TRUSTED AUTONOMOUS SYSTEMS

Submission to Commonwealth Government Consultation on Supporting Responsible AI

Organisation: Trusted Autonomous Systems Defence Cooperative Research Centre
Address: PO Box 59, Toowong (Yuggera-Turrbal land) QLD 4066

Date: 26 July 2023

About Trusted Autonomous Systems (TAS)

Trusted Autonomous Systems (TAS) is Australia's first Defence cooperative research centre and is uniquely equipped to deliver world-leading autonomous and robotic technologies to enable trusted and effective cooperation between humans and machines. Our aim is to improve the competitiveness, productivity, and sustainability of Australian industry.

Supporting Australia's defence capability

TAS, together with its participants and the Department of Defence, is focused on developing the capacity of Australia's defence industry to acquire, deploy and sustain the most advanced autonomous and robotic technology through:

- delivering world-leading autonomous and robotic defence technologies
- building innovative IP through targeted research and technology programs
- assisting Australian industry to develop new, improved and competitive autonomy technologies
- evaluating the utility of autonomous systems through capability demonstrations.

Supporting assurance and accreditation of autonomous systems

In addition to specific industry-led projects, TAS is undertaking two 'common-good' activities that have broader, non-defence applications. Through these activities TAS will:

- foster ethical and legal research including value-sensitive design
- develop policy pathways for projects and participants
- support development of Queensland air, land and marine ranges for trusted trials, test and evaluation
- establish independent, world-class certification pathways for global industry.

How we work

Trusted Autonomous Systems fosters collaboration between Australia's defence industry and research organisations and aims to increase small and medium enterprise (SME) participation in its collaborative research to improve capabilities of Australia's defence industry. Established under the Next Generation Technologies Fund, with $50 million invested over seven years, and a $15 million co-investment from the Queensland State government, TAS aims to deliver trustworthy smart-machine technologies for new defence capabilities based on advanced human-machine teaming.

For additional information on TAS, click here.

Submission in Response to the Commonwealth Government Consultation on Supporting Responsible AI

To whom it may concern,

Trusted Autonomous Systems (TAS) welcomes the opportunity to make this submission in response to the consultation on supporting responsible AI.

Through our Commonwealth and State government funded common good activities, specifically *Activity 1: Ethics and Law of Trusted Autonomous Systems* and *Activity 2: Assurance of Autonomy*, TAS is working to enhance Australian capacity for designing, manufacturing and operating safe and trusted autonomous systems across the land, air and maritime domains. These two activities are also applied where appropriate to practical autonomy projects overseen by TAS to ensure translation to real-world capability.

We note that the discussion paper supporting the consultation process provided an excellent overview of existing and proposed regulatory efforts addressing general AI, distinguished from sector-specific AI. On that basis, our submission will focus largely on regulation of sector-specific AI applications in use around Australia, specifically autonomous systems in the transport, general commercial and defence industry sectors in the land, air and maritime domains.

At a sector-specific level, our submission highlights the following key points:

1) AI-enabled autonomous technology **is in regular use in the air, maritime and land domains, for government, commercial, research and defence purposes,** and is increasingly capable and accessible.
2) **Better regulatory approaches and infrastructure** are needed to support safe and trusted design, manufacture and use.
3) Implementing a set of **overarching regulatory principles** will support harmonised, best practice domestic and international regulatory development.
4) **Whole-of government guidance** on the implementation of a risk management approach to supporting responsible AI will be needed.
5) Australia should consider adopting **targeted initiatives and programmes** designed to foster **co-design** arrangements between government and non-government organisations.

The submission also provides an overview of some of the **regulatory tools and initiatives** that TAS has led to bridge the regulatory gap. These serve as useful examples of what can be done at a sector-specific level to make incremental but important advancements.

We would also like to take this opportunity to suggest that there is need for further discussion in relation to the development of a **framework for developing responsible AI**. To this end, TAS recommends that the Australian Government consider its position on this matter, both in relation to private sector activity and government activity, through four key lenses (posed as questions, which we do not propose to answer here). Each lens offers a perspective that deliberately does not presuppose a specific model for achieving responsible AI, but instead invites

consideration of basic assumptions about what AI is used for, how it is developed and who gets to develop it. This framework is compatible with risk-based methods for managing the development and deployment of AI, and points to aspects that any regulatory regime will need to consider; but also speaks more fulsomely to issues that responsible AI policy must address. These questions are:-

- **ownership:** who owns the AI and its infrastructure?
- **process:** how is AI developed and deployed?
- **problems:** what problems are legitimate for tacking with AI?
- **properties**: what properties must a system meet or exhibit to transition from a ring-fenced laboratory to the real world?

If the review team would like to engage directly with TAS in relation to our submission, the primary contact in respect to this submission is Rachel Horne. Please find her contact details below.

Rachel Horne | General Manager – Law, Regulation and Assurance
Trusted Autonomous Systems
T: +61 (0) 411 054 462
PO BOX 59, Toowong, Qld 4066, Australia
rachel.horne@tasdcrc.com.au | tasdcrc.com.au

Submission

Key Point 1: *AI-enabled autonomous technology is in regular use in the air, maritime and land domains, for government, commercial, research and defence purposes, and is increasingly capable and accessible.*

It is important to identify at the outset that AI and autonomous technology is already in regular use because it highlights why a better calibrated regulatory approach for supporting responsible AI is needed.[1] Growing uptake of these systems is driven by the demand for safety, efficiency and sustainability and the impetus to remove humans from 'dull, dirty and dangerous' professions. This is likely to continue at an enhanced pace going forward.

Examples of 'live' autonomous technology in use include:

- uncrewed aircraft systems (UAS), commonly referred to as 'drones'
- uncrewed surface vessels (USVs) and uncrewed underwater vessels (UUVs) commonly referred to as 'autonomous vessels'
- uncrewed ground vehicles (UGVs), and
- automated vehicles or cars.

Design, testing and deployment of these systems is being supported by a range of physical and digital test and evaluation infrastructure; industry-led standards, policies and practices; and in some cases – such as the National Transport Commission Automated Vehicle Program - government-led regulatory preparedness exercises. [2]

Key Point 2: *Better regulatory approaches and infrastructure will support safe and trusted design, manufacture, and use.*

Currently, AI and AI-enabled systems in the air, land and maritime domains are being regulated by traditional 'platform-centric' regulators administering frameworks that are not generally suited to the features and risks of new technologies. As a result, regulators are being forced to address autonomous technology in a 'square peg, round hole' context that relies on flexibility mechanisms – such as exemptions – to enable design, testing and evaluation, and operation. These mechanisms are often ad hoc and opaque, and do not provide a clear, consistent or common pathway for compliance. While organisations such as TAS have worked with government, defence and industry stakeholders to develop pathways, industry standards, codes of practice and guidance material (see ***Table 1*** below), these can generally only be used to satisfy regulatory requirements 'by exception' - that is, using flexibility mechanisms such as exemptions and equivalent means of compliance.

---

[1]See *Autonomous Systems Body of Knowledge* (2022) Trusted Autonomous Systems. Available from https://www.rasgateway.com.au/maritime/resource-hub#maritime-body-of-knowledge.

[2] For example, see *Automated Vehicle Program* (2020) National Transport Commission. Available from https://www.ntc.gov.au/sites/default/files/assets/files/Automated%20vehicle%20approach.pdf.

Generally, regulation sets an acceptable minimum standard of safety, or a safety benchmark. Compliance with this safety benchmark is generally demonstrated through obtaining and maintaining a certificate, permit or other approval, which generates trust from various stakeholders including operators, other waterway users and the general public. There is certainty and consistency of both expectation and outcome. Without a clear safety benchmark or means of achieving it – that is, compliance with designated standards or requirements that is verified by third party accreditation processes - the pathway to achieving certification or approvals is obfuscated or blocked. This lack of suitable regulatory pathways can delay or prevent innovation from occurring, causes firms to move or invest offshore to seek out a more favourable regulatory environment, and - worst-case - result in safety being jeopardised.

Maritime domain: Autonomous vessel regulation

Autonomous vessels – both surface and subsurface - are being used in a range of settings.[3] While the external shell of some autonomous vessels may look similar to a crewed vessel, there are a range of key design, structural, computational and operational differences. Autonomous vessels:

- are able to operate at varying degrees of independence using components such as sensors, effectors, software and hardware
- are typically remotely controlled, supervised or monitored
- often use alternative communications arrangements and energy sources
- require different skills from persons involved in their operation and maintenance
- face different risks, such as cyber threats, or risks that manifest differently, such as a lost vessel. Accordingly, they also have different risk treatments, and
- have different testing, evaluation and trial needs.

However, in Australia, the regulatory framework that applies to autonomous vessels is the same framework that applies to traditional crewed vessels. Vessels operating for commercial, government or research purposes within Australia's Exclusive Economic Zone are generally required to comply with the following, unless otherwise exempt:

- have a certificate of survey
- be operated under a certificate of operation
- be crewed by crew holding designed certificates of competency
- be marked by a unique vessel identifier, and
- a safety management system.

In order to obtain certification, vessels and operators must demonstrate compliance with certain standards relating variously to the operation, crewing, construction, design, and equipping. Compliance is generally checked at specific intervals by an accredited marine surveyor or recognised organisation accredited relevant 'categories' of surveying. Requirements for safety management systems assume persons, including the master and crew, will be on board the vessel. Certificates are issued for up to five years, and assume that there will be little change in how or where the vessel is used. Elements of the framework – from legislation through to technical standards – are impracticable or impossible for autonomous vessels to meet, or are not relevant to autonomous vessel operations. Equally, the law is largely silent on key

---

[3] See *Autonomous Systems Body of Knowledge* (2022) Trusted Autonomous Systems. Available from https://www.rasgateway.com.au/maritime/resource-hub#maritime-body-of-knowledge.

features of autonomous vessels and their operation, as well as on effective assurance mechanisms.

The consequence is that autonomous vessels require exemptions from some or all of these requirements. This process can be inefficient, opaque, and uncertain, and leads to increased financial and opportunity costs for both vessel owners and government. It does not recognise or allow for the importance of appropriate mechanisms for conducting testing, nor does it reflect or support the strategic, regulatory and operational agility that both government, defence and industry alike need. For example, building sovereign capability in the AI-enabled autonomous space is critical for the Australian Defence Force; however, regulatory reform also supports the Australian defence industry, which needs a clear framework for undertaking crucial testing and evaluation, much of which occurs under the auspices of civilian regulation.

TAS has sought to work in a collaborative manner with government and industry alike to overcome these challenges, including developing regulatory initiatives and tools. However, strategic direction – flanked by legislative change - is needed to more fully recognise the benefits of autonomous technology, and ensure that can be tested and deployed safely, now and in the future.

Key Point 3: *Implementing overarching regulatory principles will support harmonised domestic and international best practice regulatory development.*

Total harmonisation of regulatory approach across domains or industries is unlikely to be achievable. However, a principles-based approach is an effective option that supports a risk-based future direction and can steward consistency in regulatory reform across domains, industries and sectors. This approach is also consistent with the international trend toward a principles-based approach toward AI principles and AI ethics principles, intended to guide development and use without being prescriptive or enforceable. [4] The principles proposed – advanced by Horne et al, - provide a voluntary framework to guide the development of regulation and its implementation. They are founded in the regulatory theories of risk-based regulation and anticipatory regulation, and are domain and technology agnostic. [5]

The proposed principles are:

- **Trust-centered:** Regulatory frameworks should include consideration of features of emerging technology that contribute to their trustworthiness, including explainability and robustness.

---

[4]   For example, see NATO Principles for Responsible Use of AI. Available from https://www.nato.int/cps/en/natohq/official_texts_187617.htm.

[5] Horne, Rachel et al Ten Regulatory Principles to scaffold the design, manufacture and use of trusted autonomous systems, illustrated in a maritime context. TAS '23: Proceedings of the First International Symposium on Trustworthy Autonomous Systems (July 2023) Article No 6, pages 1-12. This approach has been informed by the existing literature base and the collective experience of the authors. These principles are domain and technology agnostic and encompass learnings from diverse fields including law and regulation, systems engineering and computer science. See also NESTA (2020) *Anticipatory regulation*. Available from https://www.nesta.org.uk/feature/innovation-methods/anticipatory-regulation/

- **Collaborative:** Development and implementation of regulatory frameworks should occur collaboratively using the principles of co-design where possible.
- **Risk-based:** Regulatory frameworks should be risk-based, with requirements imposed consistent with the risk posed by the artefact and its operation. Regulatory frameworks should enable the use of objective and repeatable risk assessment methodologies.
- **Evidence-led:** Regulatory frameworks must enable decision makers to consider a variety of evidence types when assessing applications, and regulatory agencies must upskill their employees in engaging with new types of technical evidence.
- **Facilitate experimentation**: Regulatory frameworks should facilitate experimentation, also referred to as test and trialling or test and evaluation, of emerging technology with minimal requirements, building to more significant documentation and evidentiary requirements for ongoing operations.
- **Systems-focused:** Regulation of autonomous systems must take a systems approach, enabling risk methodologies and treatments that reflect systems engineering principles, including consideration of the various elements that interact to impact the risk of an autonomous system and its operation.
- **Usable:** Regulatory frameworks must be usable by a broad range of stakeholders, including both users and regulators of autonomous systems.
- **Consistent:** Regulatory frameworks, and decisions made under them, should aim for consistency of approach, including harmonisation across domains and technologies where possible, including using consistent terminology and risk assessment methods.
- **Adaptable:** Regulatory frameworks must be capable of adapting to emerging technology and new methods of operating.
- **Reviewable:** Regulatory frameworks should incorporate review procedures, for example sunsetting clauses, to ensure they remain fit for purpose.

The intent is for these principles, with further refinement, to be used as a reference point for government and non-government organisations developing regulatory frameworks for autonomous systems, and a reference point for best practice for academia and industry using, scrutinising or contributing to those frameworks. [6]

Key Point 4: *A whole-of-government process for implementation of a risk-based approach to supporting responsible AI is required to maximise the benefits it delivers.*

It is important to note that regulators generally administer and implement, rather than make, legislation and regulation. To this end, they require support from government through their responsible policy department in obtaining strategic direction and advancing legislative and regulatory amendments. Modern regulators need to look to incremental options within their purview, including increasing and upskilling personnel, embracing RegTech solutions, adopting new assurance measures and tools, and providing flexibility in a transparent manner.

---

[6] These principles are also relevant to other technology such as internet of things and connected mobility and can be used to guide regulatory development in that space as well.

Regulation within the air, land and maritime domains, particularly in the transport context, has been historically siloed and tend toward prescriptive approaches. While there has been positive movement towards more risk-based regulation, a principles-based risk management approach to supporting responsible AI will be of limited utility if it is implemented inconsistently. Guidance, training and adequate resourcing should be provided both to regulators and their responsible policy departments. This includes ensuring that the risk-based approach to supporting responsible AI is included in processes for:

- establishing new legislation or regulation, amendments, or reviews, including when establishing or amending the objects of legislation
- making or amending new technical standards, or amendments to technical standards, and
- new strategies, policies and advisory material for industry on how regulators will undertake their functions and duties in relation to AI and AI-enabled system.

We also suggest that guidance include illustrative examples of risk-based regulatory approaches and activities that can be taken in the short term, including:

- **broad-based flexibility for 'trusted partners'** such as the Australian Defence Force and other Commonwealth agencies and authorities with appropriate risk management arrangements in place
- the establishment of **quantifiable approaches to risk management**
- leveraging off **existing industry standards and codes** of practice rather than having an entrenched starting point that governments alone are responsible these documents, and
- facilitating regulatory experimentation – such as using regulatory sandboxes.

Finally, while this discussion paper focuses on Commonwealth approaches, state and territory regulatory frameworks should also support responsible AI approaches, and processes should be put in place through intergovernmental arrangements to seek consistency.

Key Point 5: *The proposed risk-based approach should be supported by targeted initiatives or programmes designed to support regulatory co-design between government and non-government organisations.*

Industry co-design and collaboration are both foundational principles of the anticipatory regulatory philosophy that guides approaches to emerging technology.[7] AI and AI-enabled systems are being designed and used around the world, with ownership of systems, and the resultant models and data, similarly dispersed. Australia's approach should rightly reflect national priorities, including sovereign capability requirements, and the importance of government in achieving intended regulatory outcomes. However, the industry-led nature of autonomous systems design and use to date – as seen in its role and expertise in the development of alternative regulatory approaches - highlights the importance of regulatory co-design between government and non-government organisations. Examples of initiatives or

---

[7]   See   NESTA   (2020)   *Anticipatory   regulation*.   Available   from https://www.nesta.org.uk/feature/innovation-methods/anticipatory-regulation/.

programmes that could be established in Australia – either at a general or sector-specific level - to maximise the impact of a risk-based approach to supporting responsible AI includes:

- A 'pathfinder-style' funding programme such as the **UK Regulatory Pioneers' Fund**. Under the Regulatory Pioneers' Fund, funding is available to industry to help regulators work with industry deal with regulatory hurdles created by new technology. For example, an existing project at Plymouth Smart Sound is underway with the aim of identifying regulatory issues associated with the use of autonomous vessels and creating a new proposed regulatory framework for them to propose to UK Maritime and Coastguard Agency (MCA). The proposed regulatory framework will then be assessed by MCA, and if successful could be expanded around the nation. This approach enables the industry who are actively involved in designing, manufacturing and using autonomous vessels to play a crucial role in designing and testing new regulatory approaches, to then feed into the MCA. [8]

- A new international **Regulation of Safe and Trustworthy AI Community of Practice** to bring together government and non-government organisations in Australia and other countries.[9] Members could be drawn from a broad base, including domestic and international regulators, peak bodies, standards organisations, industry and training organisations. A new Community of Practice would play a vital role in refining the proposed principles and contributing to their implementation in emerging domestic and international frameworks.

---

[8] See *Projects selected for the Regulators' Pioneer Fund* (2022) Available from https://www.gov.uk/government/publications/projects-selected-for-the-regulators-pioneer-fund/projects-selected-for-the-regulators-pioneer-fund-2022.

[9] See Horne, Rachel et al.

Overview of TAS regulatory tools and initiatives

TAS has sought to respond constructively to the absence of a fit for purpose regulatory framework for autonomous systems in Australia. *Table 1* sets out **some of the key tools and initiatives that TAS** has led to overcome regulatory limitations. These are designed to

- enable applicants to identify the requirements that apply to their vessel, system and operation, and how they can demonstrate compliance with those requirements, and
- assist regulators to assess applications related to the design and use of autonomous systems, as well as to enhance their understanding of autonomous and related technologies and identify potential regulatory, workforce, IT and administrative improvements needed.

Table 1: TAS regulatory tools and initiatives

Project	Outcome	Impact
Australian Code of Practice	Introduces tailored standards for autonomous vessels suitable for the Australian Operating environment.	▫ Create clarity, consistency and efficiency for operators and regulators, reducing time and resources spent on regulatory processes
Guidance materials for Australian Code of Practice	Assists users in how to apply the Code of Practice and navigate regulatory processes.	▫ Support innovation and technology development and uptake in Australia
COLREGS Operator Guidance Framework	Makes it easier to understand COLREGS, which rules apply for specific vessels and operations, the capabilities required to comply, and how to demonstrate compliance.	▫ Support development of Australian sovereign capability
RAS-Gateway	Provides an online portal for autonomous vessel stakeholders to access regulatory information, resources and support. RAS-Gateway hosts the Maritime Regulatory Pathway Finder, which enables users to identify the specific regulatory requirements and processes for a particular vessel, and navigate options for gaining approval to operate.	▫ Support development of improved regulatory frameworks, approaches and processes for autonomous technology

Conclusion

AI-enabled autonomous systems are increasing accessible and capable. They present many possible opportunities to increase safety and efficiency while decreasing risks to humans and minimising environmental impacts. However, access to these benefits is being hampered by the absence of clear regulatory pathways, making design, manufacture and use challenging and resource-intensive for regulators and the regulated community. Both government and non-government organisations can play an important role in investing in regulatory reform, including upskilling regulatory agencies and enhancing regulatory frameworks to ensure that Australian industry can innovate, in a safe and trusted way, and capture as much of the market share in this rapidly developing industry as possible. While this consultation is not defence focussed, building sovereign capability in the AI-enabled autonomous space is also critical for the Australian Defence Force, which seeks to partner with Australian defence industry to develop this capability.

The Australian Government needs to provide a whole of government policy directive regarding the priority or otherwise of supporting the uptake of AI-enabled emerging technology, together with endorsing and disseminating a set of high level regulatory principles to enable greater harmonisation in regulatory development and implementation. Investing in the exploration of other regulatory initiatives such as regulatory sandboxes, trust marks, research translation platforms, digital tools, industry codes of practice, co-regulation, and information platforms will also support Australian industry and the potential users of this emerging technology.

TAS, as a defence and government funded entity with significant expertise in the emerging technology space, is already working to support industry to use autonomous technology. We have the capacity to conduct and deliver research and practical projects to support government efforts, and would welcome the opportunity to contribute further to the Department's work in this respect.