10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

Call for views on the Cyber Security of AI

techUK Response

09/08/2024

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

About techUK

techUK represents the companies and technologies that are defining today, the world that we will live in tomorrow. The tech industry is creating jobs and growth across the UK. Over 1000 companies are members of techUK. Collectively, they employ more than 700,000 people, about half of all tech sector jobs in the UK. These companies range from leading FTSE 100 companies to new and innovative start-ups. The majority of our members are small-and medium-sized businesses.

Executive Summary

techUK commends government's commitment to building cyber resilience across the UK and we welcome and share government's ambition to improve the resilience and security of Artificial Intelligence (AI). The cyber security threat landscape is constantly evolving, and cyber-attack techniques are becoming ever more sophisticated, therefore, the security and resilience of AI systems and models needs to evolve at an even greater pace to combat these threats. Given that digital supply chains are international, and many AI developers and systems operators work across multiple countries, techUK also applauds government's ambition for an international standard for the baseline cyber security of AI, because it is critical that the UK does not take a siloed route with any potential interventions. Indeed, international standardisation is key to promoting the UK's expertise and enabling trade.

It should be noted that techUK is responding to this Call for Views on behalf of its members (that is, those member organisations that constitute some of the stakeholders identified in the Code of Practice), and not in regard to our own day-to-day operations and/or use of AI. We have also encouraged members to submit individual responses on these proposals, particularly as we are aware that there may be some commercial sensitivities that cannot be shared in an open/collective response.

techUK and its members broadly support the principles contained in the voluntary Code of Practice, with some other provisions and points noted for inclusion, or expansion upon, such as data governance and transparency. It is worth noting at the outset that, techUK's members already take the security of any digital technology and the privacy of their users extremely seriously and undertake activities that exceed many of these proposals. Many of the sectors that our members work across – for example, health, defence and financial services – have existing stringent security requirements and practices, as well as commercial incentives to deliver high standards of security for their customers.

While some members agree with government's two-part-intervention approach, the majority of those who contributed to techUK's response do not believe that there is a need for a new national code of practice for the cyber security of AI. There is, however, agreement that the Code in its current form is not mature enough for an international standard. Members have noted that the Code over-steps the definition of a code of practice: that is, by definition, it is

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

a standard as it contains requirements rather than recommendations, however, this inherently creates inconsistencies with actual standards and undermines the work of the UK's national standards body in seeking international alignment (see BS0 section 9.4.1). As a standard, it has been noted that it fails to present the processes and policies required of any substantive standard (again see BS0 for exemplar policies).

Indeed, most members agree that government should focus on encouraging the application of existing standards to AI systems and then look to identify where specific further action is required for AI systems. In this response, we point to examples of existing industry initiatives, including regulations, standards, principles, and frameworks and members would encourage government to learn lessons from these examples when reviewing potential intervention options.

In terms of the comments made on specific principles in this response, these should be taken as indicative and not an in-depth line-by-line analysis. As mentioned above, it has been highlighted that a 'code of practice' by its definition should focus on principles or recommendations only rather than prescriptive requirements; and the issues created by this code/standard approach should be borne in mind.

techUK understands that government is proposing a 'voluntary' Code of Practice at this time and is waiting for the responses to the Call for Views to decide whether to proceed with this course of action. However, given the progression of other voluntary codes of practice to legislation (namely, the Product Security and Telecommunications Infrastructure Act) because of a lack of adherence, we would encourage further clarity regarding the intended roadmap for this process and that of determining whether regulatory action is needed – including the provision of more detail around how government will monitor and evaluate uptake of the Code and its effectiveness at encouraging the outcomes that we hope to see in the AI ecosystem. In addition, the timescale of voluntary codes to regulation (if that is the route progressed) can be significant (approximately 8 years in the case of the PSTI regulations). There is, therefore – whatever the outcome of this Call for Views – a clear and pressing need to instigate an education and awareness effort to promote the cyber security of AI.

It is also important to recognise that the UK is facing a significant shortage of the skills we need to develop AI frameworks and to assure systems' safety, security and privacy. If the UK wants to be a global leader in AI, government must make it a priority to develop the skills we need to make its regime a success. Any focus on such inconsistent national codes fragments the already scarce skills base and acts as a barrier to trade based on the very high-level of use of international cyber security standards.

The last few years have been particularly busy in the policy space, for both the cyber security sector and the wider technology industry in the UK. Even larger digital companies with their own dedicated public affairs and policy teams have faced capacity challenges with the sheer volume of policy proposals to provide feedback on. In addition to this, and while we appreciate matters were outwith officials' control, the announcement of the General Election,

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

the engagement restrictions inherent with the pre-election period, attention turning to the new government announcing the Cyber Security and Resilience Bill and the summer break, have also presented engagement challenges within the timeline given for responses to the Call for Views.

techUK recognises that similar challenges have faced government colleagues at a time when the communication of how various current and proposed codes of practice overlap was crucial. techUK, therefore, believes that more work is needed to clarify how the cyber security codes of practices align, as well as how the draft Cyber Security of AI Code of complements them and other existing industry standards and recommendations. Indeed, members have raised concerns about the disconnect between different codes and the principles the government expects industry to meet, which creates ambiguity. Additionally, there is concern about the burden that this could place on various sectors and potential hinderance of sector growth. These issues are particularly significant for SMEs, so consideration should be given to what support they should be given to help their compliance to the codes.

It has been noted that the overlap between secure AI and secure software is total: creating and deploying AI is creating and deploying software. While there are specific additional AI risks and controls, the Software Vendors Code of Practice should in effect be a subset of the Cyber Security of AI Code of Practice. However, the two codes as presented are almost completely different from the principles on down. There must be a clear and simple way to use both simultaneously and yet as presented that would be a significant effort. This underlines the risks of going beyond codes of practice and into requirements where these are described differently as an extensive effort would have to be undertaken to approach them simultaneously to the other standards like ISO/IEC 27001 and NIST that are absolutely required by the market.

We would, therefore, strongly recommend that more engagement takes place once government has published its response to this Call for Views. Indeed, techUK would be very keen to support and facilitate this engagement, including on some of the specific issues raised later in this response.

Our response to the questions in the Call for Views document

Demographics

Q 1. *Are you responding as an individual or on behalf of an organisation?*

- Organisation

Q3*. [if organisation] Which of the following statements describes your organisation? (Select all that apply)*

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

- Other – techUK is the trade association for the technology sector in the UK. Our members include AI developers and cyber security providers.

Q4. *[if organisation] What is the size of your organisation?*

- Medium (50–249 employees)

Q6. *[if organisation] Where is your organisation headquartered?*

- England

Q7. *In the Call for Views document, the Government has set out our rationale for why we advocate for a two-part intervention involving the development of a voluntary Code of Practice as part of our efforts to create a global standard focused on baseline cyber security requirements for AI models and systems. The Government intends to align the wording of the voluntary Code's content with the future standard developed in the European Telecommunications Standards Institute (ETSI).*

Do you agree with this proposed approach?

- *Yes*
- *No*
- *Don't know*

[If no], please provide evidence (if possible) and reasons for your answer.

techUK and its members welcome government's focus on the cyber security of Artificial Intelligence (AI) and government's recognition that, in order to harness the benefits and opportunities that AI can offer people, society, the economy and planet, AI models and systems must be developed, deployed and operated in a secure and responsible way. Furthermore, we agree that the UK should be a key voice in the development of international standards that would underpin future regulation for AI models.

While some members are supportive of government's approach to develop a voluntary Code of Practice with a view to submitting it to ETSI to support the development of a global standard on cyber security requirements for AI models and systems, others do not see any need for a specific new national cyber security of AI code of practice. Furthermore, there is concern that the draft Code is not mature enough for an international standard. It lacks key aspects such as practical implementation recommendations (for example, case studies) and members feel that much more work is required before the Code could form the basis of the standard.

Members have also highlighted that several of the principles and measures outlined in the draft Code reflect wider software security practices, standards and frameworks that already exist – for example, NIST's Secure Software Development Framework; or ISO 42001 (AI

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

Governance), ISO 8000 (Data Quality: Data Governance) and ISO 55013 (Management of Data Assets) which together provide overarching guidance on the data within AI issues and more generally fit within the normal overall approach of ISO/IEC 27001 and 27002. There is also NIST's draft Secure Development Practices for Generative AI and Dual-Use Foundation AI Models. The majority of members agreed that government should, therefore, focus on encouraging the application of existing standards to AI systems and then look to identify where specific further action is required for AI systems and feed that in as specific example and inclusions for cyber security controls catalogues by the normal standards maintenance processes. Indeed, AI specific issues and controls could and should be built into existing security standards – all the main standards allow for extensions to risk mitigation and controls.

While the 12 principles outlined in the draft Code are sensible, much of what is contained echoes general good AI governance principles (such as testing and evaluation of models through the lifecycle, ensuring against data bias) or, indeed, simply general good cyber governance. Members have noted that the specific value-add of a cyber security principle on some of these points is unclear, with one member giving the example of principle 4.4 which seeks to ensure users are aware of prohibited use cases: while this is, of course, important, clarity is required as to whether this is referring to posting and enforcing a usage policy and terms of service; and then how this principle translates to increased cyber security.

Furthermore, while industry commends and fully supports government's drive to encourage Secure-by-Design technology and to shift the burden away from end users, there is concern that the current modular approach of layering code of practice over code of practice could become incredibly confusing. Indeed, members have noted that it is not clear yet how this draft Code interacts with other codes such as the Cyber Governance Code of Practice that government consulted on earlier this year, as well as national and international initiatives like the NIST Cyber Security Framework and the work of ISO/IEC JTC1 on cyber security and AI; and in which specific ways this effort is complementary. It would be useful for more clarity to be provided on what it means to take a modular approach and how the draft Code would be taken into consideration by regulators.

Much more detail is also required on how the draft Cyber Security of AI Code of Practice will align with a future AI Bill and requirements for industry under the new Cyber Security and Resilience Bill; and we would urge these matters are deferred pending a consultation including draft legal texts are undertaken. It will be important to avoid conflicting or double requirements for stakeholders in scope of multiple regulations; and, by definition, regulations will be the overriding factor.

Members are strongly in favour of the international harmonisation of standards, and we are pleased to see government working towards international alignment of security requirements – indeed, the UK is in a good position to push for this. One member has, however, highlighted that choosing the ETSI route could prove problematic because internationalisation must be at least acceptable to the main trading blocs and they think that it is very unlikely that the US (and China) would accept one of the European Standards

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

organisations, given their single market and trade remit. They also noted that they see most market demand around using the NIST approach, embedding AI aspects into general cyber security approaches, and therefore would suggest that a bilateral with NIST would be better.

With regards to engaging with the development of international standards. techUK would recommend that government establish a mechanism by which it can consult with industry on an ongoing basis.

Finally, in terms of the specific comments on specific principles. these should be taken as indicative and example comments of the issues created by such a code/standard and not an in-depth line by line analysis.

It has been highlighted that a consultation should be done for any such documents containing requirements and acting as a standard; and this requires the proper WTO TBT Annex 3 compliant policies, governance and formats as used by standards bodies including BSI to be in place. In particular, clear IPR policies, publication, maintenance and transparency policies, comment resolution processes, line numbered documents, accountable editors and so on to be in place.

Q8. In the proposed Code of Practice, we refer to and define four stakeholders that are primarily responsible for implementing the Code. These are Developers, System Operators, Data Controllers (and End-users).

Do you agree with this approach?

- *Yes*
- *No*
- *Don't know*

Please outline the reasons for your answer.

Broadly, techUK and its members agree that the stakeholders identified by the draft Code represent the correct groups in the supply chain, however, it is unclear what responsibility End-users will have (if any) to implement the Code. As noted in our response to Question 7, industry agrees that it is correct to shift the burden of responsibility away from the end-user for pre-trained systems deployed as intended by the developer as much as possible, so if this is the intention on implementation of the Code, this must be communicated clearly.

One member suggested that a form of responsibility model would be more appropriate/beneficial here, such as appropriate selection of a tool based on its conformity to this and other relevant codes/standards/regulation.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

Others have proposed that the category of Data Controller be removed from the Code. They noted that the use of the term is confusing, as data controllers may also be system developers, system deployers or both; although they recognise that there is some fluidity in the value chain depending on how a company is leveraging a foundation model. Outside of personal data, the main issue is data governance and that may not be led by an individual but represent the enforcement of organisational policies.

Data Controller has specific meaning in GDPR and yet it is being applied here in a broader sense, so there is concern that this will lead to considerable confusion.

One member suggested that a differentiator be included with regards to the Data Controller to make it clear that an AI-specific process is being referenced and not a data protection aspect.

Another member highlighted that, going forward, many Developers will leverage existing AI models; therefore, it may be helpful to clarify and distinguish this category from those that integrate a model and do not carry out significant customisation/fine-tuning – for example, consider changing the language in this definition to *'individuals that are responsible for creating or further training an AI model and/or system'*.

Finally, one member recommended the introduction of a fifth stakeholder with the introduction of an 'internal auditor' role to help with the governance and compliance of this Code of Practice. They noted that this could be useful for large organisations where implementation of AI systems is divided in different teams/groups; and that this fifth stakeholder could also help in supporting the 'Developers' by acting as 'a second pair of eyes' to check their adherence to requirements from other codes of practices. (Note: techUK has not had a chance to test this recommendation with SME members and, therefore, further engagement would be required on it.)

Q9. Do the actions for Developers, System Operators and Data Controllers within the Code of Practice provide stakeholders with enough detail to support an increase in the cyber security of AI models and systems?

- *Yes*
- *No*
- *Don't know*

Please outline the reasons for your answer.

If implemented, the actions contained in the Code might improve the cyber security of AI systems, however, as noted in our answer to Question 7, the draft Code does not have the level of detail and practical advice that is required for a standard, while exceeding the necessary limitations for a useful code of practice.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

One member noted that, as the draft Code references specific frameworks which organisations could theoretically review to understand how to implement the Code, it may be useful to consider a mapping exercise, where specific actions are mapped more directly to existing frameworks.

Q.10 Do you support the inclusion of Principle 1: "Raise staff awareness of threats and risks within the Code of Practice?"

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

[If No], please provide the reasons for your answer.

Q11. Do you support the inclusion of Principle 2: "Design your system for security as well as functionality and performance" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

techUK members support the inclusion of Principle 2.

Broadly, members feel that the same security principles, standards, etc, that apply to non-AI systems should apply to AI systems, although the latter may have specific risk and controls handled by the standard process of risk identification and risk mitigation via a catalogue of controls on people, process and technology. Most members have highlighted that requiring AI-based systems to adhere to existing compliance regimes, security standards and testing requirements should be the focus – as opposed to applying unique security rules and regulations for AI.

Other points to note on this principle include that it contains requirements for Developers and Data Controllers, but these stakeholder groups are not referenced in the 'Applies to' section, so this should be updated to include all relevant stakeholder groups. It has also been highlighted that the NCSC's *Guidelines for Secure AI System Development* as

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

referenced in 2.1.1 are not applicable to Data Controllers, therefore, it is recommended that the reference be moved to 2.1.

It has also been suggested that the language in 2.4. is tightened to more accurately reflect the realities of AI systems and the intended aim of the Code to improve cyber security: *'Data controllers shall ensure that the intended usage of the system is appropriate with the sensitivity of the data it was **trained** on as well as the controls intended to ensure the safety **and security** of data'.* And one member noted that it is unclear whether 2.5 is talking about where an AI triggers an action in another system – and that, if this is the case, the language could be more explicit.

One member noted that this principle also fails to handle or recognise the trade-offs that can occur – for example, between security and accessibility, or security and functionality.

[If No], please provide the reasons for your answer.

Q12. Do you support the inclusion of Principle 3: "Model the threats to your system" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

In addition to threat modelling, Principle 3 mentioned risk management; and we would recommend updating the title of the principle to include this; and in order to reflect that ongoing risk management is required; furthermore, the wording of principle 3.4 should be changed to be *'Where third-party organisations have responsibility for risks identified within an organisation's infrastructure, System Operations should attain assurance that these parties are able to **manage** the risk, **or that alternate risk mitigation approaches are used (for example, redundancy, isolation, segmentation, etc**.".*

While members broadly agree with the inclusion of the Principle, one member has highlighted that some of the requirements, such as 3.1.2 *('As part of this process, Developers shall create a document that includes a list of adversarial motivations and possible attack routes in line with those motivations. The type of attacks could include indirect attacks where attackers poison data which might later be used by, or sent to, the model'*) seem overly burdensome and documentation heavy for companies.

It has also been noted that risk identification is the fundamental starting point for international standards and the extant guidance should be followed not undermined.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

One member noted that this Principle could also address non-malicious issues such as unintended use; and over agency, where the AI strays out of its 'box' and attempts, or is tricked into, outcomes which are wider than those intended by the developer.

Q13. Do you support the inclusion of Principle 4: "Ensure decisions on user interactions are informed by AI-specific risks" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

There are phrases used in Principle 4 where the meaning is not immediately clear or defined – for example, *'decisions on user interactions'* in this principle's tile; or the word *'overreliance'* in principle 4.5. Government should provide clarification for these and avoid subjective terms wherever possible.

Furthermore, 4.2 should be updated to read *'Developers **should** take steps to **verify** that the designed controls specified by the Data Controller, **or organisational data governance policies** have **either** been built into the system **or equivalent solutions implemented**.'*

And one member noted that 4.4 should be reinforced so that, although every effort is made to make it unlikely that the AI system can be used for prohibited cases, any attempt to circumvent these controls is ultimately the responsibility of the end user.

Q14. Do you support the inclusion of Principle 5: "Identify, track and protect your assets" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

It has been suggested that 5.1 should be separated out into two points: one of asset management; the other on risk management. It has also been highlighted that 5.3 is more relevant to the AI system when it is in operation – and, again, these are general good governance, not AI-specific, issues.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

On member urged caution regarding 5.5 and having the ability to restore systems to a known good state: when restoring to an older state, you need to make sure that state is not vulnerable to new attacks. Forcing a rollback could be a step needed by an attacker to further compromise a system.

Q15. Do you support the inclusion of Principle 6: "Secure your infrastructure" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

It has been highlighted that 6.2. (on creating segregated environments to enforce sensitivity and threat boundaries) may be overly prescriptive to implement; and that the language of 6.2.2 is not clear.

It has also been noted that 6.3 is a generic operational requirement, not a development, nor an AI-specific, one.

Creating an incident management plan (6.4) is an example of the draft Code amplifying principles that apply to all software. As per our answer to Question 7, most members recommend promoting the use of existing software standards to AI systems, and identifying specific actions required for AI systems.

One member has recommended expanding the requirement 6.4 to introduce a 'disaster recovery plan' to take into account scenarios in the future where AI models are performing current human tasks autonomously. If the AI model fails for some reason, or is unable to function, there should be effective mechanisms in place for reverting to some backup approach. This is especially important for autonomous systems that are performing critical roles.

Q16. Do you support the inclusion of Principle 7 "Secure your supply chain" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

For clarification purposes, it has been noted that it may be worth changing the phrase 'suppliers' referenced in 7.1 to 'AI suppliers'.

7.2 is another example of the draft Code amplifying principles that apply to all software. As per our answer to Question 7, most members recommend promoting the use of existing software standards to AI systems, and identifying specific actions required for AI systems.

It has been highlighted that more context is needed for *'if their security criteria are not met'* in 7.3; clarification is needed for what is meant by *'integrity of their security protocols'* outlined in 7.3.1; and it has been suggested that the language of 7.3.2 should be updated to *'Data controllers should continually monitor the source of publicly available data that **is being** used for creating a model'.*

One member suggested that this principle could be amplified in some form of 'transparency as a service' – such as a Software Bill of Materials (SBOM) tailored to AI models that could be included to demonstrate provenance, etc. to make due diligence of vendors in supply chain easier.

Another member noted that complete transparency from vendors cannot always be met, so a balanced approach is required in this regard.

Q17. Do you support the inclusion of Principle 8: "Document your data, models and prompts" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

[If No], please provide the reasons for your answer.

Members agree with the inclusion of this Principle.

It has been highlighted by one member that a few of the sub-principles seem overly burdensome and documentation heavy, including 8.1 ('Developers shall document and maintain a clear audit trail of their model design and post-deployment maintenance plans.') and 8.1.1 ('Developers should ensure that the document includes security-relevant information, such as the sources of training data, including fine-tuning data and human or other operational feedback, intended scope and limitations, guardrails, cryptographic hashes or signatures, retention time, suggested review frequency and potential failure modes.') Other member(s) view the requirements to have detailed records of AI system's data

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

sources, model architectures[1], and input prompts as critical given they ensure transparency, facilitate audits, and assist in identifying potential security or bias issues.

From an AI buyers' perspective, one member noted that what is meant by transparency should be clearer in this principle – that is, whether it is to be used to guide buyers on what they are purchasing.

Another member noted that principle 8.1.1 should clarify what exactly *'cryptographic hashes or signatures'* would be used for.

Q18. Do you support the inclusion of Principle 9: "Conduct appropriate testing and evaluation" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

Members are broadly supportive of Principle 9.

Other points to highlight here include that the 'benchmarking' referenced in 9.3 is not mentioned in Principle 2; and is has also been queried if 9.2.1 is only applicable where the Developer is also the System Operator – if not, it may be a significant burden on developers to post-deployment test every customer who uses their model; and, as with any code of practice, equivalent substitution should be allowed.

Finally, it has been noted that there is a large deficit in skills in this area at the moment.

Q19. Do you support the inclusion of Principle 10: "Communication and processes associated with end-users" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

[1] Please note that 'model architecture' here means the general layout of the model (e.g. neural network, structure of the layers, etc.) and not the weights of the model.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

Members agree with the inclusion of Principle 10, however, one member has highlighted that principles 10.2 ('Developers should ensure that the organisation proactively supports affected End-users and System Operators during and following a cyber security incident to contain and mitigate the impacts of an incident. The process for undertaking this should be documented and agreed in contracts with end-users.') and 10.3.2 ('Moreover, Developers shall inform end-users of additional AI model functionality, and allow an opt-out option.') seem disproportionately burdensome on companies. Most products are not bespoke services and contracts are not negotiable, especially for cloud-based solutions with multiple customer organisations on the same contracts.

Another member was in favour of the requirement to inform end-users of additional AI model functionality – as outlined in 10.3.2. – being placed on System Operators

Furthermore, the phrase 'additional AI model functionality' in 10.3.2 should be more clearly defined.

Q20. Do you support the inclusion of Principle 11: "Maintain regular security updates for AI models and systems" within the Code of Practice?

- *Yes*
- *No*
- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

While broadly in favour of Principle 11, the requirements outlined in 11.1 should be part of the Secure Design stage. A lack of specificity has also been highlighted – principle 11.3 talks about a secure-by-design approach to system updates, citing that 'major' updates should be treated as a new model and undergo necessary testing. However, it isn't clear what constitutes a major update, and who would be monitoring this in any case. Indeed, lack of definitions are an issue across the document and members have noted that standards always have a very clear statement of scope and definitions before any other normative content.

One member made the further additional point on this principle that, in practice, rolling out new models is complex and that many versions might be in production at the same time.

Q21. Do you support the inclusion of Principle 12: "Monitor your system's behaviour and inputs" within the Code of Practice?

- *Yes*
- *No*

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

- *Don't know*

[If Yes], please set out any changes you would suggest on the wording of any provisions in the principle.

Members broadly agree that Principle 12 should be included, indeed, reliability and repeatability are important to end users. However, one member would argue that most of the obligations under this principle which requires the logging of inputs and outputs, etc. should not apply to on-device models as there are heightened privacy implications (for example, the way devices are built, and the approach taken to a person's data remaining as much as possible on their phone and not shared with the developer, etc.) Therefore, some of the other obligations here could potentially be impossible, or overly privacy-intrusive if it implies monitoring on-device activity.

More widely with regards to monitoring and evaluation references in the draft Code, it is not particularly clear how adherence will be assessed, therefore, given it is a voluntary code, industry would ask that this is not too burdensome a process. It has been highlighted that clearer scope and definitions, and better alignment with other interventions could help to ensure this.

Q22. Are there any principles and/or provisions that are currently not in the proposed Code of practice that should be included?

- *Yes*
- *No*
- *Don't know*

[If Yes], please provide details of these principles and/or provisions, alongside your reasoning.

Risk management

Members have noted that the risk classification is not particularly clear in the draft Code. We would, therefore, suggest that either Principle 3 should be expanded to more explicitly reference risk management, or a standalone principle on risk management should be created and operationalised via the normal approach in ISO/IEC 27001.

Developers and Systems Operators should be able to define how the Code should be applied to its systems/products based on risk identification and risk mitigation in context (as per international standards), rather than an obligation to apply all the requirements in the Code to, for example, all of the ML-based features in a product.

The EU AI Act sets out clear risk classifications and, if the idea is for this Code to feed into the European standardisation process, this alignment would be helpful.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

AI Safety Principles

Further clarity is needed on how the Code will interact with broader AI safety principles.

One member noted that there is no mention of managing bias or transparency/accountability and this should at least be clearly signposted to the AI Safety Regulatory Framework. Others have suggested that a prescriptive scope restricting the code of practice to security issues, with clear references to the sources that cover these broader issues, would help here.

Data Governance

Members have also highlighted that Data Governance is a key provision missing from the draft Code. Data Governance direction/guidance should identify a suggested process for identifying the need for a tool and the data suitable to be processed by the model. This should be initiated by a decision subject to appropriate organisational governance that a solution utilising AI is the *right* choice and mapped to clear requirements and outcomes. One member suggested that an improved Responsible, Accountable, Consulted and Informed (RACI) matrix used to inform a shared responsibility model would be useful in improving the flow down of recommendations to all involved in the lifecycle regarding development, maintenance, use and mitigation of risks associated with AI technologies.

Finally, members have noted that Large Language Models, where data controls have not been implemented, have exposed the broader issue of poor data governance and this aspect must be taken into consideration going forwards if we are to avoid building the next legacy technology.

Compliance

Members have noted that there is not yet any detail on how an organisation is deemed to be fulfilling the draft Code's principles, although it has also been noted that this only comes about due to the referencing of requirements that only belong in standards, as opposed to a recommendation. Indeed, government must set out more detail on how it intends to monitor and evaluate uptake of the Code and its effectiveness at encouraging the desired outcomes in the AI ecosystem.

On a more general point, it has also been suggested the principles should include wording that gives users of the Code of Practice a sense of what each principle is trying to achieve.

Finally, it has been noted that a code of practice should include a review mechanism to ensure that it remains fit for purpose through the years as technology/practices/threats develop at pace.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

Q23. [If you are responding on behalf of an organisation] Where applicable, would there be any financial implications, as well as other impacts, for your organisation to implement the baseline requirements?

techUK is not best placed to answer this question.

Q24. Do you agree with DSIT's analysis of alternative actions the Government could take to address the cyber security of AI, which is set out in Annex E within the Call for Views document?

- *Yes*
- *No*
- *Don't know*

[If No], please provide the reasons for your answer.

On the point of the draft Code being voluntary, some members have highlighted that this is unlikely to lead to consistent outcomes. Indeed, they question whether a voluntary code would have enough teeth to influence those who need its guidance the most (that is – those who are not already well-informed stakeholders). Furthermore, there are various examples where non-regulatory schemes have been under-resourced; they have failed to achieve the desired outcomes and/or have simply been ignored by affected organisations.

Good regulation – for example, the Telecommunications (Security) Act 2021 – can drive positive change in behaviours and is supported by regulators being given the expertise, capability and resource that they need to carry out their role meaningfully.

Therefore, some members would hope that the high-level security principles of this draft Code of Practice will be reflected in any future regulatory framework for AI models.

Q25. Are there any other policy interventions not included in the list in Annex E of the Call for Views document that the Government should take forward to address the cyber security risks to AI?

- *Yes*
- *No*
- *Don't know*

[If Yes], please provide details of these principles and/or provisions, alongside your reasoning.

It is important to recognise that we are facing a significant shortage of the skills we need to develop AI frameworks and to assure systems' safety, security and privacy more generally. If

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

the UK wants to be a global leader in AI, government must make it a priority to develop the skills we need to make its regime a success. Skills development should be the priority and exceptional care and due diligence should be undertaken with regards to actions that could fragment the skills base.

Furthermore, should the same path be taken as with other codes of practice – when certain principles of a voluntary code eventually became mandated through regulation (for example, such as the Product Security & Telecommunications Infrastructure Act) – the timescale can be significant (approximately 8 years in that case). There is, therefore – and whatever the outcome of this Call for Views – a clear and pressing need to instigate an education and awareness effort to promote the key issues that are contained in the draft Code.

One member has put forward that government should also consider the following interventions:

- Government should avoid a 'tick box' approach to compliance, instead promoting continuous assurance throughout a model's lifecycle. Where the risk profile necessitates, independent third-party product validation should be recommended to ensure compliance with security, safety and privacy standards.
- A novel idea that is being undertaken in the US, and could be replicated here in the UK, is the AI Cyber Challenge – a two-year competition offering c$20 million in prizes to the competitor who is able to develop the best AI-driven systems that develop secure software code. While we expect generative AI to support the development of secure software code in future, such technology is currently unreliable and still requires expert oversight and other traditional controls to ensure accuracy. Government-led challenges like this one may help to accelerate R&D in this area, so that we can all benefit from AI-driven secure software development.
- There will inevitably be competition between nations to be at the forefront of AI innovation, with regulatory approaches differing across borders. That said, we are hopeful that frameworks like the OECD's AI principles can help to minimise regulatory divergence for most countries engaged in developing AI. It is critical that the UK uses its global leadership position to promote international harmonisation, while ensuring its domestic frameworks are as aligned as possible to regulatory regimes seen in allied countries. In doing so, we recommend that the government:
    - utilises existing successful partnerships, including the 'Five Eyes' alliance and with the European Union; and ISO IEN UN-backed standards bodies;
    - invests time in developing practical outcomes with other governments, that go deeper than high-level principles – the forthcoming Global AI Summit presents an excellent opportunity to make progress in this area; and,
    - ensures that civil society and industry – who will play a central role in delivering governments' objectives – are involved in discussions from the outset.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

Q26. Are there any other initiatives or forums, such as in the standards or multilateral landscape, that that the Government should be engaging with as part of its programme of work on the cyber security of AI?

- *Yes*
- *No*
- *Don't know*

The UK Government is already engaging through the G7 in the Hiroshima Process and is expected to take part in the drafting and implementing of the OECD Pilot Project on the Hiroshima Process International Code of Conduct for Organisations Developing Advanced AI Systems. Following last year's UK-US MoU on AI, in particular establishing a cooperation framework between the UK and US AI Safety Institutes, the UK should continue global engagement in this space and, for example, enter into similar agreements with other counterparts across Europe, starting with the EU AI Office.

It has also been suggested that government should also engage with the following initiatives:

- ISO/IEC DIS 12792 – Information technology – Artificial intelligence – Transparency taxonomy of AI systems
- NIST AI Risk Management Framework
- AI Standards Search – AI Standards Hub (which includes numerous developments around IoT, ICS, SCADA).
- Via the BSI in ISO/IEC JTC1 SCs 27 and 42 (cyber and AI respectively).

Q27. Are there any additional cyber security risks to AI, such as those linked to Frontier AI, that you would like to raise separate from those in the Call for Views publication document and DSIT's commissioned risk assessment. Risk is defined here as "The potential for harm or adverse consequences arising from cyber security threats and vulnerabilities associated with AI systems".

- *Yes*
- *No*
- *Don't know*

[If yes], please provide evidence (if possible) and reasons for your answer.

Broadly, the report captures the distinct cyber threats not necessarily relevant to systems which do not use AI. That said, one member would draw particular attention to the following types of attacks:

- Training and runtime attacks: Some AI algorithms are, by design, susceptible to influence and change based on their inputs and lifetime. This presents the

techUK

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

opportunity for significant security risks, particularly from attacks which seek to manipulate training data inputs with the objective of causing the model to make incorrect assessments or produce malicious outputs.
- Data breaches: This involves extracting confidential or sensitive data from the model which has either been used in the training process or sent as inputs to the production model.
- Denial of Service (DoS): This entails degrading a model's performance to the point of unusability. As society and the economy becomes more reliant on AI systems, the impact of DoS attacks will only grow.

The democratisation of technology and its widespread availability risks inadvertent consequences too. There are a growing number of openly accessible AI frameworks that are available to software developers, who may use the frameworks without necessarily understanding the underlying mathematics, data science and associated operations or compliance issues such as IPR. This could, in turn, lead to poor outputs.

In addition, the safety implications of cyber security vulnerabilities should be more explicitly reflected, particularly as AI systems are increasingly relied on in the physical world (for example, smart cities and autonomous vehicles). Indeed, where enabled without suitable protection such vulnerabilities could be exploited to effect change in the real-world, with potentially disastrous consequences.

Q28. Thank you for taking the time to complete the survey. We really appreciate your time. Is there any other feedback that you wish to share?

- *Yes*
- *No*

[If yes], Please set out your additional feedback.

One member noted that references to legal requirements (such as Data Protection Impact Assessments) or terms with legal meaning (such as 'sensitive data') should be omitted from the Code of Practice unless they are citied accurately, with reference to the defining legislation (including its geographic scope) or further guidance specific to AI systems.

It has also been noted that the overlap between secure AI and secure software is total: creating and deploying AI is creating and deploying software. While there are specific additional AI risks and controls, the Software Vendors Code of Practice should in effect be a subset of the Cyber Security of AI Code of Practice. However, the two codes as presented are almost completely different from the principles on down. There must be a clear and simple way to use both simultaneously and yet as presented that would be a significant effort. This underlines the risks of going beyond codes of practice and into requirements where these are described differently as an extensive effort would have to be undertaken to

10 St Bride Street
London EC4A 4AD
T +44 (0) 7331 2000

techUK.org | @techUK

approach them simultaneously to the other standards like ISO/IEC 27001 and NIST that are absolutely required by the market.

More broadly on the Call for Views itself, while techUK welcomes government's commitment to the cyber security of AI models and systems, the timing of the consultation has been problematic. While we completely appreciate that events have been out with the control of officials – such as the announcement of the General Election; the pre-election period when engagement with industry was unable to take place; and uncertainty around whether the Call for Views would be granted approval to continue from the new Minister; as well as the clash with summer holidays – it has been more difficult for techUK to engage as many members as usual on this Call for Views within the given timeline.

techUK recognises that similar challenges have faced government colleagues at a time when the communication of how various current and proposed codes of practice overlap was crucial. techUK, therefore, believes that more work is needed to clarify how the cyber security codes of practices align, as well as how the draft Cyber Security of AI Code of complements them and other existing industry standards and recommendations. Indeed, members have raised concerns about the disconnect between different codes and the principles the government expects industry to meet, which creates ambiguity. Additionally, there is concern about the burden that this could place on various sectors and potential hinderance of sector growth. These issues are particularly significant for SMEs, so consideration should be given to what support they should be given to help their compliance to the codes.

We would, therefore, strongly recommend that more engagement takes place once government has published its response to this Call for Views. And techUK stands ready to support and facilitate this engagement.

Contact: Jill Broom, Head of Programme – Cyber Resilience, techUK jill.broom@techuk.org