



Wales Safer Communities Network response to: Department for Science, Innovation and Technology - A pro-innovation approach to AI regulation

Closed 21 June 2023

Response submitted via the online survey.

Consultation Questions

1. Do you agree that requiring organisations to make it clear when they are using AI would improve transparency?

Yes, as currently it could be being used without a person's knowledge and their information processed in ways they have not anticipated and would not agree with. By identifying clearly that AI is being used allows consumers, clients and other members of the general public the option to not engage with the organisation or if there is an option to turn off the AI that they are able to do so.

We feel that the patchwork of legislation makes it difficult for organisations and communities to understand and potentially adds to the feelings of mistrust, caution and fear. Providing clarity and being transparent could allow people the opportunity to be more aware and guard themselves and loved ones from any harmful AI, or may specifically look for specific AI that they trust.

2. Are there other measures we could require of organisations to improve AI transparency?

We suggest organisations use clear, concise, jargon free- clear language to share what technologies they use. This should be displayed and easily accessible on Websites, Terms and Conditions, client/ customer contracts and other key documents or platforms where people may expect to find information. There should be a need to have an AI statement if AI is used by any organisation just as a privacy policy or cookies policies are required for transparency and reassurance of compliance with both data protection and confidentiality legislation.

It should be mandatory for information to be easily accessible of the types of work and analysis that the AI is undertaking, so there is an understanding of, what, how, when and where the AI is functioning. If the AI is running from a different server or by another company or organisation on behalf of the organisation that is being engaged with that should be made clear and the safety standards and compliance for security be available easily through either organisation.

To assist in transparency, we think there could be an icon or logo that should be linked to any AI content, so that AI content is clearly identifiable. Alongside this there may also be the possibility of including a quality AI mark or similar which shows that



the Framework is being adhered to and has been checked by the regulators, providing extra assurance around the safety of the AI in use.

Staff training should be offered to all staff so they are aware of the AI used by the organisation, how they engage with it and their responsibility in ensuring it is used safely and in accordance with the law, regulators, principles and any other guidance.

Information on how to report concerns or inappropriate use of personal information should be shared with users/ customers from the outset in both the privacy policy and in a separate AI policy, similar to the interaction between privacy policies and cookie policies on many current websites.

3. Do you agree that current routes to contest or get redress for AI-related harms are adequate?

No, not at the current time. The AI landscape is constantly in a state of flux with changes occurring rapidly and beyond the control and experiences of those developing the AI, as the AI processes faster and differently information and data than most humans can. This rapid advancement means there is an increased risk of harms and potential misuse as developers work to keep track of developments and identify any potential changes in direction or breaches of legislation through unintended consequences or accident. Any breaches will risk organisational risk as well as the possibility of a declining trust in AI or at least that specific version of AI.

Organisations, criminal gangs and individuals who wish to use technologies for inappropriate/ illegal and harmful practices will continue to do so, and find new ways to exploit, coerce, intimidate, abuse and harm for their own personal gain. However, by consolidating the legislation, making responsibilities clearer then it will help to identify bad practice, poor management and ineffective use of resources easier and enable any appropriate action can be taken through civil and/or criminal courts.

The Framework appears to support regulators providing Guidance to users which will help to document the way to contest and seek to redress issues. If these issues are raised and appropriately acted upon with lessons learnt and systems adapted, it could help reduce inappropriate AI usage and develop more trust from those being asked to use it.

The Framework mentions the global interoperability and global ethical standards but there was no ethical justification within the framework to ensure regulators can be confident there is appropriate guidance and support available for customers and other users of AI.

There is already some concern being raised around the use of AI and other technologies to enable stalking and other forms of dangerous behaviours, nothing in the framework appears to try to address these elements. There is mention of developing pornography, but not the risk of inappropriate child images created using photographs and images of real children but whose bodies are then created by AI.



There is also the risk of increased polarisation where only information from one perspective is presented, giving an inaccurate reality to the user, potentially leading to further harms including radicalisation and extremism as the distorted view of the world becomes the brain's reality, thus potentially becoming a circular problem where each other feed information from each other with no opportunity to break the cycle.

There is a need for AI security and the safeguarding of children and adults at risk or other forms of harm remain a key priority and receive investment and commitment in order to manage the changing landscape and attempt to eradicate or at least limit the harm inappropriate use can have.

4. How could current routes to contest or seek redress for AI-related harms be improved, if at all?

Clear, jargon free language. Contest routes or ways to seek redress need to be clear and accessible to all.

Sign posting to support services should users/ customers have been victim of identity theft or fraud should be present on websites, social media platforms, Apps and on other places including relevant contract information and service level agreements where AI is in use.

There should be an opportunity for the regulators to have a shared learning resource so that any issues for contest and redress can be shared and managed across the sectors, sharing best practice and empowering the developers to fix issues in AI that have developed in other areas but not yet their own.

If AI is used to cause harm through harassment, abuse, exploitation and criminal activity then there may be a need to update current legislation to ensure that this is covered especially where domestic abuse, safeguarding and all forms of illegal activity but especially fraud (including identity theft) then there should be a way that it has to be turned off and all data destroyed so it cannot be used again. It is unclear from the framework if this would be possible and therefore could be a weakness.

5. Do you agree that, when implemented effectively, the revised cross-sectoral principles will cover the risks posed by AI technologies?

The cross sectoral principles are unlikely to eradicate all risks posed by AI, but by ensuring there is an interconnectivity could enable a clearer landscape and better sharing opportunities between the regulators. There is also the chance for better navigation and understanding by the regulators, organisations and individuals utilising AI and for service users and customers to understand and potentially engage with in a safe manner able to identify risks and issues so they can appropriately step back and report. There will still be the risk of those who want to use AI for bad purposes to do so, such as fraud, corruption, exploitation and theft. Therefore, anything to hinder them should be seen as a positive which a principle-based approach may be able to achieve.

Some of our more vulnerable or older members of the community often struggle to access and understand online services, or due to rural environments find



connectivity a challenge. We would want to see alternative service provision available to support these customers so they were not at a disadvantage.

The unintentional inclusion of inaccurate or misleading information is concerning and could lead to additional harm being inflicted. Whilst the autonomy section (in cases study 3.1) suggests little need for oversight we would recommend that oversight is certainly required especially when there is likely to be a distressed and anxious individual involved, potentially leading to increased or new mental health issues.

When mistakes happen there need to be safeguards in place to ensure these are identified and rectified as soon as possible allowing appropriate information to be shared with and between regulators.

6. What, if anything, is missing from the revised principles?

Whilst AI has many potential advantages in enhancing science, saving time and reducing overheads, there is a real threat that the use of technologies are impacting on industry and by default whole communities, with examples of some companies now opting to implement ChatBot services rather than employ staff in a call centre capacity which may increase poverty and homelessness. When elected Officials are opting to use free services such as ChatGBT whilst drafting speeches means that jobs are no longer available, opportunities are being lost and more existentially there is a risk that thinking is being done by programmable machines with the potential for manipulation and subtle influences in language which could have a detrimental impact on individuals, communities and society as a whole.

There appears to be little in regard to AI being used only for legal purposes, including around the use of personal data and information which developers may use to target specific groups, communities or individuals.

7. Do you agree that introducing a statutory duty on regulators to have due regard to the principles would clarify and strengthen regulators' mandates to implement our principles while retaining a flexible approach to implementation?

Yes, in principle but we think it should go beyond the regulators having a statutory duty to have due regard but that should apply to those using and developing AI as well. Whilst it may add to the bureaucratic burden on companies, without it organisations would be unlikely to comply fully or at all. It is our opinion that the only way to ensure compliance with the framework, by all except those intent on criminal activity and use of AI, is for the principles to be enshrined in legislation and not just in taken in 'due regard'.

8. Is there an alternative statutory intervention that would be more effective?

Yes, fully legislation that would enable the prosecution and demand that the principles and framework are not just thought about but adhered to. Without this it is unclear what powers the regulators will have if any to deal with any breaches whether deliberate, accidental and the potential consequences of such breaches on individuals, organisations, communities and specific protected groups.

9. Do you agree that the functions outlined in Box 3.1 would benefit our AI regulation framework if delivered centrally?

We agree in principle with the content of Box 3.1, specifically the need to ensure there is a consistent approach to Risk Management and that the escalation routes are clear for regulators.

The Education and Awareness section could be enhanced with a clearer explanation on how a coordinated messaging approach would be coordinated and how that would be delivered.

We would like to see more commitment within the framework in respect of Horizon Scanning, whilst it is appropriate to suggest “Proactively convene industry, frontier researchers, academia and other key stakeholders”- it does not set out what this would look like, would this be a regular forum, electronic communication, network, or ad hoc engagement? There is also no detail on how appropriate parties to represent would be identified.

We noted the lack of information in this section of the different working relationships within the devolved administrations. We would recommend including that this is an area where both reserved and devolved policies and practices interface. There needs to be clear join up with Welsh Government and other devolved agencies in Wales (such as NHS Wales, DEFRA etc). Under the section on Interoperability with international regulatory frameworks, it may be practical to make reference to ‘International and Devolved Administrations’ though a separate section may be more appropriate.

10. What, if anything, is missing from the central functions?

Funding and other types of resourcing information. In the impact assessment (page 38) whilst there is a large section on cost how this will be met is not transparent. In addition, there is no mention of the impact this will have within the reserved policy areas within Welsh Government and other Devolved agencies and partners in Wales.

There is no clear explanation of who will fund the central function and from where this money will come from.

The Welsh Language (Measures) Act 2010 specifies that Welsh is to be treated equally to English we would therefore expect provision to be made or at least mention within the central functions around language use and the option to choose for those living, working or visiting in Wales.

11. Do you know of any existing organisations who should deliver one or more of our proposed central functions?

Given the potential wide scope of use of AI as it develops there is the possibility that all current regulators plus those such as trading standards and other council run regulatory services may need to become involved. However, we think that this should be subject to wider conversations and dialogue to reduce any risk of duplication, confusion and enable potential economies of scale in delivering the



central functions. We would expect engagement around the central functions to take place with the devolved administrations and other public bodies involved in regulation of activity and may need to be carried out a sector at a time, such as social care and health are likely to have different needs and issues and application of the principles to manufacturers and education establishments.

12. Are there additional activities that would help businesses confidently innovate and use AI technologies?

Customers who are unsure of a product or fear their information is at risk will have a negative impact on a business. By encouraging transparency, staff training and customer information with alternative service delivery options for individuals who do not wish AI to form part of the service they receive.

Just as customers may be cautious of the technology and changes so will be businesses. Businesses are generally run by people all of whom manage change differently, it is important to remember that and to provide appropriate information and allow change to happen in the right way for that person or group of individuals or risk them closing the business due to stress or their wellbeing being impacted that impacts their ability to interact with AI or even previously systems that were working appropriately. Not all innovation may be appropriate for all people and all businesses at the same time.

12.1. If so, should these activities be delivered by government, regulators or a different organisation?

It may not always be appropriate so support and information should be available through a range of routes including peer to peer.

13. Are there additional activities that would help individuals and consumers confidently use AI technologies?

We feel the confusing landscape of legislation, risk associated with fraud of personal data and identity and the subsequent uncertainty by the public on how to respond if such criminal activities takes place could reduce confidence in AI technologies.

From a Community Safety perspective we would be keen to see an increased focus on reducing and limiting the capacity for AI to be used for illegal and criminal activities. To see education and information not only to safeguard and protect from unsafe user habits but to empower users so they are confident to use and try new technologies in a safe and secure manner and environment. An example is a third sector organisation that works in Wales who promote assistive technology for use by those with learning disabilities, they don't just ensure that they have the right equipment but they make sure it has all the safety features turned on and they carry out training sessions with the individuals on being safe online, the does and don'ts including if you do this you will be breaking the law and could go to prison. This approach could be used for AI as it develops.



13.1. If so, should these activities be delivered by government, regulators or a different organisation?

It may be appropriate to utilise a mixture but it may also be an opportunity to resource and support local trusted charities and community groups who may specialise as in the example provided to question 13.

14. How can we avoid overlapping, duplicative or contradictory guidance on AI issued by different regulators?

There may be an opportunity for the regulators to work together and agree common wording and application, but also where they will have variations. As the Framework suggests there may be a role for central governance and ownership, but it is essential all stakeholder voices are heard including devolved administrations and customer end users.

15. Do you agree with our overall approach to monitoring and evaluation?

Potentially, but it feels light on detail and on how any illegal activity will be identified and monitored. We also think it is important that there is a role for Devolved administrations and partners to be involved in discussions regarding identifying and mitigating risk.

16. What is the best way to measure the impact of our framework?

We are unsure as it may be appropriate to use different methods with the different stakeholders, and for there to be a way for the failures of the framework to also be measured. For example, if there is increased criminal activity due to failures in the application of the framework this should be measured and may involve criminal justice organisations who are otherwise not engaged in the framework environment.

17. Do you agree that our approach strikes the right balance between supporting AI innovation; addressing known, prioritised risks; and future-proofing the AI regulation framework?

Potentially, providing all stakeholders are genuinely engaged and the Framework is adaptable with known consequences for non-compliance. Whilst an iterative approach is being taken the Framework needs to be written in a way to reflect this in order to ensure it has resilience to protect it against unknown changes in the industry and ensure it is fit for purpose in a developing technological environment.

18. Do you agree that regulators are best placed to apply the principles and government is best placed to provide oversight and deliver central functions?

Yes in some instances, but it may also be appropriate for regulators to provide some practical oversight. How this will be resourced in both financial, technology and other resources needs to be identified including how this will be managed across devolved and reserved landscapes. The central functions need to be sustainability in the longer term.



L1. What challenges might arise when regulators apply the principles across different AI applications and systems? How could we address these challenges through our proposed AI regulatory framework?

Different regulators have different methods of sharing information. The Framework could unite a messaging process for all to adopt by reducing jargon and ensuring there is a shared language and terminology when AI is concerned.

Different regulators may have higher or lower requirements of their members due to the nature of the work that is being undertaken, as mentioned earlier health and social care is likely to be different to a manufacturer.

There may be an opportunity for accreditation schemes to be applied with and through the regulators to ensure consistency and provide assurance and transparency of standards of AI to customers and other stakeholders.

L2.1. Do you agree that the implementation of our principles through existing legal frameworks will fairly and effectively allocate legal responsibility for AI across the life cycle?

No, as mentioned previously there are concerns that there will be gaps. There are recent examples of consultations run by the Home Office to fill gaps which are now causing problems as they were not separately legislated for and which are being exploited by organised crime groups and individuals. Therefore, as stated previously whilst the Framework is voluntary and not in legislation there will be gaps which those who so wish to do so will exploit.

L2.2. How could it be improved, if at all?

See previous answers regarding the need for legislation to protect the public and ensure compliance wherever possible and so action can be taken where there are breaches for criminal or other exploitation or abuse activities which harm individuals, families or communities.

L3. If you work for a business that develops, uses, or sells AI, how do you currently manage AI risk including through the wider supply chain? How could government support effective AI-related risk management?

Not applicable.

F1. What specific challenges will foundation models such as large language models (LLMs) or open-source models pose for regulators trying to determine legal responsibility for AI outcomes?

There are many potential issues, such as students using AI for assignments. Quality of data and information which is being used and how data protection and confidentiality is managed where open-sourced models are used.

There could be benefits of foundation models for example in police investigations or case reviews where large amounts of information has to be processed to identify evidence. However, to enable AI developed evidence into the court systems may require legislation in itself. The responsibility we think would have to sit with the



organisation who purchase it. which could lead to negative impacts on victims and the finances of organisations should AI not work as it is intended.

F2. Do you agree that measuring compute provides a potential tool that could be considered as part of the governance of foundation models?

We hold no opinion on this matter at this time.

F3. Are there other approaches to governing foundation models that would be more effective?

We are not aware of any other approaches.

S1. To what extent would the sandbox models described in section 3.3.4 support innovation?

We are not sure it does support innovation and leaves a risk that if all that is done is encourage compliance then there will be those who will always comply and those who require an added incentive such as a legal requirement before they comply, especially if the action requires additional resourcing.

S2. What could government do to maximise the benefit of sandboxes to AI innovators?

We hold no opinion on this matter at this time.

S3. What could government do to facilitate participation in an AI regulatory sandbox?

We hold no opinion on this matter at this time.

S4. Which industry sectors or classes of product would most benefit from an AI sandbox?

We hold no opinion on this matter at this time as there are advantages and disadvantages for the different industry sectors and classes of products. The important element is the management of risk.

19. As a regulator, what support would you need in order to apply the principles in a proportionate and pro-innovation way?

Not applicable

20. Do you agree that a pooled team of AI experts would be the most effective way to address capability gaps and help regulators apply the principles?

A pooled team of AI experts may be beneficial in the short term, but as AI use expands across more sectors this is unlikely to be sustainable so how the capability and capacity gaps will be filled needs to be planned and implemented.



21. Which non-regulatory tools for trustworthy AI would most help organisations to embed the AI regulation principles into existing business processes?

We hold no opinion having not engaged with any of the AI tools to know their strengths and weaknesses

22. Do you have any other thoughts on our overall approach? Please include any missed opportunities, flaws, and gaps in our framework.

Whilst we appreciate this Framework tries to cover a large and complex landscape there is a lack of focus on penalties and consequences of non-compliance. We think this should be reconsidered due to the importance of this topic, potential risks to individuals, families and communities and the need for real consequences to be in place to ensure buy-in and compliance.

Impact Assessment Questions.

1: Do you agree that the rationale for intervention comprehensively covers and evidences current and future harms?

Yes- We agree that a central approach to pull legislation and accountability of AI development is necessary but as AI is a developing technology there may be unknown harms and unexpected consequences which may vary this response.

2: Do you agree that increased trust is a significant driver of demand for AI systems?

Potentially, there could be reluctance from consumers and business due to the threat of harm on or against both the person and/ or the organisation. However, there may be other constraints such as the cost at a time when many businesses are struggling due to the cost of living crisis and whilst the technology is still under development it could be a unnecessary expense.

3: Do you have any additional evidence to support the following estimates and assumptions across the framework?

- **The proposals will impact an estimated 431,671 businesses who adopt/consume AI products and services significantly less than the estimated 3,170 businesses who produce/supply AI products and services.**
- **Those who adopt/consumer AI products and services will face lower costs than those who produce and/or supply AI solutions products and services.**
- **Familiarisation costs (here referring to the cost of businesses upskilling employees in new regulation) will land in the range of £2.7m to £33.7m.**



- **Compliance costs (here reflecting the cost of businesses adjusting business elements to comply with new standards) will land in the range of £107m to £6.7bn.**

No

4: Do you agree with the estimates associated with the central functions?

- **The average FTE cost for a regulator is estimated to be £106k.**
- **A central AI regulatory coordination function would require 50 full time workers.**
- **A central AI regulator would require 300 full time workers.**
- **The average number of AI systems developed per small business is 2.**
- **The average number of AI systems developed per medium business is 5.**
- **The average number of AI systems developed per large business is 10.**
- **The proposals will impact an estimated 431.671 businesses who have adopted/consume AI products and services, and an estimated 3,170 businesses who produce/supply AI products and services.**

We are unable to provide specific comment, however there remains a question on the current cost analyses in a post Covid environment and during a period of inflation and a cost of living crisis, and the possible resource implications and how these will be funded.

5: Are you aware of any alternative metrics to measure the policy objectives?

No.

6: Do you believe that some AI systems would be prohibited in options 1 and 2, due to increased regulatory scrutiny?

Whilst we have no specific comments, we are concerned that there could be a risk that if increased regulatory scrutiny leads to some AI systems being prohibited that the scrutiny will be lowered. This could set a bad precedent and lead to increased risk of bad practice and harm coming through poor use of AI, so we feel the increased regulatory scrutiny is appropriate.

7: Do you agree with our assessment of each policy option against the objectives?

We agree with the information you have provided, however we think there should be a greater emphasis on the quality, compliance and safety of AI against the cost rather than the other way around.

8: Do you have any additional evidence that proves or disproves our analysis in the impact assessment?

No.