ITI Promoting Innovation Worldwide

July 26, 2023

ITI Response to the Australian Government's Discussion Paper "Safe and Responsible AI in Australia"

The Information Technology Industry Council (ITI) is the premier voice, advocate, and thought leader for the global information and communication technology industry. Founded in 1916, ITI is an international trade association that promotes public policies and industry standards that advance competition and innovation worldwide. Our members include the world's leading innovation companies, with headquarters worldwide and value chains distributed around the globe. These companies are leading Internet services and e-commerce companies, wireless and fixed network equipment manufacturers and suppliers, computer hardware and software companies, and consumer technology and electronics companies. Artificial Intelligence (AI) is a priority technology area for many of our members, who develop and use AI systems to improve technology, facilitate business, and solve problems big and small.

ITI and its members share the firm belief that building trust in the era of digital transformation is essential and agree that there are important questions that need to be addressed regarding the responsible development and use of AI technology. As this technology evolves, we take seriously our responsibility as enablers of a world with AI, including seeking solutions to address potential negative externalities and helping to train the workforce of the future. Our members are aware of and continue to take steps to understand, identify and treat the potential for negative outcomes while leveraging opportunities that may be associated with the use of AI systems.

ITI is actively engaged in AI policy around the world. In 2021, we issued a set of *Global AI Policy Recommendations*, aimed at helping governments facilitate an environment that supports AI while simultaneously recognizing that there are challenges that need to be addressed as the uptake of AI grows around the world.[1] In 2022, we also published our *Global Policy Principles for Enabling Transparency of AI Systems*[2] where we underscore that transparency is a critical part of developing accountable and trustworthy AI systems while avoiding unintended outcomes or other harmful impacts. ITI is currently working on paper containing policy considerations around foundation models and will share that in follow-up to this submission once finalized. We have also actively worked to inform the efforts of the National Institute of Standards and Technology (NIST) to foster trust in AI technology in the U.S. [3] In the EU, we have contributed to the debate following the EU AI White Paper in 2020, the publication of the EU AI Act in 2021, and throughout

[1] Our complete *Global AI Policy Recommendations* are available here: https://www.itic.org/documents/artificial-intelligence/ITI_GlobalAIPrinciples_032321_v3.pdf

[2] Our complete *Policy Principles for Enabling AI Transparency of AI Systems* are available here: https://www.itic.org/documents/artificialintelligence/ITIsPolicyPrinciplesforEnablingTransparencyofAISystems2022.pdf

[3] See ITI response to RFI on AI RMF Concept Paper here: ITI Comments on AI RMF Concept Paper FINAL.pdf

Global Headquarters
700 K Street NW, Suite 600
Washington, D.C. 20001, USA
+1 202-737-8888

Europe Office
Rue de la Loi 227
Brussels - 1040, Belgium
+32 (0)2-321-10-90

info@itic.org
www.itic.org
@iti_techtweets

trilogue negotiations. Most recently, we submitted comments for the UK AI White Paper in June 2023.

In 2023, ITI launched the AI Futures Initiative, which is comprised of a task force of AI policy and technical experts on AI from ITI's member companies, in order to produce actionable deliverables to inform the development of meaningful policy solutions that will support the responsible development, use and deployment of AI.

We are grateful for the opportunity to provide the feedback below to the Department of Industry, Science and Resources on the "Safe and Responsible AI in Australia" discussion paper (Discussion Paper). ITI stands ready to serve as a resource and partner for the Australian Government as it considers how best to address the complex set of questions that are emerging around AI governance and regulation.

General Feedback

ITI welcomes the Australian Government's efforts to boost its practices to support responsible AI. We believe that evaluating existing laws and regulatory authorities is an important first step to understanding how they apply to AI and to better inform whether additional policy is necessary. In seeking to leverage existing authorities, however, coordination across government will be key to ensure that there are no contradictory, duplicative, or otherwise onerous compliance burdens that could unintentionally undermine innovation.

If additional regulation is needed to address risks related to AI, we are supportive of the risk-based approach that Australia lays out in Box 4 on page 32 of the Discussion Paper. In seeking to construct a risk-based approach, it is important that requirements are proportionate to the level of risk. We see in Box 4 that Australia has proposed different requirements based on whether an application is 'low risk,' 'medium risk' or 'high-risk.' At a macro level, this is an appropriate way to approach regulation. We recognize this is a Discussion Paper and therefore Australia is likely still thinking through the nuances of different requirements related to each risk-level, but in seeking to move forward Australia should ensure that the goals and requirements are clear and non-prescriptive. We encourage Australia to review a recent response we provided to the U.S. National Telecommunications and Information Administration on AI Accountability, attached to our submission, which talks through some of our perspectives around concepts like transparency and assessments.

As accurately stated in the Discussion Paper, the safe and responsible deployment and adoption of AI presents significant opportunities for Australia to improve economic and social outcomes. AI driven medical diagnostics can alert doctors to early warning signs to help them treat patients. AI systems are capable of monitoring large volumes of financial transactions to identify fraud. Small and medium-sized enterprises (SMEs) can gather new insights and improve their businesses by using AI and data analytics.

It is also worth noting the many benefits AI has for society, especially in the cybersecurity space. AI plays an important role in the cybersecurity of IT in general, and of AI systems themselves, by increasing the speed and effectiveness of detecting and preventing cyber threats that can undermine or disrupt their safe use. An organization could face millions of indicators of compromise (IOC) per day and security teams demand contextual awareness and visibility from across their entire environments. Cybersecurity systems that leverage AI are able to better provide real-time analysis and prevention compared to cybersecurity systems that do not incorporate the latest technologies. Benefits provided by AI, like in cybersecurity, would contribute toward Australia's goals for fostering innovation.

Such technological advances can bring innumerable benefits to Australia. At the same time, promoting the responsible growth of AI is key while guarding against possible harmful impacts or adversarial use of these AI-enabled technologies. Recognizing the AI ecosystem is global, Australia has an opportunity to take an international leadership role in promoting advances in AI innovation.

Specific Responses to Questions Posed in the Discussion Paper

Definitions

1. *Do you agree with the definitions in this discussion paper? If not, what definitions do you prefer and why?*

We are supportive of the Australian government's alignment on the definitions of "Artificial Intelligence (AI)", "machine learning", and "algorithm" with the terminology specified in *ISO/IEC 22989 Information technology – Artificial Intelligence – Artificial intelligence concepts and terminology 5.17 AI Stakeholders roles*. ISO/IEC 22989 defines AI provider, user, customer, partner, and subject roles.

Potential gaps in approaches

2. What potential risks from AI are not covered by Australia's existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?

There already exist a handful of outcomes-based frameworks that help organizations develop AI risk management practices that account for many of the risks associated with AI systems, including assessing the security of those models. Before moving toward regulatory approaches, we encourage the Australian Government to determine whether encouraging use of these frameworks is an acceptable way forward, as organizations can leverage these frameworks to appropriately manage risk while allowing them to receive benefits of these systems. We are encouraged by the Australian Government's stated intentions not to "consolidate or replicate the development of existing general or sector-specific regulations and governance initiatives

across the Australian Government", per the Discussion Paper. We urge the Government to fully assess the landscape of domestic legislation to ensure any potential regulatory and governance measures do not conflict with current Australian law and instead focus on filling necessary gaps.

3. Are there any further non-regulatory initiatives the Australian Government could implement to support responsible AI practices in Australia? Please describe these and their benefits or impacts.

We appreciate the Discussion Paper's reference to the U.S. National Institute of Standards and Technology's (NIST) AI Risk Management Framework, which provides organizations with a holistic way to identify and manage AI risks throughout the lifecycle by recommending a set of principles and tools flexible enough for organizations to implement robust AI-risk management governance tailored to AI use. In addition to referencing international standards, Australia should consider if and how it might leverage this Framework as a part of AI policy development, as it may help organizations more easily embed holistic AI risk management into their business processes.

4. Do you have suggestions on coordination of AI governance across government? Please outline the goals that any coordination mechanisms could achieve and how they could influence the development and uptake of AI in Australia.

ITI is supportive of an approach where regulators and government agencies in Australia adopt a set of cross-sectoral principles in a way that makes the most sense for their sector(s). If Australia pursues this approach, it may be helpful for the Government to consider introducing a set of functions that will help to coordinate implementation across different government bodies.

Target Areas

6. Should different approaches apply to public and private sector use of AI technologies? If so, how should the approaches differ?

Risks related to AI are the same whether used by the public or private sector, so the rules should be the same. The public sector should lead by example.

7. How can the Australian Government further support responsible AI practices in its own agencies?

Incentivize or promote the security of AI systems as early in the development process as possible.

9. Given the importance of transparency across the AI lifecycle, please share your thoughts on: a. where and when transparency will be most critical and valuable to mitigate potential AI risks and to improve public trust and confidence in AI? b. mandating transparency requirements across the private and public sectors, including how these requirements could be implemented.

a.) As stated in the "ITI Policy Principles for Enabling Transparency of AI Systems," at the highest level, transparency is about being clear about how an AI system is built, operated, and functions. When executed well, AI transparency can help analyze outputs and hold appropriate AI stakeholders accountable. Since the term "transparency" can have multiple meanings, to the extent that a transparency requirement is focused on informing a user they are interacting with an AI system, particularly those with high-risk uses, we suggest use of the term "disclosure" instead. Furthermore, the Government should clarify the different disclosure requirements based on whether an organization is a deployer or developer.

It is challenging to suggest precisely where and when to introduce transparency measures without specific reference to the underlying objective. At a high level, we are supportive of disclosure to users when they are interacting with an AI system and several of ITI's member companies are already taking steps to provide more specific information to users about an AI system, piloting tools like model cards that include information about the intended use of the system, limitations of the model, and its level of accuracy, among other things. While it may not be necessary to provide model cards in every instance, they are a helpful way to inform users about key characteristics of an AI system so users are empowered to make decisions about if and how to use a system.

Please refer to the ITI Policy Principles document for a complete list of recommendations; we have highlighted a few of the principles below:

- **Consider what the ultimate objective of transparency requirements are.** Is it to ensure the user knows they are interacting with an AI system? Is it to provide a post-hoc explanation to users about a decision that was made and provide them with an appropriate redress mechanism should the decision negatively impact them? Is it to help researchers and developers test and validate the AI model or system? Is it to enable the AI system deployer to investigate an incident? Is it to enable and authorize regulators or third parties such as auditors to evaluate a system's safety features? Or is it something else? Understanding the answers to these questions is critical to determining the ultimate direction of the policy proposal and what approach is most appropriate to help achieve that objective.

- **Consider the intended audience of any transparency requirements and at what point of the AI system lifecycle they would apply.** Policymakers should also consider the target audience at which transparency requirements are directed, including their level of expertise. They should also consider when such requirements would apply (e.g., pre-deployment or post deployment). For example, transparency could be useful to several different audiences (e.g., regulators, consumers, developers, etc.), which will in turn influence requirements. Understanding the intended audience will also inform the type of information presented, the manner in which it is presented, and the amount of information presented. Indeed, if the purpose of a transparency requirement is to allow

a user to understand how or why a decision was made and allow for redress, that will result in a very different set of information being provided than if such information is being provided to allow a regulator to evaluate a system for safety.

- **Take a risk-based approach to transparency when considering requirements**. In devising any requirements around transparency, policymakers should consider the diversity of AI applications, what their ultimate use case will be, and if transparency would undermine the security of networks/information systems or divulge intellectual property. The level of need and desire for transparency requirements from various users may vary significantly based on the AI application, the types or sensitivity of data processed, or intended use.

- **Include clear definitions of what is meant by transparency in the context of a regulation or policy proposal.** Specifically, policymakers should include a definition of the term, to which part of the AI value chain it applies, to which stakeholders it applies, and in what context it applies. Policymakers should also clarify if the reference is to the transparency of AI systems, as opposed to algorithmic transparency, which could apply more broadly.

- **Consider there are different ways to approach transparency and improve trust, and that explainability is only one component**. There are other ways to help create trust and deploy AI responsibly, including using technical, procedural, or educational tools to ensure AI systems are fair and robust. Another method to approach transparency is to encourage certain stakeholders in the AI lifecycle to examine raw input data to understand the limitations of the dataset and account for and help manage potential bias, while respecting privacy.

- **Ensure transparency requirements do not require companies to divulge sensitive IP or source code or otherwise reveal sensitive individual data, including information that could jeopardize the security of network and information systems.**

- **Leverage voluntary international standards in order to maintain interoperability of various AI transparency requirements to the extent possible**.

11. What initiatives or government action can increase public trust in AI deployment to encourage more people to use AI?

Governments globally can play a vital role in facilitating dialogues about AI between companies that use and develop consumer-facing AI or models that impact individuals and their communities, with the intent of better-aligning them. As these stakeholders become more closely aligned, trust will grow, and AI adoption will scale.

To grow public understanding of and trust in AI, governments should:

- **Partner with or fund university programs whereby data science and other students in aligned disciplines conduct real world projects with communities in key areas of social need.** This initiative can also serve as a training function for the communities involved who learn which problems AI can and cannot solve and how to make the technology benefit them.

- **Encourage a transparent and secure ethical design approach**. ITI and its member companies support the need to consider in a fair manner the impact that AI applications may have on discrimination and agree it is something that should be considered throughout the entire lifecycle of an AI system.

- **Develop and support frameworks and guidelines that protect privacy and promote the appropriate/ethical use of data that may be used in data sets underpinning AI**. To protect personal information and support fundamental human rights, data in data sets used by AI systems may be required to be anonymized, aggregated, or otherwise de-identified such that the datasets exclude any personal information and cannot be re-identified. Doing so ensures the beneficial use of the data in training intelligent systems while protecting individual privacy and security consistent with protecting fundamental human rights. Additionally, the protection of personal privacy would be strengthened by securing AI systems themselves, which is a critical condition for the safe use of AI systems.

Implications and Infrastructure

12. How would banning high-risk activities (like social scoring or facial recognition technology in certain circumstances) impact Australia's tech sector and our trade and exports with other countries?

If Australia considers banning certain activities, any bans should be narrowly defined and targeted, and only finalized following significant industry engagement that includes discussion of other ways to mitigate the government's concerns. Overly broad bans on technology activities undermine the ability of Australian and Australia-based companies to participate in the global marketplace, which in turn disrupts the virtuous cycle of private-sector R&D investments made possible by revenues from sales of products to a diverse customer base in overseas markets.

13. What changes (if any) to Australian conformity infrastructure might be required to support assurance processes to mitigate against potential AI risks?

It depends on the process that Australia ultimately takes with regard to implementation of a risk-based approach. We note that presently for high-risk use cases Australia contemplates a requirement around external auditing for high-risk use cases. While we appreciate that this stringent obligation is limited to high-risk use cases, we urge Australia to exercise caution in mandating external or third-party audits because currently there are a combination of challenges that make this practically difficult.

For example, conducting external audits of AI systems requires significantly deep expertise in AI technology as well as the domain in which the AI system is intended to operate. While there may be some organizations that have the expertise and capacity to perform certain types of AI assessments, if external audits were to be mandated at scale, workforce expertise and capacity could become an issue. Additionally, the rapid pace of evolution of AI technology may make it challenging for external audits to keep pace. There is also no consistent or standardized framework for conducting audits of AI systems, which may mean that audits conducted in one sector differ from audits conducted in another sector. Without a clear framework, audits or assessments may fail to provide meaningful insights. We are also concerned that this could lead to a proliferation of external businesses that offer audit services but lack the necessary expertise to provide high-quality assessments. Another major concern for the private sector is the possibility of having to disclose sensitive intellectual property directly to regulatory authorities, potentially compromising trade secrets and undermining competitiveness.

More generally, if Australia was to require third-party conformity assessment and the scope of high-risk AI systems are not sufficiently narrow, there are other practical considerations that are important to consider. Conformity assessment for AI is a nascent field for which there is neither a commonly understood practice nor the established conformity assessment infrastructure. Although ISO/IEC is working to develop ISO/IEC 42001, an overarching AI management system to which a potential conformity assessment system might be bound, it is not yet complete. For this reason, there are significant practical and logistical concerns regarding precisely how third-party bodies, once identified, accredited, and designated, would carry out the task of assessing the conformity of certain high-risk AI systems. Given that tools and processes for assessing compliance in this field are still emerging, it is unclear how existing facilities would have to be transformed to perform these tasks in a timely way and with the needed skill and expertise, and what type of guidance would be needed to ensure appropriate capacity of the testing bodies. Logistical problems, including the lack of sufficient bodies that are accredited to perform such conformity assessments, may also lead to backlog for testing bodies, which could significantly slow down the adoption of certain AI technologies.

Risk-based Approaches

14. Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach?

Yes, ITI supports a holistic, risk-based approach for addressing potential AI risks. The Australian Government should seek to leverage and align any AI initiatives with existing published standards or those currently under development in international standards bodies. We recommend aligning with ISO Guide 73:2009, which provides definitions of generic terms related to risk management and recognizes that certain AI use cases can be positive and can result in opportunities for innovation.

There are several existing frameworks to address AI risks, including the U.S. NIST AI Risk Management Framework, the OECD Accountability Framework.

ISO/IEC JTC 1/SC 42 is also working to develop standards to achieve trustworthy AI goals, several of which we outline below. A more fulsome list is also available online[4].

Published	**Under Development**
ISO/IEC 22989: 2022 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology	ISO/IEC TS 17847 Information technology — Artificial intelligence — Verification and validation analysis of AI systems
ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management	ISO/IEC 42005 Information technology — Artificial intelligence — AI system impact assessment
ISO/IEC TR 24368:2022 Information technology — Artificial Intelligence (AI) — Overview of ethical and societal concerns	ISO/IEC 5259 Series on Data quality for analytics and ML
ISO/IEC 38507:2022 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations	ISO/IEC 5338 Information technology — Artificial intelligence — AI system life cycle processes
ISO/IEC TR 24027:2021 Information technology — Artificial Intelligence (AI) — Bias in AI systems and AI aided decision making	ISO/IEC PWI 17866: Artificial Intelligence (AI) — Best practice guidance for mitigating ethical and societal concerns
ISO/IEC TR 24028:2020 Information technology — Artificial Intelligence — Overview of trustworthiness in artificial intelligence	ISO/IEC TS 12791 Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks

[4] See published standards and standards under development here: https://jtc1info.org/sd-2-history/jtc1-subcommittees/sc-42/

ISO/IEC TR 24030:2021 Information technology — Artificial Intelligence — Use cases	ISO/IEC TS 6254 Information technology — Artificial intelligence — Objectives and approaches for explainability of ML models and AI systems

15. What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome?

Risk-based approaches to AI allow stakeholders to focus on responding effectively to specific harms while allowing for advancements in technology and innovation that benefit all of society. This targeted approach also translates to proportionate regulations for differing risks, ensuring compliance is not overly burdensome for industry stakeholders. This includes efforts to align common parameters and consider the scope of AI, taking a context-specific approach to governing AI, and evaluating existing laws and regulations to determine whether there are gaps requiring incremental new rules for AI.

16. Is a risk-based approach better suited to some sectors, AI applications or organisations than others based on organisation size, AI maturity and resources?

When it comes to addressing risks associated with AI, context is key. We recommend policymakers, in close consultation with industry and other stakeholders, consider how to characterize "high-risk" applications of AI, including by identifying the appropriate roles for AI developers and users in making risk determinations. In our view, an AI decision is high risk when it results in a negative outcome that could have a significant impact on people – especially as it pertains to security, health, safety, privacy, freedom, discrimination, or human rights.

In thinking about high-risk applications, focusing on "sectors" may lead to overly broad categorizations; it is important to use sufficiently targeted and well-outlined classification to ensure this criterion does not become irrelevant. We encourage developing a categorization that takes into account sector, use case, complexity of the AI system, probability of worst-case occurrence, irreversibility and scope of harm in worst-case scenarios, e.g., individual v. large groups of people, and other criteria.

19. How might a risk-based approach apply to general purpose AI systems, such as large language models (LLMs) or multimodal foundation models (MFMs)?

A risk-based approach should apply to general purpose AI systems, and obligations levied upon an AI system should be proportionate to the level of risk that it poses to health, security, safety, or fundamental human rights. Similarly, obligations for foundation models should be proportionate to the level of risk the models pose in the use case or deployment context. This is because the end use is often where risk to individuals materializes. It is also important to consider the capabilities of the model in assessing risk.

Foundation models present a unique set of considerations because these models can form the basis for many downstream AI applications. Therefore, risks could be promulgated throughout the ecosystem if they are not appropriately identified and mitigated during development. ITI is working on a paper containing policy considerations around foundation models and will share that in follow-up to this submission once finalized.

***