

The UK Algorithmic Transparency Standard: A Qualitative Analysis of Police Perspectives

Marion Oswald (Alan Turing Institute, Northumbria University),* Luke Chambers (Northumbria University), Ellen P. Goodman (Rutgers University), Pamela Ugwudike (Southampton University), Miri Zilka (Cambridge University)

7 July 2022

Acknowledgements: The authors are grateful to all research participants who gave up their valuable time to contribute to this study. The authors would also like to thank Dr Adrian Weller who provided helpful feedback on an earlier version of this report.

*Dr Marion Oswald, MBE: PI and Corresponding author: moswald@turing.ac.uk; marion.oswald@northumbria.ac.uk. Other authors in alphabetical order.

Table of Contents

| | |
|---|----|
| Executive Summary..... | 4 |
| Background to the Report – Algorithmic Transparency and Policing..... | 6 |
| The Algorithmic Transparency Standard and the Pilot..... | 8 |
| Study Methodology..... | 8 |
| Thematic Results of the Study | 9 |
| Scope and Use of the Standard..... | 10 |
| Scope: which tools are covered by the Standard? | 10 |
| At which point of the development process should the Standard be completed? | 11 |
| What level of detail is required by the Standard?..... | 12 |
| Benefits of Police Participation in the Standard | 12 |
| Building public trust and confidence..... | 12 |
| Demonstrating legitimacy and openness..... | 14 |
| Other human rights infringements..... | 15 |
| Perception Risk and Data Disclosure Issues..... | 16 |
| Perception risk..... | 16 |
| Data disclosure issues..... | 17 |
| Innovation and Commercial Sensitivities..... | 18 |
| Risk of discouraging technology development | 18 |
| Supplier responsibilities..... | 19 |
| Trade secrets and commercial sensitivity | 20 |
| Explainability, Ethical Scrutiny and Evaluation | 20 |
| Explaining the technology/explainability | 20 |
| Ethical scrutiny..... | 21 |
| Bias and accuracy testing | 22 |
| Resourcing and Implementation Concerns..... | 23 |
| Resource (human, financial) required to comply with the Standard..... | 23 |
| Risk of increasing FOI requests..... | 24 |
| Conclusions: Rewards, Risks and Challenges for the Police, and Improvements to the Standard | 24 |
| Rewards and risks of the Standard for police forces | 24 |
| Ways in which the Standard process could be amended and improved..... | 25 |

Suggested Next Steps..... 26
Annex A: Interview Guide 27
Annex B: Scope of the current Standard compared to policing use cases 28

Executive Summary

1. The UK Government's draft 'Algorithmic Transparency Standard' is intended to provide a standardised way for public bodies and government departments to provide information about how algorithmic tools are being used to support decisions. The research discussed in this report was conducted in parallel to the piloting of the Standard by the Cabinet Office and the Centre for Data Ethics and Innovation.
2. We conducted semi-structured interviews with respondents from across UK policing and commercial bodies involved in policing technologies. Our aim was to explore the implications for police forces of participation in the Standard, to identify rewards, risks, challenges for the police, and areas where the Standard could be improved, and therefore to contribute to the exploration of policy options for expansion of participation in the Standard.
3. Algorithmic transparency is both **achievable for policing and could bring significant rewards**. A key reward of police participation in the Standard is that it provides the opportunity to demonstrate proficient implementation of technology-driven policing, thus enhancing **earned trust**. Research participants highlighted the **public good** that could result from the considered use of algorithms.
4. Participants noted, however, a **risk of misperception** of the dangers of policing technology, especially if use of algorithmic tools was not appropriately compared to the status quo and current methods.
5. Participation in the Standard provides an opportunity to develop increased sharing among police forces of **best practices (and things to avoid)**, and increased thoughtfulness among police force personnel in building and implementing new tools. Research participants were keen for compliance with the Standard to become an integral part of a holistic system to **drive reflective practice** across policing around the development and deployment of algorithmic technology. This could enable police to **learn from each other, facilitate good policy choices and decrease wasted costs**. Otherwise, the Standard may come to be regarded as an administrative burden rather than a benefit for policing.
6. Several key areas for **amendment and improvement** from the perspective of policing were identified in the research. These could improve the Standard for the benefit of all participants. These include a need for clarification of the **scope** of the Standard, and the **stage of project development** at which the Standard should apply. It is recommended that consideration be given to a '**Standard-Lite**' for projects at the pilot or early stages of the development process in order to gain public understanding of new tools and applications. Furthermore, the Standard would benefit from a more substantial **glossary** (to include relevant policing terms) and additional guidance on the level of detail required in each section and how accuracy rates should be described, justified and explained in order to ensure **consistency**.
7. The research does not suggest any overriding reason why the Standard should not be applied in policing. Suitable exemptions for **sensitive contexts and tradecraft** would be required, however, and consideration given to ensuring that forces have the **resources to comply with the Standard** and to respond to the increased public interest that could ensue. Limiting the scope initially to tools on a **defined list** (to include the most high-risk tools, such as those that produce individualised risk/predictive scores) could assist in mitigating concerns over sensitive policing capabilities and

resourcing. A **non-public version** of the Standard for sensitive applications and tools could also be considered, which would be available to bodies with an independent oversight function.

8. To support police compliance with the Standard, **supplier responsibilities** – including appropriate disclosure of algorithmic functionality, data inputs and performance – should be covered in procurement contracts and addressed up front as a mandatory requirement of doing business with the police.

9. As well as contributing to the piloting of the Standard, it is recommended that the findings of this report are considered at NPCC level, by the College of Policing and by the office of the Chief Scientific Advisor for Policing, as new sector-led **guidance, best practice and policy** are developed.

Background to the Report – Algorithmic Transparency and Policing

Algorithmic and AI technologies are increasingly informing decision making in justice systems including those in the UK, Australia, Canada, and the United States.¹ Among the technologies currently in development or use across these jurisdictions are data-driven algorithms deployed by police forces for proactive and reactive crime control activities. Transparency relates closely to police accountability for data practices that could affect human rights, particularly rights in the following areas: privacy; fair trial; protection from discrimination; liberty and security; and freedom of assembly, and the capacity of the tools to entrench biases experienced by historically disadvantaged communities.²

As government bodies are expected to operate with social licence, efforts are being made to render public sector use of algorithmic tools more transparent.³ Recent evaluations of public opinion reveal that such transparency is considered vital.⁴ Studies of how people evaluate the legitimacy of the police in the UK⁵ and internationally⁶ similarly show that transparency (or explaining policing decisions to demonstrate trustworthy motives) enhances the perceived legitimacy of the police and acceptance of decisions as procedurally fair, even if those decisions are unfavourable. Transparency can improve trust in the motives of the police and enhances legitimacy and normative acceptance of police directives.⁷ For instance, a recent study in the UK found that trust and legitimacy are essential for public acceptance of live facial recognition technologies used by the police.⁸

Several jurisdictions have mandated levels of algorithmic transparency for government bodies. Canada has mandated disclosure of source code for government-owned AI.⁹ New York City requires an Algorithms Management and Policy Officer to develop a process for city agencies to report information

¹Metropolitan Police Department, 'ShotSpotter Data, Disclaimer and Dictionary' 2022 <https://mpdc.dc.gov/publication/shotspotter-data-disclaimer-and-dictionary>; Metropolitan Police, 'Facial Recognition' <https://www.met.police.uk/advice/advice-and-information/fr/facial-recognition>; Office of the Privacy Commissioner of Canada, 'Police use of Facial Recognition Technology in Canada and the way forward' 2021 https://www.priv.gc.ca/en/opc-actions-and-decisions/ar_index/202021/sr_rcmp/; New South Wales Government, 'Facial Recognition' n.d. https://www.police.nsw.gov.au/crime/terrorism/terrorism_categories/facial_recognition; Office of the Australian Information Commissioner, 'AFP ordered to strengthen privacy governance' 2021 <https://www.oaic.gov.au/updates/news-and-media/afp-ordered-to-strengthen-privacy-governance>.

² Sarah Brayne, 'Big Data surveillance: The case of policing', *American Sociological Review* (Vol. 82, No. 5) pp. 977–1008; Matthew Browning and Bruce Arrigo, 'Stop and risk: Policing, data, and the digital age of Discrimination', *American Journal of Criminal Justice* (Vol. 46, No. 2, 2021), pp. 298–316; Rebecca A. Johnson and Tanina Rostain, 'Tool for surveillance or spotlight on inequality? Big Data and the law', *Annual Review of Law and Social Science*, (Vol. 16, No. 1), pp. 453–472.

³ Pascal D. König, Georg Wenzelburger, 'The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it', *Technology in Society*, Volume 67, 2021, <https://doi.org/10.1016/j.techsoc.2021.101688>.

⁴ CDEI, 'BritainThinks: Complete transparency, complete simplicity' 2021 <https://www.gov.uk/government/publications/cdei-publishes-commissioned-research-on-algorithmic-transparency-in-the-public-sector>.

⁵ Kathryn Farrow, 'Policing the Pandemic in the UK Using the Principles of Procedural Justice', *Policing: A Journal of Policy and Practice*, (Vol. 14, No. 3, 2020), pp. 587–592.

⁶ Tom Tyler, 'Procedural justice and policing: A rush to judgment?', *Annual Review of Law and Social Science*, (Vol. 13, 2017), pp. 29–53.

⁷ Tyler (n. 6).

⁸ Ben Bradford, Julia A. Yesberg, Jonathan Jackson, and Paul Dawson, 'Live facial recognition: trust and legitimacy as predictors of public support for police use of new technology', *British journal of criminology*, (Vol. 60 No. 6), pp. 1502–1522.

⁹ Treasury Board of Canada Secretariat, Government of Canada. (2019). Directive on Automated Decision-Making. Available at: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>.

about algorithmic tools they use, and to maintain a public-facing portal where the public can access this information. This directory is relatively bare bones, providing the name of the agency reporting the tool, the tool name and usage date, and narrative descriptions about the tool's purpose and how it functions to aid the agency in making decisions.¹⁰ A number of cities have established registries of the AI applications their agencies use. AI registries ideally document 'the decisions and assumptions that were made in the process of developing, implementing, managing and ultimately dismantling an algorithm.'¹¹ Helsinki and Amsterdam were the first to publicize these registries.¹² Theirs describe what, where, and how their cities use AI applications as information about training data, bias and risk assessments, and the role of humans in using the applications.

In the UK, the House of Lords Justice and Home Affairs committee inquiry report *Technology Rules? The advent of new technologies in the justice system*¹³ found the following:

'There are no systematic obligations on individual departments, public bodies, and police forces to disclose information on their use of advanced technological solutions. It is impossible for Parliament, press, academia, those responsible for procurement and—importantly—those subject to their use to scrutinise and challenge the use of technological solutions as they cannot know who is using what, for how long, for what purpose, or with what safeguards. This risks undermining trust in the police, the justice system, and the rule of law. (Paragraph 98)'

It was recommended that:

'Full participation in the Algorithmic Transparency Standard collection should become mandatory, and its scope extended to become inclusive of all advanced algorithms used in the application of the law that have direct or indirect implications for individuals. This would have the effect of turning the collection into a register. Engaging with this register will require additional and dedicated resourcing... (Paragraph 112)'

The Government's response to the above report states that it 'will continue to pilot and gather feedback on [the Algorithmic Transparency Standard] to explore options for how to deploy the standard more broadly.'¹⁴ This study aimed to explore the police's view of the draft Algorithmic Transparency Standard and the implications of participation for police forces, and therefore to contribute to this exploration of options.

¹⁰ NYC. Algorithms Management and Policy Officer. Available at: <https://www1.nyc.gov/site/ampo/index>. Page (An algorithmic tool is defined as "a partially or fully automated computerized system that uses an algorithm or series of algorithms to turn data ("input") into a result ("output") to be used to make a prediction, determine a course of action, or otherwise influence decision-making.").

¹¹ Meeri Haataja, Linda van de Fliert & Pasi Rautio, Public AI Registers: Realising AI transparency and civic participation in government use of AI 3 (Saidot Sept. 2020), <https://uploads-ssl.webflow.com/5c8abed-b10ed656ecfb65fd9/5ff6f334b49d5444079726a79AI%20Registers%20-%20White%20paper%201.0.pdf> [<https://perma.cc/8V6B-FKH2>].

¹² City of Helsinki. (2020, 28 September). Helsinki and Amsterdam first cities in the world to launch open AI register. https://news.cision.com/fi/city-of-helsinki/r/helsinki-and-amsterdam-first-cities-in-the-world-to-launch-open-ai-register_c3204076. See also Ontario. Data Catalogue. Available at: <https://data.ontario.ca/group/artificial-intelligence-and-algorithms>.

¹³ *Technology Rules? The advent of new technologies in the justice system* (1st Report of Session 2021-22 - published 30 March 2022 - HL Paper 180 see 'Summary of conclusions and recommendations' <https://publications.parliament.uk/pa/ld5802/ldselect/ldjusthom/180/18002.htm>.

¹⁴ <https://committees.parliament.uk/publications/22773/documents/167387/default/>, para 31.

The Algorithmic Transparency Standard and the Pilot

The UK Government's draft 'Algorithmic Transparency Standard' ('the Standard') was launched in November 2021 by the Cabinet Office's Central Digital and Data Office.¹⁵ Responding to calls nationally and internationally for greater transparency in the use of algorithms and data-driven technologies as referenced above, the Government proposed that the Standard will 'promote trustworthy innovation by providing better visibility of the use of algorithms across the public sector, and enabling unintended consequences to be mitigated early on.'¹⁶

At the time of writing, the Standard consists of:

- 1) An excel/csv spreadsheet designed to be a standardised way of collecting and setting out the information;¹⁷
- 2) An Algorithmic Transparency Template mapping to the sections of the above spreadsheet, to guide public sector organisations through the information needed to complete the spreadsheet.¹⁸

It is proposed that submitted information would then be published proactively in a centralised location.

Both the process and the template documentation are currently in draft format. The Cabinet Office, together with the Centre for Data Ethics and Innovation (a UK government expert body, part of the Department for Digital, Culture, Media and Sport) have undertaken a 'piloting' process. This has consisted of working with public sector organisations willing to complete the template in respect of an algorithmic tool in development or use, and collating feedback on the process. The first two reports from the piloting process were published on 1 June 2022, alongside initial conclusions from the feedback.¹⁹ A number of police forces have engaged with this pilot.

Running alongside the pilot, our study has investigated the implications of the draft Standard for policing. We conducted non-attributable interviews with police personnel who would be engaged in the process of completing the Standard, and with commercial organisations working on policing algorithms, to investigate their views of the aims, process and content of the Standard. A guide to the questions that were explored is attached in Annex A to this report.

Study Methodology

Combined with a targeted review of literature focused on policing and public sector data analytics and transparency, we conducted semi-structured interviews with respondents from across UK policing and commercial bodies involved in policing technologies. 16 respondents participated in research interviews for the project: 10 representatives of UK police organisations and 6 representatives of commercial organisations, including large consulting firms and independent consultants. Interviews were conducted online during February and March 2022. These interviewees were approached due to

¹⁵ <https://www.gov.uk/government/news/uk-government-publishes-pioneering-standard-for-algorithmic-transparency>.

¹⁶ <https://www.gov.uk/government/news/uk-government-publishes-pioneering-standard-for-algorithmic-transparency>.

¹⁷ <https://www.gov.uk/government/publications/algorithmic-transparency-data-standard>.

¹⁸ <https://www.gov.uk/government/publications/algorithmic-transparency-template>.

¹⁹ <https://cdei.blog.gov.uk/2022/06/01/piloting-the-national-algorithmic-transparency-standard/>.

their involvement in policing technology development, and therefore are likely to be reasonably representative of police and commercial personnel who have expertise in this area. We used a semi-structured format for the interviews in order to ensure a broadly consistent line of questioning while permitting some flexibility for the research team to probe particular responses and respondent areas of expertise. An information sheet and guideline questions were sent in advance so that respondents could understand the purpose of the project and give their informed consent. The study received ethical approval via Northumbria University's ethical approval process.

A qualitative, participatory, practitioner-focused research approach was chosen to enable us to benefit from the experience of practitioners in the identification, assessment and diagnosis of contextual, operational and policy issues. Participants were approached based on their knowledge and experience of developing, using or managing algorithmic tools in the police using a purposive, selective sampling strategy. For the purposes of producing this report, we were bound by the timescales of the piloting process. Therefore, the study is exploratory only as we cannot say that data saturation was reached. However, we observed that latter interviews tended to raise similar themes to earlier interviews. We used an inductive approach to analyse interview data supported the use of NVivo. We identified recurring themes by way of a preliminary coding process, followed a more granular analysis to explore particular issues and patterns in further detail.

In this report when referring to interview data, we use anonymised codes to indicate the category of research participant as follows:

L = police/law enforcement respondent
C = commercial sector respondent

Although theoretical generalisability and in-depth insights are achievable and are key strengths of exploratory qualitative projects such as this, we would identify the following limitations to our study. We consulted with a number of forces and organisations; however, the overall range of forces consulted was limited and therefore our findings may not be generalisable to other law enforcement agencies. Not all interviewees had gone through the process of completing the Standard. Furthermore, respondents may have tended to be favourable towards transparency in the policing domain, and therefore our findings may not represent the full range of views on the Standard within policing and in particular views of Chief Officers and specialist units.

Thematic Results of the Study

Our research has revealed the following themes, issues and concerns relating to police engagement with the Standard. We have grouped these under five main themes:

- **Scope and Use of the Standard;**
- **Benefits of Police Participation in the Standard;**
- **Perception Risk and Data Disclosure Issues;**
- **Innovation and Commercial Sensitivities; and**
- **Resourcing and Implementation Concerns.**

Scope and Use of the Standard

Scope: which tools are covered by the Standard?

There was consensus amongst the interviewees that the scope of the Standard is unclear and not straightforward to interpret, in that it does not specify the range of tools and applications covered, as acknowledged by one police interviewee:

'Defining the threshold, defining when an Excel macro becomes something that we need to be transparent about, is part of the problem...much like other forces, I did find it difficult to define when this applies.'²⁰

Some interviewees found the definition too broad and were unsure if it meant to include non-operational algorithmic tools, used, for example, in HR or maintenance. In addition, interviewees were not sure if the Standard applies to all existing tools or only to new tools:

'I'm not entirely sure what they're after, whether they're just after new ones, or whether they're also after the myriad of tools that that are used by police every day.'²¹

Another police interviewee pointed out potential difficulties in filling out the Standard retrospectively:

'Do you apply retrospectively or do you draw a line in the sand and move forward. Any new technologies or any new use of algorithms we're going to disclose but we're not going to do this retrospectively.'²²

One commercial interviewee noted that the questions presented in the Standard itself suggest the scope is geared towards machine learning based tools rather than the wider range of police software:

'Some of the later questions seemed to assume that you were dealing with a system which learned from data or a system which was making decisions, but some of the things that potentially were in scope, like some of the technology we're looking at, doesn't have a learning model.'²³

Most interviewees leaned towards prioritising tools or applications having direct impact on individuals in the community, as highlighted by this police interviewee:

'In terms of policing activity, you know project insights which impact on people, which have the ability of or possibility of significantly impacting on an individual or group of individuals has to be part of this scrutiny.'²⁴

The need to tighten the scope seemed particularly important if the Standard was going to be mandated. One interviewee raised concerns that the extent and quality of participation may not be consistent across police forces:

²⁰ Project Interview with L6.

²¹ Project Interview with C5.

²² Project Interview with L5.

²³ Project Interview with C2.

²⁴ Project Interview with L10.

'If it was going to be mandated, there would need to be defined exemptions. I think that's where you may end up with disparity from force to force because as we see with the freedom of information exemptions, some forces will apply them more rigorously than others.'²⁵

At which point of the development process should the Standard be completed?

For new applications, interviewees were unsure at which point in the development process the Standard should be completed. One natural point that came up as an option was at later stages of testing, before operational deployment:

'So I think we've got a fairly classic development process anyway between development then testing and before we move into live production. And I suppose this standard would come towards the end of that testing phase. So we want to complete it before we move things into production or into live use because it's something we want to have undertaken before we use something in earnest.'²⁶

Some interviewees felt positively towards the idea of using the Standard, or a 'light-touch' version of the Standard, early in the development process in order to gain trust and public backing for new tools and applications early on. One police interviewee responded to this suggestion with the following:

'that could only be a positive thing and my immediate reaction to that was if we're considering doing something and there's an opportunity to submit a 'version light', so to speak, of a submission, and to get some early feedback around some of the implications, I think that has the potential to save a lot of hard work.'²⁷

Others were against putting information into the public domain at an early stage. As put by one of our commercial sector interviewees:

"we haven't been through an ethical board because we're not going live with it next week. It's not operational yet. It's an experiment form. It's in a controlled environment and so it hasn't even been decided that it will be operational yet. And so I would be a little bit wary about disclosing the development of something before it's actually been decided that it will be operationally used and before it's gone through the internal ethical considerations about whether it's appropriate."²⁸

Interviewees also highlighted that the Standard submission for a given application will likely need to be updated. However, the frequency at which updates would be appropriate is not clear. One commercial interviewee said this when discussing the appropriate update frequency:

'[regarding annual updates] in some cases that's going to be too slow and in some cases it's going to be too fast and the length of time has to be related to the tool that you're using, because a tool that's only used for 200 decisions a year, probably doesn't have the data to be

²⁵ Project Interview with L4.

²⁶ Project Interview with L6.

²⁷ Project Interview with L10.

²⁸ Project Interview with C2.

reviewed every year, and it's going to be very onerous. But a tool that's going to be used for a million decisions a week, probably needs reviewing more than annually.'²⁹

What level of detail is required by the Standard?

Although the Standard has structured questions, the interviewees were not clear on the level of detail required when answering the questions. When asked about how much detail they thought was appropriate, interviewees had differing opinions. Some interviewees felt that it would be appropriate to capture as much as possible within the Standard. One police interviewee explained the reasoning behind this viewpoint:

'my starting point would be that we would want to disclose as much as we possibly can because it defeats the point of being transparent if we, from the outset, actively seek to provide the minimum as opposed to providing what we can.'³⁰

Other interviewees, however, felt that the level of disclosure should be tailored to what is relevant for the general public. As articulated by one commercial interviewee:

'There will be certain pieces of information which are only relevant to certain stakeholders.'³¹

Benefits of Police Participation in the Standard

Building public trust and confidence

Some interviewees stressed that a key benefit of police participation in the Standard is that it provides the opportunity to demonstrate transparency and improve police legitimacy, crucial in England and Wales where 'policing by consent' is the prevailing model, with public trust and confidence the *sine qua non*³² of policing as acknowledged by this police interviewee:

'[...] We police by consent, we need the trust and confidence of the community in order to exercise our duties and our powers effectively and in order for us to have that legitimacy, we need to be transparent in everything that we do, not just with the use of algorithms. So I think this feeds into that broader need, whether it's the use of force, whether it's stop and search, whatever the activity is within policing. I think it's really important that we from the outset involve individuals, explain, listen to views, and reassure.'³³

This interviewee also recognised that public input is necessary. Studies of police legitimacy have found that giving people the opportunity to participate in policing decisions that can affect their lives is another key antecedent of procedurally fair treatment and perceived legitimacy.³⁴

One commercial interviewee argued that transparency was key to addressing public anxieties:

²⁹ Project Interview with C5.

³⁰ Project Interview with L10.

³¹ Project Interview with C6.

³² College of Policing, 'Professional standards: Authorised Professional Practice', 2022 <https://www.college.police.uk/app/professional-standards>.

³³ Project Interview with L10.

³⁴ Lorraine Mazerolle, Sarah Bennett, Jacqueline Davis, et al., 'Legitimacy in Policing: A Systematic Review', *Campbell Systematic Reviews*, (Vol. 9, No. 1, 2013) pp. i-147; Tyler n. 20.

'If we thought that we wouldn't disclose something because a particular individual or group of individuals might be worried that we're using it, I agree that that's absolutely the reason that we should be really onerous around that. And that's something that I would seek to address head on. If people are worried, there's an opportunity there as opposed to a risk, there's an opportunity to be open. There's an opportunity to be transparent. There's an opportunity to explain, to reduce concerns.'³⁵

Another police interviewee wondered about the Standard's cost/benefit ratio and whether compliance with the disclosure requirements of the Standard will in fact generate public and stakeholder acceptance of policing algorithms:

'If this goes ahead, then I do think there should be a review at some point of the level of effort required by forces against the actual take up from the public and interested parties, and I know these things can take a few years before they really gain ground. But, does it actually become a bit of a white elephant?'³⁶

An additional benefit of police participation in the Standard was highlighted by a police interviewee who stressed that transparency can improve the technical proficiency of policing technologies, which also serves to augment confidence and acceptance:

'Every algorithm that's produced with this sort of technology or approach in criminal justice should have a transparency standard attached to it, because that's the only way you start to gain industry standards and confidence.'³⁷

Another police interviewee similarly emphasised that compliance with the Standard could enable forces to demonstrate proficient implementation of technology-driven policing and this can enhance police legitimacy, particularly in relation to applications of policing technologies:

'So I think that the main thrust for me of the necessity of this piece of work is around the public competence and legitimacy side of how we blend technology with policing.'³⁸

The importance of transparency was reinforced by yet another police interviewee who noted that openness is particularly necessary when developing predictive algorithms:

'I don't think it would be useful to have transparency about the fact that police map crime. I would have thought that the majority of people would have expected police to be mapping crime as in crimes that have occurred - if it was about where crimes may occur in the future, that, to my mind, is worthy of going through this transparency process.'³⁹

This interviewee seems to imply that *predictive* algorithms pose greater ethical challenges. Existing research on the disparate outcomes associated with such algorithms suggest that ethnic minorities can

³⁵ Project interview with C6.

³⁶ Project Interview with L2.

³⁷ Project Interview with L8.

³⁸ Project Interview with L7.

³⁹ Project Interview with L1.

be particularly affected.⁴⁰ Providing information about efforts to prevent or address biases can help build trust and improve police relations with affected communities.

Demonstrating legitimacy and openness

An emerging theme seemed to be that openness via transparent algorithm design and implementation is a key benefit of participating in the Standard. A number of interviewees expressed this view, and in one case, it appeared to be partially influenced by experience:

‘it is essential for there to be transparency in the police use of algorithms. If I cast my mind back to prior to me starting this role, there are some things that I know now that I would never have considered before my [current] role, which has really highlighted to me how important it is to be transparent in how we’re creating the algorithms and what we’re ultimately doing with them.’⁴¹

One commercial sector interviewee, while proposing that transparency be limited particularly in relation to data practices, thought that even a limited approach would yield dividends if the information provided by police forces evolves into a repository of data-driven models and facilitates more efficient development of data-driven tools and thus potentially ‘deduplication across the different portfolios.’⁴²

The interviewee was nevertheless of the view that disclosure should be carefully targeted according to the requirements of specific stakeholders:

‘we can share probably a very low level of detail with the likes of an Ethics Committee, how we’re doing all of this analysis and where that information is coming from and how that information was collected.’⁴³

‘There will be certain pieces of information which are only relevant to certain stakeholders.’⁴⁴

As already noted, one interviewee did point out that a benefit of participating in the Standard is that it creates the opportunity for developing a repository of information on policing algorithms so that forces can learn from each other.⁴⁵

‘If there's a public register from our UK policing perspective as a single force, you then get to understand what all the other forces are doing initially, which is quite difficult to grasp at the moment and learn from their successes and failures and in between, and I think that's going to be valuable for every force.’⁴⁶

Other interviewees pointed out that the Standard can drive more reflective practice within forces as they consider potential downsides of algorithmic deployment:

⁴⁰ Danielle Ensign, Sorelle A. Friedler, Scott Neville, et al. ‘Runaway Feedback Loops in Predictive Policing’, *Proceedings of Machine Learning Research*, (Vol 81) pp. 1–12; Kristian Lum and William Isaac ‘To Predict and Serve?’, *Significance* (Vol 13), pp. 14-19.

⁴¹ Project Interview with L10.

⁴² Project Interview with C6.

⁴³ Project Interview with C6.

⁴⁴ Project Interview with C6.

⁴⁵ Project Interview with C6.

⁴⁶ Project Interview with L5.

‘Even if they're not real negative outcomes, I think identifying potential negative outcomes, which is the thing that we tried to go through with our own process, - we're like “well in the worst case scenario what's the worst possible impact of using this tool?” - and kind of running through some scenarios. So I think that might be a useful part of the standard.’⁴⁷

For police forces to realise the potential of this kind of reflective practice, one interviewee thought that the Standard alone was insufficient and more oversight would be needed:

‘It might also be possible to create a peer reviewing mechanism of sorts. I feel that without this oversight, we will be missing the main benefit of transparency; the ability to ensure that algorithms used in the public sector are up to the task, and are being built properly, with proper attention to the data, and to the drift that occurs with algorithms. Only through this mechanism will organisations and data scientists be forced to remain up to date with the rapid growth in the field of machine learning, especially in the areas of safety, data protection, fairness/bias and machine learning ethics.’⁴⁸

The reflective data and algorithmic practices described here are increasingly encouraged by researchers and others who highlight the merits of *ex ante* and *ex post* algorithm audits.⁴⁹

Other human rights infringements

In discussing the benefits of participation, alongside the potential to build public trust and legitimacy, respondents also considered the opportunity to reflect on privacy rights. For instance, one interviewee felt that participation in the Standard can encourage police forces to consider current and future privacy concerns which will intensify as key aspects of human life become increasingly digitised:

‘we have to not only look at our considerations of people’s expectations of privacy and security now but moving into the future because I think that our future selves will have an even higher expectation of privacy as more and more of our life starts to occupy a digital space.’⁵⁰

The same interviewee was of the view that addressing public apprehension towards surveillance should be the starting point for dealing with privacy issues and expectations:

‘I think that one of the first hurdles to get over is an increasing fear of surveillance in society.’⁵¹

This view underlines the potential adverse impact of discriminatory and unwarranted surveillance associated with certain policing algorithms such as live facial recognition technologies⁵² and predictive

⁴⁷ Project Interview with L6.

⁴⁸ Project Interview with C5.

⁴⁹ Shea Brown, Jovana Davidovic and Ali Hasan, ‘The algorithm audit: Scoring the algorithms that score us’, *Big Data & Society* (Vol. 8, No. 1) pp. 1-8; Inioluwa Deborah Raji, Andrew Smart, Rebecca N. White, et al. ‘Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing’, In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*: 33-44. New York: Association for Computing Machinery. <https://doi.org/10.1145/3351095.3372873>.

⁵⁰ Project Interview with L7.

⁵¹ Project Interview with L7.

⁵² Joy Buolamwini and Timnit Gebru, ‘Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification’, *Proceedings of Machine Learning Research*, (Vol 81), pp. 1–15; Peter Fussey and Daragh Murray, ‘Independent Report on the London Metropolitan Police Service’s Trial of Live Facial Recognition Technology’ Project Report. University of Essex Human Rights Centre, 2019 <http://repository.essex.ac.uk/24946/>.

models.⁵³ As already discussed, both pose human rights implications and can trigger legal action as well as negative publicity, risking the undermining of public trust and legitimacy.

Perception Risk and Data Disclosure Issues

While a considerable majority of interviewees, whether from the police and the commercial sector, could perceive some benefit from engagement with the Standard, a number raised reputational and operational considerations. We have categorised these considerations as ‘perception risk’ and ‘data disclosure issues.’

Perception risk

Interviewees highlighted the public good that could result from the considered use of algorithms. For instance, one police interviewee highlighted problems with current practices in relation to domestic abuse, and the potential for an algorithmic approach to improve those practices:

‘...quite a lot of our homicides are domestic abuse related and we're quite bad at seeing the risk quite often because the police officers are dealing with a presenting incident there and then, not the full background, and this is where I think algorithms can help understand the more contextual picture.’⁵⁴

While respondents appreciated a need to focus on the risks and possible adverse effects of algorithms, they also thought it important to consider how data analytics could improve policing and how *not* exploring algorithmic approaches could produce an unwanted result - ‘the victim may be exposed to more harm’.⁵⁵

This factor links to a concern expressed by some interviewees, as mentioned in the preceding sections - that information provided by the Standard could give an incorrect or misleading impression of the technology, particularly at an early stage, thus unduly heightening public concern, stoking protests, and inhibiting investment in innovation⁵⁶. One policing interviewee encapsulated this point as follows:

‘there's a huge risk of giving people sufficient information to get concerned, but not enough to actually satisfy themselves. It's not as bad as they think it may be... it's really quite a real risk of giving out too much, too soon before we're in a position to defend it and answer it.’⁵⁷

Discussing a shortcoming of *unlimited* disclosure, another commercial interviewee remarked that although it can enhance transparency and legitimacy, excessive disclosure can also provide false assurance to the public and paper over potentially harmful impacts:

‘My main concern stems from my feeling that there are two benefits from transparency. The first is the availability of information to people who the data relates to, and for fueling legitimacy and openness about the kinds of things that algorithms are used for. However, this type of transparency will raise comments based on the perceptions of use, rather than the quality of the work, and if we are not careful, poorly constructed algorithms could remain in

⁵³ Rashida Richardson et al. ‘Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice’ (Vol 94 New York University Law Review 2019), pp. 192–233.

⁵⁴ Project interview with L9.

⁵⁵ Project interview with L9.

⁵⁶ Project interview with C6.

⁵⁷ Project interview with L7.

use in areas where members of the public do not think there will be knock-on impacts or future problems.’⁵⁸

This interviewee highlights the important issue of meaningful disclosure or legible explanation required for adequate public understanding of the quality and impact of an algorithm. A police interviewee highlighted the importance of anticipating and preparing for questions about new technology deployments: ‘if you release information, the natural question is ‘Is it working? What are the benefits? How is it being used?’, and at the point that we release the information, we have to be able to answer those questions.’⁵⁹

This issue links directly to the discussion above around the scope and content of the Standard and the question of the point in an algorithm’s development at which the Standard should apply (and whether a ‘Standard-lite’ for early stages could be beneficial). One commercial interviewee was of the view that public transparency during a project’s developmental stages could be problematic:

‘whether it's automated or completely automated, or it's assistive or augmenting human decisions, I feel that being completely open and transparent when you're in the very early stages of development tend to be counterproductive, both from a potential IP perspective and also from a public attitude.’⁶⁰

Data disclosure issues

Understandably, given the operational space in which the police operate, interviewees discussed sensitive policing contexts and capabilities, for instance counter-terrorism and covert policing. Concerns focused upon safeguarding and effectiveness, rather than on public perception:

‘if we were to disclose something, the issue is not that people would be worried, but more that there would be an impact on an individual's safety or an impact on the ability of an organisation to target individuals effectively...I would be more than happy to give a strong voice as to why that shouldn't be disclosed, but only if it was from the perspective of significantly impeding the ability of an organisation to safeguard and prosecute.’⁶¹

It was suggested that providing the level of detail required by the Standard might allow certain algorithms to be ‘gamed’, such as algorithms that focus on solvability of crimes due to their simplicity and small number of input factors: ‘If you knew all of that [what the algorithm does], you could commit a crime, being very careful around very few pieces of evidence.’⁶² It was argued however that ‘gaming’ was not likely to be a significant issue in other contexts: ‘in domestic abuse I don't see it being gamed anyway, not effectively because emotion is too high often in these cases.’⁶³

⁵⁸ Project Interview with C5.

⁵⁹ Project interview with L10.

⁶⁰ Project interview with C3.

⁶¹ Project interview with L10.

⁶² Project interview with C5.

⁶³ Project interview with C5.

Furthermore, any requirement to disclose *all* algorithms publicly was said to risk revealing capability that the police *do not have*,⁶⁴ thus breaching the operational principle of ‘neither confirm nor deny’ applied to sensitive operational techniques.

A non-public version of the Standard, linked to other methods to ensure legitimacy such as independent oversight, was suggested as a method of tackling these concerns:

‘it feels like this would be one of those areas where you could have a effectively a covert version of the standard so that we've gone through the necessary steps and with exactly the same rigour, but it just simply isn't communicated.’⁶⁵

Other interviewees supported a tiered approach:

‘maybe different levels of information being shared from the point of view of transparency for those different audiences because it wouldn't necessarily be appropriate to share certain police methods with the general public because this could undermine their efficacy and things like that. So I see two different levels and types of transparency.’⁶⁶

Innovation and Commercial Sensitivities

Risk of discouraging technology development

Interviewees identified a potential trade-off between implementing the Standard and adopting useful new tools because the rigours of the Standard would raise barriers to innovation. Of note are two kinds of barriers. One is the burden the Standard process places on public officials. According to one commercial sector interviewee:

‘The barrier might be that people stop designing tools because they can't be bothered to go through it, which would probably be a bad thing given what we're coming across in [] where the new tool will be an awful lot better than what currently exists.’⁶⁷

A police interviewee thought that the stringent review process is ‘more likely to lead to responsible use of data’ but at the same time, might reduce recourse to ‘tried and tested solutions to safeguard vulnerable people.’⁶⁸ Of course this comment assumes the tools under review are actually tried and tested, which many of the relevant tools are not.

The second kind of burden the Standard process might place on innovation involves private sector partners:

‘I suppose very tight regulation and very high levels of regulation could potentially stifle some of the innovative work that could happen and also potentially be a bar to entry for some small startup companies because it could actually become very expensive to be able to comply.’⁶⁹

⁶⁴ Project interview with L7.

⁶⁵ Project interview with L7.

⁶⁶ Project interview with C2.

⁶⁷ Project Interview with C5.

⁶⁸ Project Interview with L5.

⁶⁹ Project Interview with C4.

The question of how technological adoption relates to underlying mission is a complicated one, interviewees noted. While it may be the case that new tools are better than the status quo, it may also be the case that foregoing new tools in favour of improving existing tools is in the public interest. For example, one commercial interviewee noted that:

‘Ultimately at the end of the day, time and investment is very, very precious and we have to balance ‘do we spend X amount of time trying to utilise this data set which could be slightly controversial or do we use that same amount of time in making much more progress with the information that we know is being actively used by the police on a daily basis?’⁷⁰

Supplier responsibilities

Among interviewees who are working with a third-party supplier to develop and implement algorithmic tools, it was apparent that the responsibilities for explaining and otherwise making the tools transparent reside jointly in the supplier and the public sector customer. The balance of responsibility may depend on how much involvement the force has in developing the tool on the back end and then how much the force tailors the tool or its implementation on the front end.

‘it’s a joint responsibility to complete the standard between the customer [and the supplier]” with “slightly more on the customer than the supplier.” The supplier can provide a “technical description how it works but I think so much of the use case and the data and the application and how decisions are taken and supported” are on the customer.’⁷¹

For a commercial, off-the-shelf tool, one police interviewee said they would expect the supplier to provide a lot of information:

‘the technology companies can create some really complex, insightful products that go beyond the understanding of the customers and the forces and I think in doing that, there's a responsibility almost to help the force understand what they're getting into and to be able to on-board them into that process in as transparent a way as possible.’⁷²

It was seen as important to define up front how to allocate responsibility for assessing and explaining model design and performance over time:

‘we also want to make sure ... [there is clear definition] around actually whose responsibility is this, how does the model, how does an algorithm perform over time for example.’⁷³

One police interviewee suggested that police forces get help from the Police Digital Service in clarifying public and private sector responsibilities for transparency-related tasks:

‘Police Digital Service is the group helping police forces as a whole to buy the right stuff so they would be a good body to manage [transparency allocation with the private suppliers]. They will probably have far more information generally available to them and more ‘in’s’ with the private companies.’⁷⁴

⁷⁰ Project Interview with C6.

⁷¹ Project Interview with C3.

⁷² Project Interview with L7.

⁷³ Project Interview with C4.

⁷⁴ Project Interview with L2.

Allocating responsibility for transparency to private sector suppliers raises another issue, discussed below, which is the possible resistance to disclosure those suppliers might exhibit due to business concerns.

Trade secrets and commercial sensitivity

Interviewees identified private sector supplier resistance to transparency as a potential obstacle to implementation of the Standard. In particular, they highlighted concerns with proprietary trade secrets. In the absence of intellectual property protection, which algorithms are unlikely to have in the UK, there is 'nothing to prevent once you publish the actual algorithm itself, somebody else just taking it and using it or selling it...Until that changes, which I don't think it will, you wouldn't be able to convince some places to actually publish what they're doing.'⁷⁵

The degree to which supplier concerns will impede transparency may depend on the attitudes of those suppliers, their market power, and the contractual arrangements that address this issue. One of the suppliers who participated in the project said that it does not 'develop any software that we won't share with our police and partners.' This interviewee acknowledged that other suppliers are different: 'They do develop code and that is commercially sensitive. That's effectively the unique selling point and they're probably not going to be willing to share that code base.'⁷⁶

Explainability, Ethical Scrutiny and Evaluation

Explaining the technology/explainability

Concern was raised as to how best to explain technical capabilities in a succinct way that the general public can readily understand. One police interviewee stated:

'So if you're somebody who doesn't really know what master data management [*an example of technical terminology describing a technology enabled discipline*] is, then it's a quite tricky starting point to try and describe what you're doing with it without telling a big, big story...where do we draw the line?'⁷⁷

Another highlighted a risk that the Standard could result in information overload for the public:

'[...] I think we've probably provided more details than the public needs at the moment, or can handle.'⁷⁸

Explainability of the technology for the police as well as the public was highlighted as a desired outcome, with one interviewee discussing the value to internal users:

'We wanted from the start to bring in approach where we know what it's doing. We know why it's making the decisions it's making, and the cops that are using it will know why it's making the decisions it's making.'⁷⁹

⁷⁵ Project Interview with C5.

⁷⁶ Project Interview with C6.

⁷⁷ Project Interview with L5.

⁷⁸ Project Interview with L5.

⁷⁹ Project Interview with C5.

The public may also overestimate police abilities or infrastructure, and transparency laying bare true capabilities may impact public trust. One interviewee explored this aspect:

‘we don't have a single view of our or people's data which most forces don't, actually. Most medium, large organisations don't. However, the public will expect us to have that and that's what we're trying to do.’⁸⁰

The issue of public trust is reflected in the discussion during the interviews. Though attitudes towards transparency were generally positive, potential damage to public trust was a risk of compliance with the Standard that stood out to policing stakeholders, illustrating concern, not only with determining which information to share, but how to share it without causing confusion to the public. The above interviewee believed that the public was likely not to understand the terminology without substantial background being offered. Furthermore, the multiplicity of terms used by different police forces could result in inconsistency in respect of the completion of the Standard. This was expanded on later in the interview, with the interviewee responding positively to a suggestion for a glossary and stating:

‘it's essential because otherwise you're just going to get 43 different languages and slight variations or preferences of how people would describe things and I would probably go a step further and have common definitions of things like a nominal or what master data management is.’⁸¹

A commercial interviewee agreed that a ‘single considered, consolidated way of communicating’ via a glossary and/or methods of communication linked to the Standard would be of benefit.⁸²

Ethical scrutiny

Despite a significant level of agreement that ethical scrutiny was important, interviewees had a range of opinions on when ethical scrutiny should be applied and the range or scope of such scrutiny.

One commercial sector interviewee described their perceptions surrounding the future of projects once submitted for ethical approval:

‘A really quick tangible example is that we know that there's cell site data that exists and we wanted to consider that information to look at where mobile phones were pinging on masts and to be able to work out county lines and the movement of drugs and the movement of victims connected to that. We didn't go down that route because we knew that ethically, you know, adhering to all of the submissions and the papers and the associated discussions, it was probably going to be vetoed and put to one side so that that stifled our innovation.’⁸³

Although these fears of ethical scrutiny shutting down projects do not reflect the reality of ethical scrutiny processes such as those operated by West Midlands Police,⁸⁴ it is worth acknowledging that these fears could have a real effect. If forces or commercial providers were to choose not to innovate

⁸⁰ Project Interview with L5.

⁸¹ Project Interview with L5.

⁸² Project Interview with C2.

⁸³ Project Interview with C6.

⁸⁴ <https://ejlt.org/index.php/ejlt/article/view/883>.

in certain areas for fear of ethical scrutiny, a self-imposed chilling effect would be an unintended consequence of that scrutiny.

Other interviewees expressed uncertainty, not with the ethical scrutiny process itself, but regarding the point at which to disclose projects for that scrutiny:

‘at what point do you present something to the Ethics Committee? Because sometimes you haven’t really got that much to talk about, but people are going to want to know “How accurate is it? Because that leads into how much we support you doing it”. But we want to talk about developing it with you to see if you even get to support developing it.’⁸⁵

Other interviewees mirrored similar sentiments:

‘I would be a little bit wary about disclosing the development of something before it's actually been decided that it will be operationally used, before it's gone through the internal ethical considerations about whether it's appropriate.’⁸⁶

The prevailing concern regarding ethical scrutiny mirrors those expressed around the Standard itself.⁸⁷ Interviewees were uncertain or unclear about the point at which they would be mandated or recommended to pursue ethical scrutiny, and at which stage(s) in technical development it would be most valuable.

Bias and accuracy testing

Testing for accuracy and bias was not a topic that was covered in all interviews. However, this was touched upon by some of the interviewees. When discussing the current methods of algorithmic development and bias/accuracy testing, one commercial interviewee stated:

‘Most algorithms that have been designed in policing aren't tested for bias. They're not designed with any transparency. Mostly the people who built them don't know what they do, let alone anybody else.’⁸⁸

The same interviewee expanded on these issues at their written follow up:

‘There will also come a point where there are multiple algorithms being designed in one place, and those algorithms unwittingly actually build upon each other. This can lead to major issues of bias created through another mechanism, and also major increases in technical debt which may have serious implications later.’⁸⁹

On a high level, interviewees involved in policing demonstrated an interest in tracking bias and accuracy in algorithmic tool. However, the practicality of achieving this goal in practice is not always straightforward. As highlighted by one interviewee:

⁸⁵ Project Interview with L9.

⁸⁶ Project Interview with C2.

⁸⁷ Written interview follow up notes from C5.

⁸⁸ Project Interview with C5.

⁸⁹ Written interview follow up notes from C5.

‘Our interest is ensuring that we are in the conversation about how to make sure [algorithmic systems are] understood in terms of any biases or anything that might be unacceptable in terms of how we process data. This is for us part of engaging in that conversation and with something like the algorithmics standard.’⁹⁰

Our interview data suggests that the interest in detecting and mitigating bias is well aligned with the introduction of the Standard.

Some questions were raised about communication of accuracy levels and the impression that these might give to the public. One commercial interviewee stated:

‘[...] precision and recall rates are defined by various bits and pieces, but ultimately it's a percentage, and when you see that percentage in isolation, you can immediately perceive it to mean something that perhaps it doesn't. So if the precision or recall rate is 60%, some people might think ‘60% doesn't sound very good at all. That means that 40% is inaccurate or it's not working’, whereas actually, if there's some kind of industry standard or expectation that gives the reader an opportunity to benchmark that [figure] against what is roundly perceived as acceptable, and exceeding expectations. The context in which that is done is also equally important.’⁹¹

This reflects comments from other interviewees regarding the public response to the published standard. In this instance, this is grounded in the uncertainty of the public to be able to interpret metrics such as model accuracy. Some interviewees were concerned that the Standard may facilitate the publication of accuracy and other performance metrics without sufficient context to allow for a robust interpretation of those numbers.

Resourcing and Implementation Concerns

Resource (human, financial) required to comply with the Standard

Of significant concern to interviewees was the resource burden of complying with the Standard. The size and type of burden imposed depends on the kind of tool that is being used and at what stage of implementation the force undertakes to fill out the template. For example, one policing interviewee observed that the burden might not be so great for a discrete predictive tool ‘because we would be the people keeping track of the accuracy over time, just to make sure, because you need to keep track of that in order to know when/if your model needs rebuilding, for example, because you get model drift or something.’⁹² This contrasts with ‘some of the more strategic stuff,’ presumably involving more departments with less defined workflow over time.⁹³

Resource constraints may be such that, depending on the context of other force requirements, may simply prevent adoption of a tool or compliance with the Standard.

‘Depending on what else is happening, for example if there was a big event, there would be change freezes everywhere and ...people being pulled in all sorts of different directions, so if it

⁹⁰ Project Interview with C1.

⁹¹ Project Interview with L10.

⁹² Project Interview with L1.

⁹³ Project Interview with L1.

coincides with a big event, for example, you might just be saying... 'No, we're sorry, we just can't do this.'⁹⁴

One commercial interviewee was less concerned with resource demands, noting that the template codifies burdens that forces should be undertaking in any case: 'Sure it takes time to write down the answers to those questions and the answers to those considerations, but we should be doing that anyway.'⁹⁵

Risk of increasing FOI requests

While it might be expected that proactive publication of information would result in a reduction of the number of freedom of information (FOI) requests received by policing bodies on the topic of algorithmic tools, interviewees were not convinced by this assumption.

One policing interviewee told us that the police were receiving:

'an increasing number of FOIs coming from a variety of different sources - media and academia asking some quite detailed questions about police approach to AI algorithms, etc. Who's using what, who's trying what? What are you thinking about using? So certainly it's getting noticed and I think that there's a huge risk of giving people sufficient information to get concerned, but not enough to actually satisfy themselves that it's not as bad as they think it may be.'⁹⁶

Another policing interviewee however regarded dealing with FOI requests as 'an accepted part of doing business' and was less concerned about the potential resourcing implications.⁹⁷

Conclusions: Rewards, Risks and Challenges for the Police, and Improvements to the Standard

Rewards and risks of the Standard for police forces

Interviewees generally thought that the rewards of a carefully tailored Standard implemented at the right stage of algorithmic development outweighed the risks, provided that the identified challenges were addressed.

These **rewards** centre on:

- Opportunities to demonstrate the legitimacy of technology use by the police and build public confidence and earned trust - interviewees highlighted the public good that could result from the considered use of algorithms;
- Increased public interest in, and understanding of, policing algorithmic tools;
- Increased sharing among police forces of best practices (and things to avoid) and therefore an opportunity to improve policy choices and decrease wasted costs;
- Increased thoughtfulness among police force personnel in building and implementing new tools, and therefore improvement in the quality of technology deployments.

⁹⁴ Project Interview with L3.

⁹⁵ Project Interview with C5.

⁹⁶ Project interview with L7.

⁹⁷ Project interview with L3.

The primary **risks** are:

- An increased misperception of the dangers of policing technology if use of algorithmic tools was not appropriately compared to the status quo and current methods;
- A possible increase in public opposition and therefore increased pressure to turn away from useful innovation;
- A potential ability to 'game' an algorithm in limited contexts, although this risk was not seen as particularly significant by all interviewees and only applicable to certain applications.

The primary **challenges** are:

- To ensure that forces have the resources to comply with the Standard and also to respond to the increased public interest that could ensue;
- To ensure that supplier responsibilities to assist the police with compliance with the Standard are factored into commercial arrangements;
- To factor compliance with the Standard into other oversight processes including independent ethics oversight and DPIAs;
- To ensure that the Standard allows sufficient scope for explaining the issues around current methods and the potential for technology to improve the status quo.

Ways in which the Standard process could be amended and improved

Analysis of our interview data suggests the following key areas for amendment and improvement of the Standard from the perspective of policing. Many of these are of more general applicability and therefore could improve the Standard for the benefit of all participants:

- Interviewees agreed that the **scope of the Standard** was unclear and required considerable clarification regarding how to decide whether a tool or application was covered by the Standard. In order to explain this issue further, the **scoring mechanism set out in Annex B** illustrates the current scope of the Standard compared to police use cases and breaks down under which conditions each use case will be included in the priority scope;
- In order to provide clarity as to which tools are in scope for policing, and to deal with the resourcing concerns, an **exclusive list of tools in scope** could be produced (to include the most high-risk tools initially, such as those that produce individualised risk/predictive scores and those which inform operational deployments or evidential stages). This would enable the application of the Standard to be ramped up over time and would avoid forces having to interpret terms such as 'complex' processing;
- Many interviewees agreed that the Standard would benefit from a more substantial **glossary** (to include relevant policing terms) and additional guidance on the **level of detail** required in each section and how **accuracy rates** should be described, justified and explained in order to ensure consistency;
- Interviewees expressed different views on the **stage of project development at which the Standard should apply and at which point(s) submissions should be revised and updated**. This would need to be decided after further discussion with policing bodies. Benefits could arise from transparency at relatively early stages as set out above. We would therefore suggest that consideration is given to producing a '**Standard-Lite**' that could be deployed by police forces for such early stage projects;

- Clarity is needed as to whether the Standard will be **mandated** for policing (either now or in the future) and if so, how **exemptions and issues of sensitive tradecraft** will be handled. Our interview data does not suggest any overriding reason why the Standard should not be applied in policing, subject to suitable exemptions and reasonable flexibility due to resourcing pressures. Limiting the scope initially to tools on a **defined list** could assist in mitigating concerns over sensitive policing contexts and capabilities;
- In addition, consideration could be given to a **non-public version of the Standard**, available for sensitive applications and tools as tightly defined, and available for review, in order to ensure legitimacy, to bodies with an independent oversight function;
- In order for the Standard to contribute to **improving the quality of policing technology**, it should be linked to **methods of oversight and promotion of best practice on a national basis**, and used to **enable police forces to learn from each other**. Otherwise, the Standard may come to be regarded as an administrative burden rather than a benefit for policing;
- To support police compliance with the Standard, **supplier responsibilities** – including appropriate disclosure of algorithmic functionality, data inputs and performance - should be covered in procurement contracts and addressed up front as a mandatory requirement of doing business with the police.

Suggested Next Steps

There are considerable benefits open to policing from engagement with the Algorithmic Transparency Standard, provided that the risks and challenges laid out above are addressed. A key reward of police participation in the Standard is that it provides the opportunity to demonstrate proficient implementation of technology-driven policing, thus enhancing earned trust. Interviewees highlighted the public good that could result from the considered use of algorithms. Integrating the Standard as an integral part of a holistic system to drive reflective practice across policing around algorithmic technology could improve effective, responsible deployment and avoid wasted costs.

As well as informing the Cabinet Office/CDEI's ongoing piloting process and revisions to the Standard, we would propose that the findings of this report are considered at NPCC level (including by the Information Management & Operational Requirements Coordination Committee (IMORCC)), by the College of Policing and by the office of the Chief Scientific Advisor for Policing, as new sector-led guidance, best practice and policy are developed.

Annex A: Interview Guide

Introduction

Interviewer to provide a short summary of the project, introduce the team and provide assurance that responses will remain non-attributable.

- Before we begin, do you have any questions about the project?
- What is your role in respect of the development and use of policing algorithms?

Transparency and algorithms

- What is your view of the need for transparency in respect of algorithms used by policing?

Completing the draft standard

- What was your role in completing the draft standard? Who else needed to be involved? Do you think that others should be involved?
- Did you understand the definition of algorithmic tools that are covered by the standard?
- Did this definition exclude any algorithms that you use within the force?
- What is your overall view of the process of completing the standard?
- When do you think is the right stage for a tool to be disclosed via the standard? How often should this be updated? What resource implications would this have?
- Which sections were hard to answer and why?
- Did you have concerns about disclosing particular information in answer to any of the questions and why?

Benefits and concerns

- What organisational concerns were expressed about publication of the standard?
- What benefits could accrue from applying the transparency standard?
- What issues or challenges could accrue from applying the transparency standard?
- What strategic, practical and policy issues will need to be considered?

Wrap-up

- Do you have any questions about the interview?
- Interviewer to reaffirm that responses are non-attributable

Thank you for your time

Annex B: Scope of the current Standard compared to policing use cases⁹⁸

An algorithmic tool is defined within the Standard as any ‘product, application, or device that supports or solves a specific problem, using complex algorithms.’⁹⁹ Tools will potentially be in scope if they involve one or more of the following: i) complex statistical analysis, complex data analytics, or machine learning - for example neural nets or deep learning, ii) potential public effect and/or iii) replace or assist human decision-making.

To illustrate the challenges of deciding whether a specific tool is currently within the scope of the Standard (or within the priority scope), we have developed the following scoring mechanism based on the existing definitions. According to the scope, all algorithmic tools are generally in scope, but there is a narrower definition of priority scope. We can score algorithmic tools according to the following:

Tools that:

1. Engage directly with the public – 3 points.
2. Involve:
 - a. Complex statistical analysis – 1 point.
 - b. Complex data analytics – 1 point.
 - c. Machine learning – 1 point.
3.
 - a. Has a potential legal, economic, or similar impact on individuals or populations – 1 point.
 - b. Affects procedural or substantive rights – 1 point.
 - c. Affects eligibility, receipt or denial of a programme – 1 point.
4.
 - a. Replaces human decision making – 1 point.
 - b. Assists or adds to human decision making – 1 point.

To be included in the priority scope, a tool needs to score 3 points or above, either all from 1, or at least one point from each 2, 3, and 4.

Additional definitions not included in the Standard:

- The standard does not define ‘complex’ with respect to complex algorithms. In this analysis, we define ‘complex’ as an action that cannot be reasonably performed by a person in less than a minute.
- We consider any application where the rules of the algorithm are derived from the data as machine learning.

⁹⁸ Current definitions/scope of the Standard can be found at <https://www.gov.uk/guidance/provide-information-on-how-you-use-algorithmic-tools-to-support-decisions-pilot-version>.

⁹⁹ <https://www.gov.uk/government/collections/algorithmic-transparency-standard#definitions>.

General observations:

- The vast majority operational tools will receive a point from 4.a or 4.b, while general use tools, e.g., internet browsers, mail client, etc., will not. It is unclear whether tools used in supporting divisions such as HR, maintenance, etc, should receive a point from 4.
- All tools that may inadvertently contain biases should receive a point from 3.a. However, it is not clear if this will be interpreted in this way by all who read the Standard. This means that the majority of operational tools will receive at least one point from 3 as well.
- Following the above, the decision whether a tool falls within the priority scope will be decided by 2, i.e., will depend on the complexity of the analysis, which is not currently well defined within the scope.

To highlight what is caught by the current scope, we give illustrative examples of algorithmic tools and analyse whether or not they fall within the scope of the Standard.

According to the scope, all algorithmic tools are generally in scope, but there is a narrower definition of 'priority scope'. We will score the algorithmic tools according to rules presented above. To be included in the priority scope, a tool needs to score 3 points or above, either all from 1, or at least one point from each 2, 3, and 4.

Use case examples:

- **Crime mapping software:**
 - Visualization of past data only – if the tool only shows points on the map where past crimes (of a certain type) have occurred, it will likely fall within the broader definition of the scope but not within the priority scope.
 - 'Hot spot' prediction – where data on past crime is used to predict where and when future crime will occur will likely be within the priority scope:
 - 1 point from 2.c as these tools are likely utilise machine learning.
 - 1 point from 3.a because the prediction of hotspots in certain neighbourhoods may impact the level of policing in those areas, impacting the local population.
 - 1 point from 4.a as these tools are likely to assist in resource allocation and patrol planning.
 - Mapping software that includes data processing but does not make predictions may be included in the priority scope if 2.a, 2.b, or 2.c applies.
- **Data infrastructure software:**
 - May fall inside the broad category but are unlikely to fall within the priority scope due to 2, unless:
 - The tool contains a 'link matching' component, i.e., data from two or more sources are combined and records are matched in an automated way.
 - Any type of automated score calculation occurs within the tool, e.g., risk of violence associated with an individual.
- **Data analytics suites:**
 - These types of tools often contain a range of functions within one tool, ranging from data-filtering to mapping, and sometimes assigning risk scores to individuals or places. If

one of the functionalities is within scope, it is not clear whether the whole tool suite is then considered within the scope or just the specific functionality.

- **Individualised risk scores**
 - Algorithmic tools used to produce risk scores will often use machine learning to generate the predications, ensuring they fall within the priority scope. However, this may not always be the case. An individual may be assigned a risk score based on simple rules, in which case these tools will fall outside the scope. As 'complex' is not well defined in the scope, this may be down to the police force's discretion.
- **Solvability tools:**
 - As above, based on the current definition it will depend on whether or not the analysis is considered 'complex'.
- **Facial recognition tools:**
 - Falls within the priority scope as they use machine learning (2.c), may be biased (3.a), and assist human decision making (4.c).
- **Chatbots:**
 - Chatbots will be included in the priority scope as they interact directly with the public (3 points from 1).