Office for Artificial Intelligence, DCMS
100 Parliament Street
London, SW1A 2BQ
Via email: evidence@officeforai.gov.uk

AI Regulation Policy Paper (CP 728): Response to call for views and evidence

Written submission from Dr Marion Oswald, MBE.

Introduction

1. The submission reflects my practical experience and research in respect of the use of AI and data analytics in the public sector, and in particular within policing and criminal justice.

2. This submission sets out my personal views and does not represent the views of Northumbria University, the Alan Turing Institute, RUSI, West Midlands PCC, West Midlands Police, or the CDEI.

Important challenges and contextual approach

3. There are two key challenges that remain to be addressed in my opinion: first, the current lack of regulatory focus on the **use of AI within an operational context** as governed by existing law; and secondly, the importance of a **combined approach to governance of AI**, inter-linking the application of law, technical and ethical standards, empowered and accountable people, and oversight and regulation. By addressing these issues, **innovation will be supported** not hindered, and legal risk reduced. This is because innovators, and those procuring AI, will be given the knowledge and confidence to determine **what is good innovation and what is not** (and so what types of AI should be developed and procured and what should not be).

Operational context

4. Codes, guidance and regulation that treat AI as a technology in a vacuum, and which use self-evident high-level themes such as 'safety' and 'fairness', will have little or no positive effect, either on confidence to innovate or on legitimacy of use. It is crucial that AI is both developed and its use regulated with **its intended or actual operational use in mind**.

5. 'Old' administrative law principles already provide a framework by which the use of AI within public sector decision-making can be considered, as I explain here: 'in each algorithmic-assisted environment, a **context-specific and nuanced approach will be required so that the information and explanations provided to aid intelligibility, or the way the result is interpreted, enable the particular public task to be fulfilled in a legitimate manner**.'

6. Furthermore, AI tools are not deployed in a vacuum but become part of an operational decision-making system. Therefore, the **laws and regulations that govern that particular decision apply to the use of AI within that decision-making context**. For example, police powers of stop-and-search and arrest require **objective grounds for reasonable suspicion** based on facts, information and/or intelligence.

7. Codes and guidance have yet to address adequately the use of facial recognition technology in these specific contexts i.e. addressing if and in what circumstances/under what conditions the output of a facial recognition tool can be treated as objective grounds for reasonable suspicion justifying a stop-and-search or arrest. The answer to this question requires **detailed and specific understanding of the probabilistic nature of AI and how it works,** including consideration of issues such as error rates, sensitivity settings, likelihood of bias and the circumstances of deployment. ('Reasonable' appears elsewhere in the law as an important concept, yet is not defined in AI terms, thus emphasising the need for interpretation and guidance.)

8. In other words, **we need to know if the AI is any good for the specific context in which it will be deployed and the legal test which needs to be satisfied**. These determinations should then feed back into the technical development process, thus improving quality.

9. Another context which illustrates these challenges is the concerted effort in recent years to embed a 'public health' approach in criminal justice, underpinned by Government priorities focused on risk, harm

prevention/reduction, vulnerabilities, and youth and domestic violence. A key element is the skilled use of data science techniques, particularly at a population level and for individualised risk assessment and prediction.

10. However, the models of preventative actions and the consequences of acting on a **'false positive'** AI output will be significantly different in a public health context as opposed to a criminal justice one. In a health context, a person may be subjected to an additional test or examination to re-check their health condition or may be given unnecessary treatment. In a criminal justice context, a person may be recorded in a police system in a high-risk category, thus affecting the accuracy of those records and how that person is treated in the future, or the person may be subjected to additional surveillance, investigation or police intervention. There is therefore no one-size-fits-all answer to whether it would be legal, ethical and effective to incorporate a particular AI tool in these contexts, despite the 'public health' badging of the aims.

11. The cross-sectoral principles proposed in the policy paper – while useful as a high-level starting point – require considerable **enhancement and expansion** in order to address **complex real-life operational scenarios**.

Inter-linking law, standards, accountability and oversight

12. As I explain in this article, **technical, statistical, legal, contextual, operational and ethical aspects** of AI-informed decision-making are closely interconnected. We need to know **what the output means in the context of the operational decision to be taken, and we must assess the implications of how it will be used in practice**.

13. Take a hypothetical example of a public sector body which has been given a power to intervene 'if a child is reasonably determined to be at high risk of harm'. The legality of its power to intervene is therefore dependent upon this assessment (which can and often is tested in court). The body might decide to use an algorithm to help it decide on risk levels by way of analysis of hospital admission reports. But unbeknown to users, the model both has a **high level of false positives** and is based on machine learning textual analysis of terms (chosen by the commercial developer) that expert users would regard as **irrelevant** to their assessment of the risk (and missing factors in other records that users would regard as relevant). **So the tool is not in fact answering the question that needs to be answered by the human decision-maker and is based on irrelevant factors**. If the body defers to the tool and therefore intervenes in respect of a child assessed wrongly due to issues with the tool, then it will be acting outside its powers (not to mention reducing services to children at real risk of harm and causing unnecessary disruption, stress and embarrassment to the child and family).

14. The three-pillar approach set out below attempts to illustrate the inter-linking of all these factors and is further discussed here:

<table>
<tr><th colspan="3">AI-informed public sector decision-making</th></tr>
<tr><td>Application of Law aided by guidance and policy for the relevant context</td><td>Standards:
Scientific standards
Ethical standards attached to personal responsibility</td><td>People: recruitment, professional development, culture, senior leadership commitment, accountability of all</td></tr>
<tr><td colspan="3">Independent oversight and advice/scrutiny: 'productive challenge'
Regulation and enforcement</td></tr>
</table>

Practical implementation

15. There are a number of practical implementation challenges raised by the above three-pillar approach. First, the **application of relevant law** is not easy, as it often requires the 'read-across' of common law principles into new contexts and the translation of key legal requirements into **suitably precise policy and guidance**. Secondly, **scientific and ethical standards** must address the specific operational challenges that will be encountered in the particular context. More attention is therefore needed to the acquisition of 'off-the-shelf' commercial applications or models

created through the combination of a number of tools, none of which have been specifically developed or evaluated for the context in question. Thirdly, organisations require a culture, led from the top, which welcomes informed challenge, supported by **empowered staff** who understand and are committed to the underlying values that the law and ethical standards represent, and are prepared to be thoughtful, and engage in professional skills development (therefore requiring bodies such as the College of Policing to provide the means for such professional development).

16. It will be crucial, as the AI environment becomes ever more complex, that **independent scrutiny and regulation** is **'end-to-end'** i.e. that scrutiny - and thus accountability - becomes a **rolling process**, from project planning/initiation to eventual operationalisation, rather than being limited to ex-post review. New models of review and scrutiny will be needed, and lessons could be learned from the proceedings of the **West Midlands Police and Crime Commissioner and West Midlands Police data ethics committee** (the first of its kind in UK policing) which is an ongoing experiment in scrutinising and advising upon AI policing projects proposed for real operational environments. A national model based on the West Midlands prototype could contribute to the assurance and monitoring required, the development of necessary policy and proactive longer-term thematic review.

References:

Marion Oswald, **'**A Three-Pillar Approach to Achieving Trustworthy Use of AI and Emerging Technology in Policing in England and Wales: Lessons From the West Midlands Model' (2022) *European Journal of Law and Technology* Vol. 13 No. 1 https://ejlt.org/index.php/ejlt/article/view/883

Marion Oswald, 'Algorithmic-assisted decision-making in the public sector: framing the issues using administrative law rules governing discretionary power' (2018) *Phil. Trans. R. Soc. A,* 376:2128 https://royalsocietypublishing.org/doi/full/10.1098/rsta.2017.0359

Dr Marion Oswald, MBE
Associate Professor, Northumbria Law School
SRA, The Alan Turing Institute, AI Programme
Independent Advisory Board Member, Centre for Data Ethics & Innovation
Chair, West Midlands Police and Crime Commissioner and West Midlands Police data ethics committee
Associate Fellow, Royal United Services Institute for Defence and Security Studies

23 September 2022