Regsoft

To the Department of Industry,

4 August 2023

Dear Sir or Madam,

Safe and responsible AI consultation

Regsoft Pty Ltd (Regsoft) welcomes the Australian Government ambition and desire to regulate the use of advanced technology in a manner that is secure, sustainable and beneficial to all Australians - citizens, businesses and the wider public.

Developing a coherent, national and industry-focussed view on the adoption and sustainable use of advanced technologies, including AI. The AI consultation comes at a time when the Australian Government is simultaneously grappling with

- linked data, and consistent data use across the public- and private sector, and
- regulation and adoption of advanced technologies, including AI across many sectors
- multiple regulation touch points across all industries as data- and digital technologies are adopted.

Regsoft understands the Government view for ethical AI, and for the ongoing development of AI that is safe, secure and reliable. A coherent common narrative that combines the approaches for Government agencies in a consistent manner is needed.

We believe that any approach to AI by the Australian government should adopt meaningful, measurable and enforceable standards that are adopted by Australian Government agencies acting as exemplars and supporting the wider use by Australian businesses.

In 2017, former US Secretary of State Madeleine K Albright noted:

> *I hope that together, we can devise innovative ideas and not fall victim to a disconnect best summed up by words I actually plagiarized from Silicon Valley, which are: "citizens are speaking to their governments using 21st century technologies, governments are listening on 20th century technology and providing 19th century solutions."*

AI is one of many 21st century issues that require 21st century solutions.

Meaningful, measurable, enforceable standards:

Standards, including test harnesses, should be developed to clarify generic terms such as "best practice" Without such, any guidance is likely to obscure poor practice rather than support improvement. The diagram from pp. 25 of the discussion paper, highlights the Australian

tendency to adopt voluntary principles, over the international approach for enforceable standards.

The AI Assurance Framework, NSW, provides an excellent example of enforceable standards: The NSW code mandates five principles that any agency must apply for AI-related projects. The document provides clear tests for ongoing validation of such projects.

A similar mandate for all ICT and data projects could be adopted federally, for example, as part of the 2-pass ICT business case funding model, and as part of a Digital Transformation Agency (DTA) review.

Government agencies as the exemplars:

Australian Government agencies have the ability to drive improved models of AI development by the purchasing power of the whole of the Australian Government. This can be developed through a combination of:

1. Establishing mandated requirements for developing AI within government agencies - such as the NSW Government example above.
2. Establishing mandated requirements for any AI that is procured by the Government - or licensed to drive minimum standards for sale of systems.

As part of this, the Australian Government can continue to develop, and publish models for AI use, adoption and training, to support best practice use of AI by example.

Demonstrability and verifiability:

AI responsibility should rest on trust, through empirically variable claims. The empirical tests to support such claims should be available for independent verification. Where AI's are in use by public organisations, the onus of proof should be higher with the verification made public.

Oversight by people:

AI (and technology) should be used to augment human intelligence, but should never be used without governance and checking. It is insufficient to have a human reviewer agree that "best practice" approaches have been adopted. The oversight should rest on demonstrable tests to ensure their objectivity.

Duty of care for those impacted by AI:

Where organisations (public or private) adopt AI services, they have a duty of care to address the impact of these systems. Accountability of the delivery of the AI technology should rest with people, and these people should be responsible for ensuring that the application of the AI is appropriate. Part of this responsibility is having access to the appropriate tools and capability to form reliable assessments of performance.

This implies that the users, buyers, suppliers and developers of an AI system all have ongoing responsibility and all have accountability for the outcomes of the AI. The responsibility for AI dis-benefits extends well beyond the development team that built or trained it.

Potential Gaps:

The scale of AI impact is exponential. Imperceptible errors in data and small errors in logic, can drive major, negative impacts, as AI systems are used to massively scale out "menial" tasks that would normally be limited by the capacity of humans to deliver them. Whilst malicious AI's grab headlines - poorly designed, and poorly tested AI's typically do far greater harm. A human auditor, or a code-reviewer, cannot hope to understand and invalidate the output of the erroneous AI, in a meaningful way, before such a system generates significant harm.

The risk model should be built upon a measure of the scale of human impact, and the ability to publicly demonstrate the outcomes of the AI. Where a public demonstration test cases, showing the efficacy of the algorithm or system cannot be provided, the system should be considered inherently high risk if it is to be used in a public setting.

Designing a machine-readable risk approach:

Software design has established approaches for risk mitigation, which are not dependent upon human intervention or human audit. This includes:

- test-driven design
- development processes - including Agile
- continuous integration, testing and improvement
- deployment validation

Regulation of AI should adopt a risk approach that is compatible with software development. This requires regulators to move away from risk matrices, and coloured risk charts, and consider test designs and test harnesses.

The work of government and regulation can focus on building, testing and supporting the public adoption of reliable test harnesses and test frameworks. The development of such test frameworks should remain open access to ensure that all AI systems can be meaningfully assessed against known, publicly available standards.

A common regulatory framework, with a clear regulatory owner:

In 2019, the Australian Human Rights Commission (AHRC) published the Ethical AI Whitepaper, building on the 2018 Human Rights and Technology Issues Paper. Half a decade later, the key observations remain valid. The AHRC recommended developing a separate regulator, and framework, to ensure industry and government adopted trustworthy technology and used it ethically.

Regsoft would encourage the development of an independent regulator, whose outputs and intellectual property are openly, and publicly available to all, and whose funding supports full open-source for test services and validation. Such an organisation would be expected to strongly engage with AI researchers and commercial entities around the world, however, it should remain independent of all such entities to ensure it is not conflicted by the commercialisation drives of such groups.

About Regsoft

Regsoft seeks to simplify and redefine the nature of regulation, using 21st century thinking to solve 21st century problems.

Regsoft Pty Ltd (Regsoft) was established in 2021. The directors of Regsoft have a combined 20 years of experience in the area of deontic, defeasible logic research, and have previously worked at the Commonwealth Scientific Industrial Research Organisation (CSIRO) [and National ICT Australia (NICTA)] undertaking research in this area. We have developed automated systems for licensing transfer, payroll and AI-related systems in health, security and digital law.

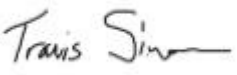
Our rules are built with useability, integration and flexibility in mind. Our vision is threefold:

1. help consumers easily achieve and prove compliance
2. help regulators automate the processing of verifying compliance to free them to focus on non-compliance
3. provide standardised schemas for information exchange between businesses and regulators

We are a small Canberra-based company, and strongly support the development of a responsible AI framework. We would like to support the ongoing work in building, implementing and delivering this ambitious goal.

Leif Hanlen, PhD
CEO Regsoft

email: leif.hanlen@regsoft.com.au



Travis Simon
CTO Regsoft

email: travis.simon@regsoft.com.au