A response to the call for evidence on 'Establishing a pro-innovation approach to regulating AI" on behalf of the *Regulation* and *Functionality* nodes of the UKRI Trustworthy Autonomous Systems Network (TAS).

Executive summary:

We welcome the Government's proposal to develop a new, coherent regulatory strategy for AI. While it maintains a sectoral focus, the development of cross-sector and cross-application principles and governance structures has the potential to create legal certainty, foster public acceptance, and facilitate responsible development of generic AI tools that are currently left unregulated. Our submission intends to discuss various aspects of the proposal, including: the design and enforcement of the regulatory framework; the context-driven and cross-sectoral principles' approach, and the coordination between regulatory bodies for coherence and monitoring. Our submission will use medical device regulation for AI as enabled medical devices (AIaMD) as a sector-specific example to illustrate our recommendations. We have structured our response around the six questions in the consultation:

1. *What are the most important challenges with our existing approach to regulating AI? Do you have views on the most important gaps, overlaps or contradictions?*
    a. The current regulatory regime, to the extent that it exists, is increasingly fragmented, with some fields (such as autonomous driving) far ahead of the discussion, while others (such as AI-based content moderation) are virtually non-existent. We are encouraged that the proposal intends to strengthen cross-regulatory coordination and oversight. Our concern is that there are challenges in how the balance between granularity and flexibility of regulation can be achieved in practice. As many AIs and their enabling technologies, such as large language models, are multipurpose applications, harmonisation with regard to common risks and emerging issues of AI technology is necessary when regulations target the design stage of the system. While sector-specific regulators have a clearer idea of the concrete risks, crucial design decisions will have already been taken before specific applications are considered. Thus even if there are application-specific requirements a regulator may deem necessary, these may be preempted by earlier design choices.
2. *Do you agree with the context-driven approach delivered through the UK's established regulators set out in this paper? What do you see as the benefits of this approach? What are the disadvantages?*
    a. One aim of the proposal is to provide both stability and flexibility of regulation. However, this assumes that regulatory bodies have the resources and expertise for capacity building. As an example, we discuss here the use of Approved Bodies in the UK medical device sector. Most standard setting bodies have no mechanism or expertise to consider ethical or legal questions in their processes. Here the government can actively encourage the relevant capacity building by fostering collaboration between academia and organisations such as the British Computing Society.

b. In light of current EU and US proposals for AI regulation, any separate UK framework will inevitably create challenges for regulatory coherence. Even if the UK regime is intended to foster innovation through a light touch regime, these efforts could easily multiply compliance burdens for industry, as developers try to accommodate potentially conflicting regulations in different markets. These negative effects will be unevenly distributed, favouring some industries and business models over others. To evaluate any success (or lack of success) of the UK proposal it needs to articulate clearly what it considers success, and which sectors, AI users, and AI-enabled business models are its intended beneficiaries.

3. *Do you agree that we should establish a set of cross-sectoral principles to guide our overall approach? Do the proposed cross-sectoral principles cover the common issues and risks posed by AI technologies? What, if anything, is missing?*
    a. The report tends to see regulation as inherently, or at least typically, hostile to innovation. This is in danger of overlooking the positive effect that robust and transparent regulation can have for innovation, customer acceptance and market success – evidenced by several digital industries recently asking for more, not less, regulation. It is important to bear in mind, therefore, that the quantity of regulation is less a concern than its quality. This requires investment in regulatory bodies to build up their expertise.
    b. "Smart regulation" is a highly desirable goal, yet the regulatory vocabulary of the proposal remains very traditional in this regard. One of the greatest opportunities for a separate UK approach to AI regulation could be the use of innovative regulatory tools. These can range from "automated compliance assessment," which utilises, in real time, the data that AI and IoT generate, or entirely novel forms of legislative drafting, such as "Law as Code" initiatives that are currently being trialled in some jurisdictions.
    c. While a key aim of the proposal is to "foster innovation", it is lacking clear criteria of what would count as "success". It also treats the sector as more homogenous than it is. As such, the proposed regime is likely to have very different impacts on UK-based AI developers that target the domestic market only, and those that aim for international exports. Both will be affected in different ways from UK companies that innovate but use AIs developed abroad. The proposal needs clarity on this issue and a more fine-grained approach that identifies its intended beneficiaries.
4. *Do you have any early views on how we best implement our approach? In your view, what are some of the key practical considerations? What will the regulatory system need to deliver on our approach? How can we best streamline and coordinate guidance on AI from regulators?*
    a. Developing AI regulation requires crucial normative decisions between conflicting values, such as equitable risk allocation. These have to be the result of open and public discussions, with ultimate responsibility lying with Parliament. While this does not preclude sectoral regulators from creating substantive and/or procedural rights to ensure the UK's approach is dynamic and agile to the risks of AI, there has to be clear democratic accountability and oversight. As formulated, the proposal (like its EU and US counterparts) risks creating democratic deficits, and could disempower Parliament.
    b. We do acknowledge the UK's approach to provide a flexible foundation for the sectoral regulation of AI based on the use of cross-sectoral principles. Nevertheless,

cross-sectoral principles require a robust baseline on the individual regulator level. Clearly articulated, minimum standards and risk-profiles acting as a baseline and/or due diligence obligation for regulatory bodies, should inform how notions of accountability and responsibility offer a shared understanding across regulatory bodies. It would help individual regulators, such as the Medicines and Healthcare products Regulatory Agency (MHRA), to "streamline risk" considering emerging contextual uses and risks of AI. In addition, the government's closer involvement in defining outcome-based regulation and proportionality should clarify when regulatory change and revision are required.

c. A key practical consideration regarding the implementation of the UK government's approach is to address cross-sectoral issues and fragmentation or duplication across regulatory bodies. The government's proposal needs to envisage frameworks that strengthen multi-agency partnerships to refine key AI areas within organisational structures. We also note that the government's effort to develop a coherent regulatory approach based on the AI Standards Hub will require knowledge-sharing from individual regulatory bodies. We outline some successful practices in the area of medical device regulation, and how common methodologies and evidence-based approaches support international standardisation.

5. *Do you anticipate any challenges for businesses operating across multiple jurisdictions? Do you have any early views on how our approach could help support cross-border trade and international cooperation in the most effective way?*
    a. The UK proposal states that, unlike its EU counterpart, ensuring free movement of AI-services and AI-enhanced goods is not a concern, and therefore permits a different approach. This underestimates the impact a UK AI framework will have on the devolved powers. Police, education and health are all likely users of AI, yet fall under devolved powers. Northern Ireland poses its additional difficulties. The precise relationship between the Act and the devolved powers needs clarification, and a UK-wide AI Act requires the involvement and support of the devolved parliaments.
6. *Are you aware of any robust data sources to support monitoring the effectiveness of our approach, both at an individual regulator and system level?*
    a. There are advantages and risks of using existing data sources in supporting an evaluation and monitoring mechanism on the system and regulator level. The UK government's proposal needs to think about shared indicators to identify areas for policy intervention on the system level. Another task is defining methodologies to monitor the effectiveness of the UK's approach on a regulatory level. Here, we use the post-market surveillance framework regarding medical device regulation to highlight the pitfalls of using incidence data and argue for more dynamic and traceable approaches for monitoring and oversight.

Response authors:

Daria Onitiu, Research Associate Governance & Regulation of Autonomous Systems, School of Law, The University of Edinburgh

Bhargavi Ganesh, PhD Student, School of Informatics, The University of Edinburgh

Burkhard Schafer, Professor of Computational Legal Theory, School of Law, The University of Edinburgh

John Downer, Senior Lecturer in Risk and Resilience, School of Sociology, Policy and International Studies (SPAIS), University of Bristol

Peter Winter, Research Associate in Regulation of Autonomous Systems, School of Sociology, Policy and International Studies (SPAIS), University of Bristol

Phoebe Li, Reader in Law and Technology, Sussex Law School, University of Sussex

Ram Ramamoorthy, Personal Chair of Robot Learning and Autonomy, School of Informatics, The University of Edinburgh

Robin Williams, Director, Institute for the Study of Science, Technology and Innovation, School of Science and Political Science, The University of Edinburgh

Stephen Gilbert, Professor of Medical Device Regulatory Science, Else Kröner Fresenius Center for Digital Health, Technische Universität Dresden, Germany

Stuart Anderson, Professor of Dependable Systems, School of Informatics, The University of Edinburgh

About the UKRI TAS Governance & Regulation Node and Functionality Node:

The UKRI TAS Governance & Regulation Node (EP/V026607/1) explores frameworks for formal and informal governance of autonomous systems, and brings together expertise in computer science and AI specialists, legal scholars, AI ethicists, as well as experts in science and technology studies and design ethnography.

The UKRI TAS Functionality Node (EP/V026518/1) is engaging with design requirements, verification and regulation of autonomous systems, considering their changing functionality and using a range of research approaches and case studies.

At the centre of the project is the Trustworthy Autonomous Systems Programme, which is funded by the UKRI Strategic Priorities Fund.

List of Abbreviations:

AAMI	Association for the Advancement of Medical Instrumentation
AI	Artificial Intelligence
AaMD	AI as medical device
BSI	British Standards Institution
CDEI	UK's Centre for Data Ethics and Innovation
CQC	Care Quality Commission
EU AI Act	EU Commission's Proposal for a Regulation on Artificial Intelligence
DRCF	Digital Regulation Cooperation Forum
FDA	U.S Food & Drug Administration
GDPR	General Data Protection Regulation
GMLP	Good Machine Learning Practice
ICO	Information Commissioner's Office
IP	Intellectual Property
LLM	Large Language Models
MAAS	Multi-Agency Advisory Service
ML	Machine Learning
MHRA	Medicines and Healthcare products Regulatory Agency
SaMD	Software as medical device
SMEs	Small and Medium Enterprises
TAS	Trustworthy Autonomous Systems

Submission:

1. **What are the most important challenges with our existing approach to regulating AI? Do you have views on the most important gaps, overlaps or contradictions?**

i. "Problematising" the UK's current regulatory landscape

1. The most important challenge in regulating AI is ensuring legal certainty for all stakeholders – developers, users and citizens it affects, as AI innovation is accelerating horizontally and vertically (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme, 2022, p. 8). This is highlighted in the Alan Turing Research Programme report, which states that '[d]eveloping an effective and inter-organisationally robust regulatory environment is crucial for mitigating the far-reaching risks associated with AI' (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme, 2022, p. 8). The report refers to 'regulatory capacity and readiness in relation to AI' as being necessary for realising a robust approach to AI innovation, as 'the UK's current approach is to regulate AI technologies and services through existing regulators' (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme,

2022, p. 8-9). We would like to build on this statement, and articulate some specific challenges of AI governance on the individual regulator level, considering medical device regulation, as well as the system level.

2. The current system is marked by a high level of fragmentation and diversification, with regulatory silos and uneven readiness across sectors. The discussion (and legislation) is very advanced in some sectors – most notably autonomous vehicles and medical devices – and for some aspects of AI, for instance data privacy. The discussions in other sectors (such as AI on Internet platforms) and around other issues (such as environmental impact) are still in their infancy. The proposal is correct that many dangers and risks can only be identified at the application level and are therefore best left to sectoral initiatives; hence a degree of "sectoral devolution" is inevitable and indeed advantageous. However, for AI regulation, this approach creates significant risks, including risks to the innovation ecosystem.

   Many of the AI technologies that create the greatest concerns are at the development stage and are application neutral. "Image recognition", for example, can be used for medical diagnosis, for identifying a criminal suspect, or for searching for an image to personalise a birthday card. These use-cases obviously have very different risk profiles. A "light touch" regime may seem more appropriate for the birthday card, but this does not ease the compliance burden for the AI provider if they also sell their services to police or the NHS, and have to demonstrate compliance appropriate to these fields. Indeed, even a light touch regime can, under these conditions, create entirely unnecessary compliance burdens if the different sectoral regulators use e.g. incompatible IT systems that manage certification, risk assessment or approval for the regulator.

   The danger of regulatory fragmentation is particularly troublesome in the technology field. Recent decades have seen an acceleration in the use of "design thinking" as a regulatory tool. Rather than focus mainly on issues of liability and redress once a harm has materialised (often after expensive litigation), charging developers to develop technologies that are "built to be lawful" from the beginning has become an internationally preeminent form of regulation ("regulation by code", or "regulation by architecture"). While popularised by the "data protection by design" maxim of data protection law, this continued a development that began with the first industrial revolution. The malleability of the digital world enabled new ways in which laws could be directly incorporated into products, in itself a substantial technological innovation. On one hand, this demonstrates how good regulation can directly incentivise innovation: to comply with this requirement, new methods and technologies had to be developed that enabled businesses to achieve their objectives while using as little data as possible, with as much security, creating collateral benefits from fraud prevention to reduced environmental impact. Both UK startup companies and established players developed new methods, services and products in response to this regulation. Indeed, given the importance of the service sector in general for the UK, and the financial service sector in particular, the UK software sector is well placed to develop innovative "RegTech" (Regulation Technology).

   Regulatory fragmentation is a potential impediment for these "by design" approaches to compliance. They can mean that the developer of a generic AI tool not only has to undergo several regulatory procedures, but also has to develop entirely different tools and pipelines for each application field. This increases costs, limits growth into new markets, and reduces competitiveness. Conversely, if regulators only become involved when a specific application

has been decided upon, crucial design decisions with normative implications will have already been taken "upstream". This can mean that the "AI architecture" can pre-empt desirable regulation, with industry developers in effect binding the hands of regulators or even Parliament.

It is therefore important to note that this approach should not preclude harmonisation on a system level. For instance, the work by the Information Commissioner's Office (ICO) regarding the 'Data Protection Risk Toolkit' underlines that regulators need to articulate the risks of AI, considering its sectoral application. However, there are emerging challenges surrounding AI – from the continuous learning / changing functionality regarding safety critical applications to the "responsibility gap" of AI systems – which still require robust guidance on the system level (Matthias, 2004). Robust standard-setting across sectors in this way can also assist industry in developing better compliance tools. Achieving the right balance between flexibility and granularity in regulatory standards is inevitably difficult in practice, considering that the current regulatory framework only gives some regulators the ability to define the risks of AI horizontally (i.e. the Competition and Markets Authority, the ICO or Ofcom), with trade-offs between the approaches. We argue however that overall, a stronger emphasis on a common-framework on the system level, to ensure cross-regulatory AI governance, is ultimately beneficial for the sector.

3. The delegation of responsibility to regulators to 'take action to support the responsible use of AI' (Department for Digital, Culture, Media & Sport, 2022, p. 5) involves specific challenges in the medical device industry, including regulation of software as a medical device (SaMD) and AI as enabled medical devices (AIaMD). It is therefore a good case study from which a general approach can benefit. In particular, it entails government processes that enable robust implementation strategies for regulators and assurances for industry. The medical device industry is an important area where responsible innovation needs to be balanced with robust standard setting, convergence and market competitiveness. The Medicines and Healthcare products Regulatory Agency (MHRA) recently launched an initiative to reform the regulatory framework of SaMD and AIaMD (MHRA, 2021). In addition, the MHRA published the government's response to consultation on the regulatory framework for medical devices in the UK (MHRA, 2022). However, this requires novel features for international regulation, which are adaptable to technological progress, such as regulatory gaps for Machine learning /Artificial Intelligence (ML/AI) devices including data updates, and product liability. For instance, the way the MHRA's plan to publish clarificatory guidance on how clinical performance evaluation methods will align with the UK government's effort to embed notions of pro-innovation and market competitiveness on an international plane is important for stakeholders working in industry (MHRA, 2022, p.123). A robust implementation strategy regarding the changing demands of software development and risk profiles for AI as a medical device (AIaMD) entails adaptive regulation and convergence between standard-setting bodies.

4. Another challenge with regard to the UK's current approach is aligning regulation on the sectoral level whilst promoting certainty for vendors. There are several initiatives for reforming medical device regulation on the international level. For instance, working groups intend to develop horizontal AI standards (The European Coordination Committee of the Radiological, Electromedical and Healthcare IT Industry, 2021, p. 43). The BSI and the Association for the Advancement of Medical Instrumentation (AAMI), supported by the

MHRI, intend to adapt existing standards to AI. Further, the MHRA engaging with the U.S Food & Drug Administration (FDA) and Health Canada intend to navigate the field of regulating AIaMD and SaMD and developed the "Good Machine Learning Practice (GMLP)" guidance (FDA, Health Canada, MHRA, 2021). All these efforts, while providing advancement in the regulation of SaMD and AIaMD, still do not provide enough assurance for regulatory alignment and assurance on the organisational level.

5. "Smart regulation" is a highly desirable goal. While there has been some innovation in the regulation field, for instance the use of regulatory sandboxes, the current way in which technology is regulated would be familiar to a lawyer in the early 20th century. A system of laws that govern liability after harm exist, together with mandatory certification schemes that decrease the likelihood of harm occurring, both supported in some cases by mandatory insurance schemes. There is scope to explore more innovative forms of regulation in addition to the substantive content of the rules. The vocabulary of the current proposal remains, in this respect, very traditional. One of the greatest opportunities for a separate UK approach to AI regulation could be the use of innovative regulatory tools. Low hanging fruit is the use of smart tools by sectoral regulators to facilitate regulatory filings and other compliance tasks. More ambitious would be the use of "automated compliance assessments" that utilise, in real time, the data that AI and IoT generate. Finally, it is also worth exploring if entirely novel forms of legislative drafting are required, such as "Law as Code". In this approach legislators (or regulators) enact a natural language version side by side with a code-based version that is then used for automated compliance assurance. This addresses the emerging democratic deficit in technology regulation that sometimes turns software developers into de-facto legislators, by leaving the translation of legal code into machine code to them. There are several tentative initiatives that are currently trialled in some jurisdictions, the UK has the knowledge and expertise to become a leader in this field.

2. Do you agree with the context-driven approach delivered through the UK's established regulators set out in this paper? What do you see as the benefits of this approach? What are the disadvantages?

i. Context-driven approach and implementation strategy

6. The advantage of the context-driven approach is that it provides a differentiated picture of AI regulation and risk-assessment, focusing on sectoral application of AI technology. In addition to a set of underlying principles for AI governance, the context-driven approach can fine-tune or optimise proportionate regulatory intervention for a given AI system. It provides both *stability and flexibility* of regulation, so long as the regulators have the necessary resources and expertise. As noted above, different applications of the same AI technology will indeed create very different risk profiles, and the expertise of sectoral regulators will be needed for framing adequate and proportionate responses.

The government's proposal assumes that regulatory bodies have the resources and expertise for capacity building. This is questionable. Even in fields where there is an existing regulator with significant expertise, such as the ICOs for the data protection implication of AI, there is a known and increasing enforcement deficit. Brexit already necessitates the replication and/or

recreation of regulatory structures in crucial fields such as medicine. Giving these regulators additional tasks to regulate AI will not be possible without significant investment. Considering the role of Approved Bodies in the medical device sector is illuminating. The MHRA recognises the implications of third-party conformity assessments and CE marking for the UK market, which includes the recognition of certificates issued by an EU Notified Body until 30 June 2023 (MHRA, last updated, 2022; compare Northern Ireland approach). This implies a risk fragmentation – considering the 'UK industry's interests in converging international standardisation' – and therefore, avenues must be designed for industry to demonstrate compliance with international standards, and for interpreting standards, which requires further guidance by the UK government, as well as the MHRA (Department for Digital, Culture, Media and Sport, 2022, p.5). Further, the way MHRA designates UK Approved Bodies against relevant requirements for UKCA marking implies regulatory capacity and assurances for industry- from clarifying implementation strategies to enforcement which provide certainty for regulatory approval.

7. We suggest a revision of the 'light-touch' approach, which implies deregulation on a system level. The Department of Business, Innovation and Skills highlights that product market deregulation might negatively influence industries' engagement with innovative activity, provided this strategy will significantly limit the level of competition in industry (Department of Business, Innovation and Skills, 2012, p. 25). Rather, we suggest agile, smart and fit-for-purpose regulation that is decentralised but supports harmonised AI governance. This can be addressed to a degree through enforcement discretion of low risk AI, as practiced already by the Information Commissioner's Office (ICO) in data protection law, or through the US FDA's flexibility strategy involving selecting priority cases and ignoring minor changes where impacts are not significant. But enforcement is only one role of regulators. The ICO in particular was not conceived just as an "enforcer" after the event, but also as a place where individual companies and entire sectors could go for guidance and support.

ii. The context-driven approach and AI innovation

8. The stated objective of the proposal is to enhance AI regulation whilst being "pro-innovation." The implicit juxtaposition of regulation versus innovation was flagged up above, as in our view it underestimates the positive role of (smart) regulation for market success. Recently, representatives for large social media platforms called for more stringent regulation, and similarly, the cryptocurrency/blockchain sector has long left its libertarian roots behind and is arguing the benefits of regulation for investors and developers. While some of these calls may be self-serving, as large incumbents find it easier to shoulder compliance costs than smaller and younger competitors, the increasing recognition that market success requires levels of public trust, which industry alone cannot generate, is convincing.

   A shortcoming of the proposal in this respect is that it treats the sector as largely homogenous. In reality, the proposed context driven and "light touch" approach is likely to affect different industries and different business models in very different ways. If the aim is to support UK businesses that *develop* AI, a separate UK approach to AI innovation risks creating burdens on those companies that aspire international reach, which will have to comply with foreign international standards in addition to those relevant for the UK market.

For example, any UK AI export will likely have to comply with the EU Commission's Proposal for a Regulation on Artificial Intelligence (EU AI Act) and/or the US Algorithmic Accountability Act. Even if the UK approach were "lighter" than that of the EU or US, in the absence of formal recognition (mutual or unilateral) of the other regime(s), it will increase compliance burdens. By contrast, UK approaches that, due to their nature, only aim at the national market – for instance an in-house development by a government department – would potentially benefit from a less arduous regime (but would risk a loss in public trust, especially if the inevitable comparisons to the other regimes are made should a problem occur). There will be different impacts on UK industries and businesses in all sectors that do not develop AI tools, but benefit from their *use*, regardless of where in the world those tools have been developed. In this instance a low-regulation environment may be attractive, but the UK risks becoming a testbed for technologies that foreign legislators and regulators still consider too risky for citizens in their home jurisdictions. This may give the edge to UK businesses and allow them to experiment more and see a steeper increase in productivity – but, again, with a much higher reputational risk should these systems then fail. Due to the 'Brussels' and 'Silicon Valley' effects, foreign AI developers are, in any case, likely to focus their development efforts on systems that comply with the EU and US rules, in order to access these much larger markets. As with the situation for UK based companies, regulatory fragmentation would then increase their costs in the UK, even if the UK system were comparatively light touch.

There are a number of strategies possible to mitigate these risks. One is a systematic use of "recognition of compliance" both between different UK sectoral regulators ("if a Human Rights impact assessment was made for police uses, this will be recognised also for AI use by social media regulators"), and also between international regulators ("EU AI act compliance automatically implies UK compliance"), which may or may not be mutual.

9. The policy document implicitly pitches "innovation" against "regulation". Historically, successful disruptive technological innovation always co-evolved with innovative approaches to regulation. The modern-day emphasis on third party safety certification grew out of experience with the then-revolutionary steam engine, which, at the time, was associated with dangers – opacity, responsibility gaps, inequitable allocation of risks vs benefits – which are eerily familiar with those being associated with AIs today. While international trade agreements, including mutual recognition of certifications, can mitigate some dangers for a separate UK approach for AI innovation, however, they may simultaneously reinforce a democratic deficit that is already present in EU and US proposals (i.e. the EU AI Act and US Algorithmic Accountability Act) for AI regulation, by delegating crucial normative decisions to regulatory bodies outside parliamentary control. International trade negotiations tend to be subject to only minimal public and parliamentary scrutiny, and are consistently lacking in transparency. We argue that key normative decisions – regarding, for instance,equitable risk allocation – should come in the form of primary legislation and after proper public debate, with clear accountability of the respective ministers and MPs. A citizen who wants to challenge decisions made by an AI operated by a UK authority, for instance, should not be bereft of remedies because a trade deal agreed to exempt the foreign developer from the duty to provide information in a UK judicial review process.

10. The government proposal sees in this sectoral, problem-specific approach a main distinction between the UK's approach and the EU AI Act's proposed "top-down framework". However, on closer inspection these differences may be on the surface level only. While it is true that

the EU AI Act starts with abstract and general definitions of AI, and cross-sectoral risk categories, individual provisions quickly carve out sectoral exceptions and rules (e.g. on facial recognition in policing). The references to different regulatory bodies, and the lists of concrete examples in the Appendices, further push towards sector- and application-specific rule-making. The UK proposal in turn starts with a discussion of sectoral and domain specific issues, but then discusses the need for coherent, cross-sectoral rules as well. In the end they may reach the same goal, if from different directions. In this case it would be particularly regrettable if UK companies would face additional compliance costs to achieve substantially identical outcomes.

Both the UK and the EU proposal struggle with application agnostic AI systems, especially if these in turn are the result of decentralised or open source initiatives. Neither system gives good answers on how to treat e.g. open source Large Language Models (LLMs), a key driver of many AI applications. It is unclear in particular how the EU AI Act would allocate design and documentation responsibilities to the users of such a model (whose development is entirely outside their control). The same issue however also arises under the UK proposal, indicating further that despite differences in structure, the two approaches cohere on substance and in practice.

The EU AI Act has to be understood in terms of its wider reform agenda to complement the General Data Protection Regulation (GDPR). The UK faces the challenge that it reforms its data protection regime at the same time as it starts developing a regulatory approach to AI. These two projects will need close co-ordination and alignment, especially as changes in the data protection regime may cause already significant disruption and uncertainty to UK businesses, in addition to the danger of losing adequacy recognition by the EU, an independent "inadequacy finding" for the purposes of the AI Act would be a serious impediment for cross-border services. This issue, obviously, would also impact the situation in Northern Ireland. While the focus of the discussion there has been on movement of goods, interruption of cross-border data exchange and AI service provision are likely to be just as harmful. The government should ask itself what the regulatory regime for a NI startup company will look like that provides e.g. image analysis services for hospitals in the Republic of Ireland and the UK mainland, and for this requires transfer for patient data.

The EU AI Act is part of a package that includes the draft Digital Services Act, the draft Digital Markets Act (with provisions on AI-relevant hardware, operating systems and software distribution); the draft Machinery Regulation; announced product liability revision relating to AI; and the draft Data Governance Act. The provisions of these Acts are not mirrored in the UK proposal, and there is the danger that by treating AI regulation in isolation, regulatory gaps will emerge.

iii. Context-driven approach and responsible innovation

11. The context-driven approach requires regulatory bodies to streamline cross-disciplinary findings for responsible innovation. The NHS AI Lab refers to external resources, including the ICO and The Alan Turing Institute guidance on explaining decisions made with AI, with regard

to AI development in health and care (ICO and the Alan Turing Institute, 2020). Regulation should be informed by best available science, both regarding technological aspects but also the potential long-term social and psychological implications. For the domain-specific regulatory bodies, this requires the consultation of external resources, as well as substantial capacity building and a nuanced assessment of input from fields (such as Human Computer Interaction (HCI) and, sociology of technology) that in the past may not have been conceived as relevant for the regulator in question.

12. A disadvantage regarding the government's proposal is that it does not articulate how the context-driven approach interacts with cross-sectoral principles and international standardisation. Global technical standards and assurance services are intended to facilitate trade in international markets. At the same time, the government intends to pursue a 'multi-stakeholder approach' in maintaining a 'values'-focused perspective on AI innovation on the international plane (Department for Digital, Culture, Media and Sport, 2022, p.12, p. 17). Here the main issues will only become apparent once more detail is provided. Often, purportedly "multi-stakeholder approaches" benefit only a small number of stakeholders, those with the resources and expertise to make their voices heard. As noted above, if key questions surrounding the implications of technology – from value-focused judgements on the impact of AI on health, safety, and human rights to the implementation of fundamental values- are to be delegated to regulatory bodies at all instead of parliamentary scrutiny, active, and fully resources, steps would have to be taken to ensure the stakeholders are fully representative and represented, and the voice of the various publics given equal room and consideration.

13. A targeted approach and context-driven approach to AI need to strike a fair balance between supporting innovation and a human-centric perspective of AI regulation. The policy document makes a direct reference to the OECD Principles on Artificial Intelligence, the context-driven approach to be 'values' focused. The risk-based approach to AI regulation is based on the idea that 'AI is a dynamic, general purpose technology and that the risks arising from it depend *principally* on the context *of its* application' (emphasis added, Department for Digital, Culture, Media and Sport, 2022, p.11). How this risk-based approach complements 'existing and distinct approaches to AI regulation' and the government's formulation of cross-sectoral principles depends on whether a context-driven approach adequately reflects *emergent risks* and *emergent contexts* regarding AI application.

    The role of advanced AI and ML approaches – giving rise to differentiated risks based on the context of their application in medical imaging, personalised medicine and/or robotic surgery – is 'bringing a paradigm shift to healthcare' (UK Taskforce on Innovation, Growth and Regulatory Reform, 2021, para 390). The NHS AI Lab recently conducted a survey examining public perception and trust regarding the role of AI to support diagnostic decisions, whereby the degree of oversight with regard to AI as decision-support in dynamic healthcare environments is an important aspect for trustworthiness and uptake of technology. Similarly, robotics in surgery raises specific issues of oversight and training in human-robot interactive settings (Beane, 2020). These are examples whereby *emerging contextual uses* can shape new concerns that need to be reflected in the cross-sectoral principles to promote trustworthiness and uptake of new technology.

The policy document mentions that 'the precise impact of this technology will vary greatly according to its context *and* application' (emphasis added, Department for Digital, Culture, Media and Sport, 2022, p.8). Here, context can be defined by the specific risk-profile of AI technology "horizontally" and within a sector (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme, 2022, p. 23). In particular, 'cross-regulatory' collaboration intends to 'create holistic understandings of the ways that AI is being used and its impacts….recognising that AI is increasingly employed in ways that cut across traditional sectoral boundaries' (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme, 2022, p. 83). This step should be clarified using specific guidance on how the context-driven approach addresses "common cross-cutting challenges" across AI sectors. For instance, the role of regulatory bodies' risk mitigation strategies is an important indicator for the government to identify common contextual challenges of AI on the systemic level.

In addition, the proliferation of AI methods in sectoral applications, including healthcare, implies cross-cutting challenges to identify emergent risks for health and safety. The Equality and Human Rights Commission's Strategic Plan 2022-2025 intends to work with 'governments, service providers and regulators' and identify 'discriminatory barriers that stop people from accessing health and social care services are understood and addressed' (Equality and Human Rights Commission, 2022, p. 21- 22). The NHS Transformation Directorate and NHS AI Lab, using the National COVID-19 Chest Imaging Database (NCCID) for AI models, currently want to understand the role of bias based on the technologies' validation process including validation data. The former Secretary of State for Health and Social Care commissioned a review for the government on how to tackle health inequalities and how design and deployment of medical devices should respect diversity considerations. These considerations are illustrative of the way *emergent risks*, such as risks of bias and lack of diversity are multi-faced and require constant negotiation between regulators to inform and clarify the relevance of cross-sectoral principles on an ongoing basis. This point could be clarified based on further guidance channelling the context-driven and risk-based approach, and how proportionate responses are 'values-focused'.

iv. Context-driven approach and legislative design:

14. The proposal states that it does not create 'a new framework of rights for individuals' (Department for Digital, Culture, Media and Sport, 2022, p.12). There are two aspects to this question. One is the creation of new substantive rights, such as a right to "minimal human contact" in a care home setting, or a right to use an AI as assistant in legal procedures. The other is the creation of new procedural rights, such as a right for class actions if an AI provider broke the law but the regulator is unwilling to act.

Closing the door on new *procedural* rights, including new forms of contestability of AI decisions, would in our view be premature. Finding new ways to get fast, cost-efficient and appropriate redress may be required, in addition to enabling citizens to enforce their rights. This can be through forms of class action litigation, or by giving some third sector organisations standing in bringing actions on behalf of citizens (in analogy to the super complaint in consumer protection law). The rules of evidence and procedure will need revisiting to ensure the right type of data, at the right level of granularity, can be produced

during litigation. This is true even for the more cautious and limited approach taken in the proposal

A different question is if the UK AI framework also needs new substantive rights. It seems sensible to exclude them from a generic regulation. However, this should not prejudge the creation of new substantive rights later on, either by the sectoral regulators or through primary legislation. This "non-prejudicial" approach to new substantive rights should be stated explicitly. Here the approach taken in the EU AI Act may pose significant dangers for the EU member states and their democratically elected governments, as it might close the door on the creation of new substantive rights by national governments, if these burden free movement of AI services or products.

3. Do you agree that we should establish a set of cross-sectoral principles to guide our overall approach? Do the proposed cross-sectoral principles cover the common issues and risks posed by AI technologies? What, if anything, is missing?

i. Cross-sectoral approach to common risks and contextual uses

15. Yes. Cross-sectoral principles provide a necessary and flexible foundation for sectoral regulation of AI and intend to complement existing regulatory regimes. Without them, the negative effects of fragmentation outlined above will be worse. They are also needed, as discussed above, to create a predictability for developers of generic AI systems where it is yet decided in which regulatory context they will eventually be used. A particular problem in this context is the regulation of generic open source tools such as large language models. The EU AI Act fails to adequately address these tools by treating them in the same way as systems developed by a single developer or company, making them potentially unusable for many contexts where they have significant benefits. Finding a way to regulate their use appropriately will also be challenging for the UK. In any case the only way to address them is through cross-sectoral rules, which however have to account for the fact that the development of these tools is highly decentralised, and their users have no influence over things such as their data accuracy, completeness or quality.

The Taskforce, on Innovation, Growth and Regulatory Reform independent report highlights the benefits of 'agile regulation' that allows for 'more forward-looking, judgement-based regulation without needing such complex and exhaustive rules for every situation [to be] set out in advance' (TIGR report, para 53). The implications of this approach are that minimum standards and risk-outcomes need to be outlined in advance and should be based on the cross-sectoral principles. In particular, the Better Regulation framework Interim guidance highlights that Parliament 'should set out only what is prohibited or the outcomes to be achieved, in plain English, and sets out any parameters within which regulators would need to operate to meet these outcomes', whereby these parameters should apply irrespective of the discretion of regulators to act upon the interpretation of cross-sectoral principles (Better Regulation Framework: Interim guidance, 2020, 3.1.6). Whether cross-sectoral principles can sufficiently provide for the parameters regarding common issues and risks of AI technology depends on how far minimum standards – on fairness, transparency, safety, security,

responsible use, and contestability – support existing rights and fundamental values irrespective of the context and use of AI technology.

16. An important aspect that is missing the cross-sectoral principles are a baseline for impact assessments on the use of AI technology. Beyond product liability and the question if an AI system functions correctly, it is also necessary to look at the human rights implications on basic principles/values for governing AI systems: such as transparency, accountability and fairness. As noted elsewhere, industrial standard setting bodies are often ill equipped to account for these wider concerns, and even regulatory bodies may lack expertise and skills to interrogate AI systems on their human rights impacts. By regulating the use of AI but not the technology itself, it is possible to create proportionate obligations for users to assess the social, legal and ethical implications of AI systems. This is facilitated by a context-specific approach based on the use and the impact of AI on individuals, groups and businesses within a particular context, and to delegate responsibility for designing and implementing proportionate regulatory responses to regulators. However, with the future direction of the UK with regards to the UK Human Rights Act uncertain (Ministry of Justice and Justice Secretary Dominic Raab, Press release, 2021; the Independent Human Rights Act Review, 2021), there is an additional danger that the government policy proposal is missing a broader regulatory ecosystem with flanking legislation that fills gaps that are otherwise inevitable. Both the EU and the US approaches benefit from such a regulatory ecosystem that create contexts for AI regulation, whereas the UK faces the challenge that these flanking laws (from the Human Rights Act to Data Protection law) are also in a state of uncertainty.

ii. Cross-sectoral principles and robust baseline on the regulator level

17. A baseline regarding impact assessments would strengthen the context-driven approach on the regulator level. For instance, regulatory reform with regard to medical device regulation intends to consider an approach that is agile and adaptable to safety considerations of the AIaMD lifecycle. The MHRA plans a work programme with several work packages, whereby WP 10 Project Glass Box stream is dedicated to translate opacity in AIaMD into safety considerations and integrating frameworks regarding interpretability of AIaMD. Clearly articulated, minimum standards and risk-profiles acting as a baseline and/or due diligence obligation for regulatory bodies, could inform how notions of accountability and responsibility offer a shared understanding across regulatory bodies and a robust framework for further risk-based criteria established by standard-setting bodies.

    The government's proposal does not stipulate clear risk-based criteria for the enforcement of cross-sectoral principles. An agile and smart approach enables regulatory bodies, including the MHRA, to deliver proportionate responses and risk-based criteria in guidance or voluntary measures. The policy document emphasises that 'regulators [should] focus on high risk concerns rather than hypothetical or low risks associated with AI' (Department for Digital, Culture, Media and Sport, 2022, p. 2). The consultation document reforming the Framework for Better Regulation further highlights that a proportionate approach is 'focused on risk and reaching the right outcome' (Department for Business, Energy & Industrial Strategy, 2021, 3.2.5) to keep up with the pace of emerging technology. Delivering proportionate and adaptable risk-based criteria to AI applications require clearly outlined priority areas – set out by regulatory bodies and government – to enable dynamic

adaptations to existing processes and create new ones and risk-based criteria. For example, the MHRA recently proposed to introduce an obligation for manufacturers to implement measures that are 'proportionate to the risk class, type of device and the size of the company, to cover any legal liability arising from adverse incidents with medical devices that they place on, or supply to the UK market' (MHRA, 2022, p. 22). The government's closer involvement to define outcome-based regulation and proportionality should clarify when regulatory change and revision of cross-sectoral principles are required. At the same time, independent expert committees, including the Regulatory Horizons Council (RHC) can assist with regard to the identification of priority areas.

18. Finally, a holistic approach to regulation entailing cross-sectoral principles requires mechanisms for regulatory oversight. The consultation document reforming the Framework for Better Regulation recognises the downside of delegating regulatory discretion to standard-setting bodies, arguing that 'this approach could lead to more uncertainty in the regulated markets and more litigation' (Department for Business, Energy & Industrial Strategy, 2021, 3.2.7). In addition, the document underlines that delegating more power to regulatory bodies 'could ultimately lead to more regulation being created overall, through mechanisms which are less responsive to public scrutiny and democratic accountability' (Department for Business, Energy & Industrial Strategy, 2021, 3.2.7). The Digital Regulation Cooperation Forum (DRCF) recently published a discussion paper to clarify the role of algorithmic audit and emphasised the 'role of governments, standards bodies, large technology companies …. and wider civil society actors', albeit the DRCF's mission would be restricted to core areas entailing 'digital regulation and online platforms' (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme, 2022,p. 75). That is why, a baseline needs to be articulated with regard to the cross-sectoral principles, to ensure robust standard-setting, as well as a comprehensive mechanism for oversight, and due diligence.

4. Do you have any early views on how we best implement our approach? In your view, what are some of the key practical considerations? What will the regulatory system need to deliver on our approach? How can we best streamline and coordinate guidance on AI from regulators?

i. Positioning risk-based and proportionate approach across regulators

19. Whilst regulatory bodies need to proactively ensure that cross-sectoral principles apply to the risk-based criteria and sectoral regulation of AI, it is the government who ensures that the 'pro-innovation' approach is embedded in regulatory practices. The government 'can offer a strong steer to regulators to adopt a proportionate and risk-based approach (for example through government-issued guidance to regulators)' (Department for Digital, Culture, Media and Sport, 2022,p. 13). In addition, the policy document highlights 'the role of technical standards and assurance mechanisms as potential tools for … enabling international trade' (Department for Digital, Culture, Media and Sport, 2022, p.18). As noted above, this emphasis on (international) standard setting bodies – whose approaches are often dominated by industry stakeholders at the exclusion of civil society – contributes to a potential democratic deficit in AI regulation. Most of the standard setting bodies have no mechanism or expertise to include ethical or legal dimensions in their processes. For the UK,

the British Standards Institution (BSI) could be supported in developing this broader expertise and include it more systematically in their standard setting activities.

ii. Streamline and coordinate guidance on AI from regulators

20. A key practical consideration regarding the implementation of the UK government's approach is to address cross-sectoral issues and fragmentation or duplication across regulatory bodies. We support the 'creation of an AI and Regulation Common Capacity Hub (ARCCH)' (Aitiken, Leslie, Ostmann, Pratt, Margetts, Dorobantu, The Alan Turing Institute Research Programme, 2022, p. 7). Its role would be coordination and assistance – in the same way in which the ICO does not just enforce data protection law, but assists businesses in achieving compliance, by providing both general information and individual guidance and feedback. This could also include the provision of Regulatory sandbox or Pre-Cert Pilot programmes in particular for Small and Medium Enterprises (SMEs).

21. In addition, we recommend that a "common AI hub" could adopt a "one-stop-shop approach" across regulators and which is represented by members from regulatory bodies, from the ICO and Equality and Human Rights Commission, the MHRA to the UK's Centre for Data Ethics and Innovation (CDEI). For instance, NICE and the Care Quality Commission (CQC), the Health Research Authority (HRA) and MHRA recently developed the Multi-Agency Advisory Service (MAAS) for AI and data-driven technologies (expected to go live in early 2023). A "common AI hub" needs to strengthen multi-agency partnership, which refine evidence-based solutions for key AI areas within organisational structures.

5. Do you anticipate any challenges for businesses operating across multiple jurisdictions? Do you have any early views on how our approach could help support cross-border trade and international cooperation in the most effective way?

22. Some of this we covered earlier in this report. Yes, there are significant dangers that a UK-only approach will leave UK businesses stranded between the "AI regulatory blocks" in the EU and the US, forcing developers to work towards multiple, mutually inconsistent standards. International trade cooperation might mitigate some of the negative consequences, but at additional costs for regulatory transparency and democratic accountability which are already impediments to creating public trust in and acceptance of AI.

23. It is important to remember that there are also multiple jurisdictions within the UK. The government proposal states that unlike its EU counterpart, ensuring free movement of AI-services and AI-enhanced goods is not a concern, and therefore permits a different approach. This underestimates the impact a UK AI framework will have on the devolved powers. Police, education and health are all likely users of AI, yet fall under devolved powers. Northern Ireland poses its additional difficulties. The precise relation between the Act and the devolved powers needs clarification, a UK-wide AI Act requires involvement and support by the devolved parliaments.

24. Being outside the EU standard setting bodies, and also outside the post-market surveillance system that the EU is proposing (EU AI Act, TITLE VIII), could mean that UK regulators, and also UK businesses, are "frozen out" of access to information about risks that have materialised, known problems or successful mitigation strategies. This could also be mitigated by more efficient community/sectoral sharing of data, systems, approaches and know-how. Intelligence about incident sharing needs to be balanced with economic interests of the AI developers, including but not limited to their IP. Sharing of information between regulators and manufacturers can assist evidence- based regulation. Regulation will be needed to create the right incentive structures for sharing information about negative events. This will likely combine legal duties, economic incentives and a trusted environment that ensures disclosed data is not misused.

25. The government's proposal emphasises the role of standards for a coherent regulatory approach. The Alan Turing Institute, supported by the British Standards Institution (BSI) and the National Physical Laboratory (NPL), intends support the UK's engagement regarding the articulation of global technical standards based on an AI Standards Hub (Department for Digital, Culture, Media and Sport, 2022; Department for Digital, Culture, Media and Sport, Office for Artificial Intelligence and Chris Philip MP, Press Release, 2022). An important consideration regarding this effort is to learn from successful practices and build methodologies supporting international standardisation. One example of successful practice in the area of medical device regulation are shared knowledge units including Quality Management Open Source Models for technical documentation. In addition, the AI Standards Hub will need to draw from specialist advice for information consolidation of international standards and medical device regulation to support industry (see for example, the work by the OpenRegulatory group). The AI Standards Hub, engaging with working groups, could streamline common regulatory ontologies, such as terminologies for adverse event reporting in medical device regulation (see, IMDRF Adverse Event Terminology Working Group, 2020). Finally, the AI Standards Hub needs to envisage a collaborative approach that includes manufacturers, regulators and other stakeholders, such as patients using medical device technology, to deliver evidence-based approaches informing international standardisation.

26. In addition, the scope of independent regulatory activity with regard to extension of SaMD and AIaMD need to be identified. The MHRA intends to develop 'access routes that build on synergies with both EU and wider global standards' to address this (MHRA, 2022, p. 6). The MHRA also aims to introduce greater transparency of regulatory decision making through updating the requirements that apply to Approved Bodies and increasing the consistency of conformity assessments…' (MHRA, 2022, p. 8). Finally, the Regulatory Horizons Council (RHC) proposes Mutual Recognition Agreements to promote 'international regulatory cooperation' (RHC, 2021, p. 24). These activities indicate that the scope of independent regulatory activity will require capacity building for both regulatory bodies, as well as the medical device industry. In particular, streamlining the regulatory approval process requires a delicate balance between standardising processes, such as post-market surveillance for specific device types, as well as maintaining a high-degree of safety for patients regarding emerging technology (see, MHRA, 2021).

6. Are you aware of any robust data sources to support monitoring the effectiveness of our approach, both at an individual regulator and system level?

27. The government's proposal intends to build on existing regulatory bodies – including their use of existing data sources – to support an evaluation and monitoring mechanism on the system and regulator level. For instance, the MHRA uses an adverse incident reporting mechanism regarding medical devices. However, an important challenge will be operationalising findings regarding the existent regulatory bodies' evaluation mechanism. In particular, there are important pitfalls in using adverse incident reporting - from using patient incidence data that is likely to contain reporting bias to lack of technical AI knowledge informing incidence reports - to evaluate the effectiveness of regulatory approaches (Tase, Buckle, Ni and Hanna, 2021, p. 137). In addition, a more-unified approach regarding terminology of incidence reporting could improve the quality of results (see for example, Basereh, Caputo, Brennan, 2021). A shared methodology to address risks of fragmentation in the quality of reporting will help the UK government to monitor the effectiveness of regulatory approaches on a system level. In particular, the UK government's proposal needs to think about shared indicators to identify areas for policy intervention (see for example, the OECD approach in Measuring the Digital Transformation: A Roadmap for the Future, 2019).

28. Another task is defining methodologies to monitor the effectiveness of the UK's approach on a regulatory level. For instance, post-market surveillance is intended to provide safety and performance data on the AI and illustrates streamlined review and risk management process at individual regulator level. However, the post-market surveillance mechanism might require new approaches for incidence reporting. The MHRA intends to investigate novel AI signal detection techniques to be more responsive regarding the use of medical technology on the ground. In addition, the MHRA intends to 'monitor trends in data from a wide range of different sources, which enables us to quickly identify safety concerns' (MHRA, Medicines and Healthcare products Regulatory Agency Delivery Plan 2021-2023, 2021). Therefore, new methodologies on the regulator level need to be dynamic and traceable considering modes of assessment (such as, robust and uniform processes, robust on high-quality data), and quantitative and qualitative evaluation, as well as be supported by a dynamic oversight mechanism. This is a challenging task as complaint data might give a handle on how regulation might need to be adapted but we would need more data on how the regulator identified where the adaptation could take place, how quickly the regulator responded and how effective the adaptation is.