SAFE AND RESPONSIBLE AI IN AUSTRALIA

SUBMISSION TO AUSTRALIAN GOVERNMENT

AUGUST 2023

Natasha Blycha
Schellie-Jayne Price
James Myint
Ty Haberland

STIRLING &ROSE

WWW.STIRLINGANDROSE.COM

STIRLING & ROSE

Director
Technology Strategy Branch
Department of Industry, Science and Resources
Industry House
Canberra ACT 2601

Dear Department of Industry, Science and Resources,

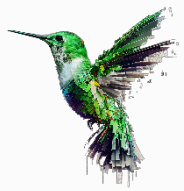
Stirling & Rose (**S&R**) welcomes the Department's discussion paper and supports the Department's commitment to Safe and Responsible AI.

Stirling & Rose is a boutique emerging technology law firm and global legal policy institute with deep experience in the subject matter of digital assets, artificial intelligence, smart legal contracts and autonomous organisations (including their sub-set decentralised autonomous organisations). We advise organisations seeking to use AI responsibly, major investors (hedge funds, VC funds, and lenders) and entrepreneurial organisations operating at the vanguard of emerging technology on major real-world application and enforcement of all the above. We have done so at an international level as the executive team of a major international law firm since the inception of the space. Our partners have backgrounds in international regulatory law and disputes, insolvency and personal property law, asset, and structured financing as well as ICT contracts and machine learning. We collaborate with the Gradient Institute, Australia's pre-eminent responsible AI institute to deliver both legal acumen and technical expertise to our clients in comprehensive AI Strategy Services. We have been invited by the Salzburg Global Seminar to facilitate the international discussions on AI and Corporate Governance at the "AI – Possibility or Peril" Corporate Governance Forum to be held in Austria in October, 2023, building on our facilitation at the Salzburg Global Finance Forum on "Global Turbulence and Financial Resilience: Implications for Financial Services and Society" held in June 2023.

We are currently leading Australia's Responsible Contracting Project in respect of smart legal contracts together with Nooriam Pty Ltd and The Commonwealth Scientific and Industrial Research Organisation (**CSIRO**) the Australian Government Agency responsible for scientific research.

We again extend a hearty congratulations and warm thanks to the Department for its work in this consultation.

Yours sincerely,



STIRLING & ROSE

Schellie-Jayne Price, Natasha Blycha and Ty Haberland.

Introduction

Stirling & Rose commends the Department in pursuing a proactive approach of public consultation as part of the Department's AI governance policy setting strategy. S&R is supportive of the comprehensive consultation exercise undertaken by the Department and set out in the Discussion Paper and welcomes the opportunity to offer our insights.

Approach to Submission Questions

Where our content touches on specific questions articulated in the Discussion Paper, we reference that question in this paper's footnotes.

1 Recommendations

Stirling & Rose recommends establishing a dedicated taskforce as soon as possible, with an operating period of 18 months, which, at a minimum:

1.1 acts as an explicitly advisory body to provide the more detailed, technical advice and guidance on equivalent international approaches—and gaps and issues in our existing domestic frameworks;

1.2 reviews and observes global AI hard law (for example in the EU and China) and guideline style soft AI law positions currently in force (for example, in Singapore and the United States) over a period of 18 months. The point of this is for Australia to learn from those other jurisdictions and gauge how the different approaches work/don't work, as well as how they need to interact with global standards, prior to implementing our own reforms;

1.3 contextualises any need for explicit AI Law within other Australian relevant regulatory and legal reforms (for example, reforms of digital assets, banking and payments and privacy);

1.4 reviews and engages with the Digital Free movement and Digital Free National Parks see here as a proportionate response to the future inability to disconnect from AI.

1.5 provides a forum for collaboration/information sharing and consultation with various legal, technical, and other stakeholders to ensure that AI frameworks can identify and respond to risks as they emerge; and

1.6 coordinates and works with relevant state and federal agency stakeholders to:

(a) review all Federal and State laws to understand how the adoption of AI and autonomous algorithms will impact the application, interpretation and enforcement of those laws in the short, medium and long term;

(b) understand what existing governance mechanisms and legal protections already in place that can mitigate the potential risks and harms of AI in a technologically neutral way; and

(c) where appropriate, takes a more active role in representing Australia in international policy forums and standard-setting contexts.

1.7 At the conclusion of an initial 18-month period, the taskforce could make its recommendations to Government for legislative reform including how urgent those reforms are.[1]

1.8 Stirling & Rose also recommends establishing an Australian Data Advisory Committee (**ADAC**) as soon as possible to consider and report back on how Australia can realise the value of Australian data as sovereign wealth in the age of AI, including consideration of the creation of Australian Strategic Data Lakes (see here) as part of Australia's overall AI policy.

2 Key Learnings and Themes from the Market

Our submission explores six key themes born of our market knowledge.

We recommend a cautious and prudent approach to AI regulation focused on leveraging existing laws and structures and making necessary changes to those existing laws and structures to address aspects of harm or negative impact. We consider that Australia should observe, evaluate and participate in global review of regulatory and similar initiatives (e.g., standard setting) in other jurisdictions to influence their development and to ascertain whether such approaches are suitable for adoption in Australia (including the hard law v soft law decision matrix). We consider there will be substantive technical policy and legal debt associated with legislating too quickly in this highly complex and emerging space.

2.1 The "othering" of AI

Recent discussion on how to "regulate AI" is often focused on the current state of AI as distributed by the big platforms/tech companies such as Microsoft, OpenAI and Google. These tools and platforms are currently controllable, identifiable, and alienable as an AI tool or platform – for example ChatGPT is an AI platform. AI as an

[1] Our leadership team was involved in the drafting of the Law Council Submission on Safe and Responsible AI in Australia. A number of our recommendations and conclusions mirror those found in that submission as a result.

alienable "other" will not be the case for long.

In working with founders, scale-ups and start-ups as well as large institutional emerging technology projects, our market exposure suggests that to "play the movie forward" is to move to a future state *where AI is embedded in everything - including us, our institutions, our consumer goods, every sector, and every industry – from our fridges to our cars, to our money*. This might take five years, this might take 15 years, it is however inevitable.

Good law is technology agnostic. Law "parametrised" by the current shape of a specific technology will be quickly found to be not fit for purpose. Any AI response must take account of the coming inability to alienate an AI tool from an underlying asset, entity, or even biological unit. We discuss this position further at 3.8 below.

2.2 Data as sovereign wealth in the near term

Data control and access is a core requirement in the governmental shepherding of responsible AI outcomes. Regulatory planning around public and private data rights is foundational to ensuring that access to (critical or relevant) data corpora by Australian legal persons is protected, and monopoly style aggregations of corpora are kept in check. We strongly recommend that Australia consider the creation and resourcing of an Australian Data Advisory Committee that works closely with, or is a committee governed by the Australian Competition and Consumer Commission (**ACCC**)

We recommend Australia establish data lake structures, infrastructure, and incentives (including data trust libraries comprising many data lakes) to leverage the increasing value of data as a strategic renewal resource for the benefit of all Australians. The governance or regulation of data may also be important as a mechanism to sanction enterprises which engage in unlawful actions. We consider that there may be a need for changes to be made to Australian Consumer Law and the *Competition and Consumer Act 2010* to engage with new monopolistic actions that are egregious to consumer rights, but do not involve aggregations or misuse of power in definable markets. We discuss data sovereignty further in section 4 below.

2.3 Role of legal and critical infrastructure

We consider that Australia like all sovereign nations, will need to invest in smart legal contract infrastructure to support the digitisation and assetisation of legal agreements as second-generation digital assets. This infrastructure will need to be developed with an eye to 2.2 above and 2.4 below. We also agree with the ALC submission at para 134:

> *"Growth in the use of AI in the operation of infrastructure is likely to be*

significant in coming years. This is an area that requires particularly wide-reaching safeguards, especially in relation to critical infrastructure. Current legislation in relation to critical infrastructure does not, in the Law Council's view, sufficiently address the operation of AI."

2.4 Governing increasingly capable AI over the coming years

We cast forward to a future where increasingly capable AIs autonomously operate enterprises in which humans invest as shareholders, contract with as suppliers and customers, are employed by and interact more broadly with. These AI Organisations (**AIO**) are expected to require diminishing human control and oversight over time and ultimately to operate entirely without human governance (e.g., board of directors) or executive management (e.g., CEO). We discuss AIOs further at 5.1 and 5.3.[2]

2.5 The Responsible Machine Problem

All existing legislation is predicated on rights, responsibilities (custody) and penalties for legal non-compliance to ultimately be placed on or with a 'person' (not a machine or an algorithm). This problem (the **Responsible Machine Problem**) will need to be addressed by the Australian government many times over the next decade and will impact all areas of the law including any decisions on the regulation of AI. We have previously covered the Responsible Machine Problem in our previous submission to the Australian Treasury (see here) and below at 5.2.

2.6 Digital-Free National Parks

If the thesis is correct that:

(a) digital disconnection is a human right, or at minimum an important human need; and

(b) we are moving from external digital connectivity devices to wearable devices, to implantable digital connectivity devices; and

(c) we are moving to ubiquitous and increasingly powerful AI, and

(d) we are moving to increasingly complex augmented, virtual and AI driven realities; and

---

[2] The language of AIO and any discussion of its legal complexity can be similarly interrogated as a discussion in respect of Autonomous Organisations (AO) or Decentralised Autonomous Organisations (DAO). We have made similar observations in respect of DAOs in a response Stirling & Rose made to the 2023 Law Commission of England & Wales – DAO Call for Evidence see here, to those we are making here in respect of AIOs.

(e) we are moving to a model of comprehensive and universal global internet coverage through the deployment of thousands of satellite constellations in Low Earth Orbit (LEO) and increasing large numbers of satellites for communication services in Geostationary Earth Orbit (GEO); and

(f) this future state of universal internet coverage will mean that there will ultimately be no location on Earth where individuals can disconnect from access to, surveillance by, or augmented reality modification without legislative protections that protect or enforce that right.

It then follows that we must at a minimum consider whether new areas or landscapes of disconnection (for example, Digital-Free National Parks coupled with Space Parks) require express legislative protections to secure their creation and ongoing operation.

Without earmarking some of our planet's national parks to remain Digital-Free, we are rapidly approaching a future state where there is no location on earth where we will be able to disconnect from access to, surveillance by, or perpetual augmented reality. The history of national parks and their legislative protections has evolved over time, driven by the need to preserve natural landscapes and wildlife for public enjoyment and future generations. The Digital-Free National Park movement is a call for a proportionate and where necessary coordinated jurisdictional legislative reform to use the existing national park network to found, develop, and run a smaller set of Digital-Free National Parks across the planet. Read the DFNP thesis here.

3 Governing AI in the Immediate Future

3.1 Complexities in regulating AI

Regulating AI presents a significant challenge due to its rapidly evolving nature, the vast diversity in its applications, the complexity and opacity of many AI systems (particularly those involving deep learning) necessitating a high degree of expertise, and Australia's relative size and influence in the geopolitical landscape in which AI operates.

3.2 Foundational models

The relatively recent emergence of foundational models introduces a further regulatory complexity. The relevant distinction between the terms is that "large language models" specifically refers to language-focused systems, while "foundation model" is attempting to stake out a broader function-based concept, which could stretch to accommodate new types of systems in the future. While current foundational models exhibit extraordinary capabilities to undertake multifarious

tasks in diverse domains, potentially transcending their creator's original intent in unforeseen ways, we are just at the start of the foundation model journey. Foundational models are expected to be increasingly pervasive as they underpin diverse products and services and influence a wide array of human endeavours. *Where AI is regulated as a technology, that law may well regulate almost everything.*

3.3 Regulator access to AI expertise

Australian regulators may face challenges in effective regulation of AI as a technology, due to constraints in technical expertise and financial resources to "buy in" the necessary expertise. This epistemic disadvantage is likely to impact not only policy formulation but also the capability to ascertain compliance and enforce the law. Irrespective of if or how Australia ultimately decides to govern AI, we recommend Australia's AI policy include urgent and robust initiatives to build AI capability including in education, business, government and the wider community.

3.4 Australia's place in global AI regulation[3]

In the global context of AI regulation, Australia's position having regard to its size and GDP, is one of a participant rather than a trendsetter. Given that Australian entities are or anticipate exporting their AI products or services, alignment or harmonisation with international regulatory frameworks and limiting additional compliance burden as far as reasonably possible is important to ensure Australian organisations can operate effectively and competitively in a global AI market.

3.5 AI governance in other jurisdictions

Existing regulatory frameworks in overseas jurisdictions differ in their approach to AI. The European Union has recently passed the AI Act which directly regulates AI (as defined in the AI Act) as a technology. The EU AI Act is expected to take effect in or around 2026. Other countries have been less keen to adopt an explicit legislative approach to AI. The United States and Singapore both have adopted what they consider more innovation-focused and sector specific approaches to regulating AI and there is no horizontal regulation of AI as a technology on the legislative horizon in these countries.

3.6 Definition of AI[4]

Against all of these complexities, there is also the persistent issue that the ontological

---

[3] Question 5.

[4] Question 1.

delineation of AI remains a subject of profound scholarly contention[5].

3.7 Technology neutral AI regulation focused on harms and risks[6]

Based on current knowledge, we consider that a technology neutral approach coupled with an impact-focused orientation is preferred to direct regulation of AI as a technology. In governing AI, outcomes should be uniformly treated irrespective of whether those outcomes are the result of human, AI or human + AI action.

An act or omission which is deemed unlawful when committed by a legal person should be equally unlawful if performed or failed to be performed by AI. This principle holds particular relevance in contexts requiring a degree of mental element or 'mens rea'. While AI, as currently conceived, lacks the consciousness and subjective intent traditionally associated with a mental element, the development and deployment of AI is nevertheless, currently, driven by the intentional acts of legal persons. This necessitates the extension of legal norms and sanctions to AI-committed acts. For example, within certain laws, this may be done by changing the definition of legal "person" or "agent" to include AI.

3.8 Future challenges in transparency & duties

Many users of AI experience AI not as AI, but as products and services in which AI may well be invisible (and non-alienable). Ubiquitous AI integration is likely to become inextricably intertwined with human and system operations thus blurring the demarcation between human and AI interaction. We foreshadow that in the future, it may be that individuals assume they are dealing with an AI in some form or degree, unless informed that they are dealing solely with a human.

There is currently something of a ground swell around the notion (particularly in discourse surrounding the EU AI Act) that consumers have a right to know, or alternatively should be informed when they are dealing with AI. This demand is at odds with the hybrid future – non-alienable AI trajectory discussed above.

To make this point another way, if AI systems (or hybrid AI systems) perform more accurately and more efficiently than a system reliant on human in the loop design – the correct expression of director or legal duties may be to disclose whether it is a human (not an AI system) that has been involved in, for example, decision making. A good analogue for this evolution is the use of technology assisted review (**TAR**) (narrow AI) in legal discovery. It is considered that the provision of legal services

---

5 Please see our comments on selected definitions in the Discussion Paper in the Appendix.

6 Question 14.

relying on human only discovery, particularly where there are high volumes of documents to review, and the failure to use TAR, would be a breach of legal duties, given that TAR is faster, more efficient, and more accurate in legal discovery[7].

3.9 Deep fakes

We call out for particular attention of the Department the burgeoning threat of deep fakes[8] to individual privacy, democratic processes, market integrity, security and societal trust. The risks and harms of deep fakes are not confined by geographical boundaries making the challenge of regulating deep fakes a global concern. The interconnected nature of today's digital world means that a deep fake created in one country can easily affect individuals, organisations or political processes in another. In additional to pursuing harmonised regulation of deep fakes, Australia's collaboration with other trusted jurisdictions to investigate and develop authentication technologies such as digital signatures, and where appropriate legal digital infrastructure that supports time stamping, to verify the origin and integrity of digital information will be an important means to protect Australians and civil society.

3.10 Recommended attributes of AI regulation[9]

An alternative option to technology specific AI law may be to adapt existing legal frameworks to account for AI. At this time, we prefer an approach to AI regulation which:

(a) is technology neutral (not explicitly directed at AI);

(b) identifies and addresses issues and gaps in regulation;

(c) builds on existing laws and structures;

(d) harmonises state, federal and international requirements as far as possible to minimise compliance costs on Australian organisations; and

(e) appropriately balances protection of individuals from harm, appropriate redress and preservation of civil society with the advantages and benefits AI innovation can offer.

---

7 McConnell Dowell Constructors (Aus) Pty Ltd v Santam Ltd & Ors [2016] VSC 734

8 Deep Fakes are created using advanced AI techniques to create convincing synthetic audio, video and other data.

9 Question 18.

3.11 Systematic review of legislation through AI-centric lens

We recommend the Australian government undertake a systemic review of all existing legislation through an AI-centric lens to identify areas where existing laws may fall short in addressing the harms and risks posed by AI.

4 Data as Australian Sovereign Wealth

4.1 Paramount importance of data in the age of AI

In the age of AI, data is of paramount importance. Without data, new AI systems cannot be trained, validated, or tested. Without data to refresh existing AI systems they will fall into error and obsolescence. Without data, realistic synthetic data cannot be generated. Data is necessary to underpin new AI and to sustain existing AI.

Accordingly, *data is a potent and highly valuable renewable resource and mechanism of influence in contemporary digital economies*. We urge Australian policymakers and enterprises to fully recognise the value of data and strategically plan to realise the increasing value and power of this asset.

4.2 Australian Data Advisory Committee[10]

In particular, we recommend an Australian Data Advisory Committee (**ADAC**) be established to consider how to capture the value of Australian data in the age of AI, including:

(a) Australian Strategic Data Lakes (**ASDL**): ADAC could evaluate the opportunity for Australia to establish strategic data lakes as part of Australia's overall AI policy. ASDLs may be either centralised or distributed virtual data lakes created through federated learning (or other privacy enhancing techniques). Australian entities and appropriately authorised international actors may draw upon the data contained in these data lakes to advance responsible AI innovation. ASDL may also be deployed as a mechanism of value and exchange, including to secure access to data lakes curated in other jurisdictions for Australian entities or to build International Strategic Data Lakes (**ISDL**) in cooperation with trusted international allies. The provision of data to ASDLs may be encouraged through soft nudges and voluntary mechanisms or may be mandated by regulation, for example if the ASDL is likely to enhance the public good or benefit (such as improving safety and environmental outcomes in Australian industry). A mandatory approach

[10] Question 3 and 4.

to data mirrors existing practices within Australia's petroleum sector, where information becomes 'open file' and accessible to all in certain circumstances[11].

The creation of ASDLs not only symbolises a shift towards a data driven economy but also represents an investment in the intellectual capital of Australia, empowering entities to generate value from Australian data, growing, attracting and retaining valuable AI skills in Australia while advancing the frontiers of AI.

(b) **Data export controls:** Given the value of data, ADAC could explore whether Australia should make a transformative shift in regulatory strategy to expand the purview of export controls to encompass data, not merely technology. Appropriate controls on the export of data signal Australia's recognition of data as a strategic asset and a source of competitive advantage in the digital economy analogous to traditional technologies.

(c) **Data as a regulatory mechanism in AI governance:** ADAC could investigate whether regulation of certain categories of data and the use of data may be a viable mechanism to exert influence over AI applications without directly imposing restrictions on AI technology itself or supressing advancement of Australian innovation. Given that data constitutes the fundamental fuel for AI systems, appropriate regulation of data may be an alternative to regulating AI directly. In contrasting the regulation of data with the regulation of AI, it becomes apparent that the former presents a relatively more manageable regulatory endeavour by operating in a comparatively mature and established knowledge domain. By considering data, a foundational element of AI, policymakers can leverage existing expertise and address a critical component of AI systems without having to fully comprehend the deep complexities of the AI itself. While this approach requires investigation, it may enable a practical, near-term effective means of influencing AI outcomes and addressing societal concerns and may be coordinated with the current review of the Privacy Act 1988 (Cth).

(d) **Data as market power:** In consultation with ACCC, ADAC could examine the role of data in markets and the emergence of data monopolies that cross traditional markets. "The ACCC considers that the role of data in future markets is likely to be significant and will be an important factor to be taken

---

11 Offshore Petroleum and Greenhouse Gas Storage (Resource Management and Administration) Regulations 2011 Parts 8 and 10.

into account in assessing the likely competitive effect of relevant mergers and acquisitions"[12]. We consider that beyond the arena of mergers and acquisitions, the ongoing acquisition of data, including the acquisition of data in the usual course of an enterprise (e.g. Google acquiring users' search data) should be examined where it provides "strong competitive advantage, creating barriers to rivals entering and expanding in relevant markets...[13]". We consider that there may be a need for changes to be made to Australian Consumer Law and the *Competition and Consumer Act 2010* to engage with data monopolies and new monopolistic actions that are egregious to consumer rights but do not involve aggregations or misuse of power in definable markets.

5 Governing Increasingly Capable AI Systems of the Future

5.1 AI Organisations (AIOs)[14]

We cast forward to a future where increasingly capable AIs autonomously operate enterprises in which humans invest as members/shareholders, contract with as suppliers and customers, are employed by and interact more broadly with. These AI Organisations (**AIOs**) are expected to require diminishing human control and oversight over time and ultimately to operate entirely without human governance (e.g., board of directors) or executive management (e.g., CEO).

5.2 Responsible Machine Problem[15]

All existing legislation is predicated on rights, responsibilities and penalties for noncompliance to ultimately be placed on or with a 'person', not an AIO. If there is no responsible 'person' in a decision system, we have a Responsible Machine Problem. The Responsible Machine Problem will need to be addressed by regulatory bodies many times over the next decade and will impact all areas of the law.

We foreshadow the potential for AIOs to have capabilities closely mimicking centralised human governance (board directors and CEOs), but with superior access to data for real-time decision making. We note that there is a continuum in the degree

---

12 Australian Competition & Consumer Commission "Digital Platforms Inquiry. Final Report" June 2019 p11.

13 Ibid.

14 Question 2.

15 Question 2.

of autonomy in AIOs, with governance delivered (at least in the near term) via a combination of AI and human decision-making. The pathway to fully autonomous AIOs will likely involve an ongoing expansion of the matters handled autonomously, that is without human intervention. We emphasise the importance of legal accountability for AIs and the need for legal personality to protect internal and external stakeholders interacting with the AIO. We note also that an AIO may be the progeny of another AIO (including many generations removed from a human founder, creator or developer).

5.3 AIOs as legal persons

Granting legal personality to an AIO should not be unconditional. We propose a list of bare requirements to recognise that an AIO is a separate entity that may functionally operate with complete independence separate from its founders/developers. These conditions include registration with appropriate agencies and a unique identifier as an AIO, evidence of de minimis functionality, sanctions for aberrant behaviour, economic reserves, adjustments for intent and protections for minority members/shareholders. The idea is to provide a balance between recognising the separate functionality of AIO as an organisation and ensuring that humans (e.g., directors and officers) are held accountable for their actions and decisions related to AI which they deploy in Corporations.

While we support personal liability being attached to any humans that direct and manage the AI, there are practical and legal challenges with relying on strict human responsibility alone. With respect to an AIO, the universe of potentially liable persons are the founders, developers, creators, or company which provided the resources for the development of the AIO. In the case of civil or criminal actions which require mental intent (such as fraud) it may be difficult or even inappropriate to attribute the requisite intent to a particular individual. Such a cause of action against a human actor may not be sufficient recourse for an injured party.

Further, a responsible human may be too far removed from an actual decision or action of the AIO to attribute fault to the human. The realistic possibility of an AIO alone being responsible for making an aberrant decision or action which founders, creators, developers are not reasonably able to foresee, (particularly when the founders creators and developers are another AIO rather than a legal person) draws a crucial distinction between an AIO and a corporation or any other traditional non-human legal personality. A corporation, partnership or unincorporated association ultimately depends on a human in the loop to make a decision, whereas the AIO may be able to independently make decisions for itself with limited or no human direction. Technically, this may be achievable without the AI being sentient or having the ability to 'think' as we currently understand 'thinking'.

For these reasons, there is merit in potentially granting legal personality to protect internal and external stakeholders interacting with the AIO. To iterate:

(a) The increasing capability of AI will mean that greater management of an organisation for day-to-day functions by AI becomes feasible and widespread;

(b) An AIO presents potential improvements over a human-led organisation. The capacity to review more relevant information quicker, undertake more efficient decision making, and to be less prone to environmental biases (e.g. improper influence and fraud) than human counterparts means that there may be productivity gains and other benefits from AIOs. Individuals and investors may prefer to be shareholders in an AI led organisation than shareholders in a human led corporation; and

(c) From the perspective of third-party protection, if the above two propositions are accepted, then the AIO itself should have legal personality to be subject to legal sanction not only for the benefit of injured parties but also to empower regulators to deter aberrant behaviour. This provides an additional avenue of relief to the extent that action against individuals is not appropriate.

However, we have two key caveats in relation to conferring legal personality to AIOs.

First, such legal personality is not unconditional, and we set out details of conditions that lawmakers may wish to impose on an AIO in exchange for conferring legal personality.

Secondly, recognising legal personality of AIOs should not abrogate personal responsibility of human actors to the extent they are involved in decision making and management. Granting legal personality to an AIO expands the recourse available to external and internal stakeholders to both the AIO and/or human actors (as appropriate), rather than a liability shielding mechanism for any one responsible person.

The liabilities of responsible human actors and the AIO should be on a non-mutually exclusive continuum that reflects the involvement of the human actors and the AIO in the actual decision making and management of the AIO. There is also a quibble that it may be incorrect to assume that developers are only ever "human persons".

A suggested list of conditions may include the following:

(d) **Registration:** All AIOs should be registered with the appropriate local and international agencies and should have their own unique identifier (such as a

registration number or digital fingerprint). We would propose that there should be a dedicated governmental authority to regulate AIOs and monitor the register.

(e) **Identification as an AIO:** All AIOs should have a signifier in their names such that they are readily identifiable as an AIO in all interactions. Much the same way companies are identified by "Ltd", "Inc", "LLC" or other signifiers of their corporate status, AIOs should also be identifiable as an AIO – for example ACME AIO Ltd.

(f) **Functional AIO test:** An AIO must be able to provide reasonable evidence of its capability to discharge de minimus functionality, i.e. those things necessary to discharge the AIO's legal obligations. For example, paying annual AIO registration fees. This evidence would need to be provided as a condition to registration of the AIO, but also possibly on a periodic and ongoing basis to confirm the continued functionality of the AIO in what is a dynamic and rapidly evolving arena where consumer harm in interacting with an AIO (or at least with an early generation AIO) may be elevated.

(g) **AIO sanctions:** Economic risk allocation does not solve the role of "skin in the game" whereby personal liability and the potential for adverse action against an individual ensures alignment of interests and actions. Even with limited recourse vehicles, there is still a responsible person (in the case of a company, directors still owe director duties to the company, and in certain circumstances to, or are personally responsible for harm inflicted on, third parties e.g., misleading and deceptive conduct).

The applicable governmental authority regulating AIOs should be conferred with powers to suspend the operation or terminate an AIO and exercise rights of forfeiture against AIO assets. While it is impossible to imprison an AIO, effective sanctions may include suspension or termination of the AIOs ability to access data which is a critical input for effective AIO operations or suspension or termination of all active operation of an AIO. To prevent AIOs "phoenixing" a sanctioned AIO, there would have to be a means of identifying the code of a phoenixed AIO from the underlying code of one that has been sanctioned.

We reiterate that even if AIOs had legal personality and could be separately held liable, it remains prudent to hold individuals accountable: (a) to the extent that the private citizens are engaged in directing the AIO; (b) to the extent that it is difficult to properly censure an AIO to ensure that there is 'skin in the game' and there is insufficient deterrence from bad behaviour.

(h) **Economic reserves:** It may be feasible to improve the economic risk of AIOs to counterparties by imposing some form of credit support. This could be in the form of actual asset backing (i.e. collateralisation) to meet potential claims or guarantees provided by individuals or legal entities of substance or insurance (such insurance to be for a sufficient level of cover for the AIOs activities and incurred obligations and with no material exclusions).

A counterargument to this is that without some form of collateralisation, an AIO represents a similar risk profile to external counterparties as corporate or other limited liability vehicles which have minimal asset backing e.g., shell companies. In which case it is a matter of caveat emptor for parties that fail to do their due diligence (assuming that the AIO is registered and identified as an AIO pursuant to conditions 1 and 2 above.).

(i) **Intent of an AIO:** As referred above, causes of actions and offences which presuppose mental intent will need to be adjusted for AIOs, given there may not be a governing 'mind' in the legally understood sense. This may involve abolishing the mental intent from such causes of actions and offences to the extent an AIO is involved or relying on a completely objective (and possibly expert determined) test to determine what the AIO should have done. Corporations attribute intent to their directors and officers. We do not think this presumption should automatically apply to AIOs, but it should be available as a continuation of our liability continuum comment above, to the extent that a founder or other human actor is a shadow director or is otherwise involved in the material decision making.

(j) **Minority AIO member/shareholder protections:** Bare minimum protections may be granted to minority AIO shareholders (e.g., no compulsory transfer or extinguishment of minority shares). This is in conjunction with the evidence required to establish a functional AIO may comprise the base governance protections to minority member/shareholders to avoid potential fraud on the minority.

Going forward, a smart legal contract can act as an AIO governance tool with greater ability to moderate the complexity of risk allocation, rules of engagement, obligations and remedies between differing members/shareholders and regulatory bodies in respect of the AIO and its individual assets.

For example, a network or constitution and some form of member/shareholder/unitholders AIO agreement, connected to the third-party legal agreements drafted and constructed as smart legal contracts (with natural language and coded instructions) could do the work of automating performance of an AIO or a

hybrid AIO with human governance rights and obligations.

If the Department would like further information on this final point, we would be delighted to provide additional written or verbal comments.

6 Conclusion

The increasing capability and multifaceted nature of AI demands an adaptable and future-focused policy response and regulatory framework.

Stirling & Rose's recommendation outlines an 18-month taskforce in Australia that will provide an advisory role in scrutinising international AI approaches and identifying domestic gaps, while observing global hard and soft AI laws to gauge efficacy and alignment with existing legal reforms. Fostering collaboration with legal and technical stakeholders, the taskforce will review Australian laws affected by AI within the existing legal protections and governance mechanisms, while also actively representing Australia in international contexts. The initiative will culminate in the formulation of legislative reform recommendations to the Australian Government.

The recommended technology-neutral, impact focused approach ensures that regulatory measures are not confined to the limitations of current AI technologies but are adaptive to the evolving landscape.

The recognition of data as a strategic asset should be a cornerstone of AI policy in Australia. We recommend the Department convene an Australian Data Advisory Committee (**ADAC**) to investigate the opportunity to establish Australian Strategic Data Lakes (**ASDLs**) and other mechanisms to enable Australia to leverage its data resources not only to support AI innovation in Australia, but also as a means of value and exchange.

Finally, we postulate a future where increasingly capable AI autonomously operates enterprises and in response, set out a rationale and indicative requirements for extending legal personality to AI Organisations (**AIOs**).

The future of AI in Australia hinges on our collective resolve to forge a path that is both progressive and conscientious, reflecting a commitment to excellence and ethical stewardship in the age of intelligent machines.

7 Appendix - Discussion Paper Definitions

(a) **Governance:** We support the definition of governance in the Discussion Paper to include regulatory and voluntary mechanisms. We consider both regulatory and voluntary mechanisms should work together to deliver safe and responsible AI in Australia.

(b) **AI:** We expect term "AI" to continually metamorphosise in response to technological advancements and the burgeoning acceptance of erstwhile AI frontiers as part of everyday life. We support the definition of AI proposed in the Discussion Paper as a helpful definition for the purposes of examining AI policy in Australia and note that it closely aligns to the OECD definition of AI.

The appropriate definition of AI nevertheless depends upon the purpose for which the definition is employed. Therefore, if regulations are created to apply or not apply depending upon the definition of AI, the definition of AI in the discussion paper may be unduly limited for future purposes having regard to the exponential advances in machine learning systems and the expectation of new techniques and capabilities. We raise three definitional considerations:

(i) **Ensemble systems:** The definition of AI is limited to systems engineered "without explicit programming". This limitation does not take into account that many AI systems are ensemble systems, that is, engineered with a combination of machine learning techniques together with some degree of explicit human executed programming.

(ii) **Human-defined objectives or parameters*:*** We foresee AI systems engineered using AI generated (rather than human-defined) objectives and parameters. Limiting the definition to "a given set of human-defined objectives or parameters" would not capture AI systems which are the progeny (i.e. engineered) by other AI systems. We do not consider there is any sensible reason to exclude from regulations AI systems which are themselves the progeny of other AI systems.

(iii) **Varying levels of automation:** Machine learning systems, for example reinforcement learning systems, are increasingly capable of acting with autonomy and adjusting to dynamic information and experience in ways that transcend the term "automation". It is unclear whether the words "designed to operate with varying levels

of automation" are intended to limit the ambit of the definition to exclude AI systems which are autonomous. If so, we do not consider there is merit in constraining the definition of AI in this way.

(c) **Automated Decision Making (ADM):** To the extent that regulation of ADM is contemplated, we consider this definition to have a very wide ambit, particularly "guide a human decision-maker through relevant facts, legislation or policy" which could capture many search engines which assist decision-makers, including in non-government settings, such as research search engines commonly used in providing legal advice or in customer service settings.