Beyond Use-Cases: A Participatory Approach to Envisioning Data Science in Law Enforcement

Caitlin Kearney
caitlin.kearney@tum.de
Technical University Munich
Munich, Germany

Jiri Hron
jirihron@google.com
Google DeepMind
London, UK

Helen Kosc
helen.kosc@sociology.ox.ac.uk
University of Oxford
Oxford, UK

Miri Zilka
mz477@cam.ac.uk
University of Cambridge
Cambridge, UK

ABSTRACT

With a rising number of law enforcement agencies facing budgetary cuts, many turn to data science in an attempt to maintain service quality with fewer resources. A number of thus adopted solutions–including facial recognition, predictive policing, and risk assessments–have been contested by researchers and journalists alike. Yet comparatively little research is done at the strategy level, which determines where data science will be deployed in the first place. In this study, we interview 40 practitioners from Police Scotland, investigating what they believe to be crucial to successfully incorporate data science in their ways of working. Bucking the external trend, the participants distanced themselves from tools like facial recognition and risk assessment. Instead of focusing on individual use-cases, their primary concerns for the future were around (i) systemic issues around data is collection and use, (ii) goal misalignment between leadership and operational levels, (iii) the fear that datafication may undervalue important aspects of policing, and (iv) appropriate ways of interaction between data science teams and operational officers. Alongside the insights particular to Police Scotland, our work reaffirms how participatory approaches can go beyond the technical, and uncover structural and political barriers to success.

CCS CONCEPTS

• **Human-centered computing**; • **Social and professional topics** → **Computing / technology policy**;

ACM Reference Format:
Caitlin Kearney, Jiri Hron, Helen Kosc, and Miri Zilka. 2024. Beyond Use-Cases: A Participatory Approach to Envisioning Data Science in Law Enforcement. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), June 03–06, 2024, Rio de Janeiro, Brazil.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3630106.3659007

FAccT '24, June 03–06, 2024, Rio de Janeiro, Brazil

ACM ISBN 979-8-4007-0450-5/24/06
https://doi.org/10.1145/3630106.3659007

1 INTRODUCTION

Faced with shrinking budgets and pressure to increase efficiency, policing agencies worldwide are increasingly embracing data science [1, 15, 23, 28, 34, 38, 57, 61, 65]. Past deployments did not come without issues, with predictive mapping, risk assessment instruments, and facial recognition tools all criticised due to low accuracy, bias, and more [2, 11, 14, 17, 18, 32, 32, 45, 48, 62, 64]. While much research has examined specific use-cases, there is a paucity of studies investigating the development of data science strategy, which determines the 'where and what' of data science deployments, potentially with little or no involvement from the practitioners and other stakeholders.

To address this gap, we interview 40 Police Scotland employees of various ranks and roles, to understand their views on, concerns about, and hopes for the *future* of data science in their organisation. The focus on the future in our interviews is motivated by findings from *sociology of expectations*, which suggest that analysing future expectations can reveal hidden assumptions influencing social processes [59]. Our goal is to explore the range of opinions, not to bestow legitimacy. Legitimacy cannot be attained without input from voices external to police, which have been studied elsewhere [24], and are subject of future work for Police Scotland.

Counter to external trends, we found our participants distanced themselves from tools like facial recognition and risk assessment. Instead, their concerns were around (i) systemic issues in data collection and use, (ii) goal misalignment between leadership and operational levels, (iii) potential undervaluation of important aspects of policing due to datafication, and (iv) appropriate ways of interaction between data science teams and operational officers. While several use-cases where data science could be useful were suggested, it is striking that most of our participants considered the *relationships and communication* between high-level decision makers, data scientists, analysts, and officers on the ground, as much more crucial to design of a successful data science strategy.

Paper structure: Section 2 introduces background on data science in policing, Police Scotland, and sociology of expectations. Section 3 discusses our research questions and methodology (incl. limitations). Section 4 presents the future expectations of our participants, and Section 5 the potential barriers to integration of data science in Police Scotland. Section 6 lists examples of the identified use-cases. Section 7 then discusses Sections 4 to 6 in light of our research questions.

2 BACKGROUND AND RELATED WORK

2.1 Data and technology in policing

Police departments have been working to incorporate data science in their organisations for over a decade [1, 15, 23, 28, 34, 38, 57, 61, 65]. Generally speaking, data science applications can be used to generate 1) organisation-level insights for leadership (e.g., demand and performance metrics) to inform high-level decision making; or 2) incident or person-specific insights used on an operational level by specific units or across the force. In the UK, adoption has been driven primarily by experimentation—often local—but many projects were abandoned shortly after initial tests, usually with no formal reason [29, 66]. Earlier efforts focused on purchasing 'off-the-shelf' solutions like PredPol [41, 43, 44], a system designed to predict where and when crime will occur, later retired by UK forces due to unsuccessful trials and public controversies [40, 48] A more recent trend is the development of custom solutions in collaboration with academia and consultancies, or using in-house resources [66]. Beyond predictive mapping, some police forces have been using data-driven tools to forecast resource demand, allocate investigative resources [37], predict re-offense risk [42], and automate facial recognition [20].

Academics took note of this innovation. Recent studies on data-driven policing technology focus on intelligence practices [12], crime analysis [46], criminal investigations [39], police surveillance activities [6, 21], and predictive tools in criminal justice [7]. Introducing modern technology in police departments does not necessarily transform police strategies and tactics for crime [35]. Studies suggest that police instead *'cherry-pick aspects of technology that are symbolically beneficial and financially affordable, and employ them to further "traditional" crime fighting'* [12, p.5] [10, 46]. Chan et al. [12] found barriers to incorporation of data science are primarily political, and that officers fear traditional policing knowledge may be impaired by leaning on—and considered inferior to—data-driven insights [6, p. 990]. Officers are aware that official police data is often subjective and not an objective measure of crime [47]. Fussey and Sandhu [21] described a culture wherein practitioners collect less data to reduce their analytical workload and the risk of wrongdoing.

2.2 Police Scotland and Data Science

Frontline officers in Scotland do not carry firearms and primarily rely on open-fist or baton if force is needed [56]. Some officers are trained with taser [55], while firearms are reserved for a fraction of specialised officers [54]. While largely response-oriented, Police Scotland also employs a community-oriented approach that includes officers working with public groups to collectively decide on how to tackle crime and 'keep people safe' [49].

In a major restructure, Police Scotland was formed in 2013 by consolidation of 8 regional forces. The merger aimed to enhance coordination and efficiency, with projected savings of 2.2 billion pounds by 2026 [50]. While economical, the reorganisation also resulted in 900 fewer officers across the nation compared to 2013. This has raised concerns about the organisation's ability to preserve community policing, and to serve the diverse needs of Scotland's regions. Technologically, the merger forced Police Scotland to work on uniting data projects and solutions across 8 historically separate divisions. At the time of the interviews, there were still some separate regional systems, and information scattered across separate databases. These systems vary in usability, with some described as old-school pen-and-paper policing in a digital format, but other newer applications reflect an effort to build for the digital age.

In the present, operational efficiency of Police Scotland suffers because officers need to switch between different siloed systems to find, record, and edit various types of information. This inefficiency incurs officer time, and concentration, and increases the risk of data inconsistencies. The problem with multiple keying and siloed searching is well known; In 2018, the announced Digital, Data, and ICT (DDICT) strategy included addressing ageing legacy infrastructure as a goal [50, p.8]. However, the DDICT effort was delayed for budgetary reasons [25], and is still ongoing. In 2023, a new Digital Strategy [53] was introduced, intending to address crime through cyber and intelligence-led policing, modernise access to services and digital tools, and enable partner collaboration. It was acknowledged that establishing a strong technical foundation is required before moving data science forward. The results of this study do not directly focus on the limitations of the current IT landscape but are highly relevant for understanding the results.

2.3 Sociology of Expectations

The methodology we use is based on findings from Sociology of Expectations (SE), a framework for making sense of the present through discussing future expectations. SE can shed light on how to prevent certain trajectories, and uncover existing inequities, discrimination, power issues, and sensitivity to uncertainty [59]. The framework provides two key concepts, *agency* and *hype*, that are useful for our study. *Agency* explores the dynamic between technology and people, examining who holds influence in decision-making, and the nature of their interaction [9]. We use agency as a framing to consider the power of participants in shaping their future with data science. SE has also shown that visions of technology can involve cycles of hype and disappointment [5]. For example, a hype phase might obscure the problems and limitations of what is sold as a promising technological solution [59]. Hype is created through high expectations, and promises made during this phase are not to be trusted because disappointment often follows [57]. Although, expectations are not formed by organisational culture alone and are drawn from media, policy, and the public arena [16].

SE was used to study practitioner perspectives in robotics [8], genomics [19], and the military [58]. The latter is a rare example of SE applied in the security space, revealing a disconnect between the innovative future Dutch military imagines, and the on-the-ground experiences of soldiers [58]. SE was also used to show people can have competing visions of the future, and simultaneously hold both negative and positive expectations about a technology [4, 5, 59].

3 METHODOLOGY

We perform a qualitative study of Police Scotland employees perspectives on present and future of data science integration in their organisation. Our aim is to answer the following questions using a series of semi-structured interviews:

- **Q1.** How does data science influence the way Police Scotland imagines the future of policing?
- **Q2.** Which areas within Police Scotland might (not) be suitable for datafication?

Q3. How do data science and community-oriented policing interact? Can there be negative impacts?
Q4. Who do our participants think should be involved in planning, vision, and design decisions, and how?

This study is a part of a larger project examining best practices around the use of data-driven tools within UK policing. The unique contribution of this paper is investigation of data science strategy. Police Scotland was invited to participate based on the interest they expressed in hearing a critical perspective on their data science strategy development process.

3.1 Semi-structured interviews

3.1.1 Design. We conducted group and individual semi-structured interviews with officers, data science practitioners, and leadership in July–September 2023. The participants were asked to imagine potential futures of policing, where data science and technology respectively (a) meet, (b) exceed, or (c) fail expectations in the next 5–10 years. Following recommendations of [59], we emphasise future expectations in our interviews to help reveal hidden assumptions influencing the involved social processes (see Section 2.3). The goal is not to predict the future, but to understand relative implications and feasibility of various policy measures.

We designed and followed a three part conversation guide (see the appendix). This guide was reviewed in advance by Police Scotland to ensure that (a) interviewers use correct and comprehensible language, and (b) the leadership understands the intent of the study. To enable franker conversations, we interviewed the leadership separately from the officers, who may be more reluctant to speak in presence of higher ranks.

Group interview sessions were organised by role, and lasted three hours in person, and two hours online. Individual interviews were one hour long. Due to a shortage of policing resources, we had to accommodate flexible schedules for some participants. Where needed, we scheduled additional time with some non-leadership interviewees. Group interviews began with an introduction to ensure that participants less familiar with data and technology had a common understanding of the session's structure. The introduction included the intended structure for the day, a re-iteration of participant confidentiality, and a short presentation about data science including use-cases in policing.

3.1.2 Participants. We interviewed 40 employees of Police Scotland. Recruitment considered the following criteria in descending order of importance: role, rank, region, and gender. The roles included:

- *Officers*: Persons who create data as part of daily policing work, or use data as part of investigative crime work.
- *Data practitioners*: Mostly civilian analysts working with data science or technology to support other roles. They often perceive leadership as their primary stakeholder, and officers as their subject and 'consumer'.
- *Leadership*: Officers and civilians responsible for overseeing other roles, vision, and strategy.

Over half of the participants came from the Greater Glasgow area. The rest were recruited from other regions, prioritising less urban areas. One in six officer participants were female, which was less than our target of one in three officers[52]. We attempted to recruit more female officers, but ultimately their policing work took priority. The on-call nature of officer work also meant we could not know if participants would arrive until a session started.

Our access was facilitated by the Data Strategy Team, who liaised all recruitment communication, and provided feedback on our initial recruitment criteria with regards to regional coverage and organisational structure. All participants were informed their participation has no associated monetary or professional performance benefits.

3.1.3 Supplemental interviews. Beyond the main interviews described above, we organised separate sessions to help us better understand the current interplay between technology and people in Police Scotland. While not the focus of this paper, these studies were invaluable in helping us follow and interpret the interviews.

3.1.4 Data collection and analysis. Interviews were held both online and in person. Online sessions were used for participants who would not be otherwise able to travel to Glasgow, but ensuring participant representation reflected location. All participants signed an Informed Consent document and reviewed a Participant Information Sheet (see the appendix) explaining how their identity will be protected and data anonymised for the study. Interviews were automatically transcribed using otter.ai (in-person) and Microsoft Teams (online), and manually corrected against the original audio. Transcripts were anonymised and brought into the qualitative coding software MaxQDA [36]. We used a constructivist grounded theory approach that involved an iterative comparative method for analysis in which codes emerge from the data [13]. The analysis was grounded in data but attentive to literature from criminology (i.e., technology and policing cultures), and science and technology studies (i.e., sociology of expectations). Further details in Appendix A.

3.2 Limitations

This study examined how Police Scotland imagines a future with data science through the eyes of its employees. We include the voices of neither the public in general, nor the marginalised communities in particular. Additional research seeking input from external stakeholders is needed. We are not aware of a planned public consultation around data science and policing in Scotland; however, we received the response that there is interest in public consultation on a project-to-project basis, especially in the case of any major technological developments.

Due to the lack of views from critical stakeholders beyond police, we deliberately do not consider legitimacy, which is an important question that deserves a separate study with a wider stakeholder involvement. This work is therefore not intended to endorse or condemn the use of any data science applications within Police Scotland or more generally. Police Scotland is characterised by its mandate around prevention and resident well-being, as well as a significant rural component, led by community-oriented policing (Section 2.2). This provides an interesting perspective and contrast to police forces in other parts of the world, but also limits transferability of the findings into other policing contexts. As external academics, we would not have the access or resources to conduct this study without the cooperation of Police Scotland. Police Scotland had the ability to review the results prior to publication and request changes. We state that minor amendments were requested, but none altered

Table 1: Breakdown of participants by role in Police Scotland.

Role	Operational	Data practitioners		Leadership
Affiliation	Officers	Desk-based & embedded officers	Civilian employee	Mixed
Number of Participants	15	9	11	5

the findings or conclusions we made. Our impression was that members of the Data Strategy team facilitating our access were willing and determined to engage on aspects relevant to data strategy. Nevertheless, a limitation of working closely with the team was overrepresentation of data practitioners as participants. This is likely because much of the data work happens within this group, and leadership encouraged participation. Research participants considered location, role, and rank, so officer level insights can be more generally applied, however senior level and leadership discussions leaned towards representation of two units, and are less representative.

We made a distinct effort to not group participants cross-rank (Section 3.1.1); however, determining and grouping based on seniority was an imperfect process. We attempted to pre-group everyone in advance, but this proved difficult with operational officers who could not confirm joining in advance. Ultimately, the levels of seniority varied between our groups, which may have impacted individuals grouped with higher ranked participants (although never their direct superior). Generally, our purely subjective perception was that participants were frank and frequently critical.

4 VISIONS OF FUTURE OF DATA SCIENCE AND TECHNOLOGY IN POLICE SCOTLAND

This section presents futures envisioned by our participants (see Section 3.1). These can be organised along two axes: (a) level of change, and (b) desirability. For level of change, participants discussed scenarios ranging from technological stagnation, to changes that would fundamentally alter the nature of their job and relationship with the public. For desirability, the described futures can be categorised as negative (Section 4.1) or positive (Section 4.2).

4.1 The *failed* futures that nobody wants

Failed futures describe how data science could go wrong, illustrating the social and technological failure modes described by our participants. Understanding undesirable futures in policing aids the analysis of what is vital for a desirable future.

4.1.1 Technological stagnation. The most mundane and collectively shared failed future is where Police Scotland continues to rely on its already outdated systems. All the current fissures widen, and police becomes increasingly unable to prevent and respond to crime. The concern is exacerbated by how much faster technology outside policing develops: *'I can see everything developing faster than we are. I can't see us improving to get to where we ultimately need to be' (Operational Officer).* Fulfilling this future would be simply *'allowing inertia to guide the data science strategy' (Civilian Employee).*

Participants identified several paths to technological stagnation:

- *Band-aids:* Tools are adopted in reactionary fashion without a strategic view of data governance and ethics.
- *Workarounds:* Instead of using new tools, existing ones are re-appropriated. An example is the ongoing practice where employees looking to find a particular person often use the Vulnerable Persons Database (VPD), instead of the slower less user friendly crime database designed for this purpose, despite the potentially less relevant results.
- *Consumer apps:* As consumer technology becomes increasingly faster and more convenient, officers may become tempted to use such unauthorised tools to save time: *'The technology I can get for free in five seconds on my phone far outstrips anything that I've used at work, for the most part. . . . [T]hat in and of itself will create tension. You will get police officers trying to use ChatGPT for various things' (Civilian Employee).*
- *Talent drain:* Insufficient organisational and infrastructural support to hire and retain the talent needed for development and maintenance of new tools. Participants shared this concern due to the ongoing lack of funding.

4.1.2 Veneer of success. In this future, Police Scotland has deployed several new data science tools, but after the initial optimism, systemic issues surface. These may include unreliability, bias, and overall lack of trustworthiness of either the data or the insights. Participants mentioned four potential paths to this future:

- *Lack of support:* A self-service tool is implemented, and extended to divisions which lack the expertise to meticulously scrutinise bias, accuracy, and relevance when developing their solutions.
- *Oversimplification:* Rigid data structures may lead to loss of nuance in translating complex situations into simplistic 'yes'/'no' form responses. Consequently, data science may yield interpretations founded on data of subpar quality. For example, in 2019, Scottish emergency call controllers were asked to change from six to four categories when assessing the caller's situation [51]. While initially celebrated, a participant mentioned that one of the new categories is too broad to properly capture urgency. The controllers thus manage prioritisation within this category manually, overriding the system assignment of resources.
- *Deprioritising prevention:* As part of community-oriented policing, Police Scotland engages in crime prevention, and assists at-risk individuals (Section 2.2). Officers felt this work can become deprioritised if policing becomes more metric-driven, as quantifying how much crime was prevented is difficult.
- *Accountability:* If a risk score was used to justify a preemptive action, accountability is unclear. This concern possibly stems from developments elsewhere—media or politics [16]—as Police Scotland does not favour such systems.

4.1.3 Insights ≠ better decisions. Here new tools are adopted and provide relevant insights for specific use cases, with shortcomings well documented. The issue lies in translating these insights into better policing outcomes. This future considers how adoption and reliance on future tools may change the officer role, reduce decision-making capacity and traditional investigative work skills, and increase data work and overall job dissatisfaction.

The following are examples of concerns relevant to this future:

- *Discretion:* Community-oriented policing within Scotland relies on officers' ability to examine incidents holistically. Officers fear they may face formal or informal pressures to follow AI suggestions instead of their judgement.
- *Over-reliance:* Officers become too reliant on new tools, and their critical thinking and core policing skills atrophy. This concern is fuelled by perceived decline in information gathering skills due to ongoing automation: *'I know that it's good for doing things quickly and streamlining, [but] I think there's a loss of quality' (Officer).*
- *Misplaced priorities:* While new tools could reduce time in redundant data entry work, there is a fear that new digital tasks will replace the old ones, and continue to encroach on time spent with the community and addressing crime. For example: *'[A] system could be used to help pre-write reports to reduce the amount of time report writing. But … are we going to have to be coming in and double checking quality?' (Operational Officer)*
- *Overload:* Automation can create more signals than there are officers and resources to cover: *'A computer can look at footage, whether that's evidence, body wall or CCTV, and … flag things which we're then directed to patrol … in reality we don't have the resources to manage that' (Operational Officer).* This belongs to a larger theme of ensuring that business resources, processes, and social values are aligned with automation.

4.1.4 Automation removing steps vital for public confidence. This last future is about impact of automation on public confidence. Our participants are also Scotland residents, concerned about how the public trust could be impaired. They brought up the following possible unintended social consequences of automation:

- *Procedural justice:* Automation may make people feel their concerns are not heard: *'It matters hugely to people that they feel like they've been given a fair hearing, that they've been treated with respect by the police and the courts … even if the person ends up getting off. In [the victim's] view, it makes a massive difference … that they feel like the process was followed in terms of their faith in the system and the outcome. And there's a huge potential to undermine that procedural justice, when you start introducing datafied elements to it' (Civilian Employee).*
- *Trust:* AI may make the public question human judgement: *'There's a risk of … machine learning taking away human judgement and … trust from the officers. [The leadership] wanted a machine to step in … so that [public] won't have concern that there's bias from the officers. But then that will look like a lack of trust in us' (Operational Officer).*
- *Emergencies:* If someone calls emergency services and is connected to an AI instead of human operator, they may feel lack of emotional understanding and support, even if the AI provides correct responses.

4.2 A *positive* future to aspire to

Prompting participants to imagine positive futures often lead to a conversation about current issues and concerns, or to revisiting an already discussed negative future in more depth. The following positive futures can therefore partly be viewed as the counterpart of participant concerns about present or future (Section 4.1).

4.2.1 Low-hanging fruit. There was a general desire for simple applications: *'[K]eep it really simple and just try and find as many opportunities to save time as possible and keep final decision making with officers' (Operational Officer).* An example is demand for an integrated nominal[1] database: *'If I could go on a system, search the name and date of birth, and have their criminal history, their VPD history, their custody history, … all on one screen, with one search, instead of having to switch between systems to do all the searches, it would be really, really impressive' (Civilian Employee).*

4.2.2 Documented limitations. Participants wanted to *'[know] our limits in relation to data science, not just with the ethics side of things, … [there is] this kind of overreach, this desire, wherever there's a problem, that we need to solve it through data science, [but it] is actually a business process [or] training thing that needs to change.' (Civilian Employee)* These limitations are both in knowing the limits of data science to solve problems and additionally communicating the limits of any given data-driven tool. The limits should be well-researched, documented, circulated, and shown in context whenever they are informing a human decision (e.g., a crime trends dashboard should explicitly state it was tested only on specific urban data). Policing expertise should guide the entire process, from data collection to interpretation and application.

4.2.3 Deliberation. Decision-makers will have ample opportunity to review data summaries and investigate further.

4.2.4 Participatory practices. Consulting and co-creating new tools to ensure they consider policing craft, operational context, and regional requirements with the representation of *'everybody that's going to put their hands on it' (Operational Officer).* Data practitioners would work with anyone impacted by the changes they implement before deploying them: *'I made some analysts come out and have a week in my area. And they were completely blown by the stuff that they didn't know, and stuff that they've been able to take back' (Operational Officer).*

4.2.5 Data awareness. Experienced officers think all their colleagues should know how data supports policing work. *'I think training for the next 10 years should be something … around data, because I've not met a lot of [officers] that will be comfortable, with what they can do with what kind of [data], or where data goes and comes from, etc.' (Operational Officer)*

5 BARRIERS TO CHANGE

Our interviews focused on future expectations (Section 3.1) not only to identify which changes our participants (do not) want, but also to uncover potential obstacles to successful adoption of new tools. Here we discuss the uncovered barriers.

5.1 Skepticism rooted in past experiences

Years of budget cuts, inconsistent IT programs, and metric-based goals have eroded trust in new technology within Police Scotland. Participants believe future resources will be limited and are concerned that AI and automation may be used primarily for cost-cutting. These views are grounded by experience, where past programs have often overpromised and underdelivered: *'[First] there's*

[1]A nominal is an individual who has a recordable offence (caution, reprimand, warning, or arrest) [3].

enthusiasm, [but] then we have a half-finished product, and it just doesn't really help in the end. …I don't think it's been too great an experience for people.' (Operational Officer). One cause of such failures was thought to be office politics, which led to effort duplication, and lack of long-term funding.

Another source of scepticism is the past culture of metrics-driven performance evaluations that prioritised numerical targets over values of community-oriented policing [26]. Our participants thought adoption of AI can swing the focus back to easily measurable outcomes and cost-cutting, rather than public safety. External controversies around inaccurate matches in DNA analysis, fingerprinting, and facial recognition have also driven negative perceptions. Overall, addressing past disappointments and aligning innovation goals with community safety is crucial for future progress.

5.2 Lack of officer input and prioritisation

While participants shared universal respect for the work of officers, some leaders and data practitioners believed they *'don't have an intelligent customer' (Leadership)*, i.e., that officers lack the knowledge needed to identify where data science can help them. A language barrier was suggested as a potential issue: *'If we speak in their terms, [officers] would understand the connection of data science to their business problems' (Leadership)*. Some thought connecting data quality with what officers care about is key; a senior officer explained how they motivate colleagues to include detailed information in the vulnerable persons database (VPD): *'[Take] a person suffering from dementia. There's a lot higher risk that they will go missing because they'll just disappear and not tell anyone. A lot more information in the VPD [can] aid a missing person inquiry in the future. …[G]iving that insight on what I'm asking the data for could be really valuable' (Operational Officer)*. Operational officers made no mention of language being an obstacle.

Other data practitioners believe the problem is in resource allocation: *'[W]e don't work with the cops on the ground. …[W]hat we are working on comes from senior managers, who are thinking on the strategic level. …I think data science techniques are probably better suited to try to solve some of these more operational problems' (Civilian Employee)*. There are additional concerns about how data practitioner resources are used. There was a concern that because the organisation's structure prioritises data governance and security, it slows the ability to effectively enact change. The fear is a downstream effect that forces the organisation to rely on third-party consultancies to bridge a resource gap that could have been avoidable. Moreover, a concern that consultancies are onboarded before consulting and working with internal talent, who might already be working towards similar goals. Which dilutes the power of the local knowledge of existing staff and also creates frustration within the existing talent.

5.3 Collaboration between teams and implementation of participatory measures

Police Scotland already has some participatory measures in place, but they are not inclusive of all types of expertise. While participants reflected positively on the activities of groups like Demand and Productivity Unit, where officers work alongside data practitioners, operational officers do not feel heard: *'I think Police Scotland sometimes fails in consulting their officers. …[A] lot of bosses sitting in a room … without the guys that really are actually out there dealing with things' (Operational Officer)*. Participants also raised that although officers embedded with data practitioners speak the language of policing and bring rich domain insights, they cannot replace direct input from operational staff in all regions of Scotland. Moreover, some in the leadership felt that the embedded officer time would be better spent in operational capacity.

Other leaders were less pessimistic about mixed teams though: *'[K]ey to the success of the team is having that blend of officers and staff …I suppose you could say these officers should be on the frontline, but actually bringing that context of how it works in the real world and having the technical team and the data science team help pull that data together to evidence it, really works well for us. …[I]t's gonna be a hard sell going forward because of the financial restrictions and you know that when we're cutting numbers of officers, we need to get them back out and operational' (Leadership)*.

Participants also reported problems connected to implementation of participatory measures. One in particular elaborated on the national drive to collect form data, which were to be used in a case against a partner agency. The form contained a misleading section, which meant officers were answering the wrong question. The issue was reported, but the feedback was ignored: *'We're getting wrong forms in, and we're getting bad data, which means it will be a bad result at the end of it. This has been flagged, but there's been no change, even though it's been raised several times' (Operational Officer)*. Beyond the obvious issue of using incorrect information to guide decisions, this ineffective mechanism to handle concerns impairs the trust officers have in the organisation's ability to resolve issues, and potentially the data.

5.4 Trustworthiness and transparency

Opinions on how to build trust in data science products varied. Some suggested to limit how much officers know about tool design, and instead focus on demonstrating how the tool will make their life easier. In contrast, officers complained about the current lack of transparency: *'We moved to a local policing model … and there was this mysterious calculator that we weren't allowed to see how it worked, but we fed in the metrics and it came back telling us exactly how many police officers should be in a subdivision. …[A]lthough figures wise [i.e., logically] it makes sense, it doesn't make sense on the ground' (Operational Officer)*. Lack of explainability itself can also raise suspicion in officers. The participant continued, *'I was reporting directly to the divisional commander and he didn't know how it worked either …I think that was probably the basis of it in that as soon as you knew how it worked, the divisional commanders would be fighting it' (Operational Officer)*.

The debate also touched on transparency, particularly how much data access should be given to officers. More senior level officers criticised the position that only strategic-level policing roles should have access, to protect officer time and avoid introducing officers to irrelevant information that could be accidentally misused. In contrast, one leader feared that restricting access runs the risk of

missing out on important insights for operational safety.

5.5 Challenges of datafication

Information is lost when a real-world situation is converted into data structures. Regional differences in particular can be difficult to capture across standardised systems, which is a problem for the many rural regions served by Police Scotland. A divisional officer explains: *'[W]e always feel that we're just lumped in with everyone else when really geography is actually the biggest factor in pretty much everything we do . . . [D]rilling down into what's going on in a small radius, [some tools] just don't have the capability to do that' (Operational Officer).*

There was also a concern about assuming everything is a technical problem, instead of asking: *'Is it actually a business process that needs to change? Is it a training thing that needs to change? Is that a conversation with partner agencies that needs to happen? Are we falling into that trap of only solving problems that we have data for, and therefore missing the bigger picture? There's a real keen sense of are we being responsible?' (Civilian Employee).* Several participants also expressed uncertainty about how to measure, track, and weigh prevention measures. The discussions emphasised the utmost importance of preventative measures, but clearly articulated a lack of certainty about how to tackle the problem.

5.6 Underdetermined accountability chain

As many individuals contribute to data collection, processing, presentation, and analysis, there is a many-hands problem without clear accountability for a misjudgement: *'[W]hen people get wrongfully arrested, because information [in a system] is wrong . . . where does the responsibility lie? . . . [N]ow you can look back and see what happened . . . and then you can deal with that through training, etc. [But if] it was all dealt with through an automated program, then who's responsible for that? Because saying that the computer made the mistake, that's not really acceptable' (Officer).* Officers and leadership were concerned that developers will not share the same level of accountability as them: *'If there is an increase in mistakes because [of] relying more on machines and not using our own judgement, then who's gonna get cited to court to explain? . . . [I]n my previous experience, it is very difficult to hold the developers of the tools accountable. It will either be officers that have been feeding in the information or using that data, or the bosses that have kind of put it together and signed off on it. And in some ways, that's unfair, because obviously [officers] are not involved in the nitty gritty"' (Operational Officer).*

5.7 Work demand and confidence influencing data quality

Leadership and data practitioners share the concern that many officers see data entry as merely box-ticking exercise. Officers and data practitioners alike expressed that because the demand for officers is high, they are forced to move from task to task quickly. This means they do not necessarily have the capacity to capture and record high quality data. Senior officers also see this as a failure in training in core policing skills and a lack of organisational support. This is tied to the amount of information they are asked to complete and to time pressure limiting their ability to engage in an 'old school' chats with the public to build trust and gather high quality information: *'[Y]ou would sit and talk to Miss Brown about her bin being stolen for two or three hours sometimes, and she would tell you everything. That's now gone. And I think that's linked to the demand [for officers] . . . [In] the 'old school' style of policing, you have actually spoken to people instead of just quickly ticking some boxes and leaving' (Officer).*

It was also discussed that officers do not always have the confidence to ask for helpful information while talking to the public because they are overly concerned about gathering sensitive data. Some felt there is too much scaremongering around what officers are permitted to collect, and that frontline officers are unaware that back office staff are tasked with sanitising any entered information that is outside the organisational guidelines. This self-indemnifying behaviour towards data collection has been observed in other UK forces [21].

6 EXAMPLES OF USE-CASES SHARED BY PARTICIPANTS

The participant imagined use-cases broadly fit into three categories: (a) investigation aides (Section 6.1), (b) information sharing (Section 6.2), and (c) efficiency enhancements (Section 6.3). Hardly anyone favoured use of predictive policing tools, such as predictive mapping or risk assessment instruments; the few who desired predictive tools wanted help with resource allocation and route planning. Most of the discussed topics would require more than data science to be addressed, as many of them are rooted in social, policy, or political issues.

We expected many interviews to centre use-cases, but our participants repeatedly returned the conversation to related structural and political issues. The presented use-cases are therefore best understood as flowing from the context of previous sections (especially Section 5). It is also worth noting that many of the use-cases can be addressed using existing technology, even though the participants were asked to think about a 5–10 year horizon.

6.1 Investigation aides

6.1.1 Unified database of all incident and crime related data. Currently, information related to different aspects of an investigation can be stored in multiple unlinked databases. This wastes officer time and motivates them to use make-do workarounds: *'[T]he crime report is . . . often written in immense detail, every footstep they took, what street they took, who they were with, what conversation they had. You've got to read all that, and then all the subsequent updates . . . [W]hen I was in CID,[2] you'd have a hard folder with documents, . . . a printout of the crime report. What I would do is type where the investigation was at, and staple it to the front of it . . . so that someone could pick it up and decide what we're doing next. Because seeing people going in, it can be three folders. . . . [You] need a couple hours to read it all ' (Desk Officer).*

6.1.2 Locations of available street cameras. During some investigations, officers can spend weeks canvassing an area. This involves surveying availability of any video footage (CCTV, ANPR,[3] etc.). The results of this time-consuming process are never saved, leading to unnecessary duplication of work: *'[Y]ou can have a house break, and two weeks later a corpse turns up. . . . [The two investigating teams] might not be asking the same questions . . . but CCTV locations are probably still the same. . . . But we're not talking to each other' (Desk*

[2] Criminal Investigation Department.
[3] Automatic Number Plate Recognition.

Officer). The participant went on to imagine a technical solution, *'I click a button . . . and it tells me I've got an ANPR camera 0.6 miles down that road, [and] maps out known CCTV locations. [With that] I can start targeting my door to door, it makes my investigation far more intelligent and efficient as well' (Desk Officer).* Another participant expanded: *'Maybe even have it showing other officers in the area, so that you know . . . there's another officer around the corner, there's five cameras here, and there's an ANPR camera' (Desk Officer).* The participants however also said such an application could be abused, and its use would have to be appropriately constricted.

6.2 Information sharing

6.2.1 Localised searchable directory of partner organisations. Officers responding to, e.g., a domestic abuse incident may offer to connect the victim with a local partner agency for support services. However, there is no searchable directory and no tool for tracking ongoing discussions with partner agencies. *'We manage our partner data on a spreadsheet or a Word document . . . that's probably five years out of date. I've been fighting for the past nine months to try and get the force to get some customer relationship management software on the books that would allow us to get all our partner details, get all our projects in one place . . . So that [for example] if I in my subdivision decide I'm going to approach it by Glasgow City Council, I access that data there and then have who I need to speak to' (Operational Officer).*

6.2.2 Real-time updates about missing persons. Time plays a crucial role when searching for missing persons. However, there is currently no system that officers could use to send information to all officers and relevant partners: *'[I wish] we could send an image to everybody. Bus companies, shopping centres, and it's real time, rather than 8 hours, 9 hours, 10 hours down the line. . . . [J]ust an automatic email that gets sent out with an image. . . . Unless you're specifically looking for that individual, and going into the correct application, you're not necessarily going to see that information' (Operational Officer).*

6.2.3 Triaging. Participants envisioned a system for triaging emergency calls, providing the human operators with accurate and relevant data. Urgent incidents would receive prompt responses, and preventative measures and partner data would be considered in the process to avoid bias in incident assessments. While difficult to achieve, this use-case was brought up repeatedly due to the high pressure on existing resources.

6.3 Efficiency enhancements

6.3.1 Automating briefings, incidents, reports. Producing briefings and reports is time intensive. Automation could alleviate this burden, although some thought quality control might be as laborious: *'"[The] system could be used to help pre-write reports and . . . reduce the amount of time and report writing. But then once again, are we gonna have to be coming in and double checking quality? . . . [I think] we have to quality check and spend that time actually correcting what's been put on certainly within 10 years. Maybe once it's really been developed, it would be less of an issue. . . . [A]nd then when these things are delivered, will there be a decision? You don't need as many officers if you've got all these tools' (Operational Officer).*

6.3.2 Automatic summarization. There is often more information than law enforcement have time to absorb during a time-sensitive scenario. An intelligent assistant could help summarise large amounts of intelligence, *'"Automated decision making doesn't have a big strong place in policing, I would say. What does have a strong view is an assistant for policing, and almost every process could benefit from that. . . . We have hundreds of millions of unstructured data files: PDFs, emails, all sorts stuff that's coming in from Internet based sources . . . Can you read the dissertations worth of information in 5 minutes, and make a sensible threat and harm assessment? No, no human can do that. Can we automate and draw out what the most relevant pieces of information are? . . . [An] officer can then decide to delve into other areas, as well as ensure that they can get the best view of vulnerability' (Leadership).*

7

7.1 Our findings with respect to the research questions (Section 3)

7.1.1 3: How does data science influence the way Police Scotland imagines the future of policing? The participants worried the most about technological stagnation (Section 4.1.1). While there was a feeling that data science integration is needed to maintain service quality Section 4.1.1, pessimism about the achievability of positive outcomes was common. This was primarily fuelled by negative past experiences linked to: (i) over-reliance on metrics; (ii) unsatisfactory design and implementation of solutions; (iii) lack of officer consultation and feedback being ignored.

Over-reliance on metrics seemed to be the key complaint of those who objected to data-science as a whole Section 5.1. These reservations stem from in the insistence of past leadership on 'achieving' quotas (e.g., speeding tickets, searches, arrests), which led to immense dissatisfaction among officers, due to perceived encroachment on their autonomy. The emphasis on metrics also led to arguably one of the biggest controversies Police Scotland has experienced, when in 2014, it was reported that *'Police Scotland frisk nine times as many people as the NYPD'* [27].

Participants further expressed suspicion that the push to incorporate data science is predominately driven by cost-cutting objectives, rather than a desire to provide better service to the public Section 5.1. This is rooted in organisational and political tensions that go beyond data science, which have also been documented in a December 2023 report published by HM Inspectorate of Constabulary in Scotland [26]. Our findings on the political nature of barriers are consistent with other forces [12]. These suspicions—combined with the experience of past disappointment and IT strategies announced only to be abandoned—led to a common feeling that the desire to incorporate new data-related technologies can be more about 'hype' than genuine progress.

7.1.2 3: Which areas within Police Scotland might (not) be suitable for datafication? In essence, officers wished for applications that empower rather than disempower them, consistent with the findings of [47, p.909]. While similar desires were also found among practitioners in other sectors [30, 31] the impact on the areas our participants perceived as having a potential to benefit vs. be harmed by data science was particularly noteworthy.

Participants with different roles highlighted different areas they thought data science could assist. Operational officers felt like assistance with information management and reduction of administrative work would be the most valuable Section 6. This included better access and management of information about crimes and

incidents, missing persons, and contacts in partner agencies. It was clear operational officers wished to preserve their ability to exercise professional judgement, and were sceptical of tools like automated risk assessment.

Leaders, on the other hand, prioritised organisational insights to enhance their understanding of demand and performance [26]. While more affected by pressures to innovate and improve efficiency, leaders were as sceptical towards predictive technologies as the officers. Public expectations and consent played a significant role: *'Everyone wants to talk about facial recognition . . . [but] we've been miles away from doing that because ultimately the public are not ready for it. Not in Scotland anyway. I mean, other forces in England, Wales has trialled it, but it's those technology solutions and automation that we don't have the consent to use from the public. If it impacts public confidence, we won't do it' (Leadership).* Public opinion was also a driver for datafication in Police Scotland, as there is a concern that the public may expect the state of IT infrastructure to be better than it is.

Finally, some participants feared datafication can erase or undervalue important aspects of their work, unless addressed by measures outside data science. The most prominent example of this was preventative work, which the officers felt is immensely valuable, but could not imagine benefiting from datafication Section 4.1.2. Another example is resource planning, which may work well in urban areas, but fail in the countryside due to lack of consideration for distances between places, limited access to specialised services, and a lack of backup resources.

7.1.3 3: How do data science and community-oriented policing interact? Reduction or loss of preventative police work is one of the key ways datafication can harm community-oriented policing. More generally, officers felt community-oriented policing can be hurt if officer time with the members of the community is impacted. Examples range from being required to follow a form rather than have a free conversation with a victim or witness, to having to spend more time on administrative tasks. The tension between datafication and community policing was illustrated by one of the interviewees: *'We had a great cop who used to go out to all our outer islands. The communities loved them, because they all knew who he was, they could all approach him, they knew how he would deal with things, and he had a way about him that on our computer systems there's no way of logging . . . [N]ewer cops coming in with the new systems and new things in place didn't see the benefit of what he did. He wasn't good in computers. He wasn't good at updating his logs, but he was a great community cop' (Operational Officer).*

While some participants suggested technology can help automatically capture data, and thus save time, no general solution was offered regarding how can community policing be preserved as data science becomes more prevalent. It was emphasised, though, that preserving community policing should be a strategic priority, and that leadership should accept that some of the key aspects of police work cannot be tracked using metrics.

7.1.4 3: Who should be involved in planning, vision, and design decisions, and how? Participants emphasised the need for collaboration between data practitioners and officers Section 5.3. Many data practitioners felt underutilised, attributing this to officers' lack of technical knowledge to recognise how data science can help them Section 5.2. This is an example of a 'knowledge deficit', a phenomenon which often appears fixable by training [63]. However, deeper sociopolitical tensions were identified as barriers for officer trust and engagement Section 5.4.

The key to fostering collaboration was felt to be officers working alongside data scientists Section 5.3. Two main strategies were proposed: (a) creating mixed units; (b) temporarily embeddings officers withing data science teams (and vice versa). While both options showed promise, some criticised the imposition on the already strained operational officer resources, and questioned the ability of embedded officers to represent all units, geographies, and roles. Frontline officers stressed officer participation should a be a strategic priority, not an afterthought, as they felt it to be key for successful implementation of data science in Police Scotland. While Police Scotland has measures for officer consultation, it does not happen consistent and connected way that is felt by officers.

7.2 Relationship to existing literature and implications

Our work strengthens the case of studies that criticise tools like predictive policing and facial recognition [2, 11, 14, 17, 18, 32, 32, 45, 48] by showing that even police officers in certain areas (Scotland) do not want them. The barriers to change (Section 5) further cast a critical light on papers which propose novel tools without considering practical issues related to real-world deployment (e.g., ignoring challenges in rural areas). The desire for tools that enable spending more time with the community but preserve discretion of judgement (Section 4.1.3) was found in other public sector areas (e.g., social work [22, 33, 60]). Other findings, such as cycles of technology hype and disappointment, are well-aligned with those in the sociology of expectations literature [5, 57, 59], but are novel within the context of policing.

8 CONCLUSION

We interviewed 40 practitioners from Police Scotland, asking about their visions of future of data science within their organisation. We found general consensus that utilisation of data science should be increased within the organisation, and that not doing so would be a failure. However, participants highlighted many risks and barriers to successful integration.

While applications that enable better access and interaction with information were universally desired, concerns exist about the potential of datafication to harm community-oriented policing through reduction in preventative work and increased administrative burdens. Officers' pessimism predominately is rooted in negative past experiences with data-driven approaches characterised by misaligned incentives, cycles of hype, and ultimately disappointment. Past over-reliance on metrics in particular has led to undermining officer's perceived autonomy, and damaged public trust. Current leadership was instead more interested in using data to provide organisational insights, but—like other participants—remains sceptical of predictive technologies.

The majority of participants felt that meaningful collaboration between data practitioners and officers is crucial to overcome existing barriers to a positive change. This requires addressing sociopolitical tensions, and ensuring consistent officer consultation from the outset. We highlight that meaningful participation must extend

beyond police itself and that future studies that include a wider spectrum of stakeholders—incl. impacted individuals, general public, non-governmental organisations, partners & partner services, and policy & law enforcement experts—are crucial.

Researcher positionality statement. We acknowledge that given our constructionist grounded-theory approach [13], there are many findings that can be co-created by the authors. Of the four authors, three were the conversation partners who conducted the sessions. The first author and interviewer draws from a background in Human-Computer Interaction and Science and Technology Studies (STS). The first author was also solely responsible for transcribing, cleansing, and coding the data. Analysis and discussions where shared amongst all authors. The second author comes from a Machine Learning background, and was not present in the conversations. The third has been embedded in post-incarceration research with release programs and recently released individuals. The supervising author comes from a machine learning discipline and has experience in conducting critical socio-technical research within the criminal justice system context.

None of the authors live in Scotland or are Scottish; each individual's background is grounded in nations with ways of policing not easily comparable to Scotland. One author comes from the US, which, in contrast, is known for armed officers, systemic racism, and a very polarising public opinion. None of the authors has direct personal ties with the police, and while our backgrounds influence our thoughts, no strong emotional ties in Scotland guided our discussions.

Appendix B contains further discussion on how our positionality affected the selection and identification of themes.

Ethical considerations statement. The authors acknowledge several ethical challenges associated with this work. First, there is little to no evidence from critical academic literature that introducing data science to law enforcement leads to better societal outcomes. In contrast, there is ample literature on potential and proven harms [2, 11, 14, 17, 17, 18, 32, 32, 45, 48, 62, 64], especially to marginalised communities. Working to inform data science strategy for policing can be considered a form of endorsement for increasing data science applications within policing. However, we felt that our familiarity and involvement with critical research allowed the work to include a significant discussion of what should not be datafied, and potential harms and risks of the adoption of more data science. Overall, we felt that conducting this research does not further encourage the adoption of data science, but does encourage the adoption of responsible practices around data science. We acknowledge the undesirability of conducting this research with police employees only. This is due to practical constraints and we highlight that broader stakeholder engagement, including community members, civil rights groups, and other organisations concerned with the potential impacts on marginalised communities is greatly needed. Additionally, we are mindful of the potential negative repercussions for participants who were asked to frankly and critically comment on their workplace and superiors. While we took measures to protect their anonymity, it may be possible for their peers presume attribution, due to the unique position of leadership participants. No participants were pressed to speak about anything they deemed compromising; however, as interviewers, we felt most participants were comfortable expressing their views rather freely. Generally, Leadership interviewees felt more constrained in expressing their views.

Adverse impact statement. We acknowledge potential adverse impacts of our research. First, the possibility of this work to be inadvertently or deliberately used to legitimise the use of data science in law enforcement. The positive futures discussed in this work are futures envisioned as positive by participants of the research, who are all employees of Police Scotland. These may not be viewed as positive futures by other stakeholders, and may not lead to positive outcomes in practice. We also stress that following the strategic measures that participants felt will promote better outcomes, such as more participatory design and development measures, is not sufficient to ensure good societal outcomes and absence of harms. We do consider these to be positive steps, however, there is many more ethical considerations and evaluation practices that were not touched upon in this work. We re-iterate that we do not endorse any specific data science tools or increasing datafication in policing more generally.

ACKNOWLEDGMENTS

The authors thank Harry Schone, Data Science Coordinator in the Chief Data Office for Police Scotland, for valuable discussions as well as coordinating and facilitating the workshops and interviews presented in this work.
This research was supported by a grant from the Capabilities in Academic Policy Engagement (CAPE) Collaboration Fund. CK acknowledges support from the Elitenetzwerk Bayern and TUM School of Social Sciences and Technology. HK acknowledges support from the Canadian Centennial Scholarship Fund (CCSF), and the Institute for Humane Studies and the Mercatus Center. MZ acknowledges support from the Leverhulme Trust grant ECF-2021-429.

REFERENCES

[1] Alexander Babuta and Marion Oswald. 2020. Data analytics and algorithms in policing in England and Wales: Towards a new policy framework. (2020).
[2] Solon Barocas and Andrew D Selbst. 2016. Big data's disparate impact. *Calif. L. Rev.* 104 (2016), 671.
[3] Jacqueline Beard. 2020. The retention and disclosure of criminal records. https://researchbriefings.files.parliament.uk/documents/SN06441/SN06441.pdf Accessed: Jan 2024.
[4] JENS BECKERT. 2016. *Imagined futures: Fictional expectations and capitalist dynamics.* Harvard University Press, [Cambridge]. http://www.jstor.org/stable/j.ctvjnrvrw
[5] Mads Borup, Nik Brown, Kornelia Konrad, and Harro van Lente. 2006. The sociology of expectations in science and technology. *Technology Analysis & Strategic Management* 18, 3-4 (2006), 285–298. https://doi.org/10.1080/09537320600777002
[6] Sarah Brayne. 2017. Big Data Surveillance: The Case of Policing. *American Sociological Review* 82, 5 (2017), 977–1008. https://doi.org/10.1177/0003122417725865
[7] Sarah Brayne and Angèle Christin. 2021. Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts. *Social Problems* 68, 3 (2021), 608–624. https://doi.org/10.1093/socpro/spaa004
[8] Svenja Breuer, Maximilian Braun, Daniel Tigard, Alena Buyx, and Ruth Müller. 2023. How Engineers' Imaginaries of Healthcare Shape Design and User Engagement: A Case Study of a Robotics Initiative for Geriatric Healthcare AI Applications. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–33. https://doi.org/10.1145/3577010
[9] Nik Brown. 2016. Contested futures: A sociology of prospective techno-science. https://doi.org/10.4324/9781315259420
[10] Morgan Burcher. 2020. *Social network analysis and law enforcement: Applications for intelligence analysis.* Springer.
[11] Bradley Butcher, Chris Robinson, Miri Zilka, Riccardo Fogliato, Carolyn Ashurst, and Adrian Weller. 2022. Racial Disparities in the Enforcement of Marijuana

Violations in the US. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. 130–143.
[12] Janet Chan, Carrie Sanders, Lyria Bennett Moses, and Holly Blackmore. 2022. Datafication and the practice of intelligence production. *Big Data & Society* 9, 1 (2022), 205395172210893. https://doi.org/10.1177/20539517221089310
[13] Kathy Charmaz. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. sage.
[14] Alexandra Chouldechova. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data* 5, 2 (2017), 153–163.
[15] Sarah L Desmarais, Samantha A Zottola, Sarah E Duhart Clarke, and Evan M Lowder. 2021. Predictive validity of pretrial risk assessments: A systematic review of the literature. *Criminal Justice and Behavior* 48, 4 (2021), 398–420.
[16] Hamid Reza Ekbia. 2008. *Artificial dreams: the quest for non-biological intelligence*. Vol. 200. Cambridge University Press Cambridge.
[17] Danielle Ensign, Sorelle A Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian. 2018. Runaway feedback loops in predictive policing. In *Conference on Fairness, Accountability and Transparency*. PMLR, 160–171.
[18] Anthony W Flores, Kristin Bechtel, and Christopher T Lowenkamp. 2016. False positives, false negatives, and false analyses: A rejoinder to machine bias: There's software used across the country to predict future criminals. and it's biased against blacks. *Fed. Probation* 80 (2016), 38.
[19] Joan H Fujimura. 2003. Future imaginaries: Genome scientists as socio-cultural entrepreneurs.
[20] Peter Fussey and Daragh Murray. 2019. Independent report on the London Metropolitan Police Service's trial of live facial recognition technology. (2019).
[21] Peter Fussey and Ajay Sandhu. 2022. Surveillance arbitration in the era of digital policing. *Theoretical Criminology* 26, 1 (2022), 3–22. https://doi.org/10.1177/1362480620967020
[22] Philip Gillingham. 2021. Practitioner perspectives on the implementation of an electronic information system to enforce practice standards in England. *European Journal of Social Work* 24, 5 (2021), 761–771.
[23] Melissa Hamilton. 2021. Evaluating Algorithmic Risk Assessment. *New criminal law review* 24, 2 (2021), 156–211.
[24] Andy Higgins. 2020. POLICING AND THE PUBLIC: UNDERSTANDING PUBLIC PRIORITIES, ATTITUDES AND EXPECTATIONS. https://www.police-foundation.org.uk/wp-content/uploads/2010/10/insight_paper_1.pdf Accessed: Jan 2024.
[25] HMICS. 2019. HMICS Annual Report 2018–19. https://www.hmics.scot/sites/default/files/publications/HMICS20190626PUB.pdf Accessed: Jan 2024.
[26] HMICS. 2023. HMICS Thematic Inspection of Organisational Culture in Police Scotland. https://www.hmics.scot/sites/default/files/publications/HMICS20231207PUB_0.pdf
[27] Paul Hutcheon. 2014. Police Scotland frisk nine times as many people as the NYPD. https://www.heraldscotland.com/news/13171034.police-scotland-frisk-nine-times-many-people-nypd/ Accessed: Jan 2024.
[28] Fieke Jansen. 2018. Data Driven Policing in the Context of Europe. *Data Justice Lab* (2018).
[29] Justice, Home Affairs Committee, et al. 2022. Technology Rules? The advent of new technologies in the justice system. *HLPaper180. Westminster: The House of Lords* (2022).
[30] Anna Kawakami, Venkatesh Sivaraman, Hao-Fei Cheng, Logan Stapleton, Yanghuidi Cheng, Diana Qing, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. Improving human-AI partnerships in child welfare: Understanding worker practices, challenges, and desires for algorithmic decision support. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
[31] Anna Kawakami, Venkatesh Sivaraman, Logan Stapleton, Hao-Fei Cheng, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. "Why Do I Care What's Similar?" Probing Challenges in AI-Assisted Child Welfare Decision-Making through Worker-AI Interface Design Concepts. In *Designing Interactive Systems Conference*. 454–470.
[32] Kristian Lum and William Isaac. 2016. To predict and serve? *Significance* 13, 5 (2016), 14–19.
[33] George Mair, Lol Burke, and Stuart Taylor. 2006. 'The worst tax form you've ever seen'? Probation officers' views about OASys. *Probation journal* 53, 1 (2006), 7–23.
[34] Cynthia A Mamalian and Nancy Gladys La Vigne. 1999. *The use of computerized crime mapping by law enforcement: Survey results*. US Department of Justice, Office of Justice Programs, National Institute of Justice.
[35] Peter K Manning. 2008. *The technology of policing: crime mapping, information technology, and the rationality of crime control*. Vol. 4. NYU Press.
[36] MAXQDA. 2024. Software for qualitative and mixed methods analysis. https://www.maxqda.com/ Accessed: Jan 2024.
[37] Kent McFadzien, Alan Pughsley, Andrew M Featherstone, and John M Phillips. 2020. The Evidence-Based Investigative Tool (EBIT): a Legitimacy-Conscious Statistical Triage Process for High-Volume Crimes. *Cambridge Journal of Evidence-Based Policing* 4, 3 (2020), 218–232.
[38] John Monahan and Jennifer L Skeem. 2016. Risk assessment in criminal sentencing. *Annual review of clinical psychology* 12 (2016), 489–513.
[39] Laura Neiva, Rafaela Granja, and Helena Machado. 2022. Big Data applied to criminal investigations: expectations of professionals of police cooperation in the European Union. *Policing and Society* 32, 10 (2022), 1167–1179. https://doi.org/10.1080/10439463.2022.2029433
[40] BBC News. 2018. Kent Police stop using crime predicting software. https://www.bbc.co.uk/news/uk-england-kent-46345717
[41] Patricia Nilsson. 2018. First UK police force to try predictive policing ends contract. https://www.ft.com/content/b34b0b08-ef19-11e8-89c8-d36339d835c0 Accessed: Dec 2023.
[42] Marion Oswald, Jamie Grace, Sheena Urwin, and Geoffrey C. Barnes. 2018. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality. *Information & Communications Technology Law* 27, 2 (2018), 223–250. https://doi.org/10.1080/13600834.2018.1458455
[43] PredPol. [n. d.]. How Predictive Policing Works. https://www.predpol.com/how-predictive-policing-works/
[44] Predpol. [n. d.]. Kent Police Use PredPol To Prevent Violent Crime. https://www.predpol.com/kent-police-use-predpol-to-prevent-violent-crime/ Accessed: Dec 2023.
[45] Rashida Richardson, Jason M Schultz, and Kate Crawford. 2019. Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *NYUL Rev. Online* 94 (2019), 15.
[46] Carrie Sanders and Camie Condon. 2017. Crime analysis and cognitive effects: the practice of policing through flows of data. *Global Crime* 18, 3 (2017), 237–255. https://doi.org/10.1080/17440572.2017.1323637
[47] Ajay Sandhu and Peter Fussey. 2021. The 'uberization of policing'? How police negotiate and operationalise predictive policing technology. *Policing and Society* 31, 1 (2021), 66–81. https://doi.org/10.1080/10439463.2020.1803315
[48] Aaron Sankin, Dhruv Mehrota, Surya Mattu, and Annie Gilbertson. 2021. Crime Prediction Software Promised to Be Free of Biases. New Data Shows It Perpetuates Them. https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them
[49] Police Scotland. [n. d.]. Jargon Buster. https://www.scotland.police.uk/about-us/how-we-do-it/strategic-planning/children-and-young-people/jargon-buster/ Accessed: Jan 2024.
[50] Police Scotland. 2020. Joint Strategy for Policing. https://www.scotland.police.uk/spa-media/crhngr0e/joint-strategy-for-policing-2020-21.pdf Accessed: Jan 2024.
[51] Police Scotland. 2020. Our new approach to call handling. https://www.scotland.police.uk/about-us/how-we-do-it/call-handling/our-new-approach/ Accessed: Dec 2023.
[52] Police Scotland. 2021/2022. Equality and Diversity Employment Monitoring Report. https://www.scotland.police.uk/spa-media/3jmlr1yq/edi_monitoring_report_21_22_9958_ar.pdf Accessed: Jan 2024.
[53] Police Scotland. 2023. Digital Strategy. https://www.scotland.police.uk/spa-media/ujamefig/digital-strategy-5-12-_accessible-version.pdf Accessed: Jan 2024.
[54] Police Scotland. 2024. Armed Policing Report 2023/24. https://www.scotland.police.uk/about-us/how-we-do-it/armed-policing-quarterly-reports/armed-policing-report-2023-24/ Accessed: Jan 2024.
[55] Police Scotland. 2024. Taser. https://www.scotland.police.uk/about-us/how-we-do-it/taser/ Accessed: Jan 2024.
[56] Police Scotland. 2024. Use of Force. https://www.scotland.police.uk/about-us/how-we-do-it/use-of-force/ Accessed: Jan 2024.
[57] Daniel Sprick. 2019. Predictive Policing in China: An Authoritarian Dream of Public Security. *Naveiñ Reet: Nordic Journal of Law and Social Research (NNJLSR) No* 9 (2019).
[58] Sofie van der Maarel, Désirée Verweij, Eric-Hans Kramer, and Tine Molendijk. 2023. "This Is Not What I Signed up for": Sociotechnical Imaginaries, Expectations, and Disillusionment in a Dutch Military Innovation Hub. *Science, Technology, & Human Values* (2023). https://doi.org/10.1177/01622439231211032
[59] Lilla Vicsek. 2021. Artificial intelligence and the future of work–lessons from the sociology of expectations. *International Journal of Sociology and Social Policy* 41, 7/8 (2021), 842–861.
[60] Katie Wassal, Carolyn Ashurst, Jiri Hron, and Miri Zilka. 2024. Reimagining AI in Social Work: Practitioner Perspectives on Incorporating Technology in their Practice. *arXiv* (2024).
[61] David Weisburd, Rosann Greenspan, Stephen Mastrofski, and J Willis. 2008. Compstat and organizational change: A national assessment. (2008).
[62] Ali Winston. 2018. Palantir has secretly been using New Orleans to test its predictive policing technology. *The Verge* 27 (2018).
[63] Brian Wynne. 1993. Public uptake of science: a case for institutional reflexivity. *Public understanding of science* 2, 4 (1993), 321.
[64] Miri Zilka, Riccardo Fogliato, Jiri Hron, Bradley Butcher, Carolyn Ashurst, and Adrian Weller. 2023. The Progression of Disparities within the Criminal Justice System: Differential Enforcement and Risk Assessment Instruments. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*.

1553–1569.
[65] Miri Zilka, Holli Sargeant, and Adrian Weller. 2022. Transparency, Governance and Regulation of Algorithmic Tools Deployed in the Criminal Justice System: A UK Case Study. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Oxford, United Kingdom) *(AIES '22)*. Association for Computing Machinery, New York, NY, USA, 880–889. https://doi.org/10.1145/3514094.3534200
[66] Miri Zilka, Holli Sargeant, and Adrian Weller. 2022. Transparency, Governance and Regulation of Algorithmic Tools Deployed in the Criminal Justice System: A UK Case Study. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Oxford, United Kingdom) *(AIES '22)*. Association for Computing Machinery, New York, NY, USA, 880–889. https://doi.org/10.1145/3514094.3534200

A ADDITIONAL DATA COLLECTION DETAILS

Our first contact with most participants was in large sessions where we introduced common use-cases of data science in policing before breaking into 3-4 person groups based on role and rank. We asked participants to silently think before describing how they imagine a negative future for incorporation of data science in Police Scotland within a 10 year time-frame, and then encouraged a discussion amongst the group. We repeated this exercise for positive futures. Later, we directed participants to think about steps needed to avoid the unwanted futures and build a desirable one. For leadership interviews, we used a similar conversation guide. As moderators, we debriefed after every session, which is how we identified if more time with a participant, or a specific role or rank, is needed.

We iteratively read the transcript to create codes along with short summaries, trying to reuse the original words of the meaning of that section. Every couple transcripts, we would group and summarise similar codes into parent codes, then introducing these parent codes when processing transcripts. The parent codes and supporting notes and memos later become the foundation of the themes. The parent codes would evolve with each transcript as new codes were created. Section 5, Barriers to Change, were discussed by multiple participants across more than one session. Some of the examples within negative futures were local to group conversation, but never to one person.

B IMPACT OF AUTHOR POSITIONALITY ON THE SELECTION AND IDENTIFICATION OF THEMES

Beyond the implications of our positionality statement, the following factors may have further influenced the themes:

(1) Our lack of personal experience with day-to-day police work may have affected which findings we highlight as noteworthy.
(2) Our prior work on algorithm use in the UK public sector may have affected the results. For example, we expected some managers to blame lack of innovation on frontline worker technophobia; this might have made us even more conscious of not marginalising practitioner perspectives.
(3) Our prior work on algorithm use in the UK public sector may have affected the results. For example, we expected some managers to blame lack of innovation on frontline worker technophobia; this might have made us even more conscious of not marginalising practitioner perspectives.
(4) Relatedly, knowing that technological innovation is often driven by the management, without sufficient input from the practitioners, likely made us more sensitive to differences in perspectives between employees of varying professional ranks.
(5) We were interested not only in how technology fails, but also in how it can be beneficial from the point of view of our participants. This could have affected the balance of positive to negative futures that came out from the interviews.
(6) We view community-oriented policing as preferable to other common approaches to law enforcement. While conscious of this inclination, it may have affected our questions related to the topic.

Police Scotland Data Science Strategy Workshop Guide & Conversation Guide

Workshop Guide

5 m	Introductions and informed consent on recording	● Objective of research ● Round table introductions describing role and declaring consent for recording	Group
10 m	Presentation	● Data being considered ● "How is data science used across the policing 'use cases" (Given by researcher)	Group
3 m	Break into groups	● No right or wrong answers, reiterate it is not related to performance evaluations	Group
10 m	Warm up - discussing day to day behaviours	● 2- 3 minutes to think silently and write down tasks you do often that are vital to your role. ● Reintroduce yourself, and discuss amongst your group which of these tasks could benefit from data, AI, or data tools.	Small Groups
30 m	Failed future	● Imagine and describe a narrative of a FAILED possible future in your group ○ What is the use case here and how is it failing? ○ What actions would need to happen for this to happen? ○ Probing question: Who loses or gains autonomy? ● Summarize your findings	Small Groups
Break			
20 m	Best future	● Imagine and describe the BEST possible outcome within the next 10 years in your group ○ What is the use case here and how is it succeeding? ○ What actions would need to happen to make this happen? ○ Who plays a role in this, how so? ● Summarize your findings	Small Groups
20 m	Concerns	In this section, we will build on the use cases we previously discussed. ● What should be considered to avoid the worst future and work towards the best	Small Groups

Police Scotland Data Science Strategy Workshop Guide & Conversation Guide

		future? ● What skills do data scientists need to learn to understand the limits of their tools in operation? ○ And, what sort of skills would the police need to develop to meet the expectations of a positive future? ● How do you imagine the relationship between officers and the data science team?	
Break			
10 m	Update	● One person from each small group to report interesting findings with larger Group. ● Opportunity for participants to react to any findings from other groups	Group
5 m	Wrap up	● Closing notes, group questions ● Is there any important information that you expected we would cover today, but did not?	Group
End Session			
20 m	Moderator debrief	● Reflections on the session recorded on Otter	Moderators

Police Scotland Data Science Strategy Workshop Guide & Conversation Guide

Semi-structured Conversation Guide

Introduction

This session will be recorded for the study. What you say will be anonymized and screened for approval before being shared. Please verbally confirm that you have received and understood the Participant Information Sheet and Informed Consent in advance and agree to this meeting being recorded now and again after we start the recording.

Warm-up

1. Personal introductions
    a. Can you describe your role for us?
    b. A typical day or week in your job?
2. For you, what is the most interesting and important question you want to take away from this study?
3. In what ways is Scotland Police community-oriented? How do you suspect this changes the data collected and the context of how it's captured?

Future expectations

We want to run you through a few questions to understand your expectations for successful and unsuccessful, reflecting on a specific use case. (Q1 - 4, Backcasting exercise)

1. What data do you have, and what do you want to do with it?
2. Could you describe in a story format a FAILED possible outcome within the next 10 years?
    a. What is the use case here, and how is it failing?
    b. Q&A
3. Could you describe in a story format a VERY SUCCESSFUL possible outcome within the next 10 years?
    a. What is the use case here, and how is it achieved?
    b. Q&A
4. What should be considered to avoid the worst future and work towards the best future?
    a. What relationship with officers needs to have with the Data Sci team to create the best future?
5. Where do you see data science being useful, are there limitations to datafication?

Relationship with officers and training

- Describe how you see the future relationship between officers and data scientists within Police Scotland?
    - In the case of issues arising in data?
    - How much contact/interaction?
    - Embedded officers in data science?
    - Regional representation?
- What type of education/training is important for this future for..
    - Officiers
    - Data scientists
    - Other roles?

Wrapping up

- Any questions or comments you have for us?

Informed Consent

Concerning the participation in an interview and the analysis of the resulting data for reports and academic publications

By signing this form, I consent to participate in a workshop/interview for **Police Scotland community policing data strategy** with the University Name. The session will be recorded and transcribed. I have read and understood the **Participant Information Sheet.** All data will be handled with strict confidentiality. I am aware and consent that quotes from the interview may be cited in publically accessible reports and academic publications, however, only in anonymized form. My identity will not be revealed, and Police Scotland will approve the report before publication. The data from this workshop/interview will be kept in a secure location on password-protected computers and servers. Only researchers for this project will have access to the data. My answers will not be used to evaluate my performance at work.

If I decide to withdraw my participation and end the workshop/interview I can do that at any point in time during the workshop/interview. In that case, the recorded file will be destroyed and not used for analysis.

Name of the participant: _____

Date of the session: _____

_____ _____

Signature of the Participant Signature of the Interviewer

Informed Consent

Concerning the participation in an interview and the analysis of the resulting data for reports and academic publications

Participant Information Sheet

Thank you for your interest in participating in our research project focused on the data strategy of the Police Scotland. This participant information sheet aims to provide you with detailed information about the study, its objectives, the data anonymization process, and potential risks associated with your participation.

Title of Study: Data Strategy of the Police Scotland: Workshops and Interviews with Data Scientists and Officers

Purpose of the Study:
The primary objective of this research project is to gain insights into the expectations, challenges, and opportunities surrounding the incorporation of data and algorithms in community-oriented policing within Police Scotland. By conducting workshops and interviews with data scientists and officers, we aim to explore various aspects of the data strategy, including data collection, analysis, utilization, and the impact on community engagement. The findings will contribute to the development of effective strategies and recommendations to enhance data-driven community-oriented policing practices.

Study Procedures:

1. Workshops: The study will involve workshops where participants, including data scientists and officers, will engage in group discussions and collaborative activities. These sessions will provide an opportunity for you to share your experiences, perspectives, and expectations regarding the integration of data and algorithms in community-oriented policing. Workshops will be conducted in a safe and inclusive environment, promoting open dialogue and respectful interaction.
2. Interviews: Individual interviews will be conducted with participants to delve deeper into specific areas of interest identified during the workshops. The interviews will be conducted by the researchers, who will ask questions aimed at gaining a better understanding of your experiences, opinions, and suggestions related to the data strategy. The interviews will be audio-recorded to ensure accurate data capture and analysis.

Data Anonymization:

Informed Consent

Concerning the participation in an interview and the analysis of the resulting data for reports and academic publications

Protecting your privacy and confidentiality is of utmost importance to us. All collected data will be treated with strict confidentiality and stored securely. Here's how we will ensure data anonymization:

1. Anonymization of Personal Information: Any personally identifiable information (e.g., names, rank) will be replaced with unique codes or pseudonyms during transcription and analysis.
2. Confidentiality Measures: The research team will handle all data with utmost care and will not disclose any personal information outside the research team. Access to data will be restricted to authorized personnel only.
3. Data Storage and Retention: Collected data will be stored in secure password-protected systems. All data will be retained for a specified period as required by the research institution's guidelines and will be securely destroyed after the retention period.

Potential Risks:
We acknowledge that participating in this research project may involve potential risks and confidentiality concerns. These include:

1. Emotional Discomfort: Discussing sensitive topics related to data strategy and policing may evoke emotional discomfort or distress. You are free to withdraw from the study at any time if you find the discussions or questions distressing.
2. Misunderstanding: This study aims to understand the expectations of incorporating data in Police Scotland. It does not promise any action on a specific idea discussed during the session.
3. Confidentiality Risks: While we take every precaution to protect your confidentiality, it is important to note that complete anonymity cannot be guaranteed. However, all efforts will be made to minimize the risk of identification by removing any identifying information during the data analysis and reporting.

Your participation in this research study is voluntary. You have the right to withdraw your consent or discontinue participation at any stage without providing a reason. Your decision will have no impact on your relationship with Police Scotland or any associated personnel.

Contact Information:
If you have any questions, or concerns, or require further information regarding this study, please feel free to contact: researcher name and email address