

Written Evidence Submitted by Imperial College London Artificial Intelligence Network

(GAI0014)

Explanatory note:

This response has been developed by Imperial College London's [AI Network](#). The AI Network is comprised of approx. 250 researchers with expertise in AI and a diverse range of its applications.

In this submission, the full names of contributing researchers will appear at the beginning of their first answer and any subsequent answers will begin with their initials. A full list of contributors is included at the foot of the document.

--

1. How effective is current governance of AI in the UK?

1.1

Dr Juan Pablo Bermúdez & Professor Rafael A. Calvo (JPB & RAC): Out of the pillars of the UK approach, the context-specificity might be the most distinctive in comparison with other approaches, particularly the EU AI Act. This pillar makes sense from the interactional perspective of design, which considers that the meaning of a technology materialises at the moment of interaction, and therefore only technology-in-context can be analysed for the purpose of ethical design.

The second pillar ("Pro-innovation and risk-based") is potentially too broad and reactive, since waiting for risks to become evident is probably the same as waiting for the harms to materialise.

1.2

Professor Danilo Mandic (DM): Given its widespread use, and the vast number of aspects, governance of AI is a very delicate question. Certainly when it comes to ethics, unfair bias, and related issues, there is a need for clear governance, at least when it comes to the use of AI in the public sector and the government.

Another issue is that of energy consumption by AI models, which is a totally unregulated area. For example, Data Centres currently use around 200 terawatt hours of energy per year, forecast to grow by around an order of magnitude by 2030.

A single training run of a recent (2020) Deep Neural Network for Natural Language Processing exceeds 1,000 MWh (more than a month of computation on clusters) (IEEE SPM, Sep 2022). This amounts to 500 tons of CO2 carbon footprint (equiv. to ca. 250 round trips from London to Seoul).

Such issues are quite pressing but are not yet addressed. So, is it ethical to spend so much energy (from fossil fuels and nuclear power plants) on various entertainment applications?

--

1(b). What are the current strengths and weaknesses of current arrangements, including for research?

1.3

Dr Rossella Arcucci (RA): AI's potential to have an impact on our economy, politics, and culture is growing as it's getting more powerful. This might go in either very well or incredibly poorly directions. AI may enable us to make scientific and technological advancements that will assist us address the world's most pressing and urgent problems including Climate Change. Big operational Centres for weather and climate forecasting worldwide (including Met Office, ECMWF, CMCC, NOAA and others) are now incorporating AI models in their predictive systems. Machine learning technologies combined with data science have shown great capabilities in accelerating physics simulations, ingesting new observations from satellites or sensors in real time and estimating physics parameters with high accuracy.

In addition, these efficient AI models have shown a big impact in terms of reduction of computational power consumption with respect to standard models, with a consequent reduction of emissions which makes the AI models more sustainable. AI models have been developed to help decision makers to reduce the risks or the effects of natural hazards, including wildfires. The use of human sensors, i.e. the use of information coming from people (e.g., from social media), has been terrific to include the human perspective in AI models and improve natural hazards forecasting.

On the other hand, badly-controlled AI systems could spell calamities for humans. Given the stakes, regulating and promoting key topics such as explainable AI, privacy and security for AI, human centric AI and physics informed AI is a high-priority that we strongly encourage to support, particularly to preserve the long-term future.

1.4 DM:

I cannot speak about the current strengths and weaknesses in government applications.

Regarding research, the very nature of research is that it must not be over-regulated. Of course, we need various ethics permissions when recording human data. Perhaps, we could also have ethics permissions for AI algorithms developed for similar sensitive areas?

--

2. What measures could make the use of AI more transparent and explainable to the public?

2.1

Dr Mark Kennedy (MK): Inspired by the medical profession's commitment to do no harm, we are developing a practical diagnostic for assessing subpopulation variance in model errors to ensure they do not land inequitably on any protected subpopulation. The idea behind this method, which we are calling SPAVA (for 'subpopulation prediction accuracy variance analysis'), is to create a simple report card that splits out model accuracy by subpopulation, showing differences from the overall as green, yellow, and red, where red indicates harms that are frequent and serious enough to recommend against using a model for predictions on that subpopulation.

Subpopulations considered include both (a) legally protected categories (race, ethnicity, religion, minors, children, infants, etc.) and (b) statistically relevant subpopulations identified by scan statistics adapted to flag clusters where accuracy deviates from population norms. Whether by SPAVA or other methods, our ambition for UK AI deployments is to see a practical scheme for certification that could, in time, earn public recognition and support from public policy. This work is joint with my colleague David Hand and collaborators at Validate AI, a community interest organisation with strong ties to government, especially through HMRC and OSR.

2.2 RA:

Education and communication is the key. Explainability and transparency in AI must be built into the environments but to make this happening, it's important to make sure the AI developers maintain a flow of communication with the public. The human decision-maker needs to be heard when developing and putting into practise an algorithm. For AI to be understood by the public, communication is essential. The output of an algorithm must be written in a language that is accessible to the general audience to achieve that. AI does a terrific job of acquiring data and transforming it into knowledge, but we must keep in mind that humans still need to have the insight and wisdom to make decisions.

2.3 Dr Davide Amato (DA):

Academic research into explainable AI obviously has a key role in guaranteeing that decisions made by AI systems are explainable. UK technical standards should encourage using “Occam’s razor” for the development of AI algorithms, as simpler AI systems will always be easier to explain than more complex ones.

In the long term, investing in secondary education is crucial in order to give the public the adequate tools to understand the basic concepts underlying AI. Including “computational thinking” in the secondary school curriculum might help with this. Although STEM subjects are important, humanities should not be disadvantaged, as future generations will need to make decisions (for instance, on Artificial General Intelligence) that must be guided by solid values. This can only take place if the public has been educated in philosophy, ethics, and the arts.

2.4 JPB & RAC:

An intelligent system should have a feature that allows it to “explain itself” in layman’s terms. Why is the decision being made? What potential conflicts with the users’ values is it implementing? Regulators could provide a list of such questions that should be answered by all systems. This requirement need not hinder innovation, and would create standards of trust that can provide a business advantage.

2.5 DM:

Other potential tools for public engagement on topics such as this include public lectures, training of primary and high-school teachers, summer camps for students, and ‘hackathons’. Each of these formats are familiar to us in the research sector and universities in particular would be well-placed to support the development of this kind of engagement programme.

3. How should decisions involving AI be reviewed and scrutinised in both public and private sectors?

3.1 Dr Hamed Haddadi (HH):

There is an increasing array of technical solutions developed both by the eXplainable AI (XAI) community, and the systems research community. These include model explainability, attestation, auditing, and verification techniques. The UK has existing advantage in this space via established research institutes (e.g., Centre for eXplainable AI, UKRI Centre for Doctoral Training in Safe and Trusted Artificial Intelligence). There are existing efforts in this space, specifically in the space of online targeted advertising and real-time bidding, where the information Commissioner’s Office has been playing an active role in engaging with the industry to ensure consumers’ privacy. Increased engagement of policymakers with these institutes and centres can enable rapid fostering of frameworks for auditable and verifiable AI.

3.2 DM:

In medical research, there is an important role for 'ethics committees'. These committees are composed of a combination of qualified professionals with relevant sectoral experience and 'lay members' of the public. Committees are tasked with reviewing medical research applications to ensure that any research that is funded meets required ethical standards. Some sort of review and scrutiny bodies similar to ethics committees could be considered a viable approach for the review of decisions about AI regulation and implementation. The involvement of 'lay' members in any such body would go a long way towards maintaining public confidence in their decisions.

--

3b. Are current options for challenging the use of AI adequate and, if not, how can they be improved?

3.3 DA:

The UK AI legal framework should provide a clear pathway for citizens to issue complaints/appeal in case they suspect of having been algorithmically discriminated. Opportunities for citizens to opt out of algorithmic decisions should be considered in the UK legal framework, especially if there is the demonstrated risk that "algorithmic discrimination" can significantly impact the livelihoods of minorities.

--

4. How should the use of AI be regulated, and which body or bodies should provide regulatory oversight?

4.1 DA:

I believe that the context-driven approach outlined in the recent white papers from the Office of AI makes sense. However, cross-sectoral principles covering common issues and risks should be issued in future legislation. These cross-sectoral principles should be in line with international legal frameworks, and in particular with the EU legal framework, in order to achieve the global "critical mass" required to enforce compliance from non-state actors. This will also enhance the UK's global standing as a country spearheading the responsible development and use of AI.

4.2 JPB & RAC:

Ethical tools should be promoted as catalysts of innovation. AI Ethics would often catalyse, rather than hinder innovation. This is particularly the case when ethical issues and stakeholder values are used in the generative phase of design. In addition, products that generate trust and transparency tend to support user engagement, and therefore business growth. Therefore, providing clear and sound evidence of trustworthiness can be critical for boosting AI product adoption. For example, to support innovation the Legal Framework should encourage the adoption of ethical tools (like value-based design methodologies, and ethics standards like IEEE 7000), not as requirements for some risky products, but as tools to boost innovation and commercial success for AI companies of all sizes.

4.3 Dr Luca Magri (LM):

AI should be regulated in publications.

Recently, we have been witnessing an exponential growth of number of publications in AI applied to scientific fields. This is partly dictated by short-term and short-sighted visions on increasing immediately citations metrics, both for authors and publishers; however, it is not

necessarily dictated by making scientific breakthroughs (which take more time and thinking). Basic requirements of the scientific method (reproducibility, repeatability, verification) are sometimes overlooked, or poorly performed, in some of AI papers for engineering/scientific applications.

Another example, in AI for scientific disciplines, review papers appear on a frequent basis, even when there is not much consolidated to review about. Review papers should be limited in number, analyse critical mass of literature from all the community (and not only a handful of papers or authors), and be critical towards limitations and directions for future work. Additionally, being invited to write a review paper used to be an honour and recognition of the author's academic stand because it was a rare event.

The current publishing model is influenced by publishers' interest (which is profit for most publishers), which is at variance with the AI interest (rigor, testing, certifiability, and discussion of limitations). The UK should be taking the lead to promote "slow AI" with a board of academics overseeing the publisher's decisions and strategy on papers in AI. Second, the UK with the help of leading academics should make an official list of "good" publishers and "not good" publishers to guide Universities in the evaluation of their academics. We should have a REF for publishers, not only for academics.

4.4 DM:

Humans have begun to "merge" with the technologies we have created.

While this is exciting and offers advantages, this also harbours many dangers regarding to the over-reliance on the "virtual worlds", social and professional skills, and the overall outlook for the future.

In this sense, regulation should be established for commercial applications of AI, that is a priority.

Less regulation is required for AI research, as this is not commercial.

4.5 RA:

Despite all the potential advantages of public-private AI partnerships in building AI regulations, the relationship can be difficult. Government's responsibility to its citizens comes first. Private businesses, meanwhile, have their shareholders as their main priority. These concerns might lead to "conflicts" of interest. For a government to be legitimate, answers to questions such as "who will oversee and safeguard the data of citizens?" or "who will reap the financial benefits of the insights they can yield?" may be essential. Given that governments bear ultimate responsibility for the safety and well-being of the people under their care, ultimate regulatory responsibility must be held and exercised by the state on behalf of its citizens.

--

5. To what extent is the legal framework for the use of AI, especially in making decisions, fit for purpose?

5.1 DA:

Citizens should be informed when AI algorithms make choices that affect them, and this should be enshrined in the UK legal framework. For example, disclaimers similar to the "Why am I seeing this ad?" icon on Alphabet products could be included on websites stating that the consumer has been targeted by an AI recommendation system. Another critical application of AI where citizens should be informed is at UK borders, where UK citizens and

residents are likely subject to algorithmic decisions that can significantly impact their livelihoods.

--

5b. Is more legislation or better guidance required?

5.2 DM:

Better guidance is always helpful. Too much legislation, at this early stage of the use of AI in commercial and government application, may prove counter-productive.

In the "Limits to Growth" report (D. Meadows and J. Randers, "The Limits to Growth: The 30-Year Update", Routledge, 2012.), to stop the climate change and various other human-created adverse effects on the planet, the current energy consumption of an average European would need to be divided by 5 by 2050 and by 10 for an average American.

The push for low-tech and low-power end devices could define a new metric - Technical Conviviality - in place of the unsustainable metric of absolute performance.

5.3 JPB & RAC:

Risk-based strategy should include monitoring for emerging risks.

We find the Framework's risk-based approach appropriate. However, focusing only on cases where "there is clear evidence of real risk" presupposes that we have solid knowledge of the risks of AI, or that we must wait for the harms to materialise. Reactive approaches like these have led regulators around the world to reacting too late, multiple steps behind the current developments.

We thus think that the Framework should include measures to actively promote research that monitors emerging AI risks. If this duty is left to individual regulators, there may be sectors that lack appropriate risk monitoring. Additionally, some AI-related risks are cross-sectoral (e.g. privacy concerns are important for consumer protection, but also for security), so some coordination across sectors would be required.

5.4 JPB & RAC:

The principle of fairness should explicitly encourage regulators to assess AI's impact on minorities.

Research has demonstrated that minorities and marginalised groups tend to bear the brunt of the risks and harms of AI. For instance, algorithmic biases tend to reinforce the biases toward populations already discriminated against, further denying them access to health, security, and other public goods. These concerns are not specific to any sector or regulator in particular: if we want to embed fairness into AI, concern for minorities should be made explicit.

--

6. What lessons, if any, can the UK learn from other countries on AI governance?

6.1 DA:

I believe that the approach centred on the protection of citizens' fundamental rights of the current EU AI act might provide a good blueprint for AI governance in the UK. This should not necessarily be seen as a zero-sum trade-off between protection of individual rights and stimulating innovation, however. Historically, legislation favouring the protection of

fundamental rights has led to periods of sustained economic growth. The UK AI governance approach should be both robust (in its protection of human rights) and nimble (in order to stimulate AI innovation).

--

Contributors:

DA: [Dr Davide Amato](#)

Lecturer in Spacecraft Engineering, Department of Aeronautics

DM: [Professor Danilo Mandic](#)

Professor of Signal Processing, Department of Electronic and Electrical Engineering

HH: [Dr Hameed Haddadi](#)

Reader in Human-Centred Systems, Department of Computing

JBP: [Dr Juan Pablo Bermúdez](#)

Research Associate, Wellbeing Technologies Lab

LM: [Dr Luca Magri](#)

Reader in Data-Driven Fluid Mechanics, Department of Aeronautics

MK: [Dr Mark Kennedy](#)

Associate Professor, Imperial College Business School, and Co-Director of the [Data Science Institute](#)

RA: [Dr Rossella Arcucci](#)

Lecturer in Department of Earth Science and Engineering, and Elected Speaker of the [Imperial College London AI Network](#)

RAC: [Professor Rafael A. Calvo](#)

Chair in Engineering Design, Dyson School of Engineering

--

Further information:

This submission was collated by the [Imperial Policy Forum](#), working in partnership with the AI Network. The IPF team supports the policy engagement work of Imperial College London researchers.

(November 2022)