards in surveillance practices, especially those enhanced by AI, is paramount to maintaining public trust and upholding democratic values. Safeguards refer to various legal, ethical, and technical measures designed to ensure that surveillance technologies are used responsibly, and that individuals' privacy and rights are protected (Ezzeddine et al., 2022). Key safeguards in this context would include strict regulatory frameworks, transparency in using surveillance technologies, and robust oversight mechanisms.

Regulatory Frameworks

The development of regulatory frameworks for AI in law enforcement is still in its early stages, with significant variations across different jurisdictions. In some countries, comprehensive laws have been enacted to regulate the use of AI surveillance, while in others, there are few, if any, legal protections in place (Akhgar et al., 2022). This lack of

consistency creates challenges for both law enforcement agencies and individuals, as it can lead to uncertainty and confusion about the rules governing the use of AI.

Key elements of effective regulatory frameworks include transparency, accountability, and the protection of individual rights. Transparency involves ensuring that the public is informed about the use of AI technologies and how their data is being used. Accountability mechanisms are necessary to hold those responsible for AI surveillance accountable for their actions, whether through legal means or other forms of oversight (Jobin et al., 2019). The protection of individual rights is also crucial, particularly in relation to privacy, due process, and the right to be free from discrimination.

Ethical Guidelines and Principles

In addition to legal regulations, there is a growing recognition of the need for ethical guidelines to govern the use of AI in law enforcement. These guidelines are often based on principles such as fairness, transparency, and respect for human dignity (Floridi et al., 2018). They provide a framework for evaluating the ethical implications of AI technologies and for making decisions about their deployment in ways that are consistent with societal values.

One of the key challenges in developing ethical guidelines for AI is ensuring that they are flexible enough to adapt to the rapidly evolving nature of these technologies. As AI continues to advance, new ethical dilemmas are likely to emerge, requiring ongoing reflection and adaptation of existing guidelines (Jobin et al., 2019). Furthermore, these guidelines must be enforceable, ensuring that they are not merely aspirational but have a tangible impact on the use of AI in law enforcement.

Research highlights the importance of these safeguards in mitigating risks associated with AI surveillance. Tufekci (2015) emphasizes the role of algorithmic transparency as a safeguard against biases in AI systems. Similarly, Lyon (2014) argues for stronger legislative oversight to ensure that surveillance practices do not infringe upon civil liberties. The development of international standards and guidelines, as discussed by Whittaker & Van der Ploeg (2018), also plays a crucial role in standardizing the ethical use of AI across borders.

Incorporating these safeguards into surveillance practices can also enhance public trust, a critical factor in the legitimacy and effectiveness of law enforcement agencies. Transparent communication about how AI technologies are used, coupled with robust oversight mechanisms, can mitigate public concerns about privacy violations and data misuse. This trust is not just about compliance with legal standards but also involves

ethical considerations, where the use of AI aligns with societal values and respects individual autonomy (Ada Lovelace Institute, 2019). Thus, implementing safeguards is not merely a technical or legal requirement, but a moral imperative that underpins the responsible use of AI in policing.

Ownership of Surveillance Tools

Ownership of surveillance tools in the context of law enforcement is a contentious issue that impacts public perception and the deployment of these technologies. Ownership determines who has the right to use, control, and access data collected through surveillance. Marx (2016) discusses different models of ownership, ranging from state-controlled to privatized systems, each presenting unique challenges and implications for privacy and accountability. The debate on ownership is closely tied to concerns over who has the power to monitor and who is being monitored. As Andrejevic (2007) points out, the concentration of surveillance tools in the hands of a few entities could lead to power imbalances and potential abuses. Therefore, discussions on ownership are critical in framing policies that promote equitable access to and control over surveillance technologies.

AI Ownership and Governance

The question of who should control and govern AI technologies in law enforcement is a contentious issue. Public debates over AI ownership often focuses on concerns about accountability, transparency, and the potential for abuse. Some scholars argue that AI should be owned and operated by public entities, such as government agencies, to ensure that these technologies are used in the public interest (Chohan & Hu, 2020). Others advocate for a more decentralized approach, where citizens have greater control over the data collected about them and the technologies used to process it (American Bar Association, 2020).

Citizen preferences regarding AI ownership vary widely, reflecting broader societal debates about the role of technology in governance. Research indicates that while some citizens trust government agencies to manage AI technologies responsibly, others are wary of state surveillance and prefer that AI be controlled by independent or private entities (Ada Lovelace Institute, 2019). The governance of AI in policing is likely to remain a complex and evolving issue, with significant implications for privacy, accountability, and public trust.

The ownership of surveillance tools in law enforcement is a complex and contentious issue that significantly influences public perception and the deployment of these technologies. Ownership determines who has the authority to use, control, and access the data collected through

surveillance, thereby shaping the power dynamics between the state, private entities, and the public (Pierson, 2019).

Marx (2016) explores different models of ownership, ranging from statecontrolled to privatized systems, each presenting unique challenges and implications for privacy and accountability. State-controlled surveillance systems are often justified on the grounds of national security and public safety, yet they can lead to concerns about government overreach and the erosion of civil liberties. On the other hand, privatized systems, where surveillance technologies are owned and operated by private companies, raise issues related to the commodification of personal data and the lack of public oversight. The concentration of surveillance tools in the hands of a few entities, whether governmental or corporate, could result in significant power imbalances and potential abuses (Andrejevic, 2007). The debate on ownership is also closely tied to concerns about transparency and accountability. Public trust in surveillance systems is contingent upon who owns and operates these technologies. If ownership is concentrated in entities perceived as untrustworthy or unaccountable, public support for surveillance initiatives may wane (Andrejevic, 2007; Zuboff, 2019; Lyon, 2007). Moreover, the lack of clear ownership guidelines can lead to conflicts over data access and control, further complicating the ethical landscape of AI surveillance (Chohan & Hu, 2020). Thus, discussions on ownership are critical in framing policies that promote equitable access and control over surveillance technologies, ensuring that these tools serve the public good rather than narrow interests.

Conclusion and need for studying citizen reactions to AI use by LEAs

The integration of AI into policing and security practices has brought about significant changes, raising important ethical, legal, and social questions. The literature reveals a complex landscape where the benefits of AI in enhancing public safety are counterbalanced by the risks of privacy invasion, discrimination, and threats to civil liberties. The incorporation of safeguards, the contentious issue of ownership, the persistent resistance to surveillance, and the varied public reactions underscore the need for a more nuanced approach to AI surveillance. As AI technologies continue to evolve, it will be essential to address these challenges proactively, ensuring that the deployment of AI in policing enhances security while upholding democratic values and individual rights. This literature review contributes to a deeper understanding of these issues, nurturing the aims and objectives of this exploratory PhD

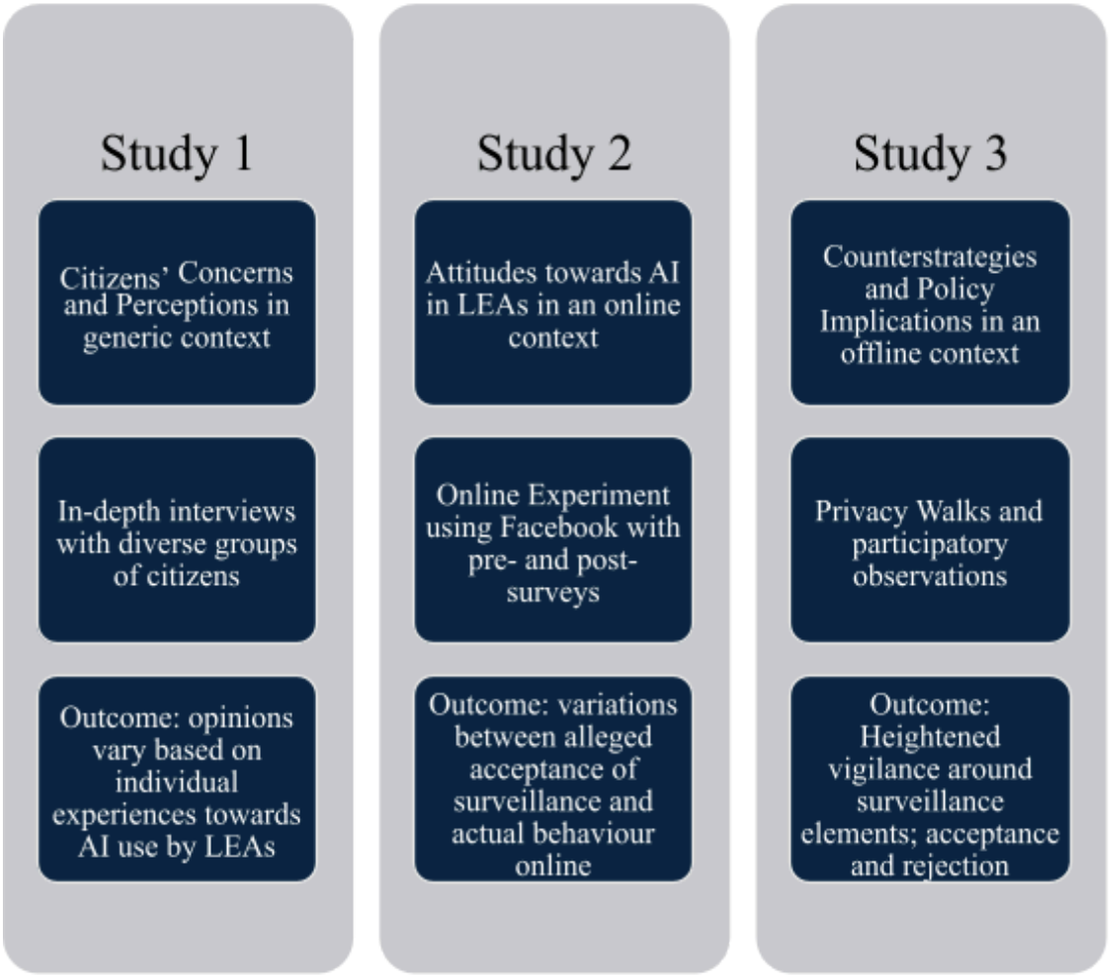
research, while providing a foundation for further research and policy development in the field of AI surveillance.

Methodology

Despite the rapid integration of AI into law enforcement efforts, significant research gaps persist regarding public perceptions and responses to AI-driven surveillance technologies. While existing literature often emphasizes the technical capabilities of AI and its potential for enhancing crime prevention (Tschider, 2018; Zuboff, 2019), it frequently neglects the nuanced understanding of citizen perspectives on privacy, ethical considerations, and the social implications of surveillance practices (Binns, 2018; Whittaker & Van der Ploeg, 2018). This oversight is critical, as the deployment of AI technologies without a comprehensive understanding of public sentiment can exacerbate distrust between law enforcement and the communities they serve (Lyon, 2020). This thesis addresses these gaps by focusing on citizen concerns about AI in surveillance, exploring their perceptions of the associated benefits and risks, examining the counterstrategies they employ in response to surveillance, and identifying policy measures that can promote ethical AI use in law enforcement while safeguarding public trust and privacy.

That is done through three main studies, briefly summarised in Figure.2 below, employing a mixed methods approach to gather rich, empirical insights.

Figure.2. General Outline and Approach of the studies



Significance of this research

By addressing these gaps, this thesis contributes to the fields of security studies and AI ethics by providing a comprehensive understanding of citizen reactions to AI surveillance. The findings reveal a spectrum of viewpoints, highlighting not only the concerns and resistance strategies of citizens but also offering practical recommendations for policymakers and law enforcement. This research emphasizes the need for transparent, ethical frameworks that align with public expectations, ultimately fostering trust between communities and technology developers. The insights generated here are crucial for informing the responsible deployment of AI in law enforcement, ensuring that technological advancements do not come at the expense of individual rights and freedoms (Crawford & Paglen, 2021). This work aims to bridge the divide between technological innovation and public accountability, thereby supporting a more equitable approach to security in the digital age.

Research Design

In this thesis, the research employs a mixed-methods approach, combining both qualitative and quantitative methodologies to investigate public perceptions, safeguards, and ownership issues related to the use of AI by LEAs. This strategy enables a nuanced understanding of the complex social, ethical, and operational dimensions of AI-enhanced surveillance, and it addresses both online and offline contexts, highlighting concerns about privacy, consent, and civil liberties (Creswell & Plano Clark, 2011).

The research design follows a sequential exploratory strategy, where qualitative data collection uncovers key themes, and quantitative methods validate these findings. This design provides both a rich narrative of citizen perspectives and empirical evidence to ensure the reliability of the findings.

Papers 1, 2, and 3 (Based on Study 1)

In the first set of studies, semi-structured interviews and Q-sorts were employed as part of research done for the AIDA project as explained in the declaration. The semi-structured interviews allowed for an in-depth exploration of citizen attitudes towards AI ownership and safeguards, while the Q-sorts helped quantify and compare these perspectives across various demographic groups (Watts & Stenner, 2012). This mixedmethods design offered insights into the broader public sentiment and specific concerns, drawing attention to the diverse viewpoints held by participants.

Paper 4 (Based on Study 2)

An online experimental design was implemented in Paper 4, which included the think-aloud method and pre- and post-task surveys. This combination of methods was selected to gain real-time insights into participants' thought processes and reactions as they interacted with AIdriven surveillance systems (Ericsson & Simon, 1984). The pre- and post-task surveys measured shifts in attitudes and opinions, revealing how interactions with AI technology influenced public perceptions.

Paper 5 (Based on Study 3)

The research culminated in Paper 5, which integrated pre- and post-task surveys with privacy walks. Privacy walks captured participants' realtime responses to AI surveillance in public spaces. This innovative methodology involved the triangulation of data through geo-mapping and recorded images, adding an ecological dimension to the research (van Es & de Lange, 2020). The use of geo-mapping allowed the researchers to analyse the impact of surveillance on participant

behaviour, such as how they navigated spaces and adjusted their routes or actions in the presence of AI surveillance tools.

Through this approach, the triangulation of data sources provided a comprehensive and context-sensitive understanding of the interactions between citizens and AI surveillance technologies, both in their conscious reflections and observable behaviours. Together, these methods ensure a well-rounded and context-specific examination of the public's response to AI in law enforcement.
This design not only supports the study's objectives of addressing public concerns but also extends the discourse on responsible AI use by LEAs, ensuring a holistic and balanced investigation of AI-driven surveillance in both online and offline environments.

Data Collection

This section outlines the methodological choices and rationale for sample selection and data collection across the papers that constitute this thesis, emphasizing the overarching strategies employed to capture diverse citizen perspectives on AI in law enforcement.

General Methodological Framework

The thesis employs a mixed-methods approach, integrating qualitative and quantitative techniques to provide a comprehensive understanding of citizen attitudes towards AI surveillance. This combination allows for the exploration of nuanced views while also facilitating broader comparisons across demographic groups.

Sample Selection

A key consideration in sample selection was ensuring diversity to reflect varied experiences and concerns. For the qualitative studies, participants were recruited from specific demographic groups relevant to their national contexts. Table.1 summarizes the choice of method, sample selection and rationale for each of the studies as published in the papers.

Table 1. Summary of methods, sample selection and rationale for the five Papers.

Paper	**Paper 1**	**Paper 2**	**Paper 3**	**Paper 4**	**Paper 5**
Method	Semi-structured Interviews	Q-Sort analysis	Scenario-based Interviews	Online Experiment with	Privacy Walks with pre- and post-task survey
				pre- and posttask survey	

Sample	Citizens from 7 EU countries and the UK	Citizens from 7 EU countries and the UK	Distinct groups within Germany, the Netherlands, Italy, and Greece	Online experiment, specifically targeted participants from Sheffield, UK	Privacy walks conducted in urban settings in the UK, also focusing on participants from Sheffield
Rationale	Wide range of cultural and social perspectives	Identify specific perspectives towards AI use by LEAs	Focused perspectives of young women, expatriates, older citizens, and IT law	Localized insights into citizen interactions with AI surveillance tools	Complementing the insights revealed in Paper 4 by exploring perceptions of AI surveillance in realworld environments

This tailored sampling strategy was designed to delve deeply into the specific attitudes towards AI ownership and its implications within different societal contexts, while also capturing localized experiences in the UK.

Qualitative Methods

Qualitative methods, including semi-structured interviews, were chosen for their ability to elicit rich, detailed insights into participants' concerns and recommendations regarding AI usage in policing. The semistructured format allowed flexibility, enabling researchers to probe deeper into critical themes such as privacy, bias, and accountability while maintaining focus on key topics. This flexibility is crucial in qualitative research, as it encourages participants to express their views in their own words, thus enriching the data collected (Auerbach & Silverstein, 2003). *Quantitative Methods*

In contrast, quantitative techniques, such as the Q-sort methodology, were utilized to systematically capture the range of opinions within the participant pool. This method allows participants to rank statements regarding AI surveillance, which facilitates the identification of shared perspectives across different demographic groups. The subsequent statistical analysis, including factor analysis, offers insights into public sentiment and highlights prevalent attitudes towards AI governance (Watts & Stenner, 2012).

Innovative Approaches

The thesis also incorporates innovative methods such as privacy walks, where participants reflect on real-world instances of AI surveillance in their environments. This experiential method not only enriches qualitative insights but also includes spatial analysis through geomapping, revealing how surveillance technologies influence individual behaviours and perceptions in urban contexts (van Es & de

Lange, 2020). By correlating observed behaviours during privacy walks with survey data, the research captures dynamic changes in attitudes towards surveillance, emphasizing the importance of context in shaping public sentiment (Klauser & Albrechtslund, 2014).

Rationale for Methodological Choices

*T*he choice of methods was guided by the research objectives, which aim to uncover both the depth of citizen concerns and the broader patterns of acceptance or resistance towards AI in law enforcement. By employing diverse methodologies, this research triangulates findings, enhancing the robustness of conclusions drawn. Each method complements the others, creating a holistic view of citizen engagement with AI technologies and their implications for privacy and civil liberties.

Ethics

Throughout the studies, ethical considerations were paramount. Informed consent was obtained from all participants, with clear explanations of the study's purpose, the nature of the data being collected, and the measures in place to protect participants' privacy. All data were anonymized prior to analysis, and participants were given the option to withdraw from the study at any time. These ethical safeguards ensured that the research was conducted with the utmost respect for participants' rights and well-being.

Data Analysis

The data analysis employed a mixed-methods approach, combining thematic and content analysis across the five papers to explore public perceptions of AI-driven surveillance comprehensively. Each paper utilized distinct yet complementary techniques, emphasizing both qualitative and quantitative methods.

In Papers 1, 2, and 3, the data from Study 1 obtained through interviews and Q-sorts were analysed using thematic analysis to identify key themes related to privacy, trust, and the legitimacy of AI in law enforcement (Braun & Clarke, 2006). The Q-sorts were further analysed using factor analysis and contextualized within participants' broader narratives, offering insights into shared attitudes and concerns (Watts & Stenner, 2012). This dual analysis allowed for a deep understanding of public opinions and their complexities beyond the quantitative data.

Paper 4 focused on content analysis of the think-aloud protocol and pre- and post-task surveys from Study 2, categorizing verbalized thoughts to reveal immediate cognitive and emotional responses to AI surveillance (Ericsson & Simon, 1993). This method provided a granular view of

participants' real-time decision-making, while surveys measured shifts in attitudes towards surveillance. Descriptive statistics quantified these shifts, while thematic analysis explored qualitative insights from openended responses.

In Paper 5, a combination of thematic (Clarke & Braun, 2017) and geomapping analysis was employed based on data collected as part of Study 3. Participants' discussions during privacy walks were transcribed and analysed to identify recurring themes about AI surveillance in real-world environments (Braun & Clarke, 2006). Geo-mapping provided a spatial analysis of participants' interactions with surveillance technologies, revealing how the visibility of surveillance influenced their behaviour, such as route selection and image-capturing patterns (van Es & de Lange, 2020).

Throughout the papers, the analysis was grounded in thematic analysis, which allowed for the exploration of both explicit concerns and deeper implicit attitudes towards AI surveillance. This rigorous coding process categorized participant responses into emerging themes, which were then grouped into higher-order categories to reflect common perspectives on AI ownership, privacy, and ethics (Nowell et al., 2017). The integration of these qualitative methods with quantitative assessments, such as surveys and geo-mapping, ensured a comprehensive and nuanced understanding of how AI surveillance is perceived and experienced in various contexts.

This approach not only provided detailed insights into the social and psychological dimensions of AI surveillance but also highlighted its ethical implications, capturing both public perception and the behaviours influenced by these technologies.

It is important to note that the methodology section in each of the below papers provides further in-depth information on sample selection, design and data analysis of the data that was specifically used in each publication.

Publications, Candidate Contributions, Reflections and Integration of Findings

All five published papers will be in the Appendices section. Below is a list of publications and the detailed contributions that I have made to each of the papers. **Paper 1**

Ezzeddine, Y., Bayerl, P.S., & Gibson, H. (2022). Citizen perspectives on necessary safeguards for the use of AI by law enforcement agencies. Transactions on Computational Science & Computational Intelligence. Springer Nature. arXiv.org. doi:10.48550/arXiv.2306.01786

Contribution:
For this paper, I led the UK data collection, analysis, and primary writing, supported by regular discussions with my DoS Saskia Bayerl and supervisor Helen Gibson, who contributed conceptually and reviewed drafts. These discussions helped shape the study's focus on citizen perspectives and safeguard mechanisms. Saskia and Helen contributed significantly to the theoretical framing and interpretation of the findings, which enriched the contextual understanding. Their input during the conceptual phase and writing stage was pivotal. **Paper 2**

Ezzeddine, Y., Bayerl, P.S., & Gibson, H. (2023a). 'Safety, privacy, or both: evaluating citizens' perspectives around artificial intelligence use by police forces. Policing and Society, 33(7), 861-876. https://doi.org/10.1080/10439463.2023.2211813

Contribution:
In this study, I was responsible for UK data collection, leading the analysis, and drafting the manuscript. Discussions with Saskia and Helen during the study analysis phase provided essential insights into shaping the research framework. Additionally, they both provided essential contributions, particularly in conceptual refinement and writing, which ensured the study's robustness.

Paper 3

Ezzeddine, Y., & Bayerl, P. S. (2024). Should everyone have access to AI? Perspectives on ownership of AI tools for security. In *International Conference on AI Research ICAIR24 Conference Proceedings*. Expected publication: December 2024.

Contribution:
For this conference paper, I led the UK data collection as part of the AIDA project, as well as the data analysis, and manuscript drafting. Saskia's extensive involvement shaped the study's theoretical approach, particularly in discussions around the ethical implications of AI ownership. Her input and review during the drafting stage were essential, providing depth to the exploration of security perspectives. This collaboration was critical in grounding the study within the wider discourse on AI governance.

Paper 4

Ezzeddine, Y., & Bayerl, P. S. (2023b). Under AI watch: Understanding online behaviours under supposed AI-surveillance. British Society of Criminology 2023 Conference Proceedings. ISSN 17759-0443. Vol. 22. https://www.britsoccrim.org/wp-content/uploads/2024/01/BSC-OnlineJournal-2023.pdf

Contribution:
In this publication, I was responsible for study design, data collection, analysis, and manuscript preparation, as this study was conducted outside of the AIDA project. Saskia's significant contributions throughout the writing and revision process, as well as her conceptual insights, along with Marjory Da Costa's input, helped shape this paper's exploration of online behaviours under surveillance.

Paper 5

Ezzeddine, Y., Bayerl, P. S. and Rodriguez, J.A. (2025). Unveiling public sentiments towards AI-driven urban surveillance: A case study from Sheffield. In *International Workshop on AI and Surveillance in Policing and Law and Order: Opportunities, Threats, Perspectives, and Cases*. Following the workshop, the chapter will be published in an edited collection within the *Routledge Studies in Surveillance* book series. Expected publication: April 2025.

Contribution:
This paper was also independently designed for this PhD, outside of the AIDA project. I was responsible for designing the study, collecting the data and analysing it. Our colleague Joan Mon Amat Rodriguez's expertise was instrumental, particularly in the design and geo-mapping analysis, enhancing the spatial interpretation of citizen reactions. Saskia also contributed significantly to the conceptual framework and manuscript, enriching the focus on urban surveillance practices and shaping the recommendations for law enforcement.

Discussion

This section critically synthesizes the key findings from the research, integrating them with the literature reviewed earlier and the insights derived from the published studies. By examining how these findings align or diverge from existing research on AI-driven surveillance, we explore their broader implications for law enforcement practices, public policy, and the development of responsible AI technologies. Furthermore, this discussion addresses the theoretical and practical implications for AI governance, citizen engagement, and the ethical use of surveillance technologies, providing recommendations for policymakers, law enforcement agencies, and technology developers to balance the benefits of AI with public trust and civil liberties.

Summary and Integration of Findings

It is becoming apparent that the integration of AI into law enforcement has significantly altered both the practice of policing and the

expectations of the public. As demonstrated across the five papers that form this thesis, the multifaceted impact of AI in surveillance reveals both opportunities for increased efficiency and operational precision in law enforcement, as well as the profound concerns of citizens regarding privacy, ownership, and the ethical use of AI in public safety. In this section, I interpret the findings of the studies, drawing out commonalities, divergences, and critical insights that align with, or diverge from, existing literature. This interpretation will aim to integrate the differing aspects addressed in this thesis to contribute to theoretical developments on AI surveillance, resistance and broader acceptance of complex technologies, as well as practical lessons on AI deployments for LEAs and AI design considerations for developers.

Table 2. Comparative Analysis of Citizen Perspectives on AI in Policing across the 3 studies

Aspect	**Paper 1 (Study 1)**	**Paper 2 (Study 1)**	**Paper 3 (Study 1)**	**Paper 4 (Study 2)**	**Paper 5 (Study 3)**
Methodology	Interviews: 111 participants across 7 countries	Interviews: 111 participants across 7 countries; contextspecific analysis	Interviews: 30 participants, 8 countries	Online experiment using Facebook: 30 UK participants	Privacy walks and surveys: 30 Sheffield residents
Main Focus	Ownership of AI tools in policing	Public attitudes towards AI in law enforcement	Perspectives on AI ownership in policing	Task difficulty and reactions to AI monitoring	Public perceptions of AI surveillance in urban settings
Participants	Diverse demographics across Europe	Varied demographics; mainly UK	Diverse demographics; included multiple countries	30 participants, average age 36, diverse backgrounds	30 Sheffield residents
Findings - Trust in Police	Trust in police ownership of AI tools was common	Police perceived as trustworthy in AI applications	Mixed trust in police vs. private entities	Trust impacted by task context; higher trust in police-related tasks	Perceptions of risk and trust affected by AI surveillance
Findings - Task Engagement	N/A	N/A	N/A	Participants engaged more with policerelated content	Engaged with identified AI tools during privacy walks

Findings - Perceived Risks	Concerns about data ownership and ethical implications	Privacy AI concerns regarding usage	Concerns about bias and accountability	Concerns about privacy and social implications	Psychological markers and perceptions of surveillance
Findings - Demographic Variations	Age and ethnicity influenced perceptions of ownership	N/A	Age, gender, and professional background impacted views	Gender differences in task perception	Variations in perceptions based on demographics
Citizen Perspectives	Support for police ownership but concerns about data misuse	Mixed feelings; some wary of AI impacts on privacy	**Five perspectives:** trust in police (1), distrust of private ownership (2), disassociation from AI (3), belief in citizen ownership (4), and uncertainty (5)	Participants felt difficulty with AI-related tasks, lower engagement in political content. High trust in police for sharing serious information.	Participants identified AI tools, showing mixed feelings; some felt morally obligated to report crimes while wary of misinformation.

The findings from this research, summarized in Table 2 above, provide important insights into public perceptions of AI-driven surveillance by law enforcement agencies (LEAs) and their broader societal implications. The research questions sought to explore the multifaceted dimensions of public engagement with AI surveillance, with particular attention to concerns over privacy, resistance strategies, and evolving attitudes towards these technologies.

1. **Perceptions of AI in Law Enforcement:** Citizens generally recognize the potential benefits of AI for enhancing security, particularly in its ability to improve efficiency and resource allocation for LEAs. However, significant concerns about privacy, ownership, and accountability persist. Participants consistently expressed apprehension regarding the lack of transparency in AI systems and the control over their data, showing a preference for decentralized ownership models where citizens have more say in how their information is handled. These findings directly address RQ.1, demonstrating that acceptance is not a simple "yes" or "no" matter but rather a complex interplay of concerns around power dynamics, trust, and control. As the results show, a more differentiated approach to measuring

acceptance is required, one that considers the various reasons for both acceptance and resistance.

2. **Safeguards and Counterstrategies:** The interviews detailed in Papers 1 and 2, along with the findings from the privacy walks in Paper 5, revealed a range of strategies that citizens employ to resist AI surveillance. These include both digital and physical methods, such as the use of privacy-enhancing tools, changes in behaviour in heavily surveilled spaces, and the avoidance of specific surveillance hotspots. The research further confirms that citizens are not passive subjects of surveillance but rather active agents who engage with and resist these technologies in various ways. This finding speaks directly to RQ.2, addressing how safeguards and counterstrategies shape public engagement with AI surveillance. It highlights the need for context-specific safeguards that account for the dynamic ways in which citizens react to and resist surveillance technologies.

3. **Evolution of Public Attitudes:** The findings also demonstrated that public attitudes towards AI surveillance evolve through direct interaction with the technology. Participants who initially had reservations became more attuned to both the ethical and practical implications of AI surveillance after experiencing it firsthand, suggesting that engagement is a key factor in shaping opinions. This is particularly evident in Papers 3 and 5, where participants' views on AI-driven surveillance shifted through deeper exposure to the technology, supporting the notion that attitudes are malleable and influenced by context and experience. This evolution in attitudes is central to RQ.3, illustrating that citizens' views are not static but evolve in response to their interactions with AI.

Overall, the findings of this thesis provide detailed answers to the research questions, contributing to the broader discourse on the societal impact of AI-driven surveillance technologies. For RQ.1, it becomes evident that acceptance is a nuanced, context-dependent phenomenon rather than a binary outcome. RQ.2 shows that citizens actively engage with AI surveillance through a range of counterstrategies, calling for more robust safeguards. Finally, RQ.3 underscores the importance of experience and engagement in shaping public attitudes, offering critical insights for future AI governance in law enforcement. The results highlight not only the ongoing concerns about privacy and control but also the evolving and active roles citizens play in resisting and shaping the future of AI surveillance.

Alignments and misalignments with the Literature on AI Surveillance

Across the studies, one of the most prominent themes is the increasing scepticism surrounding AI-enhanced surveillance and the perceived infringement on privacy and autonomy. This aligns with extensive literature on surveillance ethics, which argues that AI technologies can exacerbate existing tensions between security and privacy (Lyon, 2007; Macnish, 2021). The participants in Paper 1, where semi-structured interviews were conducted, expressed concerns about the ways AI technologies like facial recognition and biometric scanning blur the boundaries between public and private spheres. These concerns mirror the broader discourse in surveillance studies, which warns against the unchecked proliferation of AI tools without sufficient oversight (Zuboff, 2019).

What is particularly noteworthy in the findings from Papers 2 and 3 is the nuanced stance participants took when discussing AI ownership. Here, participants did not universally reject AI tools; instead, they were more concerned about who controls and accesses these technologies. Ownership of AI surveillance tools emerged as a crucial factor in shaping public opinion, with many preferring that such technologies remain under strict public control rather than private or commercial entities (Andrejevic, 2007). This finding coincides with existing research that highlights the risks of concentrating surveillance power in the hands of a few private actors, which can lead to an erosion of trust in both the technology and those who govern it (Marx, 2016).

While the general concerns about privacy align with the established discourse, one area where the findings diverge from traditional surveillance literature is in the level of conditional acceptance observed among participants. In Papers 4 and 5, which employed a mix of thinkaloud protocols and privacy walks, participants were willing to accept AI-driven surveillance under certain conditions, particularly if clear safeguards were in place and if the technology demonstrably improved public safety. This finding contrasts with the more uniformly critical view presented in earlier studies, where surveillance was often seen as an unequivocal infringement on personal liberties (Timan & Albrechtslund, 2017).

This divergence suggests a shift in public attitudes, likely due to the increasing normalization of surveillance technologies in everyday life. Citizens seem to be adapting to the presence of these technologies and, rather than outright rejecting them, are now seeking more participatory roles in the governance of AI (Wood & Thompson, 2018). This reflects a growing movement toward "surveillance citizenship," where citizens

demand a say in how surveillance is deployed and monitored (Gilliom, 2001).

Citizens' Perspectives on AI Ownership and Safeguards

One of the most significant findings from this research lies in how citizens perceive AI ownership and the necessary safeguards surrounding its use. Participants expressed varying degrees of concern over how AIdriven surveillance technologies are owned, operated, and governed. The ownership of AI technologies in law enforcement is often perceived as an issue of control and accountability, raising concerns about potential misuse by centralized authorities or private entities (Andrejevic, 2007). Ownership, in this context, is intrinsically tied to transparency, as the entity that controls AI technologies holds significant power over the collection, storage, and use of personal data (Zuboff, 2019).

Participants across all studies emphasized the importance of having clear governance structures in place, expressing scepticism over AI tools owned or operated by private corporations. Their concerns align with the broader theoretical discussions on power dynamics in surveillance, where the concentration of surveillance capabilities in a few hands is seen as a potential risk to democratic accountability and civil liberties (Macnish, 2021). While many participants supported the use of AI for enhancing security, they were particularly sensitive to how these tools are managed and operated, reflecting a nuanced understanding of the balance between security and privacy.

The concept of ownership of AI surveillance tools also intersects with the broader debate on public trust. Trust in AI technologies is shaped by who owns the systems and how transparent the processes are. Research participants often expressed a preference for public oversight or independent regulatory bodies to ensure that AI surveillance systems are used ethically and responsibly. This aligns with Constantinescu et al. (2021), who emphasize the importance of involving diverse stakeholders in AI governance to mitigate potential risks and abuses. Participants saw the need for stringent safeguards, including clear rules about data access, retention, and usage. Many voiced concerns that without robust safeguards, AI could be easily misused for political or commercial purposes.

This brings us to the role of safeguards, particularly in relation to transparency and accountability, extends beyond ownership where citizens emphasized that safeguards need to be proactive rather than reactive, ensuring that potential harms are mitigated before they occur. This highlights a critical theoretical perspective on AI governance,

where ownership and safeguards are not just practical concerns but are deeply embedded in broader discourses on surveillance, privacy, and the legitimacy of state power (Floridi et al., 2018). Participants' calls for increased transparency and stricter regulatory frameworks suggest that public trust in AI surveillance can be bolstered if safeguards are not just present but are actively enforced and continuously reviewed.

Hence, the findings from this research highlight the complex relationship between AI ownership, public trust, and the need for comprehensive safeguards. While citizens recognize the benefits of AI in law enforcement, their concerns around ownership reflect broader societal fears about surveillance overreach and the concentration of power. The need for clear, enforceable safeguards is critical in shaping public attitudes toward AI, and this research contributes to the ongoing discourse on how best to balance security, privacy, and accountability in the age of AI.

Resistance to AI Surveillance and the role of Counterstrategies

Resistance to AI surveillance has been a focal point in the literature, with a growing body of research addressing how individuals and communities challenge and evade surveillance systems. Much of this research has highlighted various counterstrategies, both digital and physical, as mechanisms of resistance. For instance, studies have demonstrated that digital tools, such as encryption and VPNs, are often employed by individuals to protect their online privacy, while physical countermeasures—ranging from avoiding surveillance-heavy areas to using clothing or makeup to obscure facial recognition—are frequently cited as methods to evade physical surveillance (Monahan, 2006; Chen et al., 2015).

Resistance to surveillance is deeply rooted in historical practices but has evolved significantly with the advent of digital technologies. The rise of AI and the internet has facilitated new forms of resistance that are more coordinated and widespread (Monahan, 2006). Encryption tools and anonymizing technologies, such as Tor browser, have become essential for individuals seeking to protect their privacy in an increasingly surveilled digital landscape (Saura et al., 2022). Additionally, several civil society organizations have developed platforms and tools to expose and challenge the use of AI surveillance by governments and corporations. For instance, organizations like the Patrick J. McGovern Foundation provide AI tools to mission-driven organizations that advocate for transparency and diverse representation in AI governance. These platforms help empower marginalized communities by raising their voices in discussions around AI governance and data rights (RAND Corporation, 2020; World Economic Forum, 2020).

The findings from Papers 1, 2, and 5 of this PhD add complexity to these traditional understandings of resistance. While the literature emphasizes digital counterstrategies like encryption, my studies reveal that these are not as commonly used in practice as one might expect. Although participants expressed concern over AI surveillance, few actively employed advanced digital tools such as VPNs or anonymizing technologies. Instead, many relied on less sophisticated methods, such as limiting their online presence or avoiding surveillance-heavy spaces in physical environments, reflecting a more passive form of resistance. This suggests a gap between awareness of digital counterstrategies and their actual implementation, potentially due to barriers such as lack of access or technical know-how.

In contrast, Paper 5 found that physical resistance strategies, particularly during privacy walks, were more immediate and prevalent. Participants altered their routes, avoided certain areas, and expressed heightened vigilance when they encountered visible surveillance technologies, aligning with the physical avoidance tactics documented in surveillance literature. Notably, these behaviours were more pronounced in environments where participants perceived the surveillance as intrusive or unjustified. In cases where surveillance was framed as necessary for public safety—such as in high-crime areas—participants were less likely to engage in overt resistance, suggesting that the perceived legitimacy of AI surveillance plays a significant role in shaping public reactions.

The policing and security context of this research also revealed unique findings regarding resistance. In Papers 1 and 3, participants expressed a complicated relationship with AI surveillance technologies in policing. On one hand, they were concerned about the erosion of privacy and potential misuse of their data. On the other hand, they acknowledged the potential benefits of AI for improving law enforcement effectiveness and public safety. This tension between privacy and security often resulted in ambivalence towards resistance strategies: while participants valued privacy, they were less likely to engage in overt resistance when AI surveillance was perceived as contributing to a safer environment. This underscores the need for law enforcement agencies to balance surveillance practices with transparent and accountable governance to reduce resistance and foster public trust.

Moreover, unlike the organized forms of resistance observed in other sectors, such as campaigns against corporate surveillance, my studies found little evidence of collective resistance efforts specifically targeting AI surveillance in policing. The absence of such movements may reflect the public's perception of AI surveillance in law

enforcement as inevitable or necessary for security, thus discouraging more organized opposition. This highlights a critical difference in how resistance manifests in different contexts, with law enforcement and security settings potentially mitigating the emergence of collective counterstrategies in favour of more individualised or passive forms of resistance.

Furthermore, the role of visibility in shaping resistance cannot be overlooked. Paper 5 revealed that participants were more likely to resist surveillance when it was visibly present, such as in the form of cameras or biometric scanners. This contrasts with more covert forms of digital surveillance, where the invisibility of data collection processes might make it harder for individuals to identify when and how they are being monitored, leading to less immediate resistance. The visible nature of AI surveillance in public spaces heightened participants' awareness, resulting in more tangible resistance behaviours like route changes and avoidance of surveillance-heavy areas.

In conclusion, the findings from this thesis extend existing research on resistance to AI surveillance by demonstrating that physical avoidance is more prevalent than digital counterstrategies in the policing context.

Moreover, the legitimacy of AI surveillance and its perceived benefits for public safety significantly mediate how citizens respond, with less resistance observed when surveillance is viewed as necessary for security. This suggests that transparency, accountability, and public engagement are crucial for reducing resistance and fostering acceptance of AI surveillance technologies in law enforcement contexts.

The Role of *Context* and Surveillance Environment

This research highlights the importance of context in shaping public perceptions and behaviours towards AI-driven surveillance. Across various public spaces, including city streets and areas where individuals frequently interact with surveillance technologies, the presence of AI surveillance tools such as facial recognition cameras and biometric scanners evokes a range of reactions. In some instances, as observed during the privacy walks, participants became more aware of their surroundings and changed their behaviour when they noticed the overt use of surveillance. This was particularly evident in areas with visible CCTV systems, where participants adjusted their routes or even altered their actions to avoid being monitored. These findings suggest that the visibility and perceived purpose of surveillance technologies significantly affect how people engage with their environment and the extent to which they perceive a threat to their privacy.

The use of geo-mapping to track participant movements coupled with responses from the post-walk surveys revealed that some participants consciously altered their routes to avoid heavily surveilled areas, underscoring how AI surveillance can shape not only behaviour but also the spatial experience of public spaces (Parviainen & Ridell, 2021). This finding is particularly important for the development of smart cities, where AI surveillance is deeply embedded in urban infrastructures, raising questions about the trade-off between technological efficiency and personal privacy.

Moreover, participants in experimental settings, particularly those involving think-aloud protocols, expressed heightened concerns when interacting directly with AI technologies. The data suggest that individuals are more likely to resist or exhibit scepticism towards AI surveillance when they feel that the technology is used without transparency or proper safeguards. This awareness, coupled with the findings from pre- and post-task surveys, demonstrates that participants' attitudes and behaviours shift in response to the presence and transparency of AI systems in public environments (Park & Yoon, 2024). This underscores the critical role of context in how surveillance technologies are perceived and the ways in which they influence daily interactions in spaces where monitoring is prevalent.

Not to forget that, in policing, context-specific AI frameworks are essential to ensuring that these systems respect the unique privacy and security needs of communities. As Jacobs (2024b) asserts, smart technologies, such as AI, should aim to unite citizens and law enforcement rather than create societal friction. Indeed, the findings from our studies highlighted how citizens' trust in AI surveillance depends heavily on the clarity of its purpose and its transparency, reinforcing Jacobs' stance that public trust is a necessary pillar for effective AI governance.

Ethical Concerns: Algorithmic Bias and Accountability

The ethical implications of AI surveillance are a recurring theme throughout the studies, particularly regarding algorithmic bias and accountability. Papers 1, 3, and 4 highlight participants' concerns about the potential for AI to reinforce existing social inequalities. This aligns with research by Buolamwini & Gebru (2018), who demonstrated that AI systems, particularly facial recognition technologies, often exhibit racial and gender biases due to the data sets used to train these algorithms. Participants in these studies voiced concerns that AI-driven surveillance could disproportionately target marginalized communities, leading to issues of social justice.

The findings of Study 1, particularly as discussed in Paper 3, which focused on citizens' perspectives on AI ownership, further underscore the need for accountability in AI governance. Participants called for transparent decision-making processes and oversight mechanisms that would allow for public scrutiny of AI surveillance tools. This reflects broader calls in the literature for accountable AI, where clear lines of responsibility are established to ensure that these technologies are used ethically and fairly (Macnish, 2021). The findings suggest that without such accountability, public trust in AI-driven law enforcement will remain fragile, limiting the potential benefits of these technologies.

Briefly said, following careful evaluation of the findings across the five papers, several key insights emerge. First, while there is a growing acceptance of AI surveillance under specific conditions, concerns about privacy, ownership, and accountability remain paramount. The findings align with broader trends in surveillance studies, particularly regarding the need for stringent safeguards and public participation in the governance of AI tools. At the same time, the findings highlight the contextual nature of surveillance, with participants' reactions varying significantly depending on whether they are interacting with AI in online or offline environments.

In light of these findings, it is evident that AI surveillance cannot be viewed solely through the lens of technological efficiency. Instead, it must be understood as a complex sociotechnical system deeply intertwined with issues of power, control, and public trust (Lyon, 2007; Macnish, 2021). The findings from the three studies conducted as part of this research—spanning citizen perspectives on ownership and safeguards, their reactions to online and offline surveillance, and their resistance strategies—highlight the critical need for more accountable AI governance.

One of the key contributions of this thesis lies in revealing how citizens' perceptions of algorithmic bias are shaped by both their direct interactions with AI systems and their broader experiences of state surveillance. The qualitative data from the first and second studies, which explored citizen perspectives across different European contexts, found that participants consistently expressed concerns about algorithmic fairness, particularly in how AI might reinforce existing social inequalities (Buolamwini & Gebru, 2018; Eubanks, 2018). This adds empirical weight to the broader theoretical discourse on AI fairness and bias, emphasizing that citizens view these issues not merely as technical challenges but as ethical imperatives that impact trust in law enforcement agencies (LEAs). The thesis makes a critical contribution by documenting these perspectives and highlighting the

disconnect between technical solutions to bias and public expectations of fairness.

Moreover, the findings from the third study demonstrate that concerns about accountability are not abstract but directly related to the perceived opacity of AI decision-making processes (Diakopoulos, 2016). Citizens are wary of "black box" AI systems and demand transparency in how surveillance decisions are made, particularly in sensitive areas such as predictive policing (Richardson et al., 2019). This aligns with broader calls in the literature for "explainable AI" but also underscores the necessity for transparency to be more than a technical feature (Brundage et al., 2020). It must involve clear communication strategies that bridge the gap between complex algorithms and layperson understanding. The research contributes to this debate by offering practical insights into how citizens want transparency to be enacted—through direct access to information, clearer accountability mechanisms, and participatory governance models where they feel empowered to voice their concerns (Floridi, 2019).

Finally, the research underscores the importance of embedding ethical safeguards into the design and deployment of AI surveillance systems (Jobin et al., 2019). The experimental methods and privacy walks used in the third study revealed that citizens' trust in AI technologies is contingent on visible and enforceable safeguards (Constantinescu et al., 2021). The pre- and post-task surveys showed a marked shift in participant attitudes when they were made aware of the limitations and controls placed on AI surveillance. This suggests that accountability cannot be a secondary consideration but must be an integral part of system design, with robust checks and balances that are visible to the public (Agarwal et al., 2020). This research provides empirical evidence that accountability mechanisms—when transparent and accessible—can significantly mitigate concerns about AI surveillance and foster greater public trust.

In conclusion, this thesis contributes to the ongoing discussion of algorithmic bias and accountability by providing a nuanced, empirically grounded understanding of how citizens perceive and respond to these issues. It offers practical recommendations for policymakers, LEAs, and technology developers to prioritize fairness, transparency, and accountability in the development and deployment of AI surveillance (Rahwan et al., 2019). By foregrounding citizen perspectives and integrating them into the broader discourse, this research bridges the gap between theoretical debates on AI ethics and the practical realities of public safety in the digital age.

Theoretical Implications

The findings across the five studies reveal substantial theoretical implications for the study of AI governance, digital surveillance, and citizen engagement, challenging and extending existing frameworks in surveillance studies, specifically within the context of policing. AIdriven policing technologies present unique ethical and operational challenges distinct from broader surveillance capitalism or consumertargeted AI systems. While much of the surveillance literature focuses on the commodification of data for economic gain, AI surveillance in law enforcement operates within a different framework. Rather than being driven by commercial interests, AI surveillance in policing is often fundamentally tied to state power, social control, and the governance of public safety. AI technologies in this context are deployed not to generate profit but to predict, prevent, and respond to criminal activities, reflecting a complex interplay between authority and civil liberties (Lyon, 2007). This dynamic creates unique ethical and operational challenges, as law enforcement agencies must balance the need for enhanced surveillance capabilities with the imperative to respect citizens' rights and maintain public trust (Babuta & Oswald, 2020).

Within the policing context, the introduction of AI challenges the existing surveillance frameworks by transforming not only the tools of surveillance but also the dynamics of citizen interaction with LEAs. The anticipatory nature of AI in predictive policing, as revealed in studies 1 and 3, highlights a future-oriented surveillance strategy that alters traditional notions of privacy and accountability. Unlike conventional surveillance, which relies on retrospective observation, AI systems in policing predict and act on potential future behaviour, raising significant concerns about pre-crime interventions and the erosion of individual autonomy. This aspect of policing, unique to AI, introduces new dimensions to power relations between the state and its citizens, as law enforcement agencies leverage predictive algorithms to control crime but may also inadvertently intensify the stigmatization of certain communities (Richardson et al., 2019).

Moreover, the policing context introduces nuanced issues around citizen resistance and counterstrategies. The findings from studies 2 and 3 demonstrate that individuals not only adapt their behaviours in response to AI surveillance but actively engage in subtle resistance. For example, physical counterstrategies such as altering one’s route or adopting specific attire to avoid detection by facial recognition cameras illustrate how citizens attempt to reclaim agency within an AI-governed space (Mann et al., 2003). This is particularly significant in the policing

domain, where citizens are acutely aware of the coercive power of surveillance and the potential implications for their civil liberties. Traditional theories of surveillance, which focus heavily on passive forms of compliance, fail to account for these active forms of resistance that are more common in high-stakes environments like law enforcement.

Another important extension of existing theoretical frameworks pertains to the balance between security and privacy in AI-driven policing. As revealed in findings of study 2, public acceptance of AI technologies is contingent on their perceived balance between improving public safety and protecting individual privacy rights. This finding aligns with broader debates in AI ethics but introduces a policing-specific dimension: citizens often accept some level of surveillance in the interest of security, but they demand stringent safeguards, particularly when AI is involved.

The study highlights the crucial role of transparency and accountability in fostering public trust—key components that are currently underdeveloped in both theoretical discussions and practical applications of AI in policing (Floridi et al., 2018). In this context, the research extends theoretical models of AI acceptance by demonstrating that acceptance is not only about transparency but also about visible, enforceable safeguards that address power imbalances between the surveilled and the surveillant.

Briefly said, the findings challenge traditional surveillance theories by demonstrating that AI policing technologies shift the balance of power in new ways, introduce predictive and pre-emptive surveillance mechanisms, and prompt diverse citizen responses that range from acceptance to active resistance. These implications contribute to the broader theoretical landscape by underscoring the need for AI governance models specifically tailored to policing, where issues of autonomy, power, and public safety intersect with concerns about privacy and civil liberties.

AI Governance and Surveillance

The issue of AI governance is at the forefront of the theoretical implications drawn from the findings. The role of AI in enhancing law enforcement capabilities raises urgent questions about how these technologies are controlled, who oversees their deployment, and how they should be regulated to ensure ethical outcomes. Existing theories of governance in technology often emphasize the need for multistakeholder involvement and public accountability (Whittaker & Van der Ploeg, 2018). However, the findings from the studies suggest a

significant gap between the current state of AI governance in law enforcement and the expectations of the public.

In Paper 3, which focused on citizens' perspectives on AI ownership, the issue of governance was particularly salient. Participants expressed concerns about the centralization of AI tools in the hands of law enforcement agencies without sufficient oversight, mirroring Marx's (2016) concerns around power imbalances in surveillance. The findings indicate a demand for more participatory forms of governance, where citizens are not only informed about AI technologies but also have a say in their deployment. This aligns with the long-standing concept of "surveillance citizenship" proposed by Gilliom (2001), which calls for greater public involvement in the governance of surveillance technologies.

Moreover, while "surveillance capitalism" predominantly refers to the commodification of personal data by private corporations (Zuboff, 2019), the context of policing AI introduces a different set of risks. In law enforcement, the issue is not solely about corporate ownership but about the governance and accountability of AI technologies in statecontrolled security frameworks. The findings from all three studies touches on citizens awareness of AI surveillance tools, even when deployed by law enforcement agencies, being reliant on technologies developed by private entities, which raises concerns over accountability, transparency, and public oversight.

To address these challenges, AI governance frameworks in policing need to evolve by incorporating specific provisions that ensure greater transparency in the partnerships between public and private entities. This could involve clearer regulations on data sharing, joint accountability mechanisms, and mandatory audits of AI systems used in public security (Babuta & Oswald, 2020) as clearly called for by participants in study 2. Moreover, governance frameworks should emphasize ethical AI practices, including the need for citizen participation in decision-making processes related to AI surveillance in policing. By engaging with communities and civil society organizations, law enforcement agencies can foster greater public trust and ensure that AI technologies are aligned with societal values and expectations (Floridi et al., 2018). Such frameworks should also include strict limitations on the types of data that can be collected and how it can be used, ensuring that surveillance tools do not infringe upon privacy or civil liberties unnecessarily.

Digital Surveillance and Algorithmic Power

The rise of AI in law enforcement introduces a new form of algorithmic power, where decision-making processes are increasingly delegated to AI systems that operate based on algorithms and data analytics (Andrejevic, 2007). The theoretical implications of this shift are significant, particularly in terms of how AI reinforces existing power structures and challenges traditional notions of accountability. The findings from Papers 1 and 4, which explored citizens' attitudes toward AI-driven surveillance, illustrate the complex interplay between algorithmic power and public trust. Participants expressed concerns about the opacity of AI systems, echoing wider debates in the literature about the "black box" nature of AI (Pasquale, 2015). The lack of transparency in how AI systems make decisions, particularly in the context of predictive policing, raises concerns about fairness, bias, and the potential for discriminatory practices (Richardson et al., 2019).

This aligns with the theoretical work on algorithmic bias, which posits that AI systems are not neutral but reflect and amplify the biases present in the data they are trained on (Buolamwini & Gebru, 2018). The findings from Paper 4, which used the think-aloud method to capture participants' real-time reactions to AI surveillance systems, provide concrete evidence of how citizens perceive these biases in action. Participants reported feeling uncomfortable with the idea that AI could target certain groups disproportionately, particularly in light of the racial and gender biases already present in AI facial recognition systems (Benjamin, 2019). This suggests that AI surveillance, far from being a neutral tool, functions as an extension of existing power dynamics, reinforcing social inequalities rather than addressing them.

The Role of *Context* in Surveillance Theory

The findings from the studies also underscore the importance of context in shaping public responses to AI surveillance. As noted in the literature on spatial surveillance, the environment in which surveillance takes place significantly influences how individuals perceive and react to it (Klauser & Albrechtslund, 2014). Papers 4 and 5, which utilized privacy walks and geo-mapping techniques, revealed that participants were more likely to resist surveillance in highly visible, public spaces, while their reactions to online surveillance were more ambivalent.

This highlights a critical gap in current surveillance theory, which often treats surveillance as a uniform phenomenon, failing to account for the spatial and contextual nuances that influence public attitudes (van Es & de Lange, 2020). The findings suggest that a more granular understanding of surveillance is needed, one that accounts for the different ways in which AI technologies are embedded in physical and

digital environments. This extends the work of Foucault (1977) on panopticism, which focuses on the ways surveillance shapes behaviour, by incorporating the idea that the spatial dynamics of surveillance play a crucial role in how individuals respond to being watched.

Moreover, the findings suggest that resistance to AI surveillance is also context dependent. Participants in Paper 5, for example, were more likely to engage in counter-surveillance strategies in offline environments, such as altering their routes to avoid cameras, than in online spaces. This supports Timan and Albrechtslund's (2017) argument that resistance to surveillance is not a uniform phenomenon but varies according to the context in which surveillance occurs. The implications for surveillance theory are profound, as they call for a more context-sensitive approach to understanding how individuals navigate and resist AI surveillance technologies.

Citizen Engagement and Resistance

Finally, the findings from Papers 1, 3, and 5 highlight the critical role of citizen engagement in shaping the future of AI-driven surveillance. The notion of resistance, both passive and active, emerges as a key theme, challenging traditional assumptions about the passive role of individuals in surveillance systems (Monahan, 2006). The privacy walks in Paper 5, for instance, revealed that citizens are not merely passive subjects of surveillance but actively reflect on and engage in strategies to evade and resist it.

This aligns with the theoretical concept of "sousveillance," where individuals turn the tools of surveillance back onto those in power (Mann & Ferenbok, 2013). The findings suggest that as AI technologies become more pervasive, citizens are becoming more aware of the need to protect their privacy and autonomy, whether through digital counterstrategies such as encryption, or physical strategies such as avoiding surveillance hotspots. This has significant implications for how we theorize citizen engagement with AI surveillance, suggesting that resistance is not only an individual act but a collective response to the growing reach of surveillance technologies.

Expanding on this, the implications extend beyond the idea of resistance as merely a reactive or defensive mechanism. Instead, these findings propose a reconceptualization of resistance as a form of regular active civic engagement. Citizens are not only resisting surveillance but are also actively shaping the discourse around privacy, surveillance ethics, and the boundaries of state power. In this context, resistance becomes an essential aspect of democratic participation, as citizens collectively assert their rights and challenge the unchecked spread of AI

surveillance technologies. This shift in thinking moves us beyond the traditional view of resistance as isolated or passive and highlights how resistance can influence the very design and deployment of AI technologies in policing.

Moreover, this collective response underscores the potential for organized movements, such as the "Reclaim Your Face" campaign or even entities such as "Big Brother Watch", a UK-based civil liberties organization, which has spearheaded efforts to push back against state surveillance initiatives. Their campaigns have led to increased public scrutiny and legal challenges against facial recognition and AI surveillance technologies used by the police to have a tangible impact on policy and regulation. As AI surveillance becomes more integrated into public spaces, these collective actions signal a shift in the power dynamics between the state and citizens. The interplay between statedriven surveillance and citizen-driven resistance creates a dynamic landscape where resistance strategies evolve alongside technological advancements, leading to new forms of counter-surveillance and civic activism. Therefore, theorizing resistance in this context means acknowledging it as both a social and political movement that challenges the balance between security and privacy while advocating for greater accountability and transparency in AI governance.

Hence, it is safe to say that the theoretical implications of the findings extend beyond the immediate context of law enforcement and surveillance, contributing to broader debates about AI governance, algorithmic power, and citizen engagement. The studies challenge existing frameworks in surveillance theory, calling for more participatory governance models, a deeper understanding of algorithmic bias, and a context-sensitive approach to surveillance. These insights offer valuable contributions to the growing field of AI ethics and governance, highlighting the need for a more inclusive and accountable approach to the deployment of AI in law enforcement.

Practical Recommendations

Understanding and Handling/Preventing Resistance

Effectively addressing citizen resistance to AI surveillance necessitates a nuanced understanding of the diverse audiences identified across the studies. The findings reveal that individuals' perceptions and responses to AI surveillance are deeply influenced by their demographic characteristics, cultural backgrounds, and personal experiences with authority (Ezzeddine et al., 2023a). Therefore, tailored communication strategies that resonate with specific groups are essential for fostering understanding and acceptance of AI technologies in policing. For

instance, engaging younger populations who demonstrate greater digital literacy may involve leveraging social media platforms to disseminate information about the ethical use of AI and its benefits for public safety. This approach aligns with the insights from Study 2, particularly as discussed in Paper 3, where participants expressed a desire for transparency in AI governance (Ezzeddine et al., 2023b).

Conversely, reaching older demographics may require more traditional communication methods, such as community meetings or printed materials, to address their concerns regarding privacy and surveillance directly. Research emphasizes the importance of transparency and public engagement, suggesting that LEAs should actively involve communities in discussions about AI surveillance policies and practices (Petersen & Taylor, 2012). Additionally, findings from studies 1 and 2 (namely the ones published in Papers 1 and 4) indicate that acknowledging citizens' fears and hesitations can reduce resistance; thus, fostering an environment where concerns can be openly discussed is vital for building trust.

Moreover, to prevent resistance, it is crucial for policymakers and LEAs to adopt a collaborative approach that includes citizen feedback in the design and implementation of AI technologies. This engagement can facilitate a sense of ownership among community members, leading to more favourable attitudes towards surveillance initiatives. As noted by Constantinescu et al. (2021), participatory governance models that emphasize citizen involvement not only enhance public trust but also contribute to the legitimacy of surveillance practices. Ultimately, the ability to communicate effectively with disparate audiences, while demonstrating a genuine commitment to ethical AI use and transparency, is fundamental in navigating the complexities of citizen engagement and mitigating resistance to AI surveillance.

Audience-focused Recommendations

The findings from this research have important practical implications for LEAs, policymakers, and technology developers. As AI-driven surveillance becomes more pervasive, there is a pressing need to balance the benefits of these technologies with the ethical concerns and public apprehensions they raise. This section provides practical recommendations that are grounded in the research findings and are aimed at fostering trust, ensuring accountability, and promoting responsible AI deployment in law enforcement. By considering the concerns and expectations of different stakeholders such as surveillance scholars, police practitioners, AI developers, psychologists, criminologists, and public service providers, these

recommendations ensure that AI technologies are deployed responsibly, ethically, and in ways that foster public trust. Integrating these findings into existing governance frameworks will not only improve the effectiveness of AI in law enforcement but also enhance the broader social acceptance of these powerful technologies.

Recommendations for Law Enforcement Agencies

For police practitioners, particularly in agencies like the Metropolitan Police (MET), the challenge lies in effectively deploying AI surveillance technologies such as facial recognition while minimizing public resistance and maintaining trust. Resistance to AI surveillance is often rooted in a lack of transparency, perceived overreach, and concerns about data privacy. To address these concerns, law enforcement agencies should engage in open dialogues with the public, ensuring that AI technology implementation is transparent, and that citizens understand the benefits, limitations, and safeguards of these technologies.

1. Mitigating Resistance Through Public Engagement

Engaging local communities early in the deployment of AI surveillance tools can significantly reduce public resistance. Involving citizens in discussions about how AI technologies will be used fosters trust and improves public acceptance. Transparency about how AI systems operate—what data are collected, and how decisions are made—is crucial for minimizing concerns (Campion, et al., 2020). Additionally, AI systems should incorporate privacy-enhancing features such as anonymization and differential privacy to address fears of surveillance overreach (Zyskind, et al., 2015).

2. Context-Specific Deployment

The findings demonstrate that facial recognition technologies and other AI surveillance tools must be adapted to specific contexts. Deploying facial recognition at large public gatherings or protests may elicit stronger resistance compared to using the technology at border controls or airports, where the public perceives a more legitimate need for such surveillance. Law enforcement agencies should conduct risk assessments to identify contexts where AI technologies can be deployed with minimal social backlash and public support is likely to be higher (Dempsey, 2020).

Additionally, building on Jacobs' (2024b) research on integrating smart technologies, AI developers and policymakers must collaborate to design AI systems that respect the unique socio-political landscapes of communities. Rather than a one-size-fits-all approach, AI systems

should be tailored to meet the varied needs of urban and rural environments, enhancing the community-oriented policing framework. By integrating AI systems that reflect community priorities and respecting civil liberties, law enforcement can gain public cooperation rather than resistance.

3. Engage with the Public in AI Governance

Finally, the studies underscore the importance of involving citizens in the governance of AI surveillance. Law enforcement agencies should establish channels for public participation in decision-making processes related to AI tool deployment. Public consultations, community forums, or citizen advisory boards could provide input on the development and use of AI technologies in policing (Wood & Thompson, 2018). Engaging with the public ensures that law enforcement is attuned to community concerns, fostering trust and legitimacy in the use of AI. This could be done through:

- **Targeted Communication and Addressing Diverse Perspectives:** Based on the findings, LEAs must recognise that public perceptions of AI surveillance are not homogeneous. Citizens have disparate perspectives on privacy, security, and trust in law enforcement. Therefore, police forces should adopt targeted communication strategies that address these differing concerns directly. For example:

- **Privacy-Sensitive Groups:** For individuals who are particularly concerned about privacy, agencies should emphasize the safeguards in place, such as data minimization, encryption, and strict access controls. Clear information on who has access to data, how long it is stored, and how it is anonymized should be provided. This could be communicated through public information campaigns, local forums, or online platforms that explain how AI technologies are being used responsibly and in line with privacy laws.
- **Security-Oriented Groups:** For those who prioritize security over privacy concerns, agencies should focus communication on the tangible benefits of AI surveillance in reducing crime and enhancing public safety. This could involve sharing case studies or statistics demonstrating how AI technologies have contributed to solving crimes, preventing threats, or ensuring faster responses to incidents in their communities.
- **Addressing Trust and Transparency Concerns:** To improve trust, agencies must go beyond merely stating that AI tools are

beneficial; they must provide demonstrable examples of how transparency and accountability mechanisms are built into the use of these technologies. Regular updates on AI performance, data audits, and public-facing reports on how AI surveillance tools are operating within the bounds of legal frameworks can help address concerns. Additionally, the establishment of independent oversight bodies to monitor AI surveillance practices can reinforce public confidence in law enforcement's responsible use of these tools (Macnish, 2021).

By tailoring their communication strategies and addressing the unique privacy and security needs of different segments of the public, police forces can reduce resistance and foster a more nuanced understanding of the role of AI surveillance in policing. This targeted approach ensures that law enforcement is not only responsive to public concerns but also proactive in building trust across diverse communities.

Recommendations for Policymakers

Based on the findings of the three studies that constitute the heart of this thesis, it is essential to propose targeted, context-specific recommendations that address the nuanced concerns around AI governance, privacy, and public trust. The complexity of AI-driven surveillance, as revealed through the diverse reactions of citizens across multiple countries, underscores the need for multifaceted solutions that are adaptable to different social, political, and cultural contexts.

1. Reforming AI Governance Frameworks for Policing

One of the central findings from the studies is the critical role of transparency and accountability in shaping public perceptions of AI surveillance. Participants across all studies expressed concerns about who controls AI surveillance systems and how data is collected and used (Ezzeddine et al., 2023a; 2023b). As such, AI governance frameworks need to be fundamentally restructured to prioritize transparency at every level. This requires:

- **Public Disclosure and Accountability Mechanisms**: LEAs should be required to publicly disclose the AI systems they deploy, how these systems function, and the specific data they collect. This aligns with the NPCC AI Covenant, which emphasizes transparency as a cornerstone for fostering public trust in AI technologies in policing (NPCC, 2023). By ensuring public disclosure, LEAs can provide citizens with the information necessary to understand AI deployments and hold law enforcement accountable for the use of such tools. Furthermore,

the AP4AI framework advocates for comprehensive accountability mechanisms, highlighting the importance of real-time audits and independent oversight. The establishment of independent bodies with the authority to audit and assess AI deployments would ensure that these technologies adhere to ethical standards and legal obligations, reinforcing public trust and preventing misuse (Akhgar et al., 2022). This mirrors the preferences expressed in Studies 1 and 2, where participants underscored the need for transparent governance and clear ownership structures for AI surveillance tools (Ezzeddine et al., 2023a; 2023b).

- **Algorithmic Accountability**: As identified in the findings, many participants were concerned about algorithmic biases in AI systems and their potential to perpetuate discrimination, particularly towards marginalized groups (Ezzeddine et al., 2023a). AI governance frameworks must therefore mandate the regular auditing of AI systems for biases and the publication of these audit results. These measures are crucial to ensuring that AI systems do not exacerbate existing inequalities, but instead contribute to fair and equitable policing (Richardson et al., 2019). Regular audits and the involvement of third-party institutions to validate the accuracy and fairness of these technologies could address these gaps.

2. Enhancing Public Engagement and Involvement

Another key finding across the studies is the need for greater citizen involvement in the design, implementation, and monitoring of AI technologies used in policing. As revealed in the findings of Study 2, public distrust towards AI surveillance is often rooted in the lack of meaningful engagement between LEAs and the communities they serve (Ezzeddine et al., 2023b). Echoing Jacobs' (2024a) argument that technological developments should prioritize community cohesion, AI developers should engage with local stakeholders throughout the development process. Incorporating feedback from diverse communities during the design phase ensures that AI tools respect public expectations and concerns. This engagement is crucial to aligning AI surveillance practices with this vision, which advocates for smart technologies that promote unity rather than division.

Therefore, the following serve as practical recommendations for policymakers:

- **Implement Participatory Governance Models**: Policymakers must create platforms for public consultation and participation in

AI governance. This aligns with the findings from Study 1, where participants stressed the importance of community involvement in determining how AI surveillance should be used (Ezzeddine et al., 2023a). By instituting participatory governance models, citizens can have a voice in shaping the regulations and policies that govern AI surveillance, fostering a sense of ownership and trust in these systems (Constantinescu et al., 2021). Regular town halls, public consultations, and digital platforms for feedback are examples of mechanisms that could facilitate this engagement.

- **Citizen Advisory Boards**: LEAs could benefit from establishing Citizen Advisory Boards that regularly review and provide feedback on the deployment of AI surveillance technologies. These boards could consist of community representatives, civil rights advocates, and technology experts who would act as intermediaries between the public and law enforcement, ensuring that the deployment of AI technologies aligns with societal values and public expectations (Babuta & Oswald, 2020).

3. Building Robust Data Privacy Protections

One of the most significant concerns expressed by participants, particularly in Papers 3 and 5, is the threat AI surveillance poses to data privacy and autonomy. The findings suggest that participants are particularly wary of data misuse and the lack of consent in data collection (Ezzeddine et al., 2023b). To address these concerns, policymakers should:

- **Strengthen Data Privacy Regulations**: There must be robust legal frameworks in place to govern how data collected through AI surveillance is stored, shared, and used. These frameworks should require explicit user consent before any personal data is collected and mandate that citizens be given the right to opt out of AI surveillance where possible. This recommendation directly responds to concerns voiced by participants in Paper 5, where privacy walks revealed heightened sensitivity to the presence of AI surveillance in public spaces (Ezzeddine et al., 2023a; 2025).

  Existing regulations, such as the European Union's General Data Protection Regulation (GDPR), can serve as a model, but they need to be tailored to address the unique challenges posed by AI in law enforcement (EDRi, 2021).

- **Right to Explanation and Redress**: As identified in the findings from Paper 4, participants felt uncomfortable with AI systems that made decisions without providing adequate explanations

(Ezzeddine et al., 2023b). Policymakers must ensure that individuals have the right to be informed about the decisions made by AI systems and the rationale behind them. This could be facilitated through "right to explanation" clauses in AI governance policies, providing individuals with the ability to challenge decisions made by AI surveillance tools. Additionally, clear legal pathways for redress should be established to allow citizens to seek justice in cases where AI surveillance systems violate their privacy or rights (Akhgar et al., 2022; Macnish, 2021).

4. Addressing Algorithmic Bias and Promoting Fairness

As the findings reveal, algorithmic bias remains a significant challenge to the fair and just application of AI in law enforcement. Participants, especially those in marginalized communities, expressed concerns that AI surveillance systems could exacerbate existing biases and reinforce discriminatory practices (Ezzeddine et al., 2023a). Policymakers need to develop governance frameworks that:

- **Incorporate Fairness-by-Design Principles**: AI systems used in policing should be designed with fairness and equity as core objectives. This requires a concerted effort to ensure that AI models are trained on diverse datasets that reflect the populations they serve. As highlighted in Paper 2, citizens raised concerns about biased data feeding into AI systems, leading to skewed results and discriminatory outcomes (Ezzeddine et al., 2023a). Policymakers should enforce regulations that require LEAs to demonstrate that their AI systems have undergone rigorous fairness testing before deployment (Babuta & Oswald, 2020).

- **Diverse Representation in AI Development**: To mitigate the risks of algorithmic bias, there needs to be greater diversity in the teams responsible for developing AI technologies used in policing.

  Diverse representation ensures that AI systems are better equipped to understand and address the needs of different demographic groups. The findings from Paper 2 underscore the importance of inclusive approaches to AI development, where underrepresented voices are considered in the creation and implementation of AI tools (Ezzeddine et al., 2023a). Ensuring diversity in AI development teams could lead to more equitable outcomes and reduce the risk of perpetuating systemic biases in policing.

5. Continuous Monitoring and Ethical Auditing

The findings across the studies also highlight the necessity of continuous monitoring and ethical auditing of AI systems used in law enforcement. Participants expressed concerns about the long-term impacts of AI surveillance on civil liberties and societal norms (Ezzeddine et al., 2023b). Therefore, policymakers must establish mechanisms for:

- **Ongoing Ethical Audits**: AI systems should undergo regular ethical audits to ensure they comply with legal and ethical standards. These audits should assess not only the technical performance of AI systems but also their broader societal impacts. As the findings from Paper 4 suggest, continuous monitoring of AI surveillance in public spaces is necessary to prevent mission creep and ensure that these technologies do not infringe on citizens' rights (Ezzeddine et al., 2023b).
- **Longitudinal Impact Assessments**: Given the evolving nature of AI technologies, it is crucial to assess their long-term implications for society. Policymakers should mandate the use of longitudinal studies to track how AI surveillance affects public trust, social behaviour, and civil liberties over time. This would provide valuable data for adjusting governance frameworks as AI technologies develop (Završnik, 2020).

Recommendations for AI Developers in Policing Domain

AI developers working within the policing sector must carefully consider the unique challenges of deploying AI systems in law enforcement, particularly in light of public concerns regarding transparency, bias, privacy, and fairness. The recommendations below draw on the findings from the thesis, emphasizing the need for AI developers to create systems that align with both ethical principles and public trust.

1. Prioritize Transparency in AI Design and Functionality

The studies demonstrated that transparency in AI surveillance is critical for fostering trust among citizens. Participants consistently highlighted concerns over the lack of transparency regarding how AI systems make decisions and the extent of data being collected. This sense of uncertainty contributed to widespread discomfort, especially regarding the invisibility of AI processes in law enforcement. As noted in Study 1, participants expressed a strong preference for clearer explanations of how AI systems function and which data sets are being utilized.

- **Developing Explainable AI (XAI):** To address these transparency concerns, AI developers should prioritize the creation of Explainable AI (XAI) systems that clearly articulate how decisions are made (Barredo Arrieta et al., 2020). This is especially relevant in policing, where AI algorithms may influence significant decisions regarding resource allocation or suspect identification. Explainable AI tools could help law enforcement officers, and the public understand how conclusions are reached, ultimately promoting a greater sense of accountability.
- **User-Friendly Interfaces for Transparency:** Beyond developing XAI systems, user-friendly interfaces should be provided to law enforcement and relevant stakeholders. These interfaces would ensure that the decision-making processes of AI systems are visible and accessible, allowing officers and policymakers to monitor and assess how data is being interpreted in real-time. Findings from Study 2 emphasized that such transparency would improve public trust and offer necessary checks on the influence of these systems in law enforcement (Gunning & Aha, 2019).

2. Address Algorithmic Bias Through Context-Sensitive AI Development

Algorithmic bias emerged as a significant issue across the studies. Many participants raised concerns about the potential for AI technologies to replicate and amplify existing societal biases, particularly those related to race, gender, or socioeconomic status. Findings from Study 3 highlighted the importance of context in AI development. Systems designed without taking into account the specific characteristics of local populations were perceived as more likely to perpetuate these biases.

Similarly, studies from 2023 (e.g., Google AI and OpenAI research) emphasize the need for domain-specific calibration to maintain fairness without sacrificing overall model accuracy. Additionally, incorporating fairness-aware machine learning models, as advocated in Mehrabi et al. (2023), has shown promise in mitigating biases in predictive policing. Context-sensitive approaches—where training datasets are localized to reflect demographic diversity—continue to gain traction, with evidence pointing to improved equity in AI decisions (Gunning & Aha, 2023; IBM Research, 2023).

- **Contextualize AI Models Based on Local Data:** AI systems used in policing should incorporate localized data that reflect the specific demographics and social conditions of the areas where they are deployed. This would mitigate the risks associated with using generalized data that may not account for local realities, which could otherwise lead to discriminatory outcomes. The research findings suggest that contextually sensitive AI models would not only improve the accuracy of predictive policing but also address concerns regarding fairness and equity (Benjamin, 2019).
- **Regular Audits for Algorithmic Fairness:** To address the risks of bias, AI developers should implement continuous fairness audits of their systems. Such audits would assess the impact of AI tools on different demographic groups and help identify any systemic biases embedded in the algorithms. This approach would ensure that AI systems remain fair and do not disproportionately affect marginalized communities. Findings from Study 4 indicated that participants expected AI systems to be regularly assessed for their fairness and social impact, suggesting that fairness audits could serve as a mechanism to address these concerns (Mehrabi et al., 2021).

3. Design Privacy-Enhancing Features into AI Systems

The protection of personal data emerged as a central theme in the studies, particularly in relation to how AI surveillance tools collect and process sensitive information. Participants expressed strong concerns about the extent of data collection, fearing that AI systems would infringe upon their privacy and autonomy. As shown in Study 3, there was a clear demand for stronger privacy safeguards and more control over personal data.

- **Privacy-by-Design Principles:** AI developers should embed privacy-by-design principles into the architecture of AI systems, ensuring that data collection is minimized and that individuals have greater control over their personal information. These privacy-enhancing features could include encryption, anonymization, and differential privacy techniques to protect against data misuse (Dwork & Roth, 2014). Findings suggest that such approaches would align with citizens' expectations for safeguarding their privacy in the face of increased surveillance technologies.

- **Decentralized Data Storage Solutions:** A key finding from Study 3 was the public's preference for decentralized models of AI governance, where citizens maintain some degree of control over their data. To meet this demand, developers should explore decentralized data storage solutions, which would limit access to personal data and allow citizens to retain ownership over their information. This would provide an additional layer of privacy protection and help address widespread concerns about data centralization (Zyskind et al., 2015).

4. Develop AI Systems That Foster Community Trust

The studies consistently underscored the importance of community trust in AI surveillance technologies. Public perceptions of AI surveillance were influenced not only by concerns about privacy and bias but also by how these systems were introduced into local communities. Building community trust is, therefore, essential for the successful deployment of AI technologies in policing.

- **Engage Communities Early in the Development Process:** AI developers should involve communities from the earliest stages of system design and deployment. By engaging local stakeholders in consultations and discussions, developers can ensure that their systems align with the specific needs and values of the communities in which they will be used. Findings from Study 1 indicated that participants desired more meaningful engagement with law enforcement agencies and AI developers, stressing that public involvement would improve transparency and foster trust (Lyon, 2014).

- **Localized AI Solutions:** The development of AI systems should be flexible enough to adapt to different local environments. For instance, participants in Study 2 suggested that AI surveillance systems designed for urban environments may not be suitable for rural areas, where concerns and behaviours around surveillance differ. Developers should therefore prioritize the design of localized AI systems that can be tailored to the specific context in which they are deployed (Lockey, 2020).

Recommendations for Surveillance Scholars

Surveillance scholars are often concerned with the societal implications of surveillance technologies and their influence on power dynamics, control, and privacy. The findings from this research highlight the need for more nuanced studies into how AI surveillance is perceived by different populations and the socio-political environments that shape

these perceptions. Surveillance scholars can further explore how AI surveillance, particularly in law enforcement, changes citizens' behaviours, and whether these technologies truly deliver on their promise of increased security. The findings also call for deeper investigations into how citizens understand and resist AI, especially in contexts where AI surveillance becomes ubiquitous, and the blurred lines between online and offline surveillance create new terrains for resistance.

Additionally, future research directions could focus on understanding the intersection of AI surveillance and digital platforms, where the constant surveillance of user behaviour shapes societal norms. Scholars can explore how these digital surveillance mechanisms reinforce power structures in ways that traditional, physical surveillance cannot, requiring updated theoretical frameworks to address the unique societal impacts of AI. For example, how do AI systems perpetuate existing societal biases, and what mechanisms of oversight can mitigate this risk? Studies should also focus on the long-term implications of such technologies on democratic freedoms, privacy expectations, and civil liberties (Lyon, 2014).

Recommendations for Psychologists/Behavioural Scientists

For psychologists studying resistance and compliance with surveillance, the findings provide rich insights into the psychological mechanisms that drive public resistance to AI technologies. Resistance is often not just a reaction to surveillance itself but also to the loss of autonomy and the perceived threat to personal freedoms.

1. **Understanding the Psychology of Resistance**: Psychologists can explore how citizens' awareness of AI surveillance influences their behaviours, both online and offline. The findings suggest that individuals alter their behaviours in the presence of AI surveillance—such as avoiding certain areas or adopting privacyenhancing technologies like VPNs. Understanding these behavioural shifts can inform the development of strategies to reduce resistance and improve compliance (Foucault, 1977).

2. **Addressing the Fear of Surveillance**: Many individuals resist AI surveillance not because they have something to hide, but because of the psychological discomfort associated with constant monitoring. Researchers can examine how to mitigate this discomfort through community engagement, education, and the development of more transparent surveillance systems that do not infringe on personal autonomy.

Recommendations for Criminologists and Sociologists

For criminologists and sociologists, the findings point to the need for a broader understanding of how AI surveillance impacts social dynamics and crime prevention.

1. **Examining the Impact of AI Surveillance on Social Behaviour**: AI surveillance alters social interactions, particularly in public spaces. Criminologists can examine how the presence of AI technologies—whether through facial recognition or predictive policing—affects community relations, especially in marginalized areas where trust in law enforcement is already low (Benjamin, 2019). Future studies should explore how AI surveillance either reinforces or disrupts the social fabric in communities subject to heavy policing.

2. **Resisting and Responding to Surveillance**: Criminologists can also explore the role of counterstrategies in resisting surveillance. The findings indicate that resistance is a collective response that may involve grassroots activism, legal challenges, and the use of privacy-enhancing technologies. Understanding how these counterstrategies evolve can help criminologists theorize new forms of resistance and propose ways law enforcement can anticipate and address these movements.

Recommendations for Public Services Considering AI Deployment

Public services, including healthcare, education, and transportation, are increasingly adopting AI technologies. The findings offer several lessons for these sectors, particularly in addressing public resistance to AI deployment.

1. **Building Trust in Public AI Systems**: Public services must prioritize building trust with citizens by being transparent about how AI systems operate and how personal data is collected and used. Public services should follow the guidelines of fairness and transparency, ensuring that AI systems are designed with ethical principles that reflect the public's expectations for privacy and autonomy (Lyon, 2014).

2. **Incorporating Feedback Mechanisms**: Public services can reduce resistance by allowing citizens to provide feedback on AI systems and incorporating these insights into system design and policy development. This participatory approach can foster greater public acceptance and ensure that AI technologies meet the needs of the communities they serve.

On a final note, the theoretical and practical recommendations provided in this discussion offer valuable insights to a range of different practitioners from law enforcement agencies to policymakers and technology developers, not excluding behavioural and social scientists, as well as public services researching and considering the use of AI.

These recommendations emphasize the need for transparency, public engagement, and robust safeguards to ensure the ethical and responsible use of AI in public safety. By addressing these concerns, stakeholders can help foster trust in AI technologies and create a future where AI surveillance, especially in law enforcement, serves both the interests of security and the protection of individual rights. Additionally, it is essential to treat resistance as a normal and expected part of AI implementation in security contexts. Instead of viewing resistance as an inherent obstacle, public services should seek to understand the nuanced and concrete reasons behind public opposition. As demonstrated in the findings, resistance is not always monolithic but shaped by specific concerns, such as privacy and data ownership, which vary across contexts and demographic groups. A more differentiated approach is needed, one that examines how resistance is expressed and what drives it, enabling policymakers and AI developers to address these concerns more effectively rather than overstating the likelihood of opposition.

Limitations and Future Work

Research Limitations

While the research presented in this thesis provides significant insights into the integration of AI in law enforcement and its impact on public perception, it is essential to acknowledge the limitations inherent in both the scope of the studies and the methodological choices made. Recognizing these limitations not only adds transparency to the research, but also paves the way for more refined investigations in the future.

One of the primary limitations is the reliance on mixed methods across different studies, with a heavy emphasis on qualitative approaches in several papers. Although semi-structured interviews, Q-sorts, thinkaloud protocols, and privacy walks allowed for in-depth exploration of participants' attitudes and perceptions, these methods are inherently limited in terms of generalizability. Qualitative research provides rich, contextual data but is often difficult to extrapolate to broader populations. The interview sample in Paper 3, which involved 111 participants across eight European countries, offers a snapshot of specific citizen groups, but the findings cannot be universally applied to

all demographic groups or geographical regions. Despite the diverse sampling, the cultural, sociopolitical, and economic contexts of these participants might have influenced their responses in ways that are not fully accounted for in the analysis.

Moreover, the specific focus on European countries limits the global applicability of the findings. AI use in law enforcement is an international issue, with differing legal frameworks, societal norms, and governmental practices influencing both the implementation of AI technologies and public responses. For instance, AI surveillance practices in authoritarian regimes versus democratic nations may elicit vastly different public reactions. Future research should aim to include a more geographically and culturally diverse sample to capture these nuances and provide a more holistic understanding of global citizen perspectives on AI surveillance.

Another limitation arises from the methodological design employed in Paper 5, which integrated privacy walks, pre- and post-task surveys, and geo-mapping. While this innovative approach allowed for real-time observation of participants' interactions with AI surveillance in public spaces, it also introduced constraints related to the ecological validity of the findings. Privacy walks are context-specific; they are dependent on the surveillance infrastructures present in the chosen environments and on participants' prior experiences with these technologies. The findings from these privacy walks may not be entirely applicable to other settings, especially those with varying levels of surveillance or different forms of AI technologies in use. Additionally, participant behaviour during privacy walks could have been influenced by the prompted knowledge that they were being observed by potentially AI-embedded surveillance tools, a phenomenon known as the Hawthorne effect (Adair, 1984), which may have led participants to alter their actions in ways that do not reflect their natural responses to surveillance.

Another notable limitation of the same study (Paper 5) stems from the study's geographic focus on Sheffield city centre, an urban environment characterized by a dense concentration of surveillance infrastructure, including CCTV cameras and AI-enhanced surveillance tools. The unique nature of this area, with its busy streets and high levels of public activity, may not reflect the experiences of individuals in rural or less densely populated areas, where surveillance infrastructure is less prevalent. This urban setting likely influenced participants' behaviour and perceptions of surveillance, particularly in relation to their awareness and interaction with AI surveillance technologies. Future research could benefit from expanding the scope of privacy walks and surveys to rural areas, allowing for a more comprehensive

understanding of how AI surveillance impacts individuals in different environments. Investigating these differences would offer important insights into the wider impact of AI surveillance across various geographic contexts.

In terms of the data analysis methods, while thematic and content analysis provided a robust framework for identifying patterns and themes across qualitative data, the reliance on these methods introduces subjectivity in interpreting participants' responses. And even though NVivo software was employed to enhance the rigor of the analysis, coding qualitative data inevitably involves researcher interpretation, which can introduce biases. The challenge lies in ensuring that these interpretations accurately reflect participants' experiences and are not unduly shaped by the researchers' preconceived notions or theoretical leanings.

Additionally, the think-aloud method used in Paper 4, though valuable in capturing real-time cognitive processes, is not without its limitations. Participants' verbalizations during the think-aloud sessions may not fully represent their thoughts or decision-making processes, particularly if they found it difficult to articulate their reasoning. Some participants may have been more comfortable verbalizing their thoughts than others, leading to variability in the quality and depth of the data collected. Moreover, the think-aloud method may have been influenced by the artificial nature of the experimental task, where participants were aware they were being asked to reflect on AI technologies, potentially leading them to focus more critically on these technologies than they might in everyday contexts.

Lastly, the study's timeframe presents another constraint. Given the rapidly evolving nature of AI technologies and their increasing integration into law enforcement, the findings presented in this thesis reflect a particular moment in time. Technological advancements, changes in legal frameworks, and shifts in public opinion could alter the landscape of AI surveillance in law enforcement in ways that are not fully captured in this research. Additionally, changes to policing practices and priorities, driven by societal shifts, political pressures, or emerging threats, may further influence the deployment and reception of AI technologies. As AI continues to develop, new ethical concerns, public perceptions, and regulatory responses may emerge, alongside adjustments in how law enforcement agencies operate. This highlights the need for ongoing research to stay abreast of these changes and their impact on both AI implementation and public trust.

Future Research Suggestions

Building on the findings and limitations of this research, several key areas for future investigation emerge. These suggestions aim to address the gaps identified in this thesis and to further contribute to the evolving discourse on AI, surveillance, and law enforcement.

First and foremost, future research should prioritize expanding the geographical and cultural scope of the studies. While the current research focused primarily on policing in a European context, particularly in the UK, it is crucial to explore how these findings transfer to or differ from other regions. The public perceptions of AI surveillance observed in this study, including concerns around privacy, transparency, and control, may manifest differently in other legal, cultural, and societal frameworks. In regions with distinct political systems, such as North America, Asia, Africa, and the Middle East, the deployment of AI in law enforcement may be influenced by different governance structures and societal norms, which could lead to varying levels of public trust, acceptance, and resistance (Crawford & Paglen, 2021). Comparative studies across these diverse jurisdictions would provide a more comprehensive understanding of how these differences shape public reactions to AI surveillance, helping to contextualize and extend the findings of this thesis.

Second, future research should explore the long-term effects of AI surveillance on public behaviour and societal norms. While this thesis focused on immediate reactions to AI technologies, particularly through privacy walks and think-aloud methods, it would be valuable to investigate how sustained exposure to AI surveillance impacts behaviour over time. Longitudinal studies could provide insights into whether individuals become desensitized to surveillance, modify their behaviours in more permanent ways, or develop new strategies for resisting surveillance. Additionally, these studies could examine the broader societal implications of AI surveillance, such as its effects on social cohesion, democratic engagement, and civil liberties.

A third avenue for future research lies in investigating the effectiveness of various safeguards and regulatory frameworks for AI surveillance. While this research highlighted public demand for transparency, accountability, and independent oversight, there remains a need to empirically test the effectiveness of these measures in practice. Future studies could assess how different regulatory models—such as the use of independent oversight bodies or algorithmic audits—affect public trust in AI surveillance and whether these measures successfully mitigate concerns about bias, discrimination, and privacy violations. Experimental designs could also be used to test the impact of specific

transparency mechanisms, such as explainable AI, on public perceptions of fairness and accountability in law enforcement.

Moreover, further research should delve deeper into the role of AI surveillance in shaping social inequalities. The findings from Paper 4 revealed concerns about algorithmic bias and the potential for AI technologies to disproportionately target marginalized groups. Future studies could investigate the extent to which AI surveillance exacerbates existing social inequalities, particularly in relation to race, gender, and socioeconomic status. These studies could adopt an intersectional approach, examining how different forms of inequality intersect to influence public experiences and perceptions of AI surveillance. This line of inquiry is critical for developing equitable AI systems that do not perpetuate systemic injustices.

In addition to the aforementioned areas, future research should also focus on the perspectives and actions of law enforcement agencies regarding AI surveillance. Understanding how police departments perceive, implement, and interact with AI technologies is critical for a holistic view of their societal impacts. Investigating the motivations behind AI design within policing—such as efficiency, crime prevention, and public safety—can reveal ethical dilemmas that arise in the deployment of these technologies. Future research should examine how police officers and decision-makers balance the need for surveillance with the ethical imperatives of privacy, civil liberties, and community trust.

Furthermore, exploring diverse ways of understanding individuals as members of communities—rather than merely data points or suspects—can foster a more humane approach to AI in law enforcement. By considering community relationships, social context, and public sentiment, researchers can identify best practices for integrating AI technologies that respect individual rights and enhance public safety. Research of this approach could lead to frameworks that encourage ethical AI use while promoting a sense of accountability and partnership between law enforcement and the communities they serve.

Finally, there is a need for more research on citizen engagement and participatory governance in AI surveillance. While the current research emphasized the importance of involving citizens in decision-making processes, future studies could explore different models of participatory governance and their effectiveness in fostering public trust and accountability. For instance, researchers could investigate how citizen advisory boards, public consultations, or crowdsourced countersurveillance initiatives impact the development and deployment

of AI surveillance technologies. These studies could provide valuable insights into how democratic processes can be integrated into the governance of AI, ensuring that these technologies serve the public good and are aligned with societal values.

In summary, future research should aim to expand the geographical and cultural scope of AI surveillance studies, investigate the long-term behavioural and societal impacts of surveillance, assess the effectiveness of regulatory frameworks, examine the role of AI in perpetuating social inequalities, and explore innovative models of citizen engagement in AI governance. By addressing these areas, future research can contribute to a more comprehensive understanding of AI surveillance and its implications for law enforcement and society.

Conclusion

Summary of Key Findings

This thesis has explored the complex interplay between AI-driven surveillance, law enforcement, and public perceptions, providing valuable insights into how citizens navigate the ethical and social implications of AI technologies. Through a series of mixed-method studies, the research has revealed both the potential benefits and significant challenges posed by AI in law enforcement, with a particular focus on privacy, trust, and resistance.

The findings indicate that while there is a general acceptance of AI technologies for law enforcement purposes, this acceptance is conditional on several factors, including transparency, accountability, and the perceived fairness of AI systems. In Papers 1 and 2, participants expressed concerns about the potential for AI surveillance to infringe on their privacy rights, particularly in the absence of clear safeguards and regulatory oversight. The study highlighted the importance of establishing robust legal frameworks and independent oversight bodies to ensure that AI technologies are used ethically and do not disproportionately impact certain groups or infringe on civil liberties.

In Paper 3, the semi-structured interviews across eight countries provided deep insights into citizens' preferences for AI ownership and governance. The findings revealed a strong demand for participatory governance models, where citizens have a say in how AI surveillance technologies are developed and deployed. This reflects broader concerns about the concentration of power in AI technologies and the need for greater public accountability.

Papers 4 and 5 further explored citizens' real-time reactions to AI surveillance through think-aloud methods, privacy walks, and

geomapping. These studies provided critical insights into how citizens navigate both online and offline surveillance environments, with participants expressing heightened concerns about facial recognition technologies and the potential for AI to be used in discriminatory ways. The use of privacy walks allowed for the capture of spontaneous, contextdriven reactions to surveillance, highlighting how the physical environment shapes perceptions and behaviours.

Overall, the research findings underscore the need for a balanced approach to AI surveillance, one that considers both the security benefits of these technologies and the ethical concerns they raise. The studies contribute to the growing body of literature on AI ethics and governance, offering practical recommendations for law enforcement agencies, policymakers, and technology developers on how to build public trust and ensure the responsible use of AI in public safety.

Final Thoughts

On a final note, and in reiteration, this research contributes to the ongoing discourse on the ethical and societal implications of AI in law enforcement by providing a nuanced understanding of citizen perspectives, resistance strategies, and preferences for governance. The findings emphasize the importance of transparency, accountability, and participatory governance in ensuring that AI surveillance technologies are deployed in ways that respect individual rights and promote public trust. By addressing the concerns raised by citizens and implementing the practical recommendations outlined in this thesis, stakeholders can help create a future where AI serves as a tool for enhancing public safety without compromising the ethical principles that underpin democratic societies.

The implications of this research transcend the immediate scope of law enforcement, contributing to broader discourses on the role of AI within society and the intricate balance between security imperatives and the protection of civil liberties. As AI technologies continue to evolve, it is essential that future research and policy development remain grounded in the values of transparency, fairness, and human dignity. Only by doing so can we ensure that AI is used in ways that benefit society as a whole.

Appendices

I. Publications

Paper 1

Citizen Perspectives on Necessary Safeguards to the Use of AI by Law Enforcement Agencies

Yasmine Ezzeddine[1]
CENTRIC
Sheffield Hallam University
Sheffield, UK
y.ezzedine@shu.ac.uk

Petra Saskia Bayerl
CENTRIC
Sheffield Hallam University
Sheffield, UK
p.s.bayerl@shu.ac.uk

Helen Gibson
CENTRIC
Sheffield Hallam University
Sheffield, UK
h.gibson@shu.ac.uk

ABSTRACT. In the light of modern technological advances, Artificial Intelligence (AI) is relied upon to enhance performance, increase efficiency, and maximize gains. For Law Enforcement Agencies (LEAs), it can prove valuable in optimizing evidence analysis and establishing proactive prevention measures. Nevertheless, citizens raise legitimate concerns around privacy invasions, biases, inequalities, and inaccurate decisions. This study explores the views of 111 citizens towards AI use by police through interviews, and integrates societal concerns along with propositions of safeguards from negative effects of AI use by LEAs in the context of cybercrime and terrorism.

Keywords: *Artificial Intelligence; Law Enforcement Agencies; Safeguards; Citizen Perspectives; Police* ***Abbreviations:*** *LEAs: Law Enforcement Agencies; AI: Artificial Intelligence.* ***Type of Submission:*** *Regular Research Paper*

Introduction

The role of Artificial Intelligence extends beyond improving the security and safety of citizens, particularly against cybercrime and terrorism, to anticipate and recognize criminals' increasing employment of AI tools (Trend Micro Research, 2020). In fact, societies are embracing new forms of reality amplified by machine learning and use of AI (Mann, 2017), where every detail of daily routines is captured, stored, and digitalized. And once this information is distributed in the system, recalling it is nearly impossible (Petersen & Taylor, 2012). Hence, advancing the measures taken for public protection is imperative for enhancing general safety and security (Macnish, 2021). Doubtless, algorithms and data analytics are playing an increasing role in all aspects of society including the policing and security services (Babuta & Oswald, 2020) to the extent that policing through social media has been explored by several Law Enforcement Agencies (LEAs) worldwide. However, the expansion of data collection efforts and AI use continues to trigger uncertainty around its ethical and moral implications (Lyon, 2002),

[1] **Corresponding Author:** Yasmine Ezzeddine – S1 1WB – y.ezzeddine@shu.ac.uk

especially with respect to recent technologies involving the web, video monitoring and algorithmic decision-making warrants the need for critical evaluation of the inevitable

psychological consequences (Stoycheff, et al. 2020) contributing to the skepticism around AI use by police.

Additionally, the conflict mounts between the facilitated admittance that these technologies offer, and the "fear of contact" emanating from alliances with independent bodies of the private sector (Trottier, 2017, p. 475), coupled with the lack of evidence around efficiency of algorithmic-based decisions, accuracy, fairness, and risks of predictive policing leading to discrimination and inequality (Bushway, 2020; Quattrocolo, 2020; Završnik, 2020).

Similarly, amongst the numerous challenges facing AI implementations for LEAs as well the private sector is to determine how to capitalize on AI capabilities in response to changing safety and security challenges while ensuring responsible use. In fact, the AP4AI[2] Framework incorporating 12 Accountability Principles of AI laid the foundation for a "healthy balance between the need to innovate practices and enhance capabilities (…) on one hand, and the legitimate expectations by citizens that police work is conducted lawfully proportionality and in pursuit of a legitimate aim" (Akhgar et al., 2022, p.5).
Nevertheless, the scarce and limited systematic knowledge around citizen perceptions of safeguards inspired us to complement the existing insights around resistance to LEAs' data collection and use of AI, while satisfying the theoretical gaps around different types of safeguards are often considered in ethical and legal perspectives but not from the societal perspective of citizens.

Hence, the central theme of this research comprises an investigation of citizen propositions of necessary safeguards that can protect them from the potential negative effects incurred in AI use by police. In other words, this study focuses on engaging with citizens as not only beneficiaries of such innovations, but also key players in legitimizing the deployment of AI tools, since in the absence of citizens' approval and support, AI implementation can face the negative implications of chilling effects (Stoycheff, 2016), fear of contact (Marx, 2009; Marthew & Tucker, 2017), countertractions against police (Bayerl et al., 2021) and even national and international movements opposing its deployment (Montag, et al. 2021; Reclaimyour-Face[3] campaign). This warrants the investigation into citizens' perspectives to AI implementation as well as their propositions of safeguards to the potential negative consequences of incorporating AI technologies into LEA security practices.

Methodology

This study adopts a qualitative approach aiming to better understand and integrate citizens' perspectives about data collection and AI use by LEAs. Therefore, semi-structured interviews were conducted to provide in-depth insights and elucidations into necessary safeguards, allowing citizens to elaborate on their subjective views and experiences.

[2] **AP4AI:** Accountability Principles for Artificial Intelligence: www.ap4ai.eu
[3] **Reclaim Your Face:** Ban Biometric Mass Surveillance! (n.d.). Reclaim Your Face. https://reclaimyourface.eu/

Sample

As part of EU funded project AIDA we have conducted in eight countries. No specific criteria in terms of demographics were sought, except for at least 16 participants from the 'general population' of each participating country (above 18 years). The rationale behind this open choice was driven by two considerations: pragmatism, facilitating access to citizens, and guided interest, allowing partners to choose groups that are of interest to them. In total, 111 interviews were conducted with 44 males and 69 females were interviewed, the youngest participant being 19 years old and the oldest 83 years old.

The sample was varied in terms of experience with cybercrime and terrorism. About 58.2% indicated that they had no personal experience with either, while 34.5% reported experience with cybercrime or other incidents online (e.g., phishing, identity fraud, hacked email account). Only 4.5% had experience with terrorism (e.g., car attack in hometown). This reflects the considerable heterogeneity of experience and safety perceptions and therefore does not seem to be biased in a specific way towards citizens with high/low experience or specific attitudes towards safety.

Data Collection

Participants were recruited through researchers handling interviews in each of the countries. Semi-structured scenario-based interviews were conducted, in the respective country language, along pre-defined themes categorized in three main topics: "understanding of AI and acceptance conditions", "perception of safety with respect to terrorism/cybercrime and societal resilience", and "citizen reactions". This paper reports the findings from topic 3 related to safeguards, obtained particularly from responses to the question of *"What should police forces do to safeguard you from negative effects of AI systems?"* Interviews were either conducted online or face-to-face and audio recorded. For all interviews, Subsequently, signed consent forms, interview recordings, and English summaries or verbatim transcript in the country's language were provided. The latter were translated to English using an online software followed by proof-reading.

Data Analysis

Our analytic approach followed thematic and content analytic principles (Auerbach & Silverstein, 2003; Krippendorff, 2004) to identify the main topics and themes in the collected data starting with open or initial coding (Charmaz, 2006). Initial codes were then clustered into high-order categories per main topic. The categories identified for safeguarding measures are presented in the findings.

Ethics

Several steps were taken to ensure data collection adhered to relevant ethics requirements.

Firstly, the study received approval by the ethics committee of Sheffield Hallam University. Secondly, the interviews started by presenting participants with an information sheet to clarify the context of the project, details of data handling, participants' rights, and legal basis for the study. Thirdly, participants were only asked for basic personal information (gender and age). They were further reminded of their right to not answer demographic questions if they

did not feel comfortable to do so, which several participants rightfully used. All data was anonymized before analysis and interpretation.

Findings

Overall, participants produced 113 recommendations for safeguards they deem necessary to protect from potential negative effects of AI. In-depth evaluations of these revealed complex attitudes that can be clustered into the following areas (see Fig.1):

1) Educational safeguards: for citizens and LEAs
2) Technical safeguards: Regular Evaluations, anonymization of collected data
3) Legal safeguards: National and international regulations and independent agencies
4) Human safeguards: selectivity in employing AI-handling staff and importance of human validation of AI findings and decisions
5) Privacy safeguards: Limited data collection and requests for consent
6) Stop use of AI: use traditional means
7) Inevitable vs. no negative effects: either not foreseeing any negative effects or assuming no safeguards can protect from them

Figure 1: Different types of safeguards suggested by interviewees (percentage of statements)

Educational safeguards and transparency

To safeguard citizens from potential negative effects of AI, participants emphasized the importance of **education of LEA staff as well as citizens**. LEA personnel should receive enhanced and ethical training on how to handle AI tools. In extension, some participants proposed that tools should only be used by experienced LEA members while stressing the need for collaborations with outside experts, e.g., "*law enforcement are not scientists, and so*

they might make a mistake and that's not good, so I wish they have like a team that they've got people who are helping them" (UK-02). Additionally, participants emphasized civic education, particularly educating children in schools around the process of data collection, the actual purposes of AI use, and the percentage of decision-making to which AI contributes. This links to another important aspect participants referred to repeatedly, namely **transparency**. This should include the sharing of positive outcomes of AI use and developing campaigns that showcase positive scenarios of AI use in criminal investigations or using an open-source platform to increase community trust.

Technical and AI specific safeguards

Technical and AI-specific safeguards were a second recurring theme., Participants referenced **regular assessments and evaluation** of AI tools (including impacts on crime rates), and the importance of controlling biases and detecting errors in algorithms to prevent their reoccurrence.

Additionally, participants raised the importance of adopting **technical safeguards** to ensure that the information is properly "anonymized", and not being used for other purposes or by other companies. Moreover, interviewees noted that criminals and terrorists also use AI tools which is why it is important that LEAs "*always trying to keep ahead*" (UK-11) of criminals by ensuring their AI tools are as up to date as possible.

Legal Safeguards: Frameworks and Policies

Participants further asked for legal frameworks, **national and international regulations** and policies that ensure "*well-defined and well-enforced limits for what LEAs are allowed to do with the data*" (NL-06). This may require preventive and punitive regulations to the misuse of AI, whether by LEAs or the private sector. Participants also stressed the need for communication between different law enforcement agencies and cooperation between different countries. In other words, AI should be monitored and supervised, preferably by an **independent agency** or unbiased third-party or government that regulates AI, enforces legislation, monitors data collected and ensures data is stored safely and within legal timeframes.

Human Safeguards: Avoiding errors and Biases

Participants further suggested **selectivity** and proper **vetting** of staff involved in roles touching on AI data collection/handling while ensuring ethnic diversity and gender equality to minimize biases. Additionally, the importance of the **human component in decision making** was perceived as crucial in monitoring and verifying AI decisions. Participants repeatedly stressed the importance of final decisions and interpretations to be done by humans since machines "cannot replicate humans". Another singular suggestion to reduce the possibility of biases was: *"use an AI system that is developed without unfair bias in accordance with the applicable laws after much research on its development and with the involvement of citizens during the research"* (GR-14).

Privacy Safeguards: Regulated Data Collection

With respects to automatic identification and random monitoring of members of society, participants expected LEAs to not invade citizens' privacy and freedoms for the sake of security. Instead, there should be **limitations on data collection**, such as "*whether there are sensitive areas of society where perhaps it shouldn't be used, particularly around our homes*"

(NL-10). Participants strongly believed that LEAs should not waste time and resources collecting information from 'everyone'; instead, they should use alternative means to obtain data. Hence, they stressed that LEAs should not collect more information than they actually need, calling for regulation of data collection and judicial control of obtained information (see

4.3). Alternatively, LEAs should ask for individuals' **consent** prior to data collection, or hide/blur faces of people not suspected in a certain footage, as well as dispose of non-relevant information as soon as possible.

Stop use of AI

The most extreme position was expressed by a small group of statements that expressed total opposition to the implementation of AI. To them, the best safeguard was "*maybe not using Artificial Intelligence*" (PT-06). Instead, for this group LEAs are expected to work harder in the traditional way.

Inevitable vs. no negative effects to require safeguards

On the contrary, a small number of statements suggest that some participants did not perceive any negative effects to be safeguarded against. Few even encouraged LEAs to expand AI use. Other viewpoints indicated that there will always be negative effects which cannot be eliminated, either because LEAs lack the expertise on how to safeguard against negative consequences effectively, or due to the impossibility to control social media platforms collecting data. Hence, safeguards would either not be needed or not possible.

Discussion

With most eyes set on AI implementation in almost all aspect of modern life, calls for frameworks, regulations and safeguards are equally arising. The UN Rights Chief stressed the importance for safeguards in face of the "undeniable and steadily growing impacts of AI technologies" and the need to "protect and reinforce all human rights in the development, use and governance of AI as a central objective" (Geris & Wellington, 2021, n.p.). This coincides with the European Parliament's Press release that stressed the importance of subjecting AI use in policing to "strong safeguards and human oversight" (European Parliament, 2021). All of this resonates with participants' statements around educational, technical, legal, and human safeguards. It appears that citizens are not entirely against AI use, in fact, they are only discreet around its implementation, and have heightened, yet justifiable, concerns around the impact of AI on their privacy and overall quality of life.

One of the unique aspects of this study is the intersection it provides between citizens' attitudes towards AI, safety perceptions and propositions of safeguards, particularly as suggested by non-AI experts. This reveals that common citizens possess a basis of knowledge and understanding of AI and the inevitable consequences incorporated in its design. However, that did not create rejection of the entire tool. In fact, citizens were supportive of AI implementation as a tool that can safeguard societies, especially in the current modern era, if it adheres to strict rules and is monitored by trained and trustworthy individuals. This was reflected in the overall number of propositions on the need for educational and legal safeguards in the implementation of AI tools by LEAs, compared to the surprisingly lower rates on privacy safeguards and requests to stop using AI altogether.

Another beneficial aspect of these propositions lies in their potential to adapt existing AI tools and shape prospective designs to account for citizen perspectives, which may in turn reduce resistance and counterstrategies to data collection and AI implementation and hence,

enhance the feasibility of information gathering while safeguarding the quality of collected data.

Other ramifications of this also involve financial implications by safeguarding data collection as a massive source of income (Deulkar & Gupta, 2018), all of which can benefit from potential application of safeguard propositions put forth by citizens participating in this study.

Limitations and Future Work

The heterogeneity of participants across countries allowed us to obtain a highly diverse set of experiences and perspectives. Yet, the small number of interviews per country inhibits an analysis of subgroups. However, this qualitative approach can outline the personal, individual perspectives of citizens across contexts which in turn, reflects the richness of citizen views, while displaying similarities in opinions and expectations alongside personal motivations and reasoning. Future research can build on these findings to include the demographic groups most and least critical towards LEAs' use of AI in any country.

Conclusion

This study provides an in-depth evaluation of citizens propositions of safeguards to LEAs' use of AI. In brief, AI use should be justified, legitimate and only used for declared purposes. Other safeguards included the avoidance of biases through appropriate AI design, supervision and legal framework, regular evaluations, transparency, along with education, training, and selectivity in assigning LEA staff handling AI tools. In addition to propositions of civic education, national and international collaborations and ensuring that AI capabilities by LEAs are up-todate and at an arm's race with those of criminals. Interestingly, apart from some concerns about facial recognition, findings reveal concerns around how AI is being deployed, rather than the mere deployment of AI tools by LEAs. With this the study provides vital insights into the varied nature of measures that citizens deem necessary as safeguards to ensure their acceptance of AI.

Acknowledgment

This aspired work is based on the EU funded research project AIDA The AIDA project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 883596 (AIDA- Artificial Intelligence and advanced Data Analytics for Law Enforcement Agencies).

References

Akhgar, B., Bayerl, P.S., Bailey, K., Dennis, R., Gibson, H., Heyes, S., Lyle, A., Raven, A., & Sampson, F. (2022). *AP4AI Framework Blueprint*. AP4AI Project Report. Available online: https://www.ap4ai.eu/node/14

Auerbach, C., & Silverstein, L. (2003). *Qualitative Data: An Introduction to Coding and Analysis.* New York: New York University Press.

Babuta, A., & Oswald, M. (2020). Data Analytics and Algorithms in Policing in England and Wales. RUSI.

Bayerl, P. S., Akhgar, B., La Mattina, E., Pirillo, B., Cotoi, I., Ariu, D., . . . Karagiorgou, K. (2021). Artificial intelligence in law enforcement: Cross-country comparison of citizens in Greece, Italy and Spain. In S. N.-R. Intelligence. Springer Nature

Bushway, S.D. (2020). "Nothing Is More Opaque than Absolute Transparency": The Use of Prior History to Guide Sentencing. *Harvard Data Science Review*, *2*(1).

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis.* Thousand Oaks: Sage.

Deulkar, D.D. & Gupta, P. (2018). A study on usage of Online Personal Information by Data Brokers.

International Research Journal of Engineering and Technology (IRJET). Vol.5. Iss.5. May 2018. Retrieved from https://www.irjet.net/archives/V5/i5/IRJET-V5I5754.pdf

European Parliament. (2021). *Artificial Intelligence in policing: safeguards needed against mass surveillance | News | European Parliament*. (2021, June 29). https://www.europarl.europa.eu/news/en/press-room/20210624IPR06917/artificialintelligencein-policing-safeguards-needed-against-mass-surveillance

Geris, E. & Wellington, V. (2021). UN rights chief calls for safeguards on artificial intelligence. *Jurist: Legal News and Commentary. (16 Sep 2021)*. Www.jurist.org. Retrieved March 29, 2022, from https://www.jurist.org/news/2021/09/un-rights-chief-calls-for-safeguardson-artificialintelligence/

Krippendorff, K. (2004). *Content Analysis : An Introduction to its Methodology.* Thousand Oaks: Sage.

Lyon, D. (2002). Everyday surveillance: Personal data and social classifications. *Information, Communication & Society, 5*(2), 242-257.

Macnish, K. (2021). Surveillance Ethics: An Introduction to an Introduction. *Ethical Issues in Covert, Security and Surveillance Research*, 9–16.

Mann, S. (2017). Big Data is a big lie without little data: Humanistic Intelligence as a human right. *Big Data & Society*, *4*(1), 205395171769155.

Marthew, A., & Tucker, C. (2017). The impact of online surveillance on behavior. In *Cambridge Handbook of Surveillance Law* (pp. 437-454). Retrieved from https://ssrn.com/abstract=3167473

Marx, G. T. (2009). A Tack in the Shoe and Taking off the Shoe Neutralization and Counterneutralization Dynamics. Surveillance & Society, 6(3), 294–306.

Montag, L., Mcleod, R., De Mets, L., Gauld, M., Rodger, F., & Pełka, M. (2021). EDRi Report: A Legal Analysis of Biometric Mass Surveillance Practices in Germany, The Netherlands, and Poland.

Petersen, J. K. & Taylor. P. (2012). *Handbook of surveillance technologies*. Crc Press.

Quattrocolo, S. (2020). Equality of Arms and Automatedly Generated Evidence. *Artificial Intelligence, Computational Modelling and Criminal Proceedings*, 73–98.

Stoycheff, E., Burgess, G. S., & Martucci, M. C. (2020). Online censorship and digital surveillance: the relationship between suppression technologies and democratization across countries. *Information, Communication & Society, 23(4), 474–490*.

Trend Micro Research. (2020). *Malicious Uses and Abuses of Artificial Intelligence.* Retrieved from https://www.europol.europa.eu/cms/sites/default/files/documents/malicious_uses_and_abuses_of_artificial_intelligence_europol.pdf

Trottier, D. (2017). “Fear of contact”: Police surveillance through social networks. *European Journal of Cultural and Political Sociology, 4(4)*, 457–477.

Završnik, A. (2020). Criminal justice, artificial intelligence systems, and human rights. ERA Forum, Vol 20, 567–583.

Paper 2

Policing and Society

An International Journal of Research and Policy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gpas2

Safety, privacy, or both: evaluating citizens’ perspectives around artificial intelligence use by police forces

Yasmine Ezzeddine, Petra Saskia Bayerl & Helen Gibson

To cite this article: Yasmine Ezzeddine, Petra Saskia Bayerl & Helen Gibson (2023): Safety, privacy, or both: evaluating citizens’ perspectives around artificial intelligence use by police forces, Policing and Society, DOI: 10.1080/10439463.2023.2211813 **To link to this article:** https://doi.org/10.1080/10439463.2023.2211813

POLICING AND SOCIETY
https://doi.org/10.1080/10439463.2023.2211813

Safety, privacy, or both: evaluating citizens' perspectives around artificial intelligence use by police forces

Yasmine Ezzeddine, Petra Saskia Bayerl and Helen Gibson CENTRIC,

Sheffield Hallam University, Sheffield, UK

ABSTRACT

Police forces are increasing their use of artificial intelligence (AI) capabilities for security purposes. However, citizens are often aware and cautious about advanced policing capabilities which can impact negatively on the perceived legitimacy of policing efforts and police more generally. This study explores citizens' subjective perspectives to police use by AI, including tensions between security, privacy, and resistance. Using Q methodology with 43 participants in the UK, Netherlands, and Germany we identified five distinct perspectives towards AI use by police forces. The five perspectives illustrate the complex, diverse viewpoints citizens exhibit with respect to AI use by police and highlight that citizens' perspectives are more complex than often portrayed. Our findings offer theoretical and practical implications for public engagement around general versus personal safety, privacy and potentials for moral dilemmas and counter-reactions.

ARTICLE HISTORY
Received 25 October 2022
Accepted 2 May 2023

KEYWORDS
Artificial intelligence; police; surveillance; safety; privacy; citizen reactions; Q methodology; resistance.

1. Introduction

In the current era, marked by the growing relevance of artificial intelligence (AI) to most aspects of life, policing has been equally touched by the implementation of AI to advance investigations, safety, and security (Fussey and Sandhu 2022, Urquhart and Miranda 2022). These advances often trigger citizen concerns about the strategies and technical tools being used as part of modern policing with legitimate concerns about the possible implications and repercussions of implementing new capabilities (e.g. Moses and Chan 2016, Fussey et al. 2021) as well as concerns about the 'gradual outsourcing of police work' (Smith et al. 2017, p. 260). Debates in this context are often framed around the notion of a 'trade-off' between privacy and security (Pavone and Esposti 2012), which in turn contributes to the spread of a narrative that suggests an 'antagonistic relationship' between police and the public (Nalla et al. 2018, p. 271).

This study aims to question this either-or approach with the objective to obtain a better understanding of the possible complexity within citizens' views around AI use by police. Our inquiry reacts to calls to refine the 'privacy versus security' narrative in the policing domain (cp. for instance, the 'Freedom AND Security 2018 – Data Protection Conference' organised by Europol[1]). As Solove claims, the current debate 'has been framed incorrectly, with the trade-off between these values understood as an all-or-nothing proposition' (Solove 2011, p. 2). In this paper, we aim to offer a realistic appreciation of citizens' perspectives and their sense-making about the complex domain of AI use by police. Our study contributes in theoretical terms to debates on citizen attitudes about

CONTACT Yasmine Ezzeddine y.ezzeddine@shu.ac.uk

policing practices by providing a new framework for the categorisation of subjective perspectives. Specifically, it offers new insights into the way individuals make sense of and aim to resolve tensions between privacy and security raised by police use of AI. This investigation purposefully went beyond questions of attitudes (i.e. acceptance or rejection) by questioning possible behavioural consequences of surveillance fears and the moral dilemmas such behaviours may entail. In practical terms, we highlight the need for a balanced stance that can acknowledge the benefits and the risks of AI to account for their societal concerns and ramifications. This answers important calls for assessing 'the complexities and uncertainties brought by novel technologies' in modern-day policing (Fussey and Sandhu 2022, p. 11).

In the subsequent sections, we outline the background to the study, followed by an explanation of the empirical approach, an exploration of the findings and their theoretical and practical implications.

1.1. AI use by police forces

Over the past years, surveillance for safety and security purposes has expanded considerably (Turner et al. 2019) driven by the adoption of new technologies by police forces. One of the most recent entries is AI, defined as 'systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals' (Committee of the Regions for Artificial Intelligence in Europe Brussels 2018, p. 1).

AI algorithms and data analytics capabilities are being adopted by police forces in various functionalities (Babuta and Oswald 2020), generally with a view to increase efficiency and reduce resource demands on policing time and personnel. For instance, in the UK, Durham Constabulary adopted the Harm Assessment Risk Tool (HART) to predict the likelihood of re-offending by criminals within two years of being released from prison to determine whether certain individuals might benefit from a rehabilitation programme (Oswald et al. 2018). In parallel, the Metropolitan Police joined South Wales police in trialling facial recognition technology to automatically identify people through CCTV, particularly at large events, for crime detection and prevention purposes (Oswald et al. 2018, Metropolitan Police and NPL 2020). AI capabilities are also considered by police for combatting serious crimes such as terrorism, child sexual exploitation or organised crime (Završnik 2020).

The call for police use of AI is often predicated by the growing complexity and globalisation of the crime landscape, specifically the possibility to predict, identify and counter new crime trends (e.g. Fussey and Sandhu 2022). An example is cybersecurity which has gained attention due to a 'technological arms race' between attacker and defender (Schneier 2012), in the sense that the former constantly seeks weaknesses to infiltrate systems and the latter aims to prevent and defend against increasingly sophisticated intrusions (Tounsi and Rais 2018). Growing sophistication and shifts in criminal modus operandi and crime patterns in smart societies (Kaufmann et al. 2019) provide a motivation for police to advance their frameworks and capabilities with a view to safeguard their operational efficiency (Jahankhani et al. 2020), and ultimately the safety of society. Generally, the use of AI capabilities by police forces are seen to hold considerable potential and benefits for safety (Morgenstern et al. 2021), evidence collection and the mitigation of threats (Lyon 2002).

Yet, advances in AI capabilities also create conflicts between their potential security benefits and concerns about their accuracy and fairness, the potential for discrimination of specific groups and the reinforcement of societal inequalities (e.g. Bushway 2020, Quattrocolo 2020, Završnik 2020, FRA 2021). This is often coupled with a perceived lack of evidence for the efficiency of algorithmic-based decisions and a 'fear of contact' emanating from alliances of police with the private sector (Trottier 2017, p. 475).

In reaction to growing societal concerns about the use of AI capabilities by law enforcement, counter movements have sprung up particularly targeting large-scale automated surveillance in public places or online – ranging from campaigns such as ReclaimYourFace (https://reclaimyourface. eu) to technological solutions to hide one's online footprint (e.g. Cover your tracks or PrivacyBadger by the Electronic Frontier Foundation[2]). In a further example, in June 2020 Amazon, IBM and Microsoft halted their sale of facial recognition software to police forces, demonstrating the impact of strong public opinions towards police use of AI (Lee and Chin 2022). Citizen reactions can thus have a considerable impact on the opportunities of police to develop and deploy AI. In consequence, understanding citizen reactions is of vital concern for police forces to retain public trust and the legitimacy of their actions.

1.2. Citizen reactions to AI use by police

Public acceptance of AI use by police has received considerable attention over the last years. Interestingly, results reveal considerable variations in public reactions. For instance, a survey with 154,195 respondents across 142 countries (Neudert et al. 2020) suggests clear regional differences whereby 49% of participants from Latin-America and Caribbean, 47% in North America and 43% in Europe considered AI to be 'mostly harmful', whereas 59% of participants in East Asia considered AI as 'mostly helpful'. AI acceptance is also influenced by demographics. A report investigating citizens' level of trust in AI in Australia revealed that young people, people with knowledge of computer science and people of higher educational levels are more positive towards AI (Lockey et al. 2020). A recent survey across 30 countries conducted by the AP4AI project,[3] which focuses specifically on AI use by police forces (Akhgar et al. 2022), found that most participants were positive towards AI for specific purposes (e.g. 90% agreed/strongly agreed to its use for the safeguarding of children and vulnerable groups, 79% to its use to prevent crime before they happen). In contrast, an earlier investigation by Amnesty International (2015) revealed strong negative attitudes towards 'government surveillance'. Although not directly focused on AI use, participants largely disagreed with their governments intercepting, storing, and analysing their data.[4]

This short selection of findings is indicative of the diversity and contradictory nature of public attitudes towards AI use by police when framed purely as acceptance versus rejection, and towards their potential for 'safeguarding' versus 'surveilling' society. Discussions about the threat of technological advances on privacy (e.g. Bradford et al. 2020) thus tend to be accompanied by arguments that citizens are willing to sacrifice some degree of their privacy for the benefit of safety (e.g. Davis and Silver 2004), which is in line with Solove (2011), who argues that the dichotomy between privacy and security is largely artificial.

Some indications exist on the underlying factors that shape the observed diversity in attitudes and reactions. Gurinskaya (2020), for instance, identified trust in the efficiency of surveillance technologies as part of a cost–benefit assessment that affects citizens' acceptance or rejection of AI use by police. On the other hand, perceived ramifications on

citizens' rights and abilities to free expression (Benjamin 2020) can be one of many factors that trigger resistance to surveillance, defined as 'disrupting flows of information from the body to the information system' (Ball 2005, p. 104). Resistance can be seen as conscious and strategic choices made by citizens when confronted with AI use by police, ranging from technical and social counterstrategies (such as the use of Electronic Frontiers Foundation tools mentioned above), to obfuscation or the decision to remove oneself from online spheres (Bayerl et al. 2021). Marx (2009) further proposed neutralisation strategies as common reactions, which he defined as a 'dynamic adversarial social dance' (p. 99) whereby opponents reciprocate in performing innovative moves in a chain reaction of surveyed versus surveyor to neutralise surveillance/ countersurveillance consecutively. In this 'social dance' citizens often exhibit moral dilemmas that are conditional to the specifics of the usage context and type of AI deployed (e.g. Carrasco et al. 2019).

Overall, these studies indicate that contextual and psychological factors contribute to shaping attitudes. However, past inquiries provide insufficient insights to allow a clear understanding of the sense-making by citizens when it comes to balancing their stance towards AI use by police forces. Investigating citizens' sense-making affords a view into the rationales and the 'checks and balances' when considering the complex issue of AI use by police which provides a foundation for observable disparities in attitudes and reactions. This exploratory study aims to obtain an understanding of the rationales for citizens' subjective positions about police use of AI with a view to untangle the rhetoric between 'safety versus privacy'. Such an exploration is particularly important in the light of the impact of expanding technology on citizens' freedoms and self-expression abilities, whereby a balanced stance is needed to equally acknowledge the benefits and the risks of AI, and to account for societal concerns and ramifications of modern police surveillance in the context of perceptions, acceptance, and resistance. This study therefore had two interlinked aims: (1) to identify subjective positions towards AI use by police beyond mere acceptance–rejection; (2) to identify the rationales and sense-making that underly disparate subjective positions towards AI use by police.

2. Methodology

This study used Q methodology in combination with interviews (Brown 1993). As a combination of quantitative and qualitative approaches, Q methodology is an exploratory approach that offers 'a means of capturing subjectivity – reliably, scientifically and experimentally' (Watts and Stenner 2012, p. 44) and as such has been applied to numerous fields in which subjective perspectives and sense-making are relevant (e.g. social and health related studies; cp. Chururcca et al. 2021, Stenner et al. 2000).

Q methodology uses a set of pre-defined statements that together represent a range of disparate positions towards the issue in questions, in our case AI use by police forces. Participants are then instructed to sort the statements into a predefined distribution according to their agreement/disagreement with each statement (for details see section on data collection below). Using a forced distribution is the standard approach and coincides with Stephenson's notion of psychological significance (Burt and Stephenson 1939) that influences participants into reflecting on the precise psychological significance to each statement.

2.1. Q statement set

The Q statement set consisted of nine statements. The statements were created based on results from previous research by the authors which explored citizen acceptance as well as surveillance reactions (Bayerl et al., 2021). The statements were chosen to represent disparate perspectives, which also integrate stances from (supportive) security and (critical) surveillance fields. The set addresses three aspects: (1) the tension between privacy versus safety considerations offering disparate options for resistance from 'lowkey' to destructive; (2) a differentiation between use of AI capabilities in online (i.e. on digitally enabled platforms) versus offline settings (i.e. real life, on the streets, in public spaces …) and (3) the tension between benefits for oneself versus others. The statements were purposefully formulated as extreme positions (using markers such as 'need to', 'never', 'totally') to elicit strong responses from participants. A pilot-test was conducted with two volunteers who did not take part in the actual study later-on. This was done to ensure that the statements were clear and elicited useful responses. The volunteers suggested adjustments to some of the statements to reduce their complexity and make them clearer. For instance, the abstract formulation 'avoid facial recognition AI' was replaced with the more concrete 'prevent AI systems from capturing my face and movements' (statement 4) and 'not to post things' was replace with 'to never post pictures or other personal information' (statement 2). The latter is also an example of strengthening statements to elicit stronger reactions ('never post' instead of 'not to post'; similarly statements 9: 'if their presence may lead' to 'if their presence leads to...'). The resulting statement set can be found in Table 1.

Table 1. Complete list of Q sort statements presented to participants for distribution.

1.	It is totally appropriate to falsify my personal information online to protect my privacy, even if it means that AI police systems fighting against cybercrime and terrorism will be inaccurate as a result.
2.	It is ok to ask my family and friends to never post pictures or other personal information about me on their social media to avoid AI police systems collecting and inspecting my information.
3.	If I have to choose between taking measures to prevent AI police systems from monitoring or using my personal movements OR contributing to safeguarding others from terrorism/cybercrime, I have a moral responsibility to put other people's safety first.
4.	If I want to prevent police AI systems from capturing my face and movements, I need to accept that I have to avoid public spaces such as street festivals or airports.
5.	I do not object to AI-systems of police monitoring my behaviour and movements online if they keep me safe from terrorism/ cybercrime, but they should never be used to monitor my life offline.
6.	AI systems by police need plentiful and accurate information from all of us to identify bad actors (terrorists, cybercriminals). Therefore, it is immoral for others to use technologies that hide/distort information that can help these systems from keeping me safe.
7.	Trying to resist or avoid AI systems by police is a bad idea, because it only means police will develop even better AI capabilities.
8.	People should stop behaving aggressively in crowds if they know AI systems are used in the area. And if they do, they should not be surprised if AI police systems flag them up as suspicious.
9.	Destroying facial-recognition cameras on my street is appropriate if their presence leads to biased over-policing of my neighbourhood.

2.2. Participants

Participants stemmed from three countries: the UK, Netherlands, and Germany. The rationale for a multi-national sample was to provide scope for the emergence of diverse opinions on AI use by police. The three countries represent similar policing approaches (i.e. an emphasis on community-led policing), while being known for disparities in the

uptake of and attitudes towards AI use of police. The selection of these three countries was also owing to the familiarity of the authors with the countries (citizens of UK and Germany, respectively, one with over a decade of experience living in the Netherlands) which assisted in the translation and interpretation of the collected data.

The study was conducted as part of an international research project. The participants were recruited by researchers in the respective countries. Country teams were given freedom to recruit a group of interest in their specific country. The only selection criterion was an age of 18 or older for reasons of ethical consent. Overall, we received information from 43 participants: 16 each from the Netherland and Germany, 11 from UK.[5] The German sample focused on young women (average age 26.3 years), the Netherlands on participants with cybersecurity expertise (seven women, nine men, average age 32.2 years), while the UK sample focused on people with a migration background (nine women, three men, average age 33.4 years), leading to 72.2% women and an average age of 30.4 years for the full sample. Table 2 shows the demographics and gender characteristics of the selected groups. As this overview shows, the overall sample has an imbalance towards women and younger people, which will be considered in the interpretation of the data.

Table 2. Characteristics of citizens interviewed per country.

Country	Group selected	Interviews conducted	Average Age	Gender distribution Women / Men	< 35 years	34–55 years	45–55 years
Germany	Women between 18–53 years of age	16	26.3	100% / 0%	81.2%	18.7%	12.5%
Netherlands	People with cybersecurity experience	16	32.3	43.7% / 56.2%	62.5%	37.5%	0
UK	Citizens with migration background	11	33.4	72.7% / 27.2%	63.6%	36.4%	27.3%
Total		43	30.4	72.2% / 27.8%	69.1%	30.9%	13.3%

2.3. Ethics

The study received approval by the ethics committee of the authors' university. Additionally, participants received an information sheet to clarify the context and legal basis of the study, details of data handling and participants' rights. This included the right to withdraw and not provide demographic information if they did not feel comfortable doing so. All data was analysed in pseudonymised form.

2.4. Data collection

Interviews were conducted by the researchers in their respective countries to ensure interviewers were familiar with the national culture and context. In the UK and the Netherlands, interviews were held in English. In Germany, statements were translated into German. The translation was validated by the second author, who is a native German speaker, and thus well-positioned to ensure that the translated statements carried the same meaning and intent as the original versions.

The Q sort interviews were conducted with each participant individually either face-to-face or online. Participants were presented with the nine statements on an A4 paper aligned with the chart represented in Figure 1. For remote participants (i.e. over Zoom or Microsoft Teams), the Q sort template was sent by email to fill out locally or

researchers would share their screen allowing participants to view the document. Participants were then instructed to fit the statements into the forced distribution according to how much they agreed/disagreed with the statement. In face-to-face interviews, participants filled in the paper form. For remote participants, the researcher filled in the statements on a local copy as the participant announced their choice. Participants were encouraged to elaborate on the rationales for their placement of statements which provided critical background information for the interpretation of sorts and factors. All interviews were audiorecorded.

2.5. Data analysis

The analysis of Q sorts allows to identify clusters of participants with similar subjective perspectives (Ellingsen et al. 2010). This process is purely exploratory, i.e. the analysis does not use any preimposed categories or features in creating the clusters. Hence, clusters (or Factors, in Q sort terminology) emerge bottom-up from the data with individuals who share similar views loading on the same Factor. The subsequent analysis of the Factors, together with participants' comments, is the core analytical measure of Q methodology (McKeown and Thomas 1988) by investigating the pattern of agreements to the items within a Factor, as well as the degree of agreement and disagreement between perspectives.

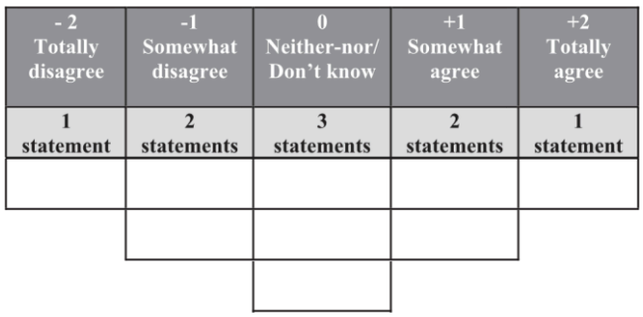


Figure 1. Q sort distribution chart used presented to participants in this study.

Using Peter Schmolck's PQ software package (Schmolck 2014), a Centroid Factor analysis was conducted. As a standard the software extracts seven Factors (Brown 1980, p. 223). Several criteria are used to determine the correct number of Factors to be extracted. Firstly, Eigenvalue (EV) analysis revealed five Factors with an EV larger than the 1.00 cutoff (see Table 4). A Scree test (Watts and Stenner 2012), as a common addition to statistical tests, was less conclusive suggesting between three to five factors, while applying Humphrey's rule eliminated two Factors due to standard errors below the cut-off point, suggesting retention of three Factors. However, retention of only three Factors would mean disregarding a considerable number of participants (9 out of 43), which would be risky, since significant viewpoints can be overlooked as a result, especially given the diversity of our sample. Following Brown (1980), we proceeded with five Factors to allow the emergence of potentially less prevalent but important viewpoints. The final solution (using Varimax rotation; complemented by manual flagging of Q sorts with loadings greater than 0.39) found that all but one Q sort loaded on the five Factors. Therefore, the 5-Factor solution was chosen representing an explained variance of 79%.

The content of the five Factors (i.e. subjective perspectives) will be discussed in the Results section. The interpretation must adopt an open-minded, careful, and comprehensive assessment of the patterns found across the perspectives. This was accomplished by examining the relative ranking of each statement to understand the

reasoning and viewpoints being reflected in each Factor. Additionally, the interpretation of Factors was crucially supported by the comments and reflections made by participants during or after completing the sort. We conducted a thematic analysis of the transcripts with a focus to understand the rationales for specific sorting decisions. Our analysis approach was primarily inductive (Patton 1990, Braun and Clarke 2006) in that we did not have predetermined themes for the potential rationales, participants may use to make their sorting decisions. Rather the rationales emerged from the comments about specific items, which could then be compared across perspectives (e.g. disparate reactions to statement 8: UK06: it's actually a good thing, knowing that AI systems are working in that way indicating acceptance due to strong safety benefits vs UK10: You shouldn't change your behaviour, just thinking you're being watched; especially when, when you're not doing anything against law indicating high value given to free expression). The analysis was done using the qualitative analysis package NVivo, and revealed complex attitudes and varied themes of acceptance, resistance and reactions to AI use by police forces.

3. Results

The subsequent sections provide a description of each perspective, followed by a comparison to draw out overlaps and specifics of each. In the descriptions, numbers such as (6: +2) indicate the statement number (cp. Table 1) and its ranking on the specific Factor. For instance, in the example (6: +2), statement 6 has been ranked in the +2 position (totally agree). The comments made by participants during or after sorting are cited in italics to support interpretation and to enhance understanding. We have also provided a title for each perspective to clarify and emphasise the core aspects of each viewpoint.

3.1. Perspective 1: 'privacy first'

Ten participants shared this viewpoint. To this group of citizens, privacy was the greatest concern. Participants considered it highly appropriate to falsify personal information online to protect their privacy, even if this means that police AI systems fighting against cybercrime and terrorism may be less accurate as a result (1: +2). As the following quote shows, this was less driven by worries about police surveillance than a general fear of revealing too much: If I feel safer behind a pseudonym, then I should be allowed to use it to protect myself against bad actors who might attack me (GER01). This was also supported by negative reactions to statements relating to the moral responsibility of sharing accurate information (6, −1) and putting other people's safety first (3: −1), further reinforced by participant NL09: I totally disagree with 6, because I personally do that, I hide or distort information like my birthday. I don't give the correct year or month.

Moreover, asking friends and family to never post anything about them (2: +1) and avoiding public spaces to prevent police AI systems from capturing their personal information (4: +1) were seen as favourable, reinforcing the Privacy First perspective also in the offline domain. At the same time, participants felt neutral towards high-level resistance measures: the destruction of facial recognition cameras and expecting people to act differently in crowds to avoid being flagged by AI systems (9, 8: 0). They further felt neutral about resistance being a bad idea (7: 0). What further distinguishes the Privacy First perspective from other positions is their strong opposition against police monitoring of behaviours online but not offline (5: −2). This was expressed clearly by participant GER02: For me there is no difference between online surveillance and surveillance offline in public places, again indicating the importance of privacy in both contexts.

3.2. Perspective 2: 'safety first'

At the opposite end stands the 'Safety First' perspective in which participants strongly agreed to AI systems by police monitoring their online behaviours for security purposes (5: +2). This viewpoint acknowledges a moral responsibility to put other people's safety above own privacy concerns (3: +1), which was also clearly phrased by UK02: People's safety is more important than anything else for me. Safety consciousness extended to asking friends and family to never post pictures about them online (2: +1) or at least be consulted. According to GER05: For me it is OK if pictures of me are posted by others, but I would always like to be asked first. Safety First participants were neutral towards the proposition that hiding or distorting information would be immoral (6: 0) and, in contrast to the Privacy First position, also had a neutral stance towards falsifying personal information online if it may lead to less accurate police systems (1: 0). As put by GER09: The statement 6 does not say for what reason people might wish to hide certain data and information. Such people could have a legitimate interest to do so or their reasons damage society. The same is true for the need to avoid public spaces to prevent police AI systems from capturing personal information (4: 0).

Conversely, Safety First proponents disagreed that people should stop behaving aggressively in crowds when AI systems are being used (8: −1). They also disagreed with the notion that trying to resist or avoid AI systems by police would lead to the development of better AI capabilities by the latter (7: −1) based on the belief that there is somewhat [of an] inevitability about it, being something that will happen in future (UK04). Similarly, these participants completely disapproved the destruction of facial recognition cameras even if they cause over-policing (9: −2). As expressed by GER16: If any policing takes place, then this happens for a good reason which is to keep the area safe. And if the result is overpolicing, then the objection should not be destruction of the cameras. This clearly emphasises acceptance and trust in AI police measures for protection.

3.3. Perspective 3: 'protective AI'

The third perspective saw considerable benefit for AI, as long as it was used for the protection of themselves or society. Unlike the previous perspectives, this group felt very positive towards AI's potential to stop aggressive behaviours in crowds giving it a preventive purpose (8: +2); e.g. GER06: this scenario seems to prevent aggressive behaviour in advance, but I do not think that every aggressive act would immediately be super suspicious. Nevertheless, I can imagine that overall, the behaviour would become much more pleasant and more respectful through the use of such AI systems. Similarly, people with a Protective AI perspective disagreed with the need to avoid public spaces to avoid AI systems (4: −1) and even more with the destruction of facial recognition cameras (9: −2) demonstrating a strong supportive view of AI. In addition, this perspective felt a moral obligation to put other people's safety first, even if this meant not masking own behaviours towards AI (3: +1). At the same time, this perspective attaches importance to privacy such as asking friends and family to not share personal information about them online (2: +1); e.g. I find my privacy should be protected and for me, in this respect, the question is what exposes my personal data. If I do not put my data online myself, then my family and friends should not make these data available (GER15). Similarly, UK08: I agree but it has nothing to do with artificial intelligence police system, it's anybody, I would not want anybody to track me or to see where I am. This indicates that generally, this

perspective appreciates privacy, however, sees a clear value in AI systems for safety purposes including as preventive measure.

3.4. Perspective 4: 'not me'

Participants with this viewpoint want 'the best of both worlds', requesting security as long as it does not infringe on their own life and privacy. Security and privacy are seen as opposing options: Because of the natural antagonism between security and privacy, the guarantee of privacy seems to reduce the level of security (GER02). In line with a positive security stance, they strongly disagreed that falsification of information is appropriate if this infringes on police AI systems (1: −2). They moreover strongly agreed with having a moral responsibility to put other people's safety first (3: +2). On the other hand, they did not find resistance against AI capabilities problematic (7: −1), while seeing the necessity to avoid public spaces to prevent police AI systems from capturing their information (4: +1) and to ask their friends and family to never post pictures of them online (2: +1). Interestingly, statements about the behaviours of others, either in terms of hiding/distorting information, aggression in public spaces or the destruction of facialrecognition cameras received neutral reactions (6, 8: 0). This suggests a focus foremost on their own personal situation, which contrasts with the Safety First perspective which focuses on security including others.

3.5. Perspective 5: 'Anti-surveillance'

The Anti-surveillance viewpoint is characterised by its clear acceptance of resistance. This group approved most strongly of asking friends and family to never post pictures of them online (2: +2), and not only for police avoidance. According to participant GER10: I would ask my family and friends not to post pictures of me for other reasons than to prevent AI systems of the police from collecting them. Moreover, this group was the only one that approved of the destruction of facial recognition cameras in areas where they may lead to biased over-policing (9: +1). This is clearly expressed by participant UK05: It doesn't matter. If it's bothering someone, they can destroy it, if it's harming them. Moreover, they did not perceive hiding/distorting of personal information online as immoral even if it infringes on security (6: −1). The Anti-Surveillance stance was also expressed in emphasising that people should be able to act as they wish without worrying about police AI systems flagging them up as suspicious (8: −1) and strong disagreement to the claim that people should avoid public spaces to prevent police AI systems from capturing their personal information (4: −2; e.g. No, I do not agree with this statement because the restriction to freedom is way too high. For me, it is not acceptable if it became impossible to walk around incognito as nobody, GER12). At the same time, they acknowledged a moral responsibility to put other people's safety before their personal privacy concerns (3: +1), which indicates that even the Anti-Surveillance position sees merit in some policing measures.

3.6. Comparison of perspectives

Table 3 provides a direct comparison of statement rankings. Comparing the five viewpoints reveals several shared and common reactions towards AI use by police, but also presents

Table 3. Statement rankings across the five factors (ordered by increasing disagreement across factors).

Q	Statement	1	2	3	4	5

2. It is ok to ask my family and friends to never post pictures or other personal information about me on their social media to avoid AI police systems collecting and inspecting my information.	+1	+1	+1	+1	+2
3. If I have to choose between taking measures to prevent AI police systems from monitoring or using my personal movements OR contributing to safeguarding others from terrorism/ cybercrime, I have a moral responsibility to put other people's safety first.	−1	+1	+1	+2	+1
6. AI systems by police need plentiful and accurate information from all of us to identify bad actors (terrorists, cybercriminals). Therefore, it is immoral for others to use technologies that hide/distort information that can help these systems from keeping me safe.	−1	0	0	0	−1
7. Trying to resist or avoid AI systems by police is a bad idea, because it only means police will develop even better AI capabilities.	0	−1	−1	−1	0
8. People should stop behaving aggressively in crowds if they know AI systems are used in the area. And if they do, they should not be surprised if AI police systems flag them up as suspicious.	0	−1	+2	0	−1
5. I do not object to AI-systems of police monitoring my behaviour and movements online if they keep me safe from terrorism/cybercrime, but they should never be used to monitor my life offline.	−2	+2	0	−1	0
9. Destroying facial-recognition cameras on my street is appropriate if their presence leads to biased over-policing of my neighbourhood.	0	−2	−2	0	+1
1. It is totally appropriate to falsify my personal information online to protect my privacy, even if it means that AI police systems fighting against cybercrime and terrorism will be less accurate as a result.	+2	0	0	−2	0
4. If I want to prevent police AI systems from capturing my face and movements, I need to accept that I have to avoid public spaces such as street festivals or airports.	+1	0	−1	+1	−2

insightful disparities in the sense-making and the challenge of balancing between privacy and safety. Essentially, the Privacy First group comprises citizens who prioritise privacy. However, they are neutral towards the destruction of facial recognition cameras, resistance, and aggressive behaviours in crowds, which signals a non-violent stance that contrasts strongly with the Anti-surveillance perspective. The Anti-Surveillance perspective encourages resistance, including but not only against police. This indicates a generalised opposition to sacrificing their freedoms of expression (both online and offline/in public spaces) as a price for safety, if needed condoning aggressive means. The Safety First perspective is strongly concerned about safety in a broad sense that includes a moral responsibility to sacrifice personal privacy concerns for the sake of other's protection. The Safety First perspective thus represents a strong collective orientation towards safety and security, which surpasses many concerns around AI use by police forces that are pronounced in other perspectives. The Not Me viewpoint can be seen as representing the other end of the spectrum. Proponents are generally in favour of AI use by police if it safeguards themselves, although preferably not on their own data. Not Me individuals thus primarily prioritise the own personal safety along with personal privacy. The Protective AI viewpoint emphasises the benefit of AI systems if applied for security purposes. At the same time, privacy concerns ranked high, while moral responsibility did not receive much attention. This perspective thus represents a narrower stance about AI with a somewhat ambiguous view that lacks a clearly integrated position.

Considering the demographic characteristics across perspectives, no single Factor was dominated by participants from a single country or group: all perspectives included representatives from all three countries and similar gender distributions (cp. Table 4). This suggests that perspectives are

Table 4. Demographic distribution of participants across the five factors (ordered by explained variance).

Factor	No. participants per Factor	Explained Variance	Eigen-value (EV)	Average Age	Gender (Male/Female)	Distribution Countries	% Participants
1	10	18%	10.6	31.4	20% / 80%	1 UK – 5 NL – 4 GER	
2	7	17%	8.1	27.1	14% / 86%	3 UK – 4 GER	

3	10	17%	6.0	30.2	30% / 70%	3 UK – 4 NL – 3 GER	4	6
14%	5.2	30.8	33.3% / 66.7%	1 UK – 3 NL – 3 GER				
5	9	13%	4.2	32.9	22.2% / 77.8%	3 UK – 3 NL – 3 GER		

Note: UK: United Kingdom; NL: Netherlands; GER: Germany.

founded on individual aspects and experiences rather than overt demographics such as country origin, gender, or age. That is, for the sense-making about AI use by police forces, other personal aspects seem more relevant than national context or membership in a specific professional or demographic group.

One person did not fall into the 5-factor solution. This participant (NL, male, 20 years) expressed views that oscillated between concerns for privacy and wanting to ethically contribute to safeguarding society. The participant was in favour of AI generally but believed that AI tools should only be owned by the police. This participant thus represents a view that wavered across factors, integrating aspects from several perspectives.

4. Discussion

This study set out to gain an understanding of citizens' perspective to AI use by police forces. The five perspectives identified in our data demonstrate the variation in citizens' viewpoints with disparate foci on general versus personal safety, privacy and potentials for moral dilemmas and acceptance of (aggressive) counter-reactions. The findings have important theoretical and practical implications by providing insights into the complexities of citizen reactions around AI use in the policing and security domain.

Crucially, our observations demonstrate the different ways in which individuals make sense of AI use by police, highlighting the checks-and-balances and moral or rationale bases for their views. For instance, citizens of the Safety First group did not oppose surveillance or AI use (online and offline), because they argue that monitoring is essential if police want to keep citizens safe, especially with the increasing challenges that police forces face with respect to social media and cybercrimes (David and Williams 2013). This resonates with the proposition of 'fear of crime' as a corner stone in community policing (Leman-Langlois 2002). Similarly, for citizens adhering to a Protective AI perspective, AI's positive outcomes – reflected in possibilities for the successful identification, screening, case linkage and other labour and time reducing functionalities (UNICRI and INTERPOL 2019) – can overcome some concerns about their privacy.

Discussions around police surveillance have been marked by notable theories drawing from Jeremy Bentham's early Panopticon (1791) that motivated numerous discussions around the costs versus benefits of surveillance (Foucault 1991, Orwell 2000), and deliberations around privacy and legality of overt and covert police surveillance (Foucault 1977, Marx 1988, Regan 1995). These were also issues, citizens sharing the Safety First and Protective AI viewpoints touched upon when stressing the need for a balanced implementation of AI to safeguard society, while ensuring basic human rights are not violated. This coincides with the 'trade-off' approach proposed by Pavone and Esposti (2012), where citizens trade, to a certain degree and in specific situations, their privacy, in exchange for enhanced security. The position is also in line with past studies which have shown that citizens are often more supportive of surveillance mechanisms than police officers would perceive (Nalla et al. 2018, Gurinskaya 2020). Crucially, only a small number of participants exhibited a complete rejection of the implementation of AI tools by police (visible in the AntiSurveillance stance).

Subtle differences emerged in terms of the balance between privacy (emphasised by the Privacy First and Anti-Surveillance groups) versus safety (emphasised by the Safety First and Protective AI groups) and foci of considerations – most notably a limited, personal conception versus a more generalised conception of safety. Moreover, we identified subtleties between stances that accept peaceable counter-reactions (e.g. not posting content) and those that accept more radical ones (e.g. destruction of cameras). These disparities in viewpoints provide explanations underlying the different positions observed in ongoing debates as well as overt citizen reactions.

Strong privacy perspectives correlate with those of various civil parties, members of academia and expert advisors who question the efficacy of AI technologies by deeming their performance 'limited' and their potential to reduce risk of algorithmic decisions as ambiguous and imprecise (Rovatsos et al. 2019), as well as those who call for in-depth evaluations and determination of the costbenefits analysis incurred on civil rights and freedoms (Benjamin 2020). These perspectives also correlate to studies revealing that individuals frequenting public and private areas are wise to, and capable of, eluding and deceiving surveillance (EDRi and EIJI 2021).

The perspectives further provide insights into the underlying strategies and reasonings for resistance. Numerous studies revealed that the public's resort to resistance and counterstrategies, in various forms of 'veillance' (Mann and Ferenbok 2013), is an effort to 'equalize' the power that the surveyor has over the surveyed (Dencik et al. 2016). This is particularly true for the 'Anti-Surveillance' stance. For some citizens, AI is best employed by police for protection and safeguarding purposes only (Protective AI group), while others support AI tools to be implemented, just not on themselves (Not Me group). The latter may be attributable to the uncertainty around AI's ethical and moral implications (Lyon 2002, DiVaio et al. 2022, Westacott 2010), which is also visible in citizens' reactions to the statements, despite the general acknowledgment of its potential for public protection.

By comparing the perceptions and viewpoints emerging from the Q sorts, it becomes apparent that the reality of attitudes and potential resistance responses towards AI use by police forces is highly complex, and that the notions of privacy and security or acceptance and rejection of AI use by police can often exist next to each other. The Not Me group can serve as a unique exemplar of the paradox of thought between safety and privacy, as these citizens want the advantages of safety and the benefits of privacy, all at once. They are generally in favour of AI use by police, which ultimately contributes to general safety, but not on their own personal data, which allows them to enjoy their own privacy. Hence, they want the best of both worlds. According to Solove (2011, p. 14), 'privacy is often misunderstood and undervalued when balanced against security'. This study revealed that citizens perceptions of privacy are much more complex than often portrayed. In fact, most individuals did not perceive AI use by police as an either/or scenario but offered differentiated arguments and contextualisation hinting towards situational, demographic, cultural and political factors.

Overall, our study provides an in-depth view on the range and complexity of attitudes justified by moral, ethical, and practical considerations around collective and personal safety and benefits of AI tools. The discrepant perspectives also explain the nature of accepted resistance (legal resistance routes versus destructive and illegal behaviours) and the personal duties and contributions towards general safety. This demonstrates that citizen perspectives towards AI use by police are much broader than often assumed, driven by reflections and propositions around acceptance, safety considerations and

moral responsibility. Understanding the underlying rationalisations expands and refines preexisting notions on surveillance and resistance and provides new pathways for the exploration of citizen reactions to the rapidly changing security environment. Our findings also shed light on the antimony between acceptance versus rejection of new technologies in the context of policing, along with the shifting attitudes towards personal privacy compared to personal and general safety.

In other words, the findings of this study not only expand on existing approaches to surveillance and resistance in the security area, but also address the gap in understanding the rationales behind the often ambiguous stance of citizens about acceptance and rejection of AI capabilities deployed by police. The contributions further include underresearched aspects, namely resistance factors and triggers that warrant the resort to counterstrategies in response to police use of AI. This research contributes to unpicking the binary narrative about 'safety versus privacy' by evaluating the rationale behind the costs and benefits of security measures, and how citizens balance privacy and security.

In practical terms, our exploration of citizens' perspectives offers a more promising avenue for police forces and policy makers to engage with public opinions and reactions. Engagements are often based on the assumption of a generalised resistance. This study helps recognise the complexity of sense-making, including benefit perspectives as well as moral and personal tensions and reflections on counter-reactions. Understanding this range and disparity of perspectives allows practitioners to better address and integrate the specific citizen concerns and expectations.

4.1. Limitations and future work

Q methodology as a qualitative interpretation method (Watts and Stenner 2012) is wellsuited to the study of subjective viewpoints by people within a specific context (Curt 1994), but it is not without limitations. The reflections and open-ended comments made by participants during and after the sorting exercise revealed that participants, although unaccustomed to the nature of a forced choice distribution, found that the Q sort challenge provided them with a unique opportunity to reflect on their opinions and stances towards AI application by police and that item rankings triggered reflection processes. Some statements contained two propositions and negative formulation (e.g. I don't) which for some participants required further clarification. This was not addressed by the participants in the pilot yet emerged during a small number of interviews. These items yielded important insights into reasons for agreement or disagreements towards specific aspects within a statement. In the rare event where participants agreed with one aspect of the statement but not the other, their rationales for focusing on one specific aspect provided crucial pointers to their sense-making.

This study did not include older participants (over 60s) and a higher number of women. We would therefore be cautious to claim that the five perspectives which emerged in this study are comprehensive of the viewpoints and perspectives around AI deployments generally. However, they do provide important insights into the complexity of reasoning around AI use by police, and the cognitive, and at times emotional, balancing acts that individuals perform. Future work can benefit from exploring these disparate viewpoints for a broader investigation into rationalisations and sensemaking around AI-based surveillance and the morality of resistance. Relatedly, the three groups across the three countries differed in important aspects, most markedly cyber-expertise, and migration experience. Although our analysis did not reveal a pattern, it cannot be ruled out that they are confounding influences. A quantitative approach to comparing perspectives

amongst demographics different groups would help to ascertain potential influencing factors as underlying reasons for disparate views.

Lastly, this study is a first exploration and should be followed up by studies that address the behavioural manifestations of subjective perspectives in its various forms, including counter-reactions towards police use of AI tools. The impact of the disparate perspectives towards AI deployments by police on actual behaviours and reactions remain important questions for future studies and the different factors that shape citizens' opinions and reactions.

5. Conclusion

Doubtless, applications of AI in policing can trigger uncertainty and scepticism around the ethical and moral ramifications of AI deployments (Feldstein 2019, Heaven 2020, McGuire 2020). In the light of on-going debates around AI implementations and the needed regulations and legislations, public opinions need to be taken seriously to allow informed decision-making about the adoption of AI for policing purposes. A pre-requisite, however, is an equally informed understanding of citizen perspectives. Current debates are often framed around binary positions of 'either security or privacy' which, as our study illustrates, is too restrictive. Our study offers a crucial window into the areas between these two extremes as expressed by citizens themselves. Generating a bird's eye view into societal concerns emanating from expanded technological advances in the security field is essential in the process of establishing legitimate means for AI applications in policing. A realistic understanding of citizen perspectives allows to account for citizen concerns adequately and ultimately to safeguard the relationship between the police and the public, whether through policy, regulations, or incorporating aspects that trigger concerns into the design and usage of AI tools.

Notes

1. https://www.europol.europa.eu/sites/default/files/documents/report_of_eden_conference_freedom_and_security_2018.pdf.
2. https://www.eff.org/pages/tools.
3. https://www.AP4AI.eu.
4. https://www.amnesty.org/en/latest/news/2015/03/global-opposition-to-usa-big-brother-mass-surveillance/.
5. The range emerged as countries were allowed to recruit a mixed sample of citizens (minimally 11) and security stakeholders (up to five). Since Q sorts are intended to capture subjective perspectives on an individual level, the interviews with security stakeholders – focusing on an organisational perspective – did not contain a Q sort. While the other countries only interviewed citizens, the UK conducted interviews with a mix of stakeholder (citizens and Civil Society Actors representing organisations engaged in fighting cybercrime and terrorism) leading to 11 citizen Q sorts for the UK.

Acknowledgements

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the European Union's Horizon 2020 Research and Innovation Framework Programme [grant agreement number 883569] as part of the AIDA project (AIDA – Artificial Intelligence and advanced Data Analytics for Law Enforcement Agencies).

References

Akhgar, B., et al., 2022. Accountability principles for artificial intelligence (AP4AI) in the internet security domain. AP4AI Project. Available from: https://www.europol.europa.eu/cms/sites/default/files/documents/Accountability_Principles_for_Artificial_Intelligence_AP4AI_in_the_Internet_Security_Domain.pdf [Accessed 12 March 2022].

Amnesty International., 2015. https://www.amnesty.org/en/latest/news/2015/03/global-opposition-to-usa-bigbrothermass-surveillance/

Babuta, A. and Oswald, M., 2020. Data analytics and algorithms in policing in England and Wales. RUSI. Available from: https://rusi.org/explore-our-research/projects/data-analytics-and-algorithms-in-policing [Accessed 22 March 2022].

Ball, K., 2005. Organization, surveillance, and the body: towards a politics of resistance. Organization, 12 (1), 89–108.

Bayerl, P.S., et al., 2021. Artificial intelligence in law enforcement: cross-country comparison of citizens in Greece, Italy and Spain. S. N.-R. Intelligence: Springer Nature, 11769.

Benjamin, G., 2020. Facial recognition is spreading faster than you realise. The Conversation. Available from: https://theconversation.com/facial-recognition-is-spreading-faster-than-you-realise-132047 [Accessed 15 February 2022].

Bentham, J., 1791. Panopticon, or, the inspection-house. Dublin, Ireland Printed: London Reprinted: T. Payne.

Bradford, B., et al., 2020. Live facial recognition: trust and legitimacy as predictors of public support for police use of new technology. SocArXiv. doi:10.31235/osf.io/n3pwa.

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. Qualitative research in psychology, 3 (2), 77–101.

Brown, S.R., 1980. Political subjectivity: applications of Q methodology in political science. New Haven, CT: Yale University Press.

Brown, S.R., 1993. A primer on Q methodology. Operant subjectivity, 16 (3/4), 91–138.

Burt, C. and Stephenson, W., 1939. Alternative views on correlations between persons. Psychometrika, 4 (4), 269–281.

Bushway, S., 2020. Nothing is more opaque than absolute transparency: the use of prior history to guide sentencing. Harvard data science review, 2 (1). doi:10.1162/99608f92.468468af.

Carrasco, M., Mills, S., Whybrew, A. and Jura, A. 2019. The Citizen's Perspective on the Use of AI in Government. BCG Digital Government Benchmarking. Retrieved from https://www.bcg.com/publications/2019/citizen-perspectiveuse-artificialintelligence-government-digital-benchmarking.aspx

Churruca, K., Ludlow, K. and Wu, W, 2021. A scoping review of Q-methodology in healthcare research. BMC Med Res Methodol, 21, 125.

Committee of the Regions for Artificial Intelligence in Europe Brussels, 2018. AI Communication from the Commission to the EU Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the regions on Artificial Intelligence for Europe. Brussels: Europe. Available from: AI Communication (europa.eu).

Curt, B., 1994. Textuality and tectonics: troubling social and psychological science. Buckingham: Open University Press.

David, S.W. and Williams, L.M., 2013. Policing cybercrime: networked and social media technologies and the challenges for policing. Policing and society, 23 (4), 409–412. doi:10.1080/10439463.2013.780222.

Davis, D. and Silver, B., 2004. Civil liberties vs. security: public opinion in the context of the terrorist attacks on America. American journal of political science, 48 (1), 28–46. doi:10.1111/j.0092-5853.2004.00054.

Dencik, L., Hintz, A., and Cable, J., 2016. Towards data justice? The ambiguity of anti-surveillance resistance in political activism. Big data & society, 3 (2). doi:10.1177/2053951716679678.

Di Vaio, A., Hassan, R., and Alavoine, C, 2022. Data intelligence and analytics: A bibliometric analysis of human–Artificial intelligence in public sector decision-making effectiveness. Technological Forecasting and Social Change (174), 121201. doi:10.1016/j.techfore.2021.121201.

EDEN Europol Conference Report, 2018. Freedom AND security killing the zero-sum process #kill0sum. Available from: https://www.europol.europa.eu/sites/default/files/documents/report_of_eden_conference_freedom_and_security_ 2018.pdf.

Ellingsen, I., Størksen, I., and Stephens, P., 2010. Q methodology in social work research. International journal of social research methodology, 13 (5), 395–409.

European Digital Rights and Edinburgh International Justice Initiative, 2021. The rise and rise of biometric mass surveillance in the EU report. Brussels: EDRi and EIJI. Available from: https://edri.org/wpcontent/uploads/2021/11/EDRI_RISE_REPORT.pdf [Accessed 3 April 2022].

Feldstein, S., 2019. The global expansion of AI surveillance. Carnegie Endowment for International Peace. Available from: https://www.carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847/ [Accessed 17 April 2022].

Foucault, M., 1977. Discipline and punish: the birth of the prison. New York: A Division of Random House, INC. Vintage Books.

Foucault, M., 1991. Discipline and punish: the birth of the prison. London: Penguin Books.

Fussey, P., Davies, B., and Innes, M., 2021. 'Assisted' facial recognition and the reinvention of suspicion and discretion in digital policing. The British journal of criminology, 61 (2), 325–344. doi:10.1093/bjc/azaa068.

Fussey, P. and Sandhu, A., 2022. Surveillance arbitration in the era of digital policing. Theoretical criminology, 26 (1), 3– 22. doi:10.1177/1362480620967020.

Getting the future right: artificial intelligence and fundamental rights. European Union Agency for Fundamental Rights, 2021. Available from: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2021-artificial-intelligence-summary_en.pdf.

Gurinskaya, A., 2020. Predicting citizens' support for surveillance cameras. Does police legitimacy matter? International journal of comparative and applied criminal justice, 44 (1–2), 63–83. doi:10.1080/01924036.2020.1744027.

Heaven, W., 2020. Predictive policing algorithms are racist. They need to be dismantled. MIT Technology Review. Available from: https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racistdismantledmachine-learning-bias-criminal-justice [accessed April 4, 2022].

Jahankhani, H., et al., 2020. Policing in the era of AI and smart societies. Champ: Springer.

Kaufmann, M., Egbert, S., and Leese, M., 2019. Predictive policing and the politics of patterns. The British journal of criminology, 59 (3), 674–692. doi:10.1093/bjc/azy060.

Lee, N. and Chin, C., 2022. Police surveillance and facial recognition: why data privacy is an imperative for communities of color. Brookings. Available from: https://www.brookings.edu/research/police-surveillance-and-facial-recognitionwhydata-privacy-is-an-imperative-for-communities-of-color/.

Leman-Langlois, S., 2002. The myopic panopticon: the social consequences of policing through the lens. Policing and society, 13 (1), 43–58. doi:10.1080/1043946022000005617.

Lockey, S., Gillespie, N., and Curtis, C., 2020. Trust in artificial intelligence: Australian insights. The University of Queensland and KPMG Australia. Available from: https://www.assets.kpmg/content/dam/kpmg/au/pdf/2020/public-trust-in-ai. pdf [Accessed 10 March 2022].

Lyon, D., 2002. Everyday surveillance: personal data and social classifications. Information, communication & society, 5 (2), 242–257. doi:10.1080/13691180210130806.

Mann, S. and Ferenbok, J., 2013. New media and the power politics of sousveillance in a surveillance-dominated world. Surveillance & society, 11 (1/2), 18–34.

Marx, G., 1988. Undercover: police surveillance in America. Berkley: University of California Press.

Marx, G., 2009. A tack in the shoe and taking off the shoe neutralization and counter-neutralization dynamics. Surveillance & society, 6 (3), 294–306.

McGuire, M., 2020. The laughing policebot: automation and the end of policing. Policing and society, 31 (1), 20–36.

McKeown, B. and Thomas, D., 1988. Q methodology: quantitative applications in the social sciences. London: Sage.

Metropolitan Police and NPL, 2020. Metropolitan police service live facial recognition trials. Available from: https://www.met.police.uk/SysSiteAssets/media/downloads/central/services/accessing-information/facialrecognition/metevaluation-report.pdf

.

Morgenstern, J., et al., 2021. 'AI's gonna have an impact on everything in society, so it has to have an impact on public health': a fundamental qualitative descriptive study of the implications of artificial intelligence for public health. BMC public health, 21 (1), 40. doi:10.1186/s12889-020-10030-x.

Moses, B.L. and Chan, J., 2016. Algorithmic prediction in policing: assumptions, evaluation, and accountability. Policing and society, 28 (7), 806–822. doi:10.1080/10439463.2016.1253695.

Nalla, K.M., Gorazd, M., and Maja, M., 2018. Assessing police–community relationships: is there a gap in perceptions between police officers and residents? Policing and society, 28 (3), 271–290. doi:10.1080/10439463.2016.1147564.

Neudert, L.M., Knuutila, A., and Howard, P., 2020. Global attitudes towards AI, machine learning & automated decision making. Implications for involving artificial intelligence in public service and good governance. Oxford: University of Oxford: Oxford Commission on AI & Good Governance.

Orwell, G., 2000. 1984 nineteen eighty-four. Berkley: Penguin Random House.

Oswald, M., et al., 2018. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'experimental' proportionality. Information & communications technology law, 27 (2), 223–250. doi:10.1080/13600834.2018. 1458455.

Patton, M.Q., 1990. Qualitative evaluation and research methods. 2nd ed. New Bury Park, CA: Sage.

Pavone, V. and Esposti, S., 2012. Public assessment of new surveillance-oriented security technologies: beyond the trade-off between privacy and security. Public understanding of science, 21 (5), 556–572. doi:10.1177/ 0963662510376886.

Quattrocolo, S., 2020. Equality of arms and automatedly generated evidence. In: Lorena Bachmaier Winter, ed. The artificial intelligence, computational modelling and criminal proceedings. Italy: Springer International Publishing, 73–98. Regan, P., 1995. Legislating privacy: technology, social values, and public policy. Chapel Hill: University of North Carolina Press.

Rovatsos, M., Mittelstadt, B., and Koene, A., 2019. Bias in algorithmic decision making. The University of Edinburgh: Centre for Data Ethics and Innovation. UK Government. https://www.gov.uk/government/publications/landscapesummariescommissioned-by-the-centre-for-data-ethics-and-innovation.

Schmolck, P. 2014. PQMethod Software Package [Computer Software]. Retrieved from http://schmolck.org/qmethod/downpqwin.htm

Schneier, B., 2012. How changing technology affects security. IEEE security & privacy magazine, 10 (2), 104–104. doi:10.1109/msp.2012.39.

Smith, G., Moses, B.L., and Chan, J., 2017. The challenges of doing criminology in the big data era: towards a digital and datadriven approach. The British journal of criminology, 57 (2), 259–274. doi:10.1093/bjc/azw096.

Solove, D., 2011. Nothing to hide: the false tradeoff between privacy and security chapter in NOTHING TO HIDE: THE FALSE TRADEOFF BETWEEN PRIVACY AND SECURITY. New Haven and London: Yale University Press.

Stenner, P., Dancey, C., and Watts, S., 2000. The understanding of their illness amongst people with irritable bowel syndrome: a Q methodological study. Social science and medicine, 51 (3), 439–452.

Tounsi, W. and Rais, H., 2018. A survey on technical threat intelligence in the age of sophisticated cyber attacks. Computers & security, 72, 212–233. doi:10.1016/j.cose.2017.09.001.

Trottier, D., 2017. 'Fear of contact': police surveillance through social networks. European journal of cultural and political sociology, 4 (4), 457–477. doi:10.1080/23254823.2017.1333442.

Turner, E., Medina, J., and Brown, G., 2019. Dashing hopes? The predictive accuracy of domestic abuse risk assessment by police. The British journal of criminology, 59 (5), 1013–1034. doi:10.1093/bjc/azy074.

United Nations Interregional Crime and Justice Research Institute (UNICRI) and International Criminal Organization (Interpol), 2019. Artificial intelligence and robotics for law enforcement. UNICRI and Interpol. Available from: http://www.unicri.it/artificial-intelligence-and-robotics-law-enforcement [Accessed 20 April 2022].

Urquhart, L. and Miranda, D., 2022. Policing faces: the present and future of intelligent facial surveillance. Information & communications technology law, 31 (2), 194–219. doi:10.1080/13600834.2021.1994220.

Watts, S. and Stenner, P., 2012. Doing Q methodological research: theory, method, and interpretation. London: Sage.

Westacott, E., 2010. Does surveillance make us morally better. Philosophy now, 79, 6–9.

Završnik, A., 2020. Criminal justice, artificial intelligence systems, and human rights. ERA forum, 20, 567–583. doi:10.1007/s12027-020-00602-0.

Paper 3

"Should everyone have access to AI? " Perspectives on Ownership of AI tools for Security

Yasmine Ezzeddine*, Petra Saskia Bayerl

Sheffield Hallam University, Sheffield, UK

*y.ezzeddine@shu.ac.uk; p.s.bayerl@shu.ac.uk

Abstract: Given the widespread concerns about the integration of Artificial Intelligence (AI) tools into security and law enforcement, it is natural for digital governance to strive for greater inclusivity in both practice and design (Chohan and Hu, 2020). This inclusivity can manifest in several ways, such as advocating for legal frameworks and algorithmic governance (Schuilenburg and Peeters, 2020), allowing individuals choice, and addressing unintended consequences in extensive data management (Peeters and Widlak, 2018). An under-reflected aspect is the question of ownership, i.e., who should be able to possess and deploy AI tools for law enforcement purposes. Our interview findings from 111 participants across seven countries identified five citizens viewpoints with respect to AI ownership of security-related AI: (1) Police and police-governed agencies; (2) Citizens who disassociate themselves; (3) Entities other than the police; (4) All citizens including themselves; and (5) No one or Unsure. The five clusters represent disparate perspectives on who should be responsible for AI technologies, as well as related concerns about data ownership and expertise, and thus link into broader discussions on responsibility for security, i.e., what deserves protection, how and by whom. The findings contribute theoretically to digitalization, smart technology, social inclusion, and security studies. Additionally, it seeks to influence policy by advocating for AI development that addresses citizen concerns, thereby mitigating risks, social, and ethical implications associated with

AI. Crucially, it aims to highlight citizens' concerns around the potential for malicious actors to exploit ownership of such powerful technology for harmful purposes.

Keywords: Artificial Intelligence; Ownership; Citizens; Law Enforcement Agencies; Police

1. Introduction

In the domain of security and policing, the integration of Artificial Intelligence (AI) tools presents both unprecedented opportunities and ethical considerations. At the crux of these advancements lies a pivotal question: who should hold ownership of these AI tools, and thus who owns the responsibility for security?

Expanding the definition of AI is crucial, considering recent entries defining AI as "systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals" (European Commission, 2020). With AI tools rapidly expanding across society, including in the security and policing domain, digital governments are seeking more inclusive dynamics in practice and design to ally valid citizen concerns (Chohan and Hu, 2020). This inclusivity can manifest in various ways, from calls for dedicated legal frameworks to algorithmic governance and better accounting for unintended consequences in large data management (Schuilenburg and Peeters, 2020; Peeters and Widlak, 2018).

An underexplored aspect in this regard is AI ownership. Generally, AI in the security and policing domain is conceptualised as police-owned capability. However, the ongoing privatisation and personalisation of security (for instance privately owned door cameras with AI capabilities) opens the field to a much more fluid landscape of ownership. This study explores the perceptions and preferences of citizens regarding the ownership of AI tools, particularly in policing and law enforcement contexts. Understanding public stances on this matter is crucial, as they reflect societal values on societal understandings about responsibilities for security (Bayerl et al., 2022) and can therefore contribute to shaping policies governing AI deployment and ownership.

In the realm of AI and societal implications, the discourse on ownership of security-related AI capabilities has been a focal point in academic and policy discussions. Numerous studies have explored AI deployment landscapes. However, these discussions often overlook citizen perspectives regarding ownership of security-related AI. While scholars and policymakers deliberate on governance models and ethical frameworks, citizen voices remain underrepresented in these discussions (Floridi and Taddeo, 2016).

Understanding citizen perspectives is pivotal for shaping inclusive, ethically sound, and socially acceptable AI deployment strategies for AI in security and policing. Citizens' concerns, and preferences play a fundamental role in determining legitimacy, trustworthiness, and societal acceptance of AI systems used by law enforcement and security agencies (Crawford and Calo, 2016). Thus, this article aims to fill the gap in

existing discourse by elucidating the significance of citizen perspectives in defining the preferred ownership of AI tools within security domains. By amplifying these insights, this study underscores the importance of inclusive governance frameworks that prioritize the amalgamation of citizen opinions, ensuring responsible and beneficial AI integration into society.

AI Ownership in Security and Policing Contexts

The deployment of AI tools within the security domain such as predictive policing algorithms (e.g., PredPol) and facial recognition systems (e.g., Clearview AI) highlight the complexities of AI ownership and showcase the intersection of technological innovation, legal frameworks, and societal implications. The discourse on AI ownership revolves around control, accountability, and responsibility for actions and decisions executed by these intelligent systems. Floridi (2019) argues that ownership extends beyond possession to include responsibility for AI actions, such as biases, errors, and ethical implications. This aligns with broader discussions on accountability and the necessity for transparency and oversight in AI deployment (Kroll et al., 2017).

The adoption of AI technologies such as automated license plate readers (ALPRs) and crime mapping tools, raises both concerns and opportunities. The potential of AI to augment law enforcement capabilities and optimize resource allocation is juxtaposed with apprehensions regarding privacy infringement, fears of biases, and the erosion of discretion in decision-making (Aloisi and Gramano, 2020; Mittelstadt et al, 2016). The ownership and deployment of these tools by police raises debates around balancing security with societal values and whether ownership should be exclusive to law enforcement or more distributed (Orwell, 2000).

Aim of this Study

This research aims to explore the perspectives of citizens regarding the preferred owner(s) of AI tools for policing and law enforcement applications. Through semi-structured interviews, this study seeks to elucidate public perceptions, concerns, and preferences concerning AI ownership, contributing to informing policy frameworks and ethical guidelines governing the deployment and ownership of AI tools in security domains. Understanding citizens' perspectives towards ownership of AI tools is crucial, especially given the legal and moral implications involved (Robaey, 2015; Hayes et al, 2020). By understanding the rationale behind citizens' viewpoints on access and ownership of AI policing capabilities, this study aims to contribute to the theoretical and social context in which security opportunities align with community needs and perspectives, leading to potential endorsement of a virtuous implementation of AI within policing.

Methodology

Semi-structured interviews were conducted in eight different countries (UK, Netherlands, Italy, Spain, Portugal, Czech Republic, Germany, and Greece), involving 111 participants. Participants were recruited based on specific group specifications relevant to each of the

partner countries as part of the AIDA2020[4] joint project. The interviews focused on citizens' attitudes towards AI use by law enforcement agencies (LEAs), and namely ownership. Interviews were chosen as a qualitative approach to better understand and integrate citizens' perceptions towards AI ownership, following a structured theme of scenario-based interview questions. The data collected from these interviews underwent thematic and content analysis to identify main themes and patterns. Participant responses were coded and clustered into high-order categories, allowing for the emergence of common perspectives reflecting preferences for AI ownership in different contexts. This analysis, performed using NVivo's qualitative data analysis software, enabled the exploration of citizen perspectives on AI ownership, contributing to the broader discourse on AI ownership within security contexts.

Participants

Participants were recruited by researchers in the eight participating countries. The selection allowed free choice of the citizen group to allow partner countries to choose groups that they considered relevant and of interest in their national context. A total of 111 individuals participated. Germany focused on young women (18-25 years, n=16), Czech Republic (n=10) focused on young people in general between 18-25 years old, while Italy addressed older citizens (65+ years). The Netherlands focused on expatriates (n=16) with experience in the Cybersecurity field. In the UK, 11 individuals with a migration background were recruited. Spain (n=20) and Portugal (n=6) chose participants who are familiar with AI, while Greece (n=16) decided to adopt an open selection of participants of diverse occupations and disciplines. Table 1 shows the demographic and gender distributions of the selected groups.

Country	Number of Participants	Group specifications	Gender distribution Women / Men	Average age
Czech Republic (CZ)	10	Young people between 18-25 years old	60% / 40%	23.4
Germany (DE)	16	Young women between 18-25 years old	100% / 0%	26.3
Greece (GR)	16	Experts in IT Law and IP law	78.57% / 24.4%	33.25
Italy* (IT)	16	Older citizens (65+)	50% / 50%	72.8
Netherlands (NL)	16	Expatriates	43.7% / 56.2%	26.3
Portugal (PT)	6	People with limited knowledge of AI	50% / 50%	44.7

[4] **AIDA2020: Artificial Intelligence and Advanced Data Analytics for Law Enforcement Agencies. https://www.project-aida.eu/**

Spain (SP)	20	AI experts	40% / 60%	37.5
UK	11	Citizens with migration background	72.7% / 27.2%	33.4
Total	*111*		*62.2% / 37.8%*	*38.3*

Table 1. Demographic characteristics of participants per country

Data Collection

A total of 111 semi-structured interviews were conducted addressing citizens' attitudes towards AI use by LEAs. While the first part of the interview focused on overall questions of acceptance and acceptance conditions, the second part offered participants the opportunity to reflect more specifically on potential ethical dilemmas of AI use and possible resistance. This paper focuses on one part of the interview, particularly the part where participants were asked about whether AI capabilities should be limited to police only, or would they want to have access to such tools themselves.

The interviews were conducted either face-to-face or online by researchers in each participating country to allow participants to react to questions in their own language. Therefore, all participating countries provided participants with the information sheet, the informed consent, and the interview guidelines in the language of the respective country, except for the Netherlands where the researchers chose to share the documents with the participants in English instead of Dutch since they interviewed expatriates. Follow-up questions, prompts and comments were made by interviewers in each country to encourage participants to elaborate on the rationales for their choices. All interviews were audio-recorded. Some were transcribed as summaries, others were transcribed verbatim, and all the data was anonymized before analysis.

Data Analysis

The transcripts and the summaries obtained in the native countries' languages were translated to English using a designated translation software which was followed by close proof-reading. The English transcripts/ summaries were used for the data analysis. Our analytic approach followed thematic (Auerbach and Silverstein, 2003) and content analysis principles (Krippendorff, 2004) for the purpose of identifying main themes and patterns in the data. First, the answers were coded in cycles, starting with open or initial coding (Charmaz, 2006), followed by clustering into high-order categories for each main topic. Thematic analysis was used to allow for the thorough evaluation of the statements made by each participant which revealed common perspectives whereby specific ownership was preferred by participants, depending on the situation, and was justified by different rationales. This coding was performed using NVivo's qualitative data analysis computer software package. As a second step, participant's responses were thoroughly reviewed, coded, and assessed for similarities, then clustered under common sub-themes. This process is largely exploratory, whereby the analysis does not rely on any predefined categories or features in creating the clustered perspectives.

Ethics

This study has been approved by the ethics committee of the authors' affiliated university. Additionally, participants were informed of the context and legal basis of the study, the details of data handling and their rights through the information sheet and informed consent form which the participants had to sign prior to the interview. The right to withdraw and to opt out from providing demographic information was also explicitly stated in the above-mentioned forms. All data was analysed in pseudonymized form.

Results

Analysis of Perspectives

The approach revealed five disparate perspectives towards the preferred ownership of AI capabilities for security. Figure 1 shows the percentages of respondents who favoured each form of AI owner. Below, we provide an in-depth analysis of these responses, citing participant comments in italics to clarify their decision-making process. Each perspective is summarized with a descriptive title highlighting key aspects.

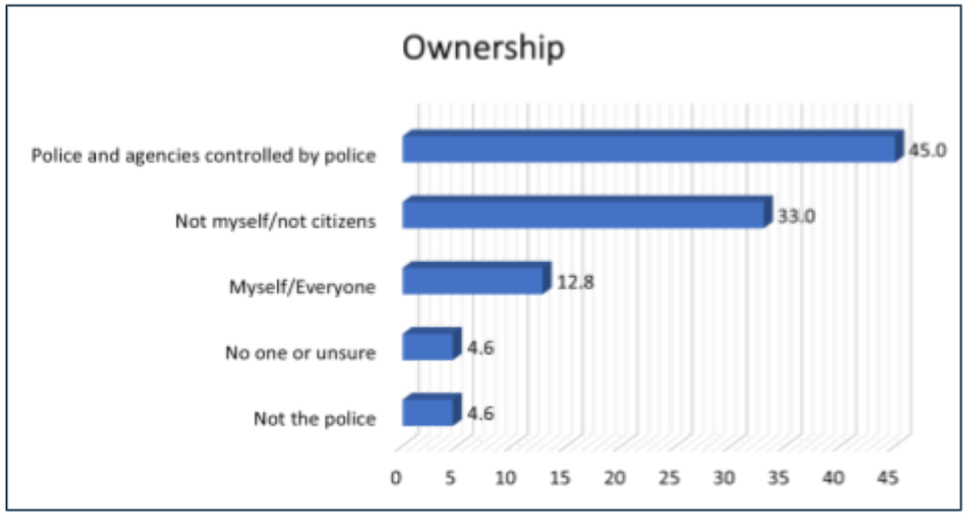


Figure.1: Clustered Perspectives on preferred owner(s) of security-related AI
(in percentages of participants in the sample)

Perspective 1: Preference for Police, LEAs, and Government Agencies

This was the most common perspective, with most participants preferring LEAs/Police to own and control AI tools. They cited public safeguarding, reducing costs, and proactively preventing crimes as key reasons. The trust in LEAs stems from their established role for safeguarding society and their existing access to confidential information, coupled with their training to handle extreme cases and the biases of AI.

Participants believe that police and LEAs have the necessary skills, competence, and ethical obligations to use AI responsibly compared to private entities. Participants specifically emphasized the role of police in protecting citizens: "This must be with the police; they must protect us" (PT-05). Concerns about private entities mishandling data surpassed worries about police use of personal data, reflecting a general acceptance of the police's role in safeguarding. They trust police forces to protect private information

and use AI for public safety. As NL-13 said, “I trust the police with this kind of information.”

Some participants suggested extending AI ownership to other government agencies, intelligence units, independent auditors, academia, and corporations with appropriate vetting and accountability measures. Centralized authorization under LEA supervision was also suggested, reflecting the belief that protecting the population is a core police mission (PT-05).

Perspective 2: Preference against Police Ownership

A smaller number of participants opposed police ownership of AI, citing the need for broader monitoring and evaluation to ensure effective use. They doubted the police's expertise and transparency in handling AI for crime prevention. GR-12 remarked, “For police I don’t know how they will use it, and I am a bit overwhelmed knowing that.” This group thus represents an opposing view to Perspective 1. The two opposing perspectives illustrate the importance of trust in police and perceptions of competence as basis for AI ownership.

Perspective 3: Preference for no ownership by citizens (including themselves)

This perspective indicates a general rejection of AI ownership by citizens, citing their lack of qualifications to handle such tools. This included participants themselves as well as others. As IT-11 stated, “It would be a waste of time if I had some computer tools to protect myself because I wouldn't be able to handle them, honestly.”

Others felt unsafe with such information, preferring police to handle AI for protection (PT04). Participants compared AI ownership to gun ownership, fearing misuse: “I don't think it should be something that goes out to just anyone because that turns out to be a bit like gun ownership” (PT-01). They were concerned about AI becoming a weapon in the wrong hands (NL-04).

Perspective 4: Preference for Everyone (including themselves)

Contrary to Perspective 3, some participants supported citizen ownership of AI tools for personal safeguarding, especially against cybercrimes. As IT-4 stated, “For cybercrime, I think that all of the people that wanted something to protect themselves from cybercrime should be allowed to have those.” However, some suggested limiting AI ownership to themselves to avoid misuse by others (GR-16). This perspective thus suggests that citizen ownership may be warranted for very specific purposes, while also expressing fear of mistrust in other citizens’ intentions.

Perspective 5: Preference for No one to own AI / unsure about preferred Owner

The final perspective argued against anyone owning AI tools due to concerns about data accuracy, algorithm transparency, and potential misuse (SP-03). Some participants expressed uncertainty about potential AI owners and preferred not to elaborate. This perspective indicates generalised concerns about the viability of AI for security, which translated into a position that no one should own such tools.

Interpretation of viewpoints

Participants from various countries shared similar reactions toward police use of AI, though there were notable differences in reasoning. Supporters of perspective 1 believed that only police should own AI tools, citing their role in public safety and trustworthiness with personal data. Some also supported AI use by other LEAs and government bodies for better oversight and objective evaluation. Conversely, participants in perspective 2 opposed police use of AI due to concerns around their capability to manage AI technologies and the transparency of AI-driven decisions. Perspective 3 revealed strong opposition to AI ownership by the public in fear of inadequate knowledge and training, linking it to the risks of public gun ownership in the U.S. In contrast, perspective 4 advocated for public access to AI tools, provided proper training and ethical guidelines are in place. These participants emphasized the need for public involvement in evaluating AI use in policing to ensure transparency. Perspective 5 included a minor group entirely opposed to AI use by anyone, preferring traditional non-AI technologies that have been effective so far.

Discussion

In our contemporary world, ownership spans both material items and intellectual properties, with laws protecting these rights to ensure owners can control and benefit from them (Hayes et al., 2020). Ownership of AI includes the rights to possess, use, manage, and benefit from these tools, along with associated responsibilities (Robaey, 2015; Honoré, 1961). The prevalence of AI tools necessitates exploring their ownership to ensure adherence to human rights and privacy laws, which could enhance societal acceptance of AI in policing (Ezzeddine et al., 2022).

This paper examines citizens' views on AI tools' ownership, discussing these perspectives in relation to debates around AI governance in policing. The findings touch on legal and ethical implications, including roles, responsibilities, expertise, accountability, and public acceptance (Carrasco et al., 2019; Neudert et al., 2020), as well as public willingness to trade privacy for safety (Pavone & Esposti, 2012).

Participants' preferred AI owner correlated with roles, responsibilities, and benefits to the public. Many expressed frustrations over data used for personalized ads without consent, while fewer were concerned about police accessing the same data. Significant concerns included AI's lack of autonomous moral operation and biases in police decision-making (Farina et al., 2020). This correlates to ongoing ethical discussions and recent research highlighting the importance of ownership models that prioritize ethical principles and societal values to foster acceptance (Nemitz, 2018).

Opinions on police ownership were divided. Some trusted police ownership due to their safeguarding roles, while others doubted police expertise and transparency. Some even suggested extending AI ownership to other government and professional entities for oversight (Martin, 2019). This debate aligns with broader discussions around transparency and need for police digitalization strategies (Gundhus et al., 2022). It also reiterates the importance of trust in LEAs and its significant impact for public support of

AI in policing. In this context, participants specifically highlighted the importance of police legitimacy and accountability in accessing personal data, reflecting expectations that police actions should prioritize national security (Tyler & Huo, 2002).

Moreover, participants' views often balanced privacy and security. Some accepted police AI use but mistrusted others, suggesting stringent regulations to prevent misuse (Jones & Haggerty, 2021). The debate on privacy versus safety remains critical, with calls for ethical considerations in AI deployment (Lyon, 2002; DiVaio et al., 2022). Some saw AI in surveillance as potentially invasive, linking it to mass data collection and bulk data analysis (Albrecht, 2020). This contrasts with historical arguments defending surveillance for improved security (Bentham, 1791) by highlighting that excessive monitoring could reduce trust in law enforcement (Yesberg et al., 2021).

Participants also recognized the significant influence of corporations on AI development and regulation. They were more critical of AI in targeted advertising by corporations than to its use by police, citing trust in law enforcement's accountability and safeguarding principles (Ezzeddine et al., 2022).

In terms of practical implications, the findings suggest options for differentiated ownership models, as well as the need for robust legal and ethics oversight, and community engagement to address preferences for AI tool ownership in law enforcement. These measures could enhance transparency, trust, and shared responsibility, aligning AI use with societal values and priorities. Below we list some concrete options to increase citizen support of AI use for security purposes:

- Differentiated Ownership Models: Exploring disparate ownership models involving law enforcement and citizen representatives (e.g., acknowledging data origins and dependencies in safety production) to enhance transparency and trust. This could include shared ownership models and public-private partnerships to ensure no single entity dominates (Crawford et al., 2019; Eubanks, 2018). Ownership should reflect data origins and AI's impact on communities, ensuring those most affected have a say in decision-making (O'Neil, 2016).

- Robust legal and Ethical Oversight: Implementing stringent ethics guidelines and independent oversight mechanisms to ensure compliance and prevent misuse. This involves developing comprehensive legal guidance and ethics standards focusing on privacy, bias mitigation, and accountability, with input from diverse stakeholders (Floridi, 2019). Alternatively, independent bodies auditing AI tools can ensure compliance with legal, professional and ethics standards through regular reports (Whittaker et al., 2018).

- Community Engagement: Promoting community-led forums for citizen input in AI tool ownership and control decisions. This includes facilitating citizen assemblies and public consultations to ensure community input is integrated into AI

governance, aligning it with public values (Zuboff, 2019). Additionally, creating community-led oversight committees to monitor AI use and advocate for necessary changes can promote education and transparency (Benjamin, 2019).

In summary, the study highlights the need for differentiated considerations on AI ownership and deployment and the importance of citizen engagement to ensure trust and accountability. Future research should explore deeper rationalizations around AI ownership preferences, focusing on roles, responsibilities, expertise, accountability, and the balance between costs and benefits.

As for limitations, our sample was skewed younger and included more women than men, suggesting the need for broader participant demographics in future research. Additionally, some participants felt unqualified to comment on AI ownership, indicating a need for inclusive discussions accounting for all perspectives. Future studies should address these issues to provide a comprehensive understanding of public perceptions and factors influencing preferences for AI ownership in policing.

Conclusion

This study critically examines citizens' perspectives on legitimate ownership of AI technologies, emphasizing the balance between rights and duties of AI implementers, particularly in policing and security. It suggests a need for public involvement in AI tool implementation, accountability, and transparency in data processing and decision-making (Vestby and Vestby, 2019). Moral responsibilities are highlighted, with citizens seen as potential owners, stressing the need for ethical governance and trust in AI applications (Lawrence et al., 2018). The research suggests active public involvement in decisionmaking to align ownership structures with societal values and ethical considerations (Pavone and Esposti, 2012). This approach enhances understanding of AI acceptance and trust, emphasizing inclusive governance and ethical frameworks (Benjamin, 2020; Ferguson, 2017).

Our findings align with ethical AI principles proposed by entities like the European Commission, addressing public concerns about accountability, legitimacy, and privacy (Ezzeddine et al., 2022). They underscore the importance of training, skills, and ethical principles for AI regulation, regardless of ownership (Albrecht, 2020). Citizens in the study exhibited diverse, instance-based perspectives on AI in policing, guided by roles, responsibilities, and a balance of costs versus benefits, rather than outright rejection (Angwin et al., 2016). This diversity indicates the need for differentiated communication and engagement with citizens about the deployment of AI capabilities for security that acknowledges multiple owners – not only of AI but also of the responsibility to secure society (Bayerl et al., 2022; Terpstra, 2009). By involving the public and ensuring transparency, LEAs can integrate AI technologies while maintaining ethical standards and public trust, reflecting citizens' inquisitive mindset towards ethically guided AI deployments (Yesberg et al., 2021).

Funding acknowledgement: **This work was supported by the European Union's**

Horizon 2020 research and innovation program under grant agreement No 883569 as part of the AIDA project (AIDA - Artificial Intelligence and advanced Data Analytics for Law Enforcement Agencies).

References

Albrecht, H.J. (2020). Data, Data Banks and Security. European Journal of Security Research, 5(1), pp. 5-23. doi:10.1007/s41125-019-00062-9.

Aloisi, A. and Gramano, E. (2020). Artificial Intelligence is Watching You at Work. Digital Surveillance, Employee Monitoring and Regulatory Issues in the EU Context. Comparative Labor Law and Policy Journal, pp. 95-121.

Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016). Machine Bias. There is software that is used across the county to predict future criminals. And it is biased against blacks. [online] Available at: https://www.propublica.org/article/machine-bias-risk-assessmentsin-criminal-sentencing [Accessed 15 February 2022].

Auerbach, C.F. and Silverstein, L.B. (2003). Qualitative Data: An Introduction to Coding and Analysis. NYU Press.

Bayerl, P.S., Butot, V., & Jacobs, G. (2022). Produktion urbaner Sicherheit aus Bürgerperspektive. In: D. Wehe, H. Siller (eds.), Handbuch Polizeimanagement, Springer Nature, pp. 1-18.

Benjamin, G. (2020). Facial recognition is spreading faster than you realize. The Conversation. [online] Available at: https://theconversation.com/facial-recognition-isspreading-faster-than-you-realise-132047 [Accessed 15 February 2022].

Benjamin, R. (2019) Race After Technology: Abolitionist Tools for the New Jim Code,

Polity. Bentham, J. (1791). Panopticon, or, The Inspection-House. Dublin: T. Payne.

Carrasco, M., Mills, S., Whybrew, A. and Jura, A. (2019). The Citizen's Perspective on the Use of AI in Government. BCG Digital Government Benchmark. [online] Available at: https://www.bcg.com/publications/2019/citizen-perspective-use-artificial-intelligencegovernment-digital-benchmarking.aspx [Accessed 15 February 2022].

Charmaz, K. (2006). Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. London: Sage Publications.

Chohan, S.R. and Hu, G. (2020). Strengthening digital inclusion in e-government: Cohesive ICT training programs intensify digital competency. Information Technology for Development, 28(1), pp. 1-23. doi:10.1080/02681102.2020.1841713.

Crawford, K. and Calo, R. (2016). There is a blind spot in AI research. Nature, 538(7625), pp. 311-313.

Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E. and Whittaker, M. (2019) AI Now 2019 Report, AI Now Institute.

Di Vaio, A., Hassan, R. and Alavoine, C. (2022). Data intelligence analytics: A bibliometric analysis of human–AI interaction in public sector decision-making effectiveness. Technological Forecasting and Social Change, 174, 121201. doi:10.1016/j.techfore.2021.121201.

Eubanks, V. (2018) Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor, St. Martin's Press.

European Commission. (2020). EU Economic and Social Committee Regulation on AI in Europe, 237 final. Brussels: European Commission. [online] Available at:

https://www.njb.nl/umbraco/uploads/2018/7/COM-2018-237-F1-EN-MAIN-PART-1.PDF [Accessed 18 February 2022].

Ezzeddine, Y., Bayerl, P.S. and Gibson, H. (2022). Citizen Perspectives on Necessary Safeguards for the Use of AI by Law Enforcement Agencies. arXiv.org. doi:10.48550/arXiv.2306.01786.

Farina, M., Zhdanov, P., Karimov, A. and Lavazza, A. (2022). AI and society: A virtue ethics approach. AI & Society. doi:10.1007/s00146-022-01545-5.

Ferguson, A.G. (2017). Policing Predictive Policing. Washington University Law Review, 94(5), pp. 1109-1189.

Floridi, L. (2019). The Logic of Information: A Theory of Philosophy as Conceptual Design. Oxford University Press.

Floridi, L. and Taddeo, M. (2016). What is data ethics? Philosophical Transactions of the Royal Society A, 374(2083), 20160360.

Gundhus, H.O., Talberg, N. and Wathne, C.T. (2022). From discretion to standardization: Digitalization in police organizations. International Journal of Police Science & Management, 24(1), pp. 27-41. doi:10.1177/14613557211036554.

Hayes, P., van de Poel, I. and Steen, M. (2020). Algorithms, values, and justice in security. AI & Society, 35, pp. 533-555. doi:10.1007/s00146-019-00932-9.

Jones, R. and Haggerty, K.D. (2021). AI policing: A white paper. Surveillance & Society, 19(3), pp. 331-337.

Krippendorff, K. (2004). Content Analysis: An Introduction to its Methodology. Thousand Oaks; London; New Delhi: Sage Publications.

Kroll, J.A., et al. (2017). Accountable Algorithms. University of Pennsylvania Law Review, 165(3), pp. 633-705.

Lawrence, D., Peterson, B. and Thompson, P. (2018). Community views on Milwaukee's police body-worn camera program. Justice Policy Center, Urban Institute.

Lyon, D. (2002). Everyday surveillance: Personal data and social classifications. Information, Communication & Society, 5(2), pp. 242-257. doi:10.1080/13691180210130806.

Martin, G. (2019). Public attitudes towards police use of facial recognition technology. Police Quarterly, 22(3), pp. 349-368.

Mittelstadt, B.D., et al. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.

Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180089. doi:10.1098/rsta.2018.0089.

Neudert, L.M., Knuutila, A. and Howard, P. (2020). Global Attitudes Towards AI, Machine Learning & Automated Decision Making: Implications for Public Service and Good Governance. University of Oxford: Oxford Commission on AI and Good Governance.

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

Orwell, G. (2000). 1984 Nineteen Eighty-Four. Introduced by Thomas Pynchon. England: Penguin Classics.

Pavone, V. and Esposti, S. (2012). Public assessment of new surveillance-oriented security technologies: Beyond the trade-off between privacy and security. Public Understanding of Science, 21(5), pp. 556-572. doi:10.1177/0963662510376886.

Peeters, R. and Widlak, A. (2018). The digital cage: Administrative exclusion through information architecture – The case of the Dutch civil registry's master data management. Government Information Quarterly, 35(2), pp. 175-183.

Robaey, Z. (2015). Looking at moral responsibility for ownership: A way to deal with hazards of GMOs. Journal of Agricultural and Environmental Ethics, 28(1), pp. 43-56. doi:10.1007/s10806-014-9517-8.

Schuilenburg, M. and Peeters, R. (2020). Algorithmic society. Algorithmic Society, pp. 1-15. doi:10.4324/9780429261404-1.

Terpstra, J. (2009a). Citizen involvement in local security networks. Security Journal, 22, 156–169.

Tyler, T.R. and Huo, Y.J. (2002). Trust in the Law: Encouraging Public Cooperation with the Police and Courts. Russell Sage Foundation.

Vestby, A. and Vestby, J. (2019). Machine Learning and the Police: Asking the Right Questions. Policing: A Journal of Policy and Practice. doi:10.1093/police/paz035.

Whittaker, M., Crawford, K., Dobbe, R. and Fried, G. (2018) AI Now 2018 Report, AI Now Institute.

Yesberg, J., Brunton-Smith, I. and Bradford, B. (2021). Police visibility, trust in police fairness, and collective efficacy: A multilevel structural equation model. European Journal of Criminology. doi:10.1177/14773708211035306.

Zuboff, S. (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, PublicAffairs.

 www.britsoccrim.org
ISSN 17759-0043; Vol 22
Postgraduate Paper

Under AI Watch: Understanding Online Behaviors under Supposed AI-Surveillance[5]

Yasmine Ezzeddine[6] and Petra Saskia Bayerl[6]

Abstract

Artificial Intelligence (AI) systems being capable of mimicking human intelligence to perform tasks have raised legitimate concerns around ethical and societal implications of their implementation. Despite the fast-paced reproduction of ethical principles to ensure safe and accountable deployment, it would be irrational to consider these sufficient. The adoption of these frameworks heavily relies on citizens' acceptance to the content and the approach of AI implementation. This study focuses on evaluating citizens' behaviours in reaction to assumed AI in online spaces, the factors that trigger rejection and potential changes in behaviour, including potential counteractions. Using an online experiment on Facebook, 30 participants were asked to perform eight tasks, accompanied by think-aloud methodology, under the assumption of AI-surveillance. The findings provide a detailed understanding of the types, reasons, and rationales for agreeing or disagreeing to conduct tasks under assumed AI-surveillance within reallife settings.

Keywords: Online surveillance; Artificial Intelligence; Law Enforcement Agencies; Resistance

Introduction

[5] For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

[6] Yasmine Ezzeddine is an experienced Researcher with a demonstrated history of working in higher education and on multi-national projects. Skilled in areas of Criminal Intelligence, Forensic Sciences, Research Management, Security, Policing and Applied Psychology. She is a PhD student researching Artificial Intelligence use in police surveillance in the UK.

[6] P. Saskia Bayerl is Professor of Digital Communications and Security at CENTRIC, Sheffield Hallam University. Her research interests lay at the intersection of human-computerinteraction, privacy, and transparency management. She holds master's degrees in psychology, linguistics and organisational dynamics from Germany and the US, and a PhD from TU Delft, Netherlands.

Surveillance involving systematic monitoring and data collection for purposes of influence or security (Lyon, 2007), are often portrayed as a double-edged sword capable of ensuring protection against crimes on the one hand and facilitating devastating attacks on the other. For instance, surveillance tools, being a source of sensitive information, could

---

cause serious damage when compromised. Recent history is marked by notable attacks targeting industrial facilities such as Trojan Black Energy in 2015 (BlackEnergy, 2021), the WannaCry ransomware attack in 2017 (Mohurle and Patil, 2017) and the Conti ransomware attack on the Costa Rican government (Datta and Acton, 2022). Traditionally, the Panopticon theory (Bentham, 1791) motivated numerous discussions around the costs vs. benefits of surveillance (Foucault, 1991; Orwell, 2000), mostly painting a sinister picture of citizens rejecting surveillance, especially when linked to Law Enforcement Agencies (LEA) (Fussey and Sandhu, 2020). This emphasizes the need to understand the psychological consequences of these technologies in security and criminological domains (Chan and Moses, 2016).

Nevertheless, there seems to be scarce research investigating resistance and counterstrategies to police surveillance and AI-use, particularly on online social platforms. This is where our keen interest in assessing complex perspectives around AI-use in police surveillance stems from, coupled with a curiosity to observe the practical implications of different attitudes in a live experience of online interaction under supposed AI monitoring. The novelty of this research lies in its aim to bridge an important gap between attitudes and behaviours exhibited when assuming AI monitoring of social platforms. The specific design of our study expands knowledge by connecting different disciplines and theoretical frameworks such as self-surveillance (Timan and Albrechtslund, 2015), and factors triggering potential resistance to online monitoring, whether by police or private entities.

In an era dominated by smart technologies that are “profoundly transforming social life, identities and relations” (Smith et al., 2017, p.259), it is crucial to investigate people’s interactions and rationales of merging their physical and virtual existences, which equally contribute to the breadcrumbs constituting their digital footprint (Laufs and Borrion, 2022). The aim is to observe the influence of AI-driven monitoring on citizens’ engagement with different content types on social media. Based on research about the influence of surveillance on behaviour (Ezzeddine et al., 2023), we seek to evaluate when citizens would draw the line for police online monitoring, triggered by which factors, if any, and for what purposes. Briefly said, we aim to answer the following research question: What triggers resistance in citizens in response to online surveillance by police compared to other entities?

Methodology

Approach

The approach consisted of an online experiment where participants were repeatedly reminded of potential AI-use while performing a series of tasks using their own personal Facebook accounts. Facebook was chosen as it is still one of the most dominantly used platforms (Snelson, 2016), making it “a potentially rich source of qualitative data for researchers” (Franz et al., 2019, p.1). We observed participants' behaviours across three

contexts: *Animals World* page, *Debate UK Politics* and *Yorkshire: Crime and Incidents*, where they were reminded of AI online monitoring by police and third parties. This was accompanied by verbalisation to collect concurrent insights into the effect of the contextual manipulations.

Participants

Individuals over 18 years old, who have a Facebook account and were willing to use it to engage with the experimental tasks, were recruited through online advertisements, LinkedIn, and flyers. Participants were also approached through direct contact (in-person or by email) based on referrals (snowball-sampling). The recruitment information contained detailed explanations about what to expect, time needed to complete the experiment (45 minutes to an hour), and the incentive participants received for their time (£20 Love2Shop voucher).

A total of 30 participants agreed to take part (13 women, 17 men) with an average age of 36 years. All participants had (at least) an undergraduate degree, 12 of them identified as members of an ethnic minority and 21 of them worked in areas related to security. Table 1 provides details on the sample.

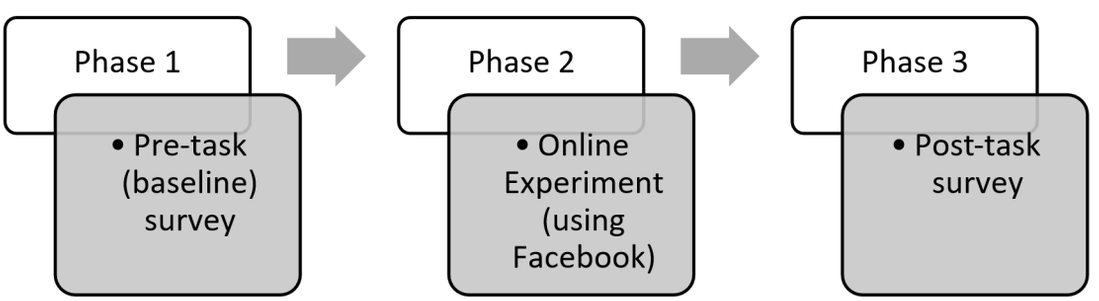
Table 1: Demographic Information of Participants

Participant s	Average Age	18-34 years	>35 years	Gender distribution Women / Men	Ethnicity minority/majority	Educational Level Univ. Degree/Master's	Occupation Securityrelated/nonsecurity related
30 participants	36.3	70%	30%	43.3 / 56.7%	40%/53.3% 6.7% prefer not to say	93.3% /6.7%	26.67%/70% 3% prefer not to say

Data collection

The study was conducted remotely in three phases (Figure.1) using MS Teams. Before the experiment, participants received an email with the Information Sheet and Consent Form, while Consent was obtained prior to scheduling the meeting. Phase 1 was a pre-task survey with ten baseline multiple-choice questions on self-rated knowledge of AI and social media activity. In addition, participants were introduced to the *think-aloud methodology* (Ericsson and Simon, 1993) using a short YouTube video, which they were then asked to practice by describing an event that happened to them recently. In Phase 2, the online experiment was conducted.

Figure 1. Phases of the Study



Participants were asked to share their screen and conduct eight tasks on each of three preselected Facebook pages, while verbalizing their thoughts (Güss, 2018). Figure.2 shows the three Facebook pages that were used in this experiment: first, "*Animals World*" for animal lovers, second, "*UK Debate Politics*" for UK politics, and lastly "*Yorkshire Crimes and Incidents"* on crimes and police updates for Yorkshire County. The rationale behind choosing these three distinct content types was to ensure diversity and to address different sharing habits (Lottridge and Bentley, 2018), as the level of user interaction in disparate online public environments can vary (Burbach et al., 2020).

Figure 2. The three Facebook pages used in the Experiment (Phase 2)

The eight tasks and their sequence in the session are shown in Table.2. They increased in difficulty, starting with joining the page, followed by inviting someone to join, reacting to a preselected post, commenting on that post, sharing it to their newsfeed, sharing it to others (via Facebook Messenger or WhatsApp), and finally, creating a post and sharing an image on the page. These tasks were chosen based on popular engagement means on Facebook and were pilot tested for complexity prior to the main study. All participants were presented with the pages and the tasks in the same order.

Table 2. List of tasks to perform on each of the Facebook pages used in Phase 2

Tasks to perform on each page	Join the page
	Invite others to join the page
	React to the post
	Comment on the post
	Share the post to newsfeed
	Share the post with others (Messenger…)
	Create a Post
	Share an Image

During Phase 2, participants were constantly reminded of AI-algorithms running in the background of Facebook to monitor online interactions and of their right to refuse performing any of the tasks. They were further reminded to verbalise rationales (Ericsson and Simon, 1993) behind their decisions when doing/refusing to do a task.

In Phase 3, participants completed a post-task survey requesting basic demographic information (i.e., age, gender, being a member of an ethnic minority/majority, securityrelated profession, and crime victimisation experience) and a ranking for the eight tasks according to perceived difficulty.

It is crucial to highlight that the monitoring was not simulated, and no algorithms were fabricated to be running in the background to collect any interactions. Instead, the participants would agree to sharing their screen and for the session to be recorded for interpretation and analysis. This approach was chosen to allow observation of participants' real-time reactions under normal conditions, to encourage a revelation of genuine and unrestrained version of their 'true selves' (O'Connor and Madge, 2017).

Ethical considerations

The study has received ethics approval from the Ethics committee at the researcher's university which was granted after providing a clear plan mitigating aspects of confidentiality, voluntary participation, anonymity of data and avoidance of any physical or psychological risks to participants. Specifically, the Information Sheet and Consent Form provided detailed information to participants about voluntary participation, use of personal accounts and right to withdraw. The material was drafted in line with the ethical guidelines set by the British Society of Criminology (2015).

Data analysis

The findings presented here are based on participants' ranking of the tasks from least difficult (1) to most difficult (8) and the verbalisation of thoughts (cp. Charmaz, 2006), which showcase the frequency of engagement and the verbalised thoughts expressed by

participants while performing the tasks. SPSS (IBM Corp. 2021) was used to cluster the data from the demographic questions and the difficulty rankings in the post-task survey. Analyses consisted of comparing ranking frequencies across tasks, investigation of engagement levels and ranking decisions for core demographic variables. These analyses used Friedman's test and Mann-Whitney-U test to accommodate for the non-parametric, non-normally distributed nature of the data (Hart, 2001). These tests can assess whether there are consistent shifts or changes in ranks across the different groups without assuming normal distribution (Conover, 1999).

The video-recorded sessions were transcribed verbatim. An in-depth qualitative analysis was conducted on the transcripts using Nvivo (QSR Int. 2020). Thematic analysis was applied to evaluate the underlying themes/patterns that emanate the think-aloud protocol (Clarke and Braun, 2013). This helped in the interpretation of subjective viewpoints through verbalised thoughts justifying participants' choices and behaviours.

This mixed data analysis approach offered a holistic opportunity for cross-validation of results though "*convergence"* or "*confirmation"* (Morgan, 1998, p.365) of findings from two distinct approaches, allowing for triangulation from monitoring of real-time behaviours, difficulty rankings, and verbalised thoughts (Güss, 2018).

Results

In this section, the combined findings from quantitative rankings and qualitative insights from participant's verbalised thoughts will be presented as direct quotes preceded by participant code (e.g., P01, indicating participant 1). A median split for age groups was used with 35 years being the cut off.

Comparing task difficulty

The Related-Samples Friedman's two-way analysis of Variance by Ranks (Table.3) revealed clear differences in difficulty rankings: '*join the group*' was overall ranked as easiest (ranked 14 times as 'least difficult'; mean rank: 2.20), followed by '*react to the post*', (ranked 'least difficult' 13 times; mean rank: 2.35). In contrast, '*share an image*' (mean rank: 7.23) and '*create a post*' (mean rank: 7.30) were deemed as 'most difficult' (cp. Table.3).

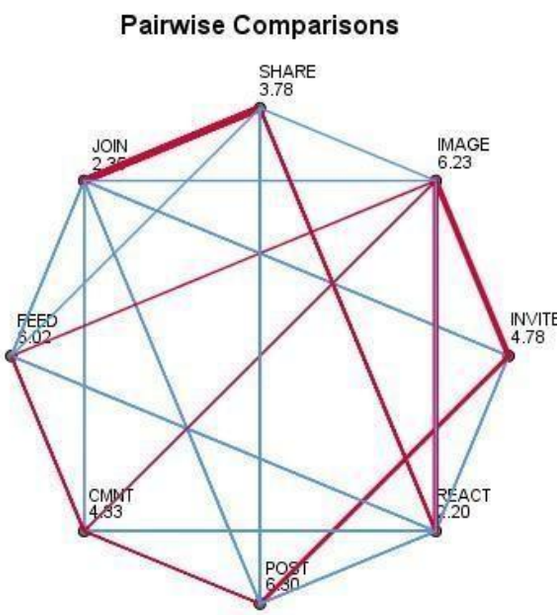
Table 3. Task difficulty as ranked by participants.

Task	Rank difficulty Least (1) to Most (8)	Number of times ranked as such by participants	% of time being ranked as such by participants	Mean Rank
Join	1	14	46.7	2.20
React	2	12	40	2.35
Share with others	2	9	30	3.78
Invite	3	8	26.7	4.78
Comment	4	12	40	4.33
Share image	5	9	30	7.23
Create post	7	10	33.3	7.30
Feed	8	9	30	6.02

A pairwise comparison test was used to reveal linkages between difficulty rankings of tasks.

The highest correlation was found between the '*join'* and '*share to friends'* tasks, followed by '*share an image'* and '*invite'* (Figure.3). This suggests that tasks are rated considerably differently, with difficulties for *'join'* and '*share with other'* ranked significantly lower compared to '*share to friends'* and '*share an image'*.

Figure.3: Pairwise Comparison of correlations between tasks rankings



These observations broadly confirm the ranking analysis in that *joining the page* was ranked as least difficult across all participants, followed by *'reacting to the post'* and

'sharing to others' via /private channels (Direct Message, WhatsApp…). '*Commenting'* and '*inviting others to join'* ranked fourth and fifth, indicating medium difficulty. Of higher difficulty emerged '*sharing an image'* (ranked 6th), while '*creating a post'* (7th) and *'sharing to newsfeed'* (8th) were ranked as the most difficult tasks.

Rationales for disparate difficulty rankings

The think-aloud data enabled an understanding of the reasons for varying levels of difficulty in performing the eight tasks. This started with reasons participants gave for ranking the '*sharing to newsfeed'* as highly difficult, which were often attributed to practical reasons rather than to security/privacy concerns. For instance, *P12: "[I] wouldn't share on my news feed, just because I live in a different area, and I do not think this will be helpful. Otherwise, I would be happy to share it"* or *P25: "I would not share the crime news because I only have 50 people on Facebook and they do not live in the UK, so I don't think it will help."*

Similarly, the lower levels of interaction with '*create a post*' or '*share an image'* tasks on the *Crime and Incidents* page were attributed to Facebook not being perceived as the proper platform to share serious cases: *"The reason why I would not create a post on this page is because I would rather go to the police directly with the information that I have"* (P17). Hence, decisions for not sharing were frequently based on usefulness considerations, triggering resistance to tasks and leading to higher difficulty rankings, e.g., "*how helpful it is to share this post since it can support the police investigation" (P14)*. This aligns with research around benefits vs. drawbacks and purposes of private information sharing online by users of social networks (Syn and Oh, 2015).

The low engagement with '*react'* and '*comment'* tasks on the *Crime and Incidents* page were further attributed to the use of emojis, or generally reacting or commenting on serious news, as being "immoral" and "unethical"; e.g., *P05: "I think it is inappropriate to react to such sad news. Like even if you react with a sad face, or write condolences, it is not going to change anything".*

Interaction levels did not change markedly despite constant reminders of AI-use. Also, no participant opposed or denied the suggestion of AI-tools monitoring Facebook or similar platforms, regardless of their perceived level of AI knowledge. Rather, participants seemed to accept that AI tools are used to monitor online environments. For instance, *P27* referred to

"AI-surveillance of passive scrolling" as a potential marketing precursor for when a person is not interacting with a post but is spending considerable time on it.

However, different content types resulted in disparate interaction patterns showing higher engagement with the *Yorkshire Crime and Incident* page, compared to the political page. Remarkably, most participants were more likely to conduct the same tasks

when related to the police-related page, than when the feed was linked to other entities. Moreover, several participants were willing to perform tasks that they would not normally engage with, based on

'having nothing to hide' from the police. *P14* argues that they are more inclined to do these tasks on a policing page, because of the oversight and safeguarding efforts that they expect from them. Participants' comments reveal some form of moral obligation to engage and share information that is quite 'serious', compared to political news or debates that can jeopardize their relationships with people of different views. Hence, they *"wouldn't want to be a part of an echo chamber" (P25).* This was coupled to a lack of trust in the admins/members of Facebook political groups, which was openly expressed by *P19: "I do not share on my feed any political posts, because you never know who the real members of that page are".*

Further, most participants were quite aware of the risks of tailored advertising where some even argued that the reason for not wanting to engage with certain posts was to avoid being

"*bombarded with similar posts and suggestions on my feed!"* (P13). This concern was often stronger than fear of police monitoring of online behaviours. It coincides with previous research on increased privacy concerns due to intrusive online marketing strategies (Dwivedi et al., 2021), which are shown to have a negative influence on online public engagement (Wang and

Herrando, 2019). Participants further expressed their concerns about what "*friends would say if they saw a kitten post shared on the newsfeed*" (*P11*). This suggests that the sharing task was deemed difficult due to social surveillance concerns, which recurred as a potent reason for refusing tasks.

Another less prominent theme was around fear of spreading misinformation/disinformation by disseminating non-trustworthy information/fake news. This suggests that the societal impact of misinformation on members of the public extends beyond influencing opinions and beliefs (Olan et al., 2022) to affect behaviours and online engagement. As *P29* states: *"I would share the crime post to support the investigation but first I would check the source of the information, if I can find a more credible source, like a government or Home office request, I will share that one".* This aligns with concerns around the lack of trust in the social platform itself. In fact, some participants even reported needing the "*government to protect us from unlawful data collection by third parties selling our data and taking advantage of fine prints on websites and social media*" *(P30)*. It may be that the assumption of trustworthiness of a post on a policerelated page contributed to the increased engagement with tasks on the *Yorkshire Crime and Incident Page*. Still, some participants preferred using an external sharing option (e.g., sharing via WhatsApp…) instead of sharing the post on their Facebook page, e.g., *"I would not share using Facebook options but take a screenshot and send it externally on other apps, or maybe show them the page" (P09).*

This coincides with tendencies to achieving a balance between sharing or hiding personal information (Pavone and Esposti, 2012) and the "complex, often ambiguous and sometimes intangible trade-offs" of posting information (Acquisti et al., 2016, p. 462).

Our findings thus align with discussions around balancing privacy rights and moral responsibilities towards public safeguarding and debates around personal information sharing vs. protecting oneself online (Ebina and Kinjo, 2021).

Overall, participants' verbalisations identified eight disparate themes, which can explain why certain tasks were considered more difficult than others:

1. **Awareness of digital footprints:** concerns around being "too visible" online.
2. **Privacy Protection:** concerns about own privacy if conducting a task.
3. **Social Surveillance and Peer Perception:** concerns around what their network and friends would think about what they post/share.
4. **Engagement depending on content types:** individualistic perspective towards acceptance vs. rejection of specific tasks based on content.
5. **Engaging in unusual actions on police-related feed:** accepting to do tasks online that they would not normally engage with
6. **Misinformation/disinformation concerns:** reluctance in sharing posts that might spread fake news.
7. **Moral obligations:** commitments to interacting with posts that might potentially lead to the arrest of a criminal for instance.
8. **Inevitability of online surveillance:** acceptance of constant online monitoring, regardless of monitoring body.

These eight disparate rationales address four broader types of concerns that impacted participants' behaviours, namely: awareness of others watching and judging their behaviours (themes 1-3), impact of the context on which behaviour occurs, including participants' trust in the organisations running the Facebook page (themes 4 and 5), concerns about potential consequences of online behaviours for others (themes 6 and 7), and feeling of unavoidability of surveillance (theme 8).

Comparison for gender differences in task fulfilment

The independent-samples Mann-Whitney U test across tasks (Figure.4) shows that '*create a post'* was perceived as more difficult by women than by men ($U=0.017$, $p<0.05$). Interestingly, women, who expressed rejection of online engagement (through lower engagement and higher difficulty rankings), were mostly concerned about privacy intrusions that bring

"unnecessary attention" to their profiles online. These concerns overlapped with longstanding discussions around online users exposing themselves to online/offline risks through private information sharing on social platforms (Gupta and Dhami, 2015). The fact that in our sample only women raised this issue correlates with suggestions of gender influences on perceptions of privacy. For instance, Rowan and Delinger (2014)

show a higher rate of women reporting concerns about collection of location-based data compared to men.

Comparison for age differences in task fulfilment

Overall, older participants (>35 years) completed more tasks per page than participants in the younger age group (35 years or younger): 61.4% compared to 47.1%. Younger participants reported being more cautious about sharing personal opinions/preferences on Facebook, because it made them "*more visible*". They preferred using Facebook "*invisibly*" instead of for self-expression. This was best put by *P17: "My purpose for using Facebook is different. I use it to keep tabs on friends and family and not to express my interests".* One participant admitted to previously sharing personal opinions when they were younger but not anymore: "*I used to do that when I was a bit younger, but now I don't like people knowing what I do or how I think.*

I don't feel the need to share my opinions, food, or holiday destinations" (*P20*).

This coincides with existing theories around younger generations' privacy preferences (Blank et al 2014). For instance, the Pew Report (2013) shows that young adults (18 to 29) are keener on limiting private information sharing online and proactively updating their privacy settings (Boyd and Hargittai, 2010). Also, research shows that older adults using Facebook/Instagram seem to rely on these platforms to compensate for the lack of social activity and face-to-face interactions in their daily lives (Sheldon, 2021). Age may thus impact how individuals behave under surveillance, as they have disparate goals for their Facebook usage.

Possible effect of security-related profession on task fulfilment

Participants with a security-related profession showed only one variation, which was a higher reluctance to engage with political content (3.3% engagement with *UK Politics* page compared to 56.3% with *Yorkshire Crime and Incident* and 40.4% with *Animals World* page). In fact, only one participant working in a security-related profession was willing to engage with the preselected political post. *P17* attributes this to fear of leaving "*political breadcrumbs on the internet that can affect my job applications to positions in the same field*".

Possible effect of crime victimization experience on task fulfilment

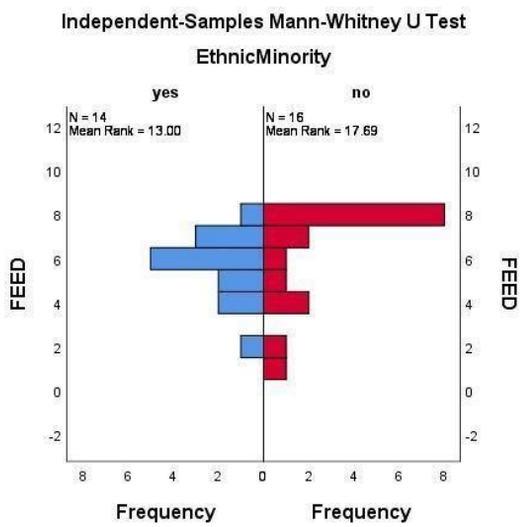
A Mann-Whitney U test results comparing participants with and without crime victimization experiences (referring to any type of crime: financial, theft, fraud, assault…) revealed that participants who identified as victims ranked the tasks of "*invite, comment, share to feed"* as more difficult compared to non-victims ($U=0.093$, $p<0.009$; $U=0.09$, $p<0.5$; $U=0.07$, $p<0.05$, respectively). Recurrent comments suggest a general reluctance amongst this group to create a post or share a picture on any of the pages, regardless of content type. This was attributed to fear of exposing themselves online and attracting "*too much attention*". Especially *'creating a post'* was deemed as a difficult task, which all participants who identified as crime victims refused. These observations coincide with suggestions that crime victimisation can lead to 'victim sensitivity' (Gollwitzer et al., 2015), fears of exploitation (Rothmund et al., 2015) and

being more reluctant towards putting themselves under the spotlight (Worsley et al., 2017).

Possible effect of ethnic minority status on task fulfilment

A Whitney-U Test revealed a single difference between participants identifying as ethnic minority members vs. ethnic majority members, namely with respect to '*sharing to newsfeed*'. Participants who did not identify as members of an ethnic minority ranked this task as more difficult (cp. Figure.4; $U=0.015$, $p<0.05$). A bivariate correlation analysis (Trauth, 2007) showed a lower frequency of interacting with political posts for people who identified as ethnic minority members (17% interaction). In addition, ethnic minority participants were more worried about '*creating a post'* or '*sharing an image'* on the *Yorkshire Crimes and Incidents* page or any police-related page/group on Facebook. This is not necessarily linked to a fear or mistrust of police. Instead, they explained their reluctance with fears of "*sharing wrong information that can lead to misinforming the police about serious cases*" *(P21).* These participants favoured using official channels to report or crime or to inform the public about serious news.

Figure.4: Variations in ranking of Share to Feed task between ethnic minority vs. majority.



Discussion

This study explores reactions to assumed online surveillance through AI, comparing three different surveillance contexts. The exploratory mixed-design nature of this study revealed the complexity of making sense of AI-use by LEAs and other entities online, with a specific focus on motivations for personal online engagement and resistance. The findings reveal complex factors that contribute to shaping citizens' perspectives and their online engagement that were largely framed under themes of inevitability of online surveillance, impact of online context and content, concerns about potential consequences of own online behaviours for others and social surveillance concerns. Some of these aspects were coupled with a sense of moral obligation to contribute to public safeguarding efforts.

Our findings expand existing knowledge on surveillance consequences by questioning longstanding notions around privacy models, fear of police monitoring, resistance and change in behaviours and revealing factors in citizens' experiences that shape their opinions, behaviours, and decision-making. This study thus constitutes an important exploration into individuals' rationales when engaging with online content under assumed AI-surveillance.

Our findings show that, in the modern era, individuals' awareness of their 'digital footprints' can lead them to perceive tasks with the most visible footprints (i.e., sharing to newsfeed, creating own posts) as 'most difficult' (Sujata et al. 2016). Yet, individuals performed more tasks on the policing-related page than on private entities' pages, suggesting that individuals may in fact feel more comfortable with police surveillance than surveillance by other entities (e.g., privacy companies). This suggests that long-defended notions of citizens fearing police surveillance (Trottier, 2017) may have changed, or may at least be more varied than often assumed. Additionally, individuals largely seemed to accept that AI-tools are used to monitor online environments, suggesting a sense of inevitability in their attitudes towards AIsurveillance.

The data further imply demographic variations that indicate that various demographic aspects may shape citizens' engagement and/or resistance to online AI-monitoring in disparate ways. Specifically, personal and demographic factors, including crime victimization or securityrelated jobs seem to shape choices for engaging/refusing to engage with certain tasks on Facebook. Our study thus illustrates the need for highly context-specific investigations to understand individual reactions to online surveillance.

Our paper also contributes to methodological innovations by enabling a deeper exploration of numerical findings with contextual insights from participants' verbalized thoughts. It demonstrates the potential for using social media platforms not only for data collection but also for real-time qualitative insights, showcasing the adaptability of mixed methods in contemporary research settings. The paper clearly outlines the integration of both quantitative and qualitative data, demonstrating transparency in methodology and analysis. This contributes to methodological rigor where similar mixed-method studies, when appropriately designed and executed, can enhance the generalizability and transferability of findings.

Our approach is especially valuable to understand actual online behaviours and reaction to assume AI-surveillance, in preference to the prevalent study of attitudes such as concerns or acceptance. The identified rationales provide an important foundation to explain decisions and online behaviours which are invaluable in understanding citizens' perspectives to AI-driven online surveillance. This demonstrates that mixed approaches, in the controlled setting of an online experiment, have proved to be ideal for investigating complex behaviours such as surveillance reactions.

Limitations and future research

Future research can benefit from exploring additional demographic groups, for instance, in terms of age and education. Our sample did not include older participants (over 64 years) nor individuals without a university degree. Including such groups may lead to additional perspectives. Moreover, this study has a restricted sample size. While the sample is substantive for the thematic analysis of the think-aloud protocol, statistical analyses are by necessity more restricted. A replication in larger samples could usefully test and validate our findings, particularly on potential group differences and impact on online context/content. Additionally, this study was conducted only with UK citizens. Extending participation beyond the UK would allow for a comparative approach to reveal whether factors such as disparate cultures, political environments and police perceptions play a role in shaping citizens' stances and reactions towards AI-use by LEAs in online surveillance.

For our study we chose an experimental setting that foregrounded conscious reflection and explanations of behaviours by participants. This introduced some artificiality and behaviours that participants would 'normally' not do, which resonates with Hawthorne's theory where participants exhibit increased performance when watched (McCambridge et al., 2014). It is, however, exactly this ability, to understand what is 'normal' versus 'nonnormal' and why, that can be considered as the main strength of our approach. It allowed us to not only observe overt patterns of behaviour, but also unearth the underlying reasons for these patterns. Further research would be valuable to explore 'unscripted' behaviours and reactions online.

References

Acquisti, A., Taylor, C., and Wagman, L. (2016). The economics of privacy, *Journal of Economic Literature*, 52(2), pp.442–492. https://doi.org/10.1257/jel.54.2.442.

Bentham, J. (1791). *Panopticon, or, The inspection-house*. Dublin, Ireland Printed: London Reprinted: T. Payne.

BlackEnergy APT Attacks: *What is BlackEnergy?* Available online: https://www.kaspersky.com/resourcecenter/threats/blackenergy. [Accessed on 22 June 2023].

Blank, G., Bolsover, G. and Dubois, E. (2014). A New Privacy Paradox: Young People and Privacy on Social Network Sites (August 13, 2014). Prepared for the Annual Meeting of the American Sociological Association, 17 August 2014, San Francisco, California., Available at SSRN: https://ssrn.com/abstract=2479938 or http://dx.doi.org/10.2139/ssrn.2479938 Boyd, D. and Hargittai, E. (2010). Facebook privacy settings: Who cares?, *First Monday*, *15*(8). doi:https://doi.org/10.5210/fm.v15i8.3086.

British Society of Criminology. (2015). 'Statement of Ethics', *British Society of Criminology*. https://www.britsoccrim.org/ethics/

Burbach, L., Halbach, P., Ziefle, M. and Valdez, A. (2021). 'Questions of Scientific Research on Violence and Inequality Applied to Women', *Journal For Educators, Teachers And Trainers*, 12(01). doi:10.47750/jett.2021.12.01.018.

Chan, J. and Moses, B. L. (2016). 'Making Sense of Big Data for Security', *British Journal of Criminology*, 75, pp.299-317. https://doi.org/10.1093/bjc/azw059.

Charmaz, K. (2006). Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. Thousand Oaks: Sage.

Clarke, V. and Braun, V. (2013). 'Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning', *The Psychologist*, *26*(2), pp.120-123.

Conover, W.J. (1999). Practical Nonparametric Statistical. 3rd Edition, John Wiley and Sons Inc., New York, pp.428-433.

Datta, P.M., and Acton, T. (2022). 'Ransomware and Costa Rica's national emergency: A defense framework and teaching case'. *Journal of Information Technology Teaching Cases*, *0*(0). https://doi.org/10.1177/20438869221149042

Dwivedi, Y.K., Ismagilova, E., Hughes, D.L., and Carlson, J. (2021). 'Setting the future of digital and social media marketing research: Perspectives and research propositions', *International Journal of Information Management*, *59*(1), 1–37. Sciencedirect. https://doi.org/10.1016/j.ijinfomgt.2020.102168

Ebina, T. and Kinjo, K. (2021). 'Paradox of choice and sharing personal information', *AI and Society* 38(1), pp.121 – 132. https://doi.org/10.1007/s00146-021-01291-0

Ericsson, K. A. and Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.

Ezzeddine, Y., Bayerl, P.S. and Gibson, H. (2023). 'Safety, privacy, or both: evaluating citizens' perspectives around artificial intelligence use by police forces', *Policing and Society*, 33(7), pp.861-876. doi: 10.1080/10439463.2023.2211813

Foucault, M. (1991). *Discipline and Punish: The Birth of the Prison*. Penguin Random House UK.

Franz, D., Marsh, H.E., Chen, J.I., and Teo, A.R. (2019). 'Using Facebook for Qualitative Research: A Brief Primer', *Journal of Medical Internet Research*, *21*(8), e13544. https://doi.org/10.2196/13544

Fussey, P. and Sandhu, A. (2022). 'Surveillance arbitration in the era of digital policing'. *Theoretical Criminology*, *26*(1), pp. 3–22. https://doi.org/10.1177/1362480620967020

Gollwitzer, M., Süssenbach, P., and Hannuschke, M. (2015). 'Victimization experiences and the stabilization of victim sensitivity', *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00439

Gupta, A. and Dhami, A. (2015). 'Measuring the impact of security, trust and privacy in information sharing: A study on social networking sites', *Journal of Direct Data Digital Marketing Practices*, **17**, 43–53. https://doi.org/10.1057/dddmp.2015.32

Güss, C. D. (2018). 'What Is Going Through Your Mind? Thinking Aloud as a Method in Cross-Cultural Psychology', *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01292

Hart, A. (2001). 'Mann-Whitney test is not just a test of medians: differences in spread can be important', *BMJ*, *323*(7309), pp.391–393. https://doi.org/10.1136/bmj.323.7309.391

IBM Corp. (2017). IBM SPSS Statistics for Windows, Armonk, NY: IBM Corp. Available at: https://hadoop.apache.org.

Laufs, J. and Borrion, H. (2022). 'Technological innovation in policing and crime prevention: Practitioner perspectives from London'. *International Journal of Police Science and Management*, *24*(2), pp.190209. https://doi.org/10.1177/14613557211064053

Lottridge, D. and Bentley, F. R. (2018). "Let's hate together: how people share news in messaging, social and public networks," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18* (New York, NY: ACM), pp.60:1–60:13. doi:

10.1145/3173574.3173634 Lyon, D. (2007). Surveillance Studies: An Overview. Cambridge: Polity Press.

McCambridge, J., Witton, J., and Elbourne, D. R. (2014). 'Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects', *Journal of clinical epidemiology*, *67*(3), pp.267–277. https://doi.org/10.1016/j.jclinepi.2013.08.015

Mohurle S. and Patil, M.A. (2017), 'Brief study of wannacry threat: Ransomware attack 2017', *International Journal of Advanced Research in Computer Science,* 8, pp.1938–1940.

Morgan, D.L. (1998). 'Practical Strategies for Combining Qualitative and Quantitative Methods: Applications to Health Research', *Qualitative Health Research*, *8*(3), pp.362–376. https://doi.org/10.1177/104973239800800307

O'Connor, H. and Madge, C. (2017). 'Online interviewing'. In: Fielding N, Lee R, Blank G (eds) *The SAGE Handbook of Online Research Methods*. London: SAGE Publications, pp.416–434.

Olan, F., Jayawickrama, U., Arakpogun, E.O. *et al.* (2022). 'Fake news on social media: the Impact on Society', *Information System Frontiers.* https://doi.org/10.1007/s10796-022-10242z

Orwell, G. (2000). *1984 Nineteen Eighty-Four.* Introduced by Thomas Pynchon. England, Penguin Classics.

Pavone, V. and Esposti, S. (2012). 'Public assessment of new surveillance-oriented security technologies:
Beyond the trade-off between privacy and security', *Public Understanding of Science, 21(5)*: pp.556– 572. https://doi.org/10.1177/0963662510376886

QSR International Pty Ltd. (2020) Nvivo (released in March 2020), https://www.qsrinternational.com/nvivoqualitative-data-analysis-software/home

Rothmund, T., Gollwitzer, M., Bender, J. and Klimmt, C. (2015). 'Short- and long-term effects of virtual violence on cooperation and social trust', *Media Psychology,* 18, pp.106–133. doi:10.1080/15213269.2013.841526

Rowan, M. and Dehlinger, J. (2014). 'Observed Gender Differences in Privacy Concerns and Behaviors of Mobile Device End Users', *Procedia Computer Science*, 37, pp.340–347. https://doi.org/10.1016/j.procs.2014.08.050.

Sheldon, P., Antony, M.G., and Ware, L. J. (2021). 'Baby Boomers' use of Facebook and Instagram: uses and gratifications theory and contextual age indicators', *Heliyon*, *7*(4), e06670. https://doi.org/10.1016/j.heliyon.2021.e06670

Smith, G.J.D., Moses, B. L., and Chan, J. (2017). 'The Challenges of Doing Criminology in the Big Data Era: Towards a Digital and Data-driven Approach', *The British Journal of Criminology*, *57*(2), pp.259–274. https://doi.org/10.1093/bjc/azw096.

Snelson, C.L. (2016). 'Qualitative and mixed methods social media research', *International Journal of Qualitative Methods, 15*(1). doi: 10.1177/1609406915624574.

Sujata, J., Saxena, S., Tanvo, G. and Shreya. (2016). 'Developing Smart Cities: An Integrated Framework', *Procedia Computer Science*, *93*, pp.902–909. Sciencedirect. https://doi.org/10.1016/j.procs.2016.07.258

Syn, S.Y. and Oh, S. (2015). 'Why do social network site users share information on Facebook and Twitter?', *Journal of Information Science*, *41*(5), 553–569. https://doi.org/10.1177/0165551515585717

Timan, T. and Albrechtslund, A. (2015). 'Surveillance, Self and Smartphones: Tracking Practices in the Nightlife', *Science and Engineering Ethics*, *24*(3), pp.853–870. https://doi.org/10.1007/s11948-0159691-8

Trauth, M.H. (2007). *Bivariate Statistics*. In: *MATLAB® Recipes for Earth Sciences*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72749-1_4

Trottier, D. (2017). '"Fear of contact": Police surveillance through social networks', *European Journal of*

Cultural and Political Sociology, 4(4), pp.457–477. https://doi.org/10.1080/23254823.2017.1333442

Wang, Y. and Herrando, C. (2019). 'Does privacy assurance on social commerce sites matter to millennials?', *International Journal of Information Management, 44(2),* pp.164–177. Worsley, J.D., Wheatcroft, J.M., Short, E., and Corcoran, R. (2017). 'Victims' Voices: Understanding the Emotional Impact of Cyberstalking and Individuals' Coping Responses', *SAGE Open*, *7*(2). https://doi.org/10.1177/2158244017710292

Paper 5

Unveiling Public Sentiments Towards AI-Driven Urban Surveillance: A Case Study from Sheffield

Yasmine Ezzeddine*; Petra Saskia Bayerl – CENTRIC – Sheffield Hallam University

Joan Rodriguez-Amat - University of Sheffield

Introduction

The integration of AI technologies into offline surveillance systems has become a focal point in contemporary discussions about public safety, privacy, and ethics. As cities worldwide increasingly adopt smart technologies, AI-driven surveillance has emerged as a powerful tool for law enforcement agencies. These technologies offer the potential to enhance public safety by predicting and preventing criminal activities, monitoring public spaces in real-time, and improving response times to incidents (Zuboff, 2019; Fussey & Murray, 2020). However, the deployment of AI in urban settings also raises critical concerns about the erosion of privacy, the potential for biased decisionmaking, and the implications for civil liberties (Crawford & Calo, 2016; Amoore, 2020). The balance between leveraging AI for public safety and safeguarding individual rights has become a key issue for policymakers, law enforcement, and the public alike (Smith & Miller, 2022).

In the context of smart cities, where AI technologies are deeply embedded in the infrastructure, understanding public perceptions and interactions with these systems is crucial (Kitchin, 2021). The effectiveness and legitimacy of AI-driven surveillance largely depend on public acceptance and trust (Webster, 2019). If the public perceives these technologies as intrusive or biased, it could lead to a loss of confidence in law enforcement and a broader societal backlash against AI adoption (Zarsky, 2016). Conversely, if AI surveillance is seen as a fair and effective tool for enhancing security, it could bolster public trust and support for law enforcement initiatives (Andrejevic, 2020). This paper aims to contribute to the discourse on AI and surveillance in policing by exploring how individuals perceive AI-driven surveillance in the context of a smart city, with a specific focus on Sheffield. By examining public attitudes, this study seeks to inform the development of policies and practices that balance technological innovation with ethical considerations (Urquhart & Miranda, 2021).

AI in Policing and Law Enforcement

Recent developments in AI have introduced numerous opportunities for enhancing public service efficiency and effectiveness, particularly in policing and law enforcement (Stahl, 2021). AI applications in law enforcement can be utilized before, during, and after a crime, enabling predictive policing through the analysis of historical crime data to forecast criminal behavior and prevent crimes (van Brakel, 2021; Brayne, 2021). Moreover, surveillance cameras equipped with image recognition technologies can identify and respond to crimes in real-time, recognizing violent situations, dangerous objects, suspicious vehicles, and individuals (Mau, 2023; Ferguson, 2017). These advancements, while offering significant potential for crime prevention and public safety, also raise important ethical and legal questions about the implications and consequences of AI in policing, particularly concerning privacy, accountability, and bias (Ljungberg, 2022; Eneman & Jansson, 2022). This workshop seeks to address these critical issues by fostering a multidisciplinary dialogue on the challenges and opportunities that AI presents in law enforcement contexts.

Methodology, Aspired Results and Case Study

Through a carefully designed combination of qualitative and quantitative research methods, this study aims to systematically map public perceptions of AI-driven urban surveillance. The study's methodology is specifically structured to delve deeply into how AI technologies are perceived by the public, particularly within the context of a smart city like Sheffield. To achieve this, the study recruited 30 Sheffield residents to participate in a three-tiered research process, beginning with privacy walks. During these walks, participants were tasked with identifying and photographing what they perceive to be AI-embedded surveillance tools in their urban environment. This initial phase serves as a crucial step in understanding the physical and psychological markers that citizens associate with AI surveillance, providing a tangible basis for further discussion and analysis.

Following the privacy walks, participants completed surveys designed to capture their immediate reactions, concerns, and expectations regarding the AI technologies they identified. These surveys are crucial for gathering quantifiable data on public sentiment, which will be further explored in the final phase of the study: focus groups. In these focus groups, participants reflect on their experiences during the privacy walks and discuss their broader views on AI surveillance. This qualitative data will offer rich, nuanced insights into how people construct and interpret the concept of privacy in the age of AI, and how they perceive the risks associated with these technologies.

The methodological purpose of this study is twofold: first, to generate a detailed, empirical understanding of public perceptions of AI-driven surveillance in an urban setting; and second, to explore the broader implications of these perceptions for the integration of AI into policing practices. By triangulating data from privacy walks, surveys, and focus groups, the study aspires to produce a comprehensive analysis of how AI surveillance is experienced by citizens, and how these experiences shape their trust in law enforcement and their views on privacy.

The outcome of this whole approach can serve as a large case study offering valuable perspectives on the organizational consequences of AI in policing, the construction and management of risks, and the implications for privacy and democratic rights (Amoore, 2014; Ball & Webster, 2019).

Workshop Relevance and Emerging Questions

This study aligns closely with the themes and questions posed by the International Workshop on AI and Surveillance in Policing and Law and Order, particularly in relation to the aspects of ‘threats’ and ‘perspectives.’ Specifically, the methodology and aspired findings of this research can help address two key questions highlighted by the workshop: the construction and meaning of privacy in relation to AI in policing, and the construction and management of risks in relation to the AI Act. These questions are crucial for understanding how AI technologies are reshaping the landscape of policing and the broader implications for both public safety and individual rights (Wright, 2022; Fussey & Sandhu, 2022).

The study employs a mixed-methods approach, integrating qualitative and quantitative research to explore these issues in depth. By engaging 30 Sheffield residents in a three-tiered methodological process—including privacy walks, pre- and post-surveys, and focus groups—the research aims to gather comprehensive data on public perceptions of AI-driven urban surveillance (Webster, 2019). The qualitative components, such as interviews and focus groups, offer a nuanced understanding of how individuals conceptualize privacy within the context of AI surveillance, helping to uncover the underlying concerns and values that shape public attitudes toward AI in policing (Webster et al., 2022). The quantitative data, derived

from surveys, will be instrumental in mapping public concerns about the risks associated with AI surveillance, such as algorithmic biases, data misuse, and potential overreach by the state (Ljungberg, 2022).

By analyzing this data, the study will provide crucial insights into how the public perceives the risks and privacy implications of AI in policing and how these perceptions align with or diverge from the regulatory frameworks proposed by the AI Act (Edwards, 2022). These findings are directly related to the workshop's key questions and will be vital for informing the development of policies that promote accountability, legitimacy, and public confidence in AI-driven surveillance. In doing so, the research will contribute not only to the theoretical discussions at the workshop but also offer practical recommendations for policymakers and law enforcement agencies on navigating the ethical and regulatory challenges posed by AI surveillance (Urquhart & Miranda, 2021; Eneman et al., 2022).

Contributions to the Workshop

This research will make a significant contribution to the workshop by offering both theoretical insights and empirical evidence on AI-enabled surveillance in policing contexts. The study's case study on Sheffield will serve as a concrete example of how AI technologies are being integrated into urban environments, providing valuable perspectives on the real-world implications for public safety, privacy, and community trust. By focusing on public perceptions and experiences, the research will illuminate the societal impact of AI in policing, which is essential for shaping informed discussions on both the opportunities and threats associated with these technologies.

In addressing the workshop's themes of "threats," "perspectives," and "opportunities," the study will explore the ethical and regulatory challenges posed by AI surveillance, particularly in light of the evolving legal frameworks such as the AI Act. These insights will be critical for understanding how AI can be utilized in law enforcement while managing the associated risks, such as potential infringements on privacy and the exacerbation of biases in policing practices (Webster & Ball, 2019). The research findings will contribute to broader discussions on how AI can be harnessed responsibly in policing, ensuring that technological advancements are balanced with the need for accountability, legitimacy, and the protection of democratic rights (Eneman & Jansson, 2022).

By linking the findings to the workshop's key themes, this paper will help foster critical discussions and reflections on the future of AI in policing. It aims to contribute to the development of policies and practices that maximize the opportunities presented by AI technologies while carefully navigating the ethical and regulatory challenges they pose. This contribution is expected to resonate with the workshop's goal of generating and sharing new knowledge on the implications and consequences of AI in law enforcement, ultimately supporting the creation of a more equitable, transparent, and trusted policing landscape.

References

Amoore, L. (2020). Cloud Ethics: Algorithms and the Attributes of Ourselves and Others. Duke University Press.

Andrejevic, M. (2020). Automated Media. Routledge.

Ball, K., & Webster, C. (2019). The Surveillance Assemblage. Routledge.

Brayne, S. (2021). Predict and Surveil: Data, Discretion, and the Future of Policing. Oxford University Press.

Connon, I., McGarry, A., & McAuliffe, E. (2022). The Surveillance State: Understanding AI, Big Data, and Privacy. Routledge.

Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. Nature, 538(7625), 311-313.

Edwards, L. (2022). Regulating AI: The European AI Act and its Implications. Cambridge University Press.

Eneman, M., & Jansson, A. (2022). Digital surveillance in the public sector: Ethical implications and societal impacts. Surveillance & Society, 20(3), 312-326.

Ferguson, A. G. (2017). The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement. NYU Press.

Fussey, P., & Murray, D. (2020). Independent report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology. Human Rights and Big Data Project.

Kitchin, R. (2021). Data Lives: How Data Are Made and Shape Our World. MIT Press.

Ljungberg, J. C. (2022). Public perceptions of AI surveillance: Ethical dilemmas and regulatory challenges. Information & Communications Technology Law, 31(3), 189205.

Mau, S. (2023). AI and the Future of Law Enforcement: Technological, Legal, and Ethical Challenges. Journal of Law, Information & Technology, 29(2), 203-220.

Murray, A. (2021). Legal uncertainties in AI and the challenges for law enforcement. Journal of Law and Technology, 17(2), 155-174.

Smith, M., & Miller, R. (2022). AI and Public Safety: Balancing Benefits and Risks. Journal of Ethics and Technology, 14(1), 76-89.

Stahl, B. C. (2021). Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies. Springer.

Urquhart, L., & Miranda, D. (2021). The AI Act and its implications for AI-driven surveillance: A socio-legal perspective. Computer Law & Security Review, 41, 105556.

van Brakel, R. (2021). Predictive Policing: Data and Algorithms in Law Enforcement. Routledge.

Webster, C. (2019). The surveillance assemblage and public trust: An analysis of contemporary issues. Surveillance Studies, 11(2), 123-139.

Webster, C., Ball, K., & Webster, C. (2022). The future of AI in policing: Perspectives on risks, privacy, and ethics. Journal of Surveillance & Society, 20(4), 415-431.

Wright, D. (2022). The AI Act: Regulating AI and Surveillance in Policing. International Journal of Law and Information Technology, 30(1), 45-62.

Zarsky, T. (2016). The trouble with algorithms: An international perspective on AI and data privacy. Harvard Journal of Law & Technology, 29(1), 124-145.

Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.

II. References

Ada Lovelace Institute. (2019). Algorithmic accountability: Safeguarding democracy and human rights in the age of AI. Ada Lovelace Institute. https://www.opengovpartnership.org/wpcontent/uploads/2021/08/executive-summary-algorithmic-accountability.pdf

Agarwal, N., Xu, K., Zeng, H., & Larson, K. (2020). Algorithmic transparency in decision making systems: Challenges, open questions, and path forward. *Journal of Artificial Intelligence Research*, 67, 993–1020. https://doi.org/10.1613/jair.1.12175

Agarwal, R., Joshi, K., & Gupta, S. (2023). Advances in AI for modern policing: Generative AI and real-time decision-making. International Journal of Artificial Intelligence Research, 47(2), 55-72. https://doi.org/10.1016/j.ijair.2023.102984

Akhgar, B., Awford, J., MacFeely, S., Staniforth, A., Franco, C., Hughes, M., & Horsman, G. (2022). Accountability principles for artificial intelligence (AP4AI): Towards responsible AI use in law enforcement and security. Sheffield Hallam University. https://shura.shu.ac.uk/31123/9/AkhgarAP4AI_Accountability_Principles%28VoR%29.pdf

Andrejevic, M. (2007). iSpy: Surveillance and power in the interactive era. University Press of Kansas.

Auerbach, C. F., & Silverstein, L. B. (2003). Qualitative data: An introduction to coding and analysis. New York University Press.

Babuta, A., & Oswald, M. (2020). Data analytics and algorithmic bias in policing. *RUSI Journal, 165*(5-6), 12-24. https://doi.org/10.1080/03071847.2020.1850728

Ball, K. (2002). Elements of surveillance: A new framework and future directions. *Information, Communication & Society, 5*(4), 537-560.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Benjamin, R. (2020). The costs of AI in policing: Racial inequality in predictive policing and face recognition technologies. AI Now Institute.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Brundage, M., Avin, S., Clark, J., Eckersley, P., Garfinkel, B., & Toner, H., et al. (2020). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. *arXiv preprint arXiv:2004.07213*. https://doi.org/10.48550/arXiv.2004.07213

Brunton, F., & Nissenbaum, H. (2015). *Obfuscation: A user's guide for privacy and protest.* MIT Press.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 1-15.

Campion, J., Grégoire, C., & Walker, B. (2020). Beyond compliance: A framework for designing responsible AI in law enforcement. *AI & Society*, 35(1), 1-15. https://doi.org/10.1007/s00146-019-00906-5

Ceyhan, A. (2012). Surveillance as biopower. In K. Ball, K. D. Haggerty, & D. Lyon (Eds.), *Routledge Handbook of Surveillance Studies* (pp. 38–45). Routledge.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. SAGE Publications.

Chen, H., Schroeder, J., Hauck, H. V., & Heuermann, A. (2015). Digital counterstrategies: Internet-based resistance tactics to online surveillance. *International Journal of Cyber Criminology, 9*(2), 97-110.

Chohan, S. R., & Hu, G. (2020). Strengthening digital inclusion in egovernment: Cohesive ICT training programs intensify digital competency. *Information Technology for Development, 28*(1), 1-23. https://doi.org/10.1080/02681102.2020.1841713

Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology, 12*(3), 297-298. https://doi.org/10.1080/17439760.2016.1262613

Constantinescu, M., Voinea, C., Uszkai, R., & Vică, C. (2021). Understanding responsibility in responsible AI: Dianoetic virtues and the hard problem of context. *Ethics and Information Technology, 23*(4), 803-817. https://doi.org/10.1007/s10676-021-09616-9

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications.

Dattatray Deulkar, D., & Gupta, P. (2020). A study on usage of online personal information by data brokers. *Journal of Data and Information Science, 5*(2), 80-100.

Davis, A. (2021). Privacy and predictive policing: Examining the impacts of AI on law enforcement transparency and accountability. *Ethics in Information Technology, 23*(2), 127-139.

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56–62. https://doi.org/10.1145/2844110

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4), 211–407. https://doi.org/10.1561/0400000042

EDRi. (2021). Reclaim your face: Europe-wide campaign against biometric mass surveillance. European Digital Rights. https://edri.org

EHRC: Equality and Human Rights Commission. (2020). Facial recognition technology and law enforcement. https://www.equalityhumanrights.com/en

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. The MIT Press.

Ezzeddine, Y., & Bayerl, P. S. (2023a). Under AI watch: Understanding online behaviours under supposed AI-surveillance. *British Society of Criminology 2023 Conference Proceedings, 22*. https://www.britsoccrim.org/wp-content/uploads/2024/01/BSC-OnlineJournal-2023.pdf

Ezzeddine, Y., & Bayerl, P. S. (2024). Should everyone have access to AI? Perspectives on ownership of AI tools for security. In *International Conference on AI Research ICAIR24 Conference Proceedings*. Expected publication: December 2024.

Ezzeddine, Y., Bayerl, P. S., & Gibson, H. (2022). Citizen perspectives on necessary safeguards for the use of AI by law enforcement agencies. arXiv.org. https://doi.org/10.48550/arXiv.2306.01786

Ezzeddine, Y., Bayerl, P. S., & Gibson, H. (2023b). Safety, privacy, or both: Evaluating citizens' perspectives around artificial intelligence use by police forces. *Policing and Society, 33*(7), 861-876. https://doi.org/10.1080/10439463.2023.2211813

Ezzeddine, Y., Bayerl, P. S., & Rodriguez, J. A. (2025). Unveiling public sentiments towards AI-driven urban surveillance: A case study from Sheffield. In *International Workshop on AI and Surveillance in Policing and Law and Order: Opportunities, Threats, Perspectives, and Cases*. Routledge: Studies in Surveillance book series. Expected publication: April 2025.

Ferguson, A. G. (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement*. NYU Press.

Floridi, L. (2019). *The logic of information: A theory of philosophy as conceptual design*. Oxford University Press.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., & Vayena, E. (2018). AI4People—An ethical framework for a good AI

society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Foucault, M. (1977). *Discipline and punish: The birth of the prison*. Pantheon Books.

Gates, K. A. (1996). *Electronics and surveillance in the workplace: Dilemmas and solutions.* Cornell University Press.

Gilliom, J. (2001). *Overseers of the poor: Surveillance, resistance, and the limits of privacy*. University of Chicago Press.

Gunning, D., & Aha, D. W. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 40(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Gunning, D., & Aha, D. W. (2023). Explainable artificial intelligence (XAI): Trends, challenges, and future directions. Artificial Intelligence Journal, 306, 1-15. https://doi.org/10.1016/j.artint.2023.103591

Haggerty, K. D., & Samatas, M. (Eds.). (2010). *Surveillance and democracy.* Routledge.

Harvey, D. (2020). *The anti-capitalist chronicles.* Pluto Press.

IBM Research. (2023). Fairness-aware machine learning models: A guide to building equitable AI systems. Retrieved from https://www.ibm.com/research

Isaac, J. (2021). Coding, braiding, and transmissions: Resistance to surveillance in hairstyling practices. *Journal of Cultural and Media Studies, 15*(2), 134-156. https://doi.org/10.1080/17495724.2021.456789

Jacobs, G. (2024a). Communities, technology, and surveillance: The impact of digital tools on citizen engagement. *Surveillance & Society, 22*(2), 1-16. https://www.erudit.org/en/journals/survsoc/2024-v22-n2survsoc09422/1112225ar.pdf

Jacobs, G. (2024b). Technology and citizen engagement in the context of AI governance. Paper presented at the *EASST 4S 2024 Conference*. https://nomadit.co.uk/conference/easst-4s2024/paper/85055

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Jones, M. (2021). Community responses to police drones: Privacy implications and community trust. *Technology and Society Magazine, IEEE, 40*(2), 74-80. https://doi.org/10.1109/MTS.2021.3074623

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). SAGE Publications.

Lee, M., & Lee, J. (2019). Understanding public perceptions of biometric technology: A survey study. *Security Journal, 32*(4), 342-358. https://doi.org/10.1057/s41284-019-00192-w

Lyon, D. (2007). *Surveillance studies: An overview*. Polity Press.

Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society, 1*(2). https://doi.org/10.1177/2053951714541861

Macnish, K. (2021). *Surveillance ethics: An introduction*. Emerald Publishing Limited.

Magalhães, J. C., & Sánchez, M. (2021). Surveillance and privacy in the pandemic: How increased surveillance is affecting freedom and autonomy.

Journal of Digital Culture & Society, 7(1), 101–120. https://doi.org/10.14361/dcs-2021-0106

Mann, S., & Ferenbok, J. (2013). New media and the power politics of sousveillance in a surveillance-dominated world. *Surveillance & Society, 11*(2), 18-34. https://doi.org/10.24908/ss.v11i1/2.4456

Mann, S., Nolan, J., & Wellman, B. (2003). Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. *Surveillance & Society, 1*(3), 331-355. https://doi.org/10.24908/ss.v1i3.3347

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2023). Survey on bias and fairness in machine learning. ACM Computing Surveys, 55(3), 1-35. https://doi.org/10.1145/3531142

Monahan, T. (2006). Counter-surveillance as political intervention. *Social Semiotics, 16*(4), 515-534. https://doi.org/10.1080/10350330601019766

Monahan, T. (2011). Surveillance as cultural practice. *The Sociological Quarterly, 52*(4), 495-508. https://doi.org/10.1111/j.1533-8525.2011.01215.x

NIST Face Recognition Vendor Test (FRVT). (2022). Face recognition accuracy under demographic variation. National Institute of Standards and Technology (NIST). Retrieved from https://www.nist.gov

Norris, C., & Armstrong, G. (1999). *The maximum surveillance society: The rise of CCTV.* Berg.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1), 1-13. https://doi.org/10.1177/1609406917733847

OpenAI Research. (2023). Transformer-based models for law enforcement applications: Opportunities and challenges. OpenAI Technical Reports. Retrieved from https://www.openai.com

Park, K., & Yoon, H. Y. (2024). Beyond the code: The impact of AI algorithm transparency signaling on user trust and relational satisfaction. *Public Relations Review, 50*(5), 102507. https://doi.org/10.1016/j.pubrev.2024.102507

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information.* Harvard University Press.

Petersen, S. M., & Taylor, C. (2012). Predictive policing and surveillance. *Surveillance & Society, 10*(2), 138-156. https://doi.org/10.24908/ss.v10i2.4435

Pierson, C. A. (2019). Data ownership and data sharing practices. In D. Poff & A. Michalos (Eds.), *Encyclopedia of business and professional ethics* (pp. 1-5). Springer. https://doi.org/10.1007/978-3-319-23514-1_315-1

RAND Corporation. (2020). AI and civil society: Empowering organizations with AI tools. https://www.rand.org/research

Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review Online, 94*, 15-55.

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2022). Assessing behavioral data science privacy issues in government artificial intelligence deployment. *Government Information Quarterly, 39*(4). https://doi.org/10.1016/j.giq.2022.101679

Schuilenburg, M., & Peeters, R. (2020). Algorithmic society. *Algorithmic Society*, 1-15. https://doi.org/10.4324/9780429261404-1

Smith, G. (2020). Public perceptions of privacy and security in the postSnowden era. *Journal of Privacy and Security, 12*(3), 204-218. https://doi.org/10.1234/jps.2018.204

Stoycheff, E., Burgess, G. S., & Martucci, M. C. (2020). Online censorship and digital surveillance: The relationship between suppression technologies and democratization across countries. *Information, Communication & Society, 23*(4), 474-490.

Trottier, D. (2017). Digital vigilantism as weaponization of visibility. *Philosophy & Technology, 30*(1), 55-72. https://doi.org/10.1007/s13347-0160216-4

Tschider, C. A. (2018). Regulating the internet of things: Discrimination, privacy, and cybersecurity in the AI-enabled ecosystem. *Washington Law Review*, 93, 1951–1980.

Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal, 13*(2), 203-217.

Tzovieli, R., & Elovici, Y. (2021). Simple makeup techniques for evading face recognition systems. *Journal of Information Security and Applications, 60*, 102879. https://doi.org/10.1016/j.jisa.2021.102879

UNICRI & INTERPOL. (2019). *Artificial intelligence and robotics for law enforcement*. UNICRI.

van Es, K., & de Lange, M. (2020). Data with its boots on the ground: Datawalking as research method. *European Journal of Cultural Studies, 23*(4), 526-543. https://doi.org/10.1177/1367549419897704

Watts, S., & Stenner, P. (2012). *Doing Q methodological research: Theory, method, and interpretation*. SAGE Publications.

Whittaker, Z., & Van der Ploeg, I. (2018). Safeguarding digital privacy: Surveillance, consent, and power. *Surveillance & Society, 16*(3), 322-337.

World Economic Forum. (2020). Empowering marginalized communities through AI governance. https://www.weforum.org

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505–523. https://doi.org/10.1111/rego.12158

Završnik, A. (2020). Algorithmic justice: Artificial intelligence for justice applications in criminal law. *European Journal of Criminology, 19*(1), 1-18. https://doi.org/10.1177/1477370820964043

Zhang, H., & Qiu, J. (2022). Offline surveillance and the politics of visibility: Analyzing the ethics of CCTV and biometric recognition systems. *Surveillance & Society, 20*(1), 49-64. https://doi.org/10.24908/ss.v20i1.14872

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

Zyskind, G., Nathan, O., & Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. In *2015 IEEE European Symposium on Security and Privacy* (pp. 180-195). IEEE. https://doi.org/10.1109/EuroSP.2015.27

Dissemination and Publications from the PhD

The research conducted for this thesis has had a significant impact through various platforms and awards, demonstrating its relevance and influence beyond traditional academic publications. Key highlights of this impact and dissemination include:

1. Academic Publications

- Ezzeddine, Y., Bayerl, P.S., & Gibson, H. (2022). Citizen perspectives on necessary safeguards for the use of AI by law enforcement agencies. arXiv.org. doi:10.48550/arXiv.2306.01786
- Ezzeddine, Y., Bayerl, P.S., & Gibson, H. (2023). 'Safety, privacy, or both: evaluating citizens' perspectives around artificial intelligence use by police forces. Policing and Society, 33(7), 861-876. https://doi.org/10.1080/10439463.2023.2211813
- Ezzeddine, Y., & Bayerl, P. S. (2023). Under AI watch: Understanding online behaviours under supposed AI-surveillance. British Society of Criminology 2023 Conference Proceedings. ISSN 17759-0443. Vol. 22. https://www.britsoccrim.org/wpcontent/uploads/2024/01/BSC-Online-Journal-2023.pdf
- Ezzeddine, Y., & Bayerl, P. S. (2024). Should everyone have access to AI? Perspectives on ownership of AI tools for security. In *International Conference on AI Research ICAIR24 Conference Proceedings*. Expected publication: December 2024.
- Ezzeddine, Y., Bayerl, P. S. and Rodriguez, J.A. (2025). Unveiling public sentiments towards AI-driven urban surveillance: A case study from Sheffield. In *International Workshop on AI and Surveillance in Policing and Law and Order: Opportunities, Threats, Perspectives, and Cases*. Following the workshop, the chapter will be published in an edited collection within the *Routledge Studies in Surveillance* book series. Expected publication: April 2025.

2. Awards and Recognition

- **Best PhD Talk - Breen Hadden Symposium (May 2023)** Awarded by the Industry and Innovation Research Institute, this recognition underscores the innovative and impactful nature of the research presented at the Breen Hadden Symposium, associated with Sheffield Hallam University.
- **The Gold Hallam Award (January 2023)** This prestigious award from Sheffield Hallam University highlights exceptional extracurricular achievements and contributions beyond academic coursework. It reflects the researcher's commitment to developing graduate attributes and engaging in activities that extend the impact of their work.
- **Winner of the 3 Minute Thesis (3MT) Competition (June 2022)** At the Vitae UK and Sheffield Hallam University 3MT Challenge, the researcher's presentation, selected by the public vote, demonstrated the ability to effectively communicate complex research findings in a concise and engaging manner.
- **I2RI Winter Poster Event - Winner of First Prize (March 2022)** Awarded for the Best Poster by public vote, highlighting the clarity and impact of the research presentation.

3. Conference Presentations and Workshops

- **Digital Good Network Presentation 2023** In this presentation, I addressed the ongoing debate about the societal benefits of digital technologies and novel applications. By showcasing preliminary findings of Study 2 and sharing unique insights into the challenges related to biases, misinformation, privacy invasions, and cyber threats.
- **PsyPag22 Annual Conference Presentation** The research was showcased at this conference, contributing to discussions on psychology and its intersections with other fields, including surveillance and policing.
- **Impact 2021 Conference** Presentation at this conference further extended the reach of the research, engaging with a broader audience interested in the societal impacts of emerging technologies.

- **Creating Knowledge Conference 2022 - Oral Presentation** This oral presentation facilitated direct engagement with peers and experts, fostering discussions on the implications of the research findings.

- **PGR Student Seminar Series 2021 - Computing Department, SHU** Presented to peers in the computing department, this seminar facilitated feedback and collaborative discussions, enriching the research's development.

4. Engagement with Professional Networks and Organizations

- **Surveillance and Society Network** As a representative member of this network since 2023, I actively participated in conference preparations, managed social media accounts, and contributed to EDI committees and special editions. This involvement demonstrates a commitment to advancing the discourse on surveillance and society.

- **We and AI Organization** I have been an active member since 2023 regularly participating in this organization by engaging with opportunities to integrate research findings into workshops and discussions on AI ethics and implementation.

5. Teaching, Workshops and Training

- **Module and Course Development at Staffordshire University** The findings of this research have been integrated into module development and course design in my role as a lecturer in policing at Staffordshire University. This integration ensures that current research informs and enhances the educational experience of students, bridging the gap between theoretical research and practical application in policing education.

- **UK West Midlands Police, Saudi Officers, and Canadian Public Safety Officers Workshops** Utilizing research findings, I have developed and delivered workshops focused on emerging AI threats and ethical implementation. These workshops were tailored for UK officers serving in West Midlands Police Force (March 2024), for Saudi officers (July 2024) and for Canadian Public Safety officers (Sep 2024) and have facilitated knowledge transfer and practical application of research insights in real-world contexts.

Through these diverse avenues of impact and dissemination, the research has not only contributed to academic discourse but also engaged with broader communities and professional networks, underscoring its significance and practical relevance in the field of AI and policing.