Ada Lovelace Institute Response to the European Commission 'White Paper on Artificial Intelligence – A European Approach'

The Ada Lovelace Institute is an independent research and deliberative body with a mission to ensure data and AI work for people and society.

This document is submitted as part of our response to the consultation on the European Commission's 'White Paper on Artificial Intelligence – A European Approach'. We address three themes with six recommendations. These recommendations are based in our research programmes Rethinking Data, The Algorithmic State and Identities & Liberties.

Risk-based framework

Reject high risk as the sole determinant of new regulation

The risk-based framework, as presented in the White Paper, results in a two-tier distinction of 'high risk' and 'lower risk'. The former would have added regulatory requirements and the latter only voluntary guidelines. This approach has three significant limitations:

1) **No consideration of impermissible uses of AI**

   Some applications of AI technology may be incompatible with our fundamental rights and societal values. In categorising AI by risk alone, the White Paper regulatory framework does not provide for how impermissible uses of AI should be evaluated. There should be a clear list of prohibited uses of AI and a duty to regularly review and update them.

2) **Regulating 'low risk' AI as if it is 'no risk' AI**

   In the White Paper, those developing or deploying AI that is not categorised as 'high risk' are only subject to voluntary guidelines, with the possible incentive of a 'trustworthy AI' stamp. This is insufficient to ensure safety and rights, but also risks making lower risk AI applications sound like no risk AI to the public. All AI applications, by their nature, raise questions of bias (in datasets and algorithms), power, and unintended or unexpected consequences. As such, no application is without risk and so suggesting 'no risk' AI exists is misleading. The approach outlined in the White Paper effectively leaves AI that falls outside the 'high risk' category unregulated, and does not account for broader societal and individual effects that might be experienced.

3) **Risk as a flawed, unsustainable model**

   The deployment of the concept of risk, let alone 'high risk', is problematic given the lack of a clear and accepted understanding of what the term means, particularly in the concept of autonomous systems. The White Paper's risk labelling model overlooks the evolving nature of AI technology and applications. An AI system may have been developed for one application considered 'lower risk', but later be applied to another considered 'high risk'. An AI technology that appeared low risk at the time of implementation could later be discovered to have vulnerabilities, or, when combined with other technological developments, become higher risk. Technology and AI systems have a time-variant nature of risk. Potential changes over time make a risk-based model insufficient. It also does not account for the differences in the types of risks that technology produces for different people and in different contexts. Embedding technology so intimately within the very fabric of our society implies a new set of challenges (and potentially failures) that narrow, risk-based, approaches are unsuitable to tackle.

In response to these concerns, we recommend the European Commission:

- Begins by adopting the proposed regulatory requirements for 'high risk' AI for all AI**.**

- Considers that some AI uses will be impermissible from the outset and that, for some of the highest risk technologies and applications, moratoriums provide an opportunity to establish whether or not there is a justifiable case for their use while minimising harm to the public.

AI in the public sector

Prioritise purpose over promotion

To build 'an ecosystem of excellence', the White Paper recommends 'promoting the adoption of AI by the public sector', working towards an 'Adopt AI programme'. While there is notable potential for AI to benefit the public sector, this unqualified promotion fails to reflect the complexity and caution required.

The use of AI in the public sector fundamentally affects the relationship between citizens and the state, and between people and services. To date, algorithmic decision-making systems in public services have often been developed in piecemeal, localised and invisible ways, affecting critical decisions like citizens' access to benefits and social care.[1] In many cases, they have been developed without scrutiny, common approaches, or an overarching vision for the role data and AI should play at the front line. As we highlighted in our analysis of the European Data Strategy,[2] European member states, their regions and structures of public service, will vary in preparedness, access to skills and infrastructure, and public sector practice that will impact the suitability of using AI tools.

We recommend that the Commission make it clear that AI applications may not be suited for all areas of the public sector and avoid universal, indiscriminate promotion of AI in the public sector. As we've commented in our response to the EU Data Strategy, more data does not in itself guarantee better decision making. We should also look for the quality of data, across three axes: quality of information, quality of process and quality of governance. Moreover, we recommend that the Commission, along with all those considering use of AI in the public sector, first seek to articulate the *purpose(s), proportionality and necessity* of that use. The adoption of AI in the public sector cannot happen for its own sake, or merely to improve European AI capabilities – it must be to the benefit of the public that the sector serves.

We recommend the White Paper prioritises communicating specific purposes for which it thinks AI may support the public sector, over indiscriminate promotion, and be clear about the risks and challenges involved.

[1] For example, Henley, J. and Booth, R. (2020). Welfare surveillance system violates human rights, Dutch court rules. *The Guardian* [online]. Available from: https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules [Accessed 12.6.20].
[2] Pavel, V. (2020). The EU Data Strategy: three key questions. *Ada Lovelace Institute.* Available from: https://www.adalovelaceinstitute.org/the-eu-data-strategy-three-key-questions/ [Accessed 12.6.20].

Designate the public sector as a high risk AI sector

The proposed risk-based regulatory framework for AI identifies sectors and applications that warrant the designation of 'high-risk' AI. While it commendably suggests that the sectors should be specifically and exhaustively listed, the examples given suggest only 'parts of the public sector' with examples such as 'asylum, migration, border controls and judiciary, social security and employment services'.

We recommend *all* public sector uses of AI meet the sector criteria for being, at minimum, 'high risk'.

Investigations into the use of algorithmic decision systems in the public sector, particularly in the US where much research has focused, have raised major concerns around the treatment of individuals: health insurance systems have incorrectly flagged fraud, blocking essential medication; algorithms have entrenched racial bias in policing decisions or sentencing; access to homeless accommodation has required acquiescence to highly intrusive monitoring.[3] Where AI is used to make judgements (detecting hate speech for example) or to predict outcomes (at-risk children; policing, etc.) its performance may not be much better than using simple regression analyses of a few factors,[4] while being far more intrusive. Even tools with a high level of accuracy may have pernicious impacts through reinforcing or clustering vulnerabilities; encoding historic social or structural inequalities and discrimination into permanence; or enhancing surveillance and eroding privacy. The datafication of individual lives and society must not result in a loss of human agency and freedom through mass experimentation within AI systems.

The types of requirements outlined in the White Paper for 'high risk' AI are essential to any successful public sector AI deployment and, as such, must be mandatory. They do not, however, eradicate risks of AI in the public sector and require further development of meaningful mechanisms for accountability and redress.

Ensure transparency and scrutiny of power dynamics in public-private partnerships

In proposing a new public-private partnership in AI, data and robotics for the Commission, the White Paper must recognise there are still many unanswered questions about what constitutes fair public-private partnerships for data and AI, and the risks involved.

In association with Understanding Patient Data, we have conducted public deliberation, by convening citizen juries, on public-private partnerships within the context of health data sharing in the UK. Our findings show that the public has expectations of accountability, transparency and public participation in decision making on use of data. They also emphasise the importance of ensuring that the benefits from such partnerships are fairly

[3] Eubanks, V. (2018). *Automating Inequality.* St Martin's Press, UK edition; O'Neil, C. (2016). *Weapons of Math Destruction*. Penguin, UK edition.
[4] Narayanan, A. (2019). How to recognise AI snake oil. Slides delivered 18th November, MIT. Available from: https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf [Accessed 12.6.20].

and equitably distributed, especially when they relate to marginalised and minority groups, as well as to regional inequalities.[5] **To that end, we recommend the White Paper provide greater detail on the proposed public-private partnerships – on the activities, data and knowledge sharing, potential benefits, risks and their distribution across member states and their populations.**

Public-private partnerships, particularly around AI that may be used in frontline services or policy decision making, shift the role of private companies in statecraft and may result in a new reliance on private tech infrastructure, often by large industry players. This creates new risks of private control over critical public infrastructure and vendor lock-in – whether through technology, or through reliance on technology knowledge and expertise. **We therefore recommend the terms of public-private partnerships are free, fair, open for scrutiny and sustainable. They should also be a continued issue for discussion and research.**

Biometric technologies, identity and liberty

Place a moratorium on the use of remote biometric identification

The White Paper recognises remote biometric identification[6] as an inherently high-risk AI application, as an intrusive surveillance technology with risks to fundamental rights. In laying out specific requirements for remote biometric identification it states that, under existing EU data protection legislation and the Charter of Fundamental Rights, 'AI can only be used for remote identification purposes where such use is duly justified, proportionate and subject to adequate safeguards.' But, in calling for a broad European debate on what, if any, justifications and appropriate safeguards may be, the Commission recognises that it does not yet have an answer to when AI may be used for remote identification purposes.

Many have called for a ban on biometric technologies that can be used for mass surveillance[7] and, as the White Paper recognises, it is possible that there are no justifiable uses of these technologies. Our report, *Beyond Face Value*, surveyed public attitudes to facial recognition (a technology that can be used for remote biometric identification) in the UK in June 2019. It found the majority supported companies voluntarily pausing sales of facial recognition technology to police (50%) and schools (70%) to allow for further public

[5] Patel, R. (2020). The foundations of fairness for NHS health data sharing. *Ada Lovelace Institute.* Available from: https://www.adalovelaceinstitute.org/the-foundations-of-fairness-for-nhs-health-data-sharing/ [Accessed 12.6.20].

[6] In the White Paper remote biometric identification is defined as when the identities of multiple persons are established with the help of biometric identifiers (fingerprints, facial image, iris, vascular patterns, etc.) at a distance, in a public space and in a continuous or ongoing manner by checking them against data stored in a database. We dispute that this should be limited to public spaces.

[7] EDRi. (2020). Ban biometric mass surveillance! *European Digital Rights (EDRi).* Available from: https://edri.org/blog-ban-biometric-mass-surveillance/ [Accessed 12.6.20].

consultation.[8] Similarly, public attitudes research in France also shows public preference for time for the consideration and development of clear protective frameworks before deployment of facial recognition.[9]

We recommend that the Commission be explicit in placing a moratorium (a time-limited ban) on the use of remote biometric identification to enable the proper debate and evidence building required to establish if there are justifiable applications with sufficient safeguards. Many industry organisations have already recognised the disproportionate impact and effects such technologies have on people, and have committed to pausing the sale, development and deployment of some of these technologies. These companies include IBM, Microsoft and Amazon.

This moratorium should not be limited to remote biometric identification in public spaces. We have already seen the limitations of this in the UK, where the use of facial recognition technology for remote biometric identification around King's Cross in London caused public uproar and condemnation.[10] However, this was conducted on private property that was free to access by the public – there is no longer a neat delineation between public and private spaces for a citizen navigating a city and the risks for fundamental rights remain in both spaces. There are also concerns about the use of remote biometric identification in the home and workplace.[11]

We therefore recommend a moratorium on remote biometric identification in its entirety.

Be informed by public deliberation in biometrics consultation

Addressing questions around the justification for, proportionality of, and appropriate safeguards around the use of biometrics requires a wide-ranging public debate. This should involve all stakeholders: legal experts, data specialists, philosophers, policymakers and – crucially – members of the public. In the UK, the Biometrics Commissioner has argued how biometrics technologies fall between the scope of established oversight mechanisms and outside of current understandings for what is acceptable for these kinds of technologies. In May 2020, he called directly for greater public debate on these issues.[12]

---

[8] Ada Lovelace Institute. (2019). Beyond face value: public attitudes to facial recognition technology. Ada Lovelace Institute. Available from: https://www.adalovelaceinstitute.org/wp-content/uploads/2019/09/Public-attitudes-to-facial-recognition-technology_v.FINAL_.pdf [Accessed 12.6.20].

[9] Renaissance Numerique. (2019). Reconnaissance faciale: ce que nous en dissent les Français. *Renaissance Numerique.* Available from: https://www.renaissancenumerique.org/ckeditor_assets/attachments/445/rn-analyse-reconnaissancefaciale.pdf [Accessed 12.6.20].

[10] Murgia, M. (2019). London's King's Cross uses facial recognition in security cameras. *Financial Times* [online]. Available at: https://www.ft.com/content/8cbcb3ae-babd-11e9-8a88-aa6628ac896c [Accessed 12.6.20].

[11] Ajunwa, I., Crawford, K. and Schultz, J. (2017) Limitless worker surveillance. *California Law Review* 105: 735. Available from: https://dx.doi.org/10.15779/Z38BR8MF94 [Accessed 12.6.20].

[12] Wiles, P. (2020). Biometrics Commissioner's address to the Westminster Forum: 5 May 2020. *Gov.uk.* Available from: https://www.gov.uk/government/speeches/biometrics-commissioners-address-to-the-westminster-forum-5-may-2020 [Accessed 12.6.20].

Earlier this year, we convened the Citizens' Biometrics Council, a broad and diverse group of 55 members of the UK public to deliberate on biometrics technologies.[13] This project has been adapted during COVID-19 lockdown to find ways of engaging citizen voices and perspectives through virtual media. We envisage it will convene face-to-face following the lockdown period, and it is now due to generate its findings and recommendations later this year. There is, however, much emerging research on the impacts biometrics technologies are having on people and society at this moment in time, and our interim findings already show that people have concerns around how these technologies. Concerns are centred around how they can miscategorise individuals' identities, how they create erroneous and discriminatory outcomes, particularly for marginalised groups, how they exclude certain groups who are often already marginalised, and how they can reduce individuals' agency while giving powers to public and private institutions.[14] What public debate ultimately deems proportionate, justified and responsible must account for these (and other concerns). And while this debate occurs, a moratorium is necessary to ensure the severe risks of these technologies are minimised.

We welcome a broad European debate on the specific circumstances, if any, which might justify the use of remote biometric identification and safeguards they would require. We seek clarity on the form and role this debate will take, and recommend the inclusion of public deliberation as part of any meaningful consultation process.

---

adalovelaceinstitute.org
@AdaLovelaceInst
hello@adalovelaceinstitute.org

Ada Lovelace Institute
28 Bedford Square
London
WC1B 3JS
+44 (0) 20 7631 0566

Part of the Nuffield Foundation
Registered charity 206601

[13] Ada Lovelace Institute. (2020). Citizen's biometrics council. *Ada Lovelace Institute.* Available from: https://www.adalovelaceinstitute.org/our-work/identities-liberties/citizens-biometrics-council/ [Accessed 12.6.20].

[14] Patel, R. and Peppin, A. (2020). Making visible the invisible: what public engagement uncovers about privilege and power in data systems. *Ada Lovelace Institute.* Available from: https://www.adalovelaceinstitute.org/making-visible-the-invisible-what-public-engagement-uncovers-about-privilege-and-power-in-data-systems/ [Accessed 12.6.20].