7 June 2019

COMMUNICATIONS ALLIANCE LTD
www.commsalliance.com.au

Artificial Intelligence
Strategic Policy Division
Department of Industry, Innovation and Science
GPO Box 2013
Canberra, ACT, 2601

Submitted via email only: artificial.intelligence@industry.gov.au

Dear Sir / Madam,

RE: Artificial Intelligence – Australia's Ethics Framework

Communications Alliance welcomes the opportunity to provide this submission in response to the *Artificial Intelligence – Australia's Ethics Framework* Discussion Paper released by the Department of Industry, Innovation and Science (DIIS).

Communications Alliance is the primary communications industry body in Australia. Its membership is drawn from a wide cross-section of the communications industry, including carriers, carriage and internet service providers, content providers, digital platforms, search engines, equipment vendors, IT companies, consultants and business groups.

Its vision is to provide a unified voice for the telecommunications industry and to lead it into the next generation of converging networks, technologies and services. The prime mission of Communications Alliance is to promote the growth of the Australian communications industry and the protection of consumer interests by fostering the highest standards of business ethics and behaviour through industry self-governance. For more details about Communications Alliance, see http://www.commsalliance.com.au.

General remarks:

Communications Alliance applauds the Department for undertaking this initiative and the associated consultation. Artificial Intelligence (AI) has profound – and not yet fully understood – implications for the national and global economies, all sectors of industry and the lives of all Australians.

The technology will manifest in countless different applications that interact with humans at many junctures of their life. Its pace of development is hard to predict and the end-point is unknown. AI is equally laden with disruptive opportunity and scope to generate risk.

The framework that is being developed 'only' covers civilian AI. While this is understandable given the formidably difficult task of creating regulation and essentials tools for AI, it is important to note that even greater risks – and therefore challenges – lie in the use of AI for military purposes. Those risks and challenges are even more daunting when considering the power of weapons of mass destruction and the potential for those, in combination with the use of AI, to fall into hands (including through cyber security breaches/hacks) that care less about ethical considerations.

The definition of AI that the report adopts is relatively broad and is likely to be subject to debate. Based on this definition, it will be difficult to discern when a certain activity or technology constitutes AI – a difficulty that would likely arise with most, if not all, definitions of

AI. For example, would the use of one piece of data to determine what is being done with the next piece fit the definition of an AI learning system? Or would a filtering component in the depths of a sensor device, which determines what information is of interest, be within the scope of AI as defined?

The scope of the definition is important because it would appear to encompass many AI applications where the proposed ethical framework is not relevant, e.g. where AI is used to deliver machine optimisation such as network security monitoring, fraud detection or machine optimisation functions. An AI system designed to learn the patterns of use and/or interference on a radio network and to optimise the use of radio spectrum to maximise speed and reliability with limited interconnection and agency does not require an ethical framework.

The use of AI as a machine optimisation solution or as part of such a solution is comparable with the use of conventional software. For cases such as this it would perhaps be more appropriate to provide suggestions or guidelines for buyers of such solutions in order to optimise selection and implementation rather than an ethics framework?

A related issue is that the ethical principles suggested by the Paper anticipate that AI will be deployed to make decisions or judgement with potential impact on individuals. However, AI may be narrow AI in machine-learning, deep-learning or general AI. A sophisticated system is likely to be a combination of different systems. There are many important potential factors associated with function, use and deployment where the use of AI will have no individual or social impact. For example:

- AI can be used in an automated environment for a functional purpose;
- AI can be used as a tool for human problem solving, scientific modelling, to design solutions and/or to produce a creative work; and
- AI may be isolated, connected to other systems and/or have a degree of agency.

In our view, the Paper would benefit from an introduction discussing the different types of AI, the different ways it can be implemented and how the ethics framework will have relevance having regard to function, use and deployment.

In addition to these definitional problems, formulating principles that can be translated into useable and adaptable tools across a multitude of scenarios poses real challenges.

Industry regards the Discussion Paper as constructive and balanced in tone, with a useful summary of existing domestic law and regulation as it pertains to AI, as well as international frameworks.

We also agree with the Department's view that "Fit for purpose, flexible and nimble approaches are appropriate for the regulation and governance of new and emerging digital technologies."[1] Given the rapid evolution of technologies and the increasing convergence of areas such as privacy protections, consumer data rights, AI, cyber security, IoT etc., we believe that it would be wise to carefully analyse existing frameworks and regulations and how those might accommodate evolving new technologies rather than defaulting to the creation of new regulatory frameworks which may be adding unnecessary complexity and cost. Also, any regulatory intervention only ought to be contemplated when there is a proven failure of markets to produce the desired outcome.

[1] p.4, *Artificial Intelligence, Australia's Ethics Framework (A Discussion Paper)*, Department of Industry, Innovation and Science

The Proposed Core Principles

The proposed principles sound reasonable and attractive when viewed through a 'high-level lens', but begin to falter under scrutiny as to how they can be applied in practice.

Principle 1 is generally sound, but calculating 'net-benefits' will often be very difficult and subject to opinion and/or interpretation.

It is also likely that research on AI will precede – often quite significantly in advance of any application of the research results – any benefits that may be generated from such AI. To avoid research being unintentionally constrained, it is important that any cost-benefit analysis has a sufficiently long-term outlook and does not decouple the (costly) process of research into AI from its future benefit-generating application.

Principle 3 should recognise that international and the various tiers of domestic law will not necessarily be in accord.

Principle 4 on privacy protection uses the term 'people's private data'. This ought to be amended to 'personal information' in order to align with Australian privacy law.

Principle 6, regarding transparency & explainability, is problematic. At what point does an algorithm impact someone such that it would trigger a requirement for disclosure? Such a decision is fairly obvious in cases where an algorithm has a substantial impact on an individual (e.g. related to health, finance, etc.). However, in many everyday cases – such as the use of AI to make automatic adjustments to a camera-phone's exposure settings – it may be extraneous to the user whether an algorithm has been used or not.

The concept of explainability is also likely to pose its own challenges given the complexities of the subject matter. Privacy regulation, which requires a similar approach, demonstrates the challenges of describing complex concepts for consumers who are understandably not equipped or inclined to read lengthy legal explanations. It can be expected that AI, its use and consequences for individuals pose even greater challenges with respect to explainability. It would appear that the application of the principle requires a fair degree of flexibility to account for the vast variety of AI applications and situations in which users could be subject to it.

Across the various domains (safety, accountability, etc.), it might be useful to insert some nuance with regard to open source technologies and off-the-shelf solutions where developers may not be interacting directly with third parties applying these technologies in particular use cases or their eventual end users.

Principle 7 relating to contestability could perhaps be amended to specify that the impact in question must be a significant or substantial one, in order to trigger the principle. Otherwise there is a risk that insignificant impacts may act as a trigger to require costly and complex processes that will damage innovation while generating only minimal gains for consumers.

Principle 8 regarding accountability, is not realistic. As presently worded, it suggests that any person or organisation involved in the creation of an open source model or API that ends up being used in an AI system should be identifiable and accountable for the impacts, even if they were unintended.

It will be impossible, however, for the developer to predict or even find out all the ways in which AI models they have created will be used, particularly if the model had been made available on an open source basis. While they can take steps to be responsible (e.g. by providing guidance on how the model was created/target uses, and include warnings of potential risks) the actual use cases are not something within their control or even visibility.

It appears that accountability for the impacts that were reasonably foreseeable at the time of the creation/release/application of the AI would constitute a more practical approach.

Security Principle: We strongly recommend the addition of a principle focusing on the security of AI. AI will pose a significant challenge from a cyber security perspective as large volumes of centralised data create a 'honeypot' that is likely to be targeted by criminal actors. In addition, the power of AI systems is likely to present an attractive target for those who seek to exert control through the use of AI and who wish to manipulate AI systems. AI itself may also facilitate very complex cyber attacks against companies and Government organisations.

It can be argued that securing the powerful AI that we create must be part of an ethical consideration rather than a mere commercial implication or prerequisite to applying other principles, such as the privacy protection principle.

Creating such a security principle would also align with other international principles, such as the OECD Principles.

A Toolkit for Ethical AI

It would be useful if the document provided further information about the difference between 'impact assessments' and 'risk assessments'. Both would seem to involve looking at the risk of potentially negative impacts (which can't be evaluated without looking at specific groups).

Is the main difference that the impact assessment goes further in documentation such that it becomes auditable, and include description of steps being taken to mitigate a risk?

The examples of risk assessment frameworks at section 7.2 of the Discussion Paper are very useful. If it was possible to create similar examples relating to impact assessments, this would be a helpful addition to the document.

The Toolkit suggests the use of experts to review AI to ensure compliance with ethical principles and/or legislation. While this may be a useful approach in theory, it seems that it would be very difficult in practice to scrutinise the AI without in-depth expert knowledge of the intellectual property (IP) that underlies the AI which the owners of that IP will be very reluctant to disclose to third parties even if they are independent.

The Paper points to "the […] need for consistent and universal guidelines, applicable across various industries using technology that is able to make decisions significantly affecting human lives."[2] This seems very ambitious. It is doubtful that 'one size fits all' principles could be developed without creating unnecessary assessments not relevant for the AI system at hand. It might be more practicable to start with sector-specific guidelines and then explore the possibility to generalise if similarities emerged.

A further suggestion is to test the practical applications of principles against use cases beyond Government services, to for e.g. use of expert systems for network operations, or transport logistics optimisation.

Evolution of the Framework and supporting material:

Undoubtedly, AI will evolve rapidly and potentially beyond our current imagination. As the technology and our understanding evolve, the Framework equally ought to be sufficiently flexible to evolve and adapt. It is, therefore, important to resist the temptation to be overly prescriptive at this early stage of the evolutionary cycle.

The Framework's Toolkit references the importance of best practice guidance, standards and certifications. It will be critical to involve industry early and heavily in the development (and

[2] p.36, *Artificial Intelligence, Australia's Ethics Framework (A Discussion Paper)*, Department of Industry, Innovation and Science

updating) of such additional tools. Where necessary, industry-specific material ought to be created.

We look forward to further engaging with your Department and other relevant stakeholders over this important topic.

Please contact Christiane Gillespie-Jones (c.gillespiejones@commsalliance.com.au) or myself if you have any questions or wish to discuss.

Yours sincerely,

John Stanton

Chief Executive Officer
Communications Alliance