Supporting Responsible AI: Discussion Paper

Consultation on Safe and responsible AI in Australia

Getty Images appreciates the opportunity to provide our comments in relation to this important consultation and commend the Australian Government for supporting safe and responsible AI practices. Getty Images is an established and respected member of the global media, a content owner and marketplace for visual content and, we maintain and operate a significant presence in Australia.

As a content creator and a company that represents hundreds of thousands of content creators, we believe AI can produce commercial and societal benefit; however, we believe that appropriate safeguards need to be put in place. More specifically, we believe the development and deployment of AI needs to be balanced with long-standing copyright protections, the rights of creators who represent a sizeable share of our overall economic output, the privacy rights of individuals and the public's fluency in fact. Although the discussion paper recognizes that existing regulations such as copyright, intellectual property, and privacy laws as a starting point for considering how best to ensure appropriate safeguards are in place, it fails to fully recognize the existential threat that AI poses to the creative and media industries. Our responses below are intended to highlight these issues and suggest additional strategies for building public trust in AI technologies and systems and for protecting the rights of human creators.

Potential Gaps in Approaches

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

While existing Australia regulatory approaches including approaches to copyright, intellectual property and privacy laws can help mitigate some risks to the creative sectors, they fall short of providing human creators sufficient protection. More specifically, Australia's existing regulatory approaches fall short of explicitly requiring consent in the collection of individual and private information, as well as the creative

works of individuals that are so vital in training robust AI systems (including but not limited to; images, videos, music and books). For many creatives, royalty payments associated with the use of their work is their sole source of income and is their primary way to recouping the considerable investments they make to create. Allowing AI developers to train generative AI models without obtaining authorization circumvents this revenue stream and discourages the generation of new creative works.

As AI developers rush to develop commercial technologies that rely on rich, large and quality data sets, there is a temptation to ignore third party rights, and it is well documented that irresponsible developers have used protected data for training without authorization or consent. This practice disrespects fundamental human rights, and in the context of generative AI systems that have the ability to generate derivative works, can compete with the protected work used for training.

One solution that can help mitigate these risks it to Impose a statutory duty of care on AI developers, combined with a government certification scheme, for when sourcing input data requiring: (i) that fundamental human rights are respected, including the rights to enjoyment of property (such as copyright and other intellectual property rights) and privacy; and (ii) compliance with certain minimum accountability and transparency criteria. For training data including copyright works, such duty/scheme should articulate certain minimum standards relating to due diligence to identify the copyright owners and to secure authorization therefrom (failing which any unauthorized data should not be used). Any certificates could be renewable annually and subject to audit against risk of revocation to protect against non-compliance.

In addition, a regulatory framework that requires transparency and provides clear avenues for rights holders to seek remedy when protected content is used as training data without consent could help promote ethical and responsible AI development. More specifically, in the context of foundation models used in generative AI systems, providers of such models should be required to document and make publicly available a detailed summary of protected content used as training data similar to the regulation proposed in Article 28 of the EU AI Act.

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

Promotion of industry developed transparency tools that help users of AI systems to know when they are interacting with AI systems and, specifically, to identify outputs of generative AI tools, e.g. through adoption of metadata and/or watermarking technologies.

Developers and deployers of generative AI systems should be obligated to meet transparency requirements regarding the output of their models. This includes indicating when content is generated or modified by AI and implementing technical standards. For example, attaching meta-data that identifies AI generated or modified content in the IPTC DigitalSourceType field of that content. (see https://www.iptc.org/news/iptc-releases-draft-of-digital-source-type-vocabulary-to-support-synthetic-media/). Transparency requirements should also include disclosures related to any protected data used for training or finetuning generative foundational models.

Responses Suitable for Australia

5. Are there any governance measures being taken or considered by other countries (including any not discussed in this paper) that are relevant, adaptable and desirable for Australia?

Harmonisation is a hard thing to do but is necessary. The Australian government needs to make it a priority to work with its counterparts in the EU, US and other key digital allies to develop basic international norms and standards including basic transparency standards. When working with international partners, Australia should support IP rights and continue its legacy of vigorously protecting IP and human rights on the international stage.

The EU AI Act is a good model to follow especially in how it handles transparency standards. There is a need for transparency requirements similar to proposed EU AI Act, i.e. in relation to both (i) composition and source of training data that includes copyright works at the input stage; and (ii) the labelling of outputs of generative-AI tools to guard against risks posed by deepfakes. Such transparency standards will help promote responsible and ethical innovation and provide chouse for consumers in a fair and balanced manner.

Target Areas

9. Given the importance of transparency across the AI lifecycle, please share your thoughts on:

a. where and when transparency will be most critical and valuable to mitigate potential AI risks and to improve public trust and confidence in AI?

Transparency is critical to mitigate against the risk of violating third party IP and privacy rights and the development of unfair and bias models. Without transparency obligations, providers of models can effectively hide such violations inside a backbox and it will be very difficult for the public to identify and understand models that are bias.

b. mandating transparency requirements across the private and public sectors, including how these requirements could be implemented.

One way to mandate transparency requirements is to require both private and public sector organisations to keep auditable records of all training data sets used including how the data was sourced.. Such auditable records should be available in a suitable manner for the intended audience(s) so that IP owners of such data may check whether appropriate consent has been obtained for the use of their content and users of the AI tools will be able to understand the make-up of the training data (which could be helpful for understanding questions of bias) and whether consent has been obtained as may be required by IP owners (which could help them to evaluate any associated legal risk before they use the tool). This in turn provides greater visibility to consumers, who can make informed decisions as to whether they are consuming and using outputs of training data that has been accompanied by any warranties and indemnities. Greater protection can be provided to consumers by requiring such AI developers to provide limited warranties in relation to their outputs.

The government should Introduce a statutory duty and/or certification scheme for when sourcing training data to encourage the design and development of AI systems that are more explainable, trustworthy and sustainable. Such duty/scheme should explicitly acknowledge where there is a legal requirement to license training data containing copyright works and be accompanied by educational programmes regarding to the sourcing training data.

The government should prioritise the protection of IP rights by obligating transparency when IP protected works are used to train generative AI. Such transparency will enable rights holders to seek remedies under existing law.

There needs to be obligations of transparency throughout the entire AI life cycle. Mechanisms for protecting rightsholders whose copyright works are used as training data without their consent, credit or compensation.

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

The idea of government sponsored AI sandboxes that are beacons for ethical and responsible AI development.

The government should require transparency on all training sets utilized and require permission of IP owner where such models or any derivative models have the potential to be commercially deployed. The government must work with rightsholders to identify sources of licensable data for use in the sandbox and could seek to make introductions to AI developers. Getty Images would be keen to explore participating in the sandboxes.

The government should ensure that compliance with sandbox standards will enable Australian companies to be able to market their AI tools globally by requiring all participants to meet high transparency standards. The government should award certification to AI developers who participate and are able to demonstrate that they have adhered to guidance regarding the licensing of data that includes copyright works.

Due to the importance of protecting IP Rights and preventing private data from being misused, an AI sandbox that facilitates the ethical development of generative AI foundational models is critical in making sure that transparency tools are incorporated by design. More specifically, an AI sandbox for the development of generative AI models that are intended to create or modify images could help make sure that the rights of copyright owners are respected and that creators receive adequate renumeration for the use of their works.

Risk Based Approaches

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

Regulators will need to avoid the temptation to create exemptions for open-source developers that enable them to circumvent legal requirements regarding the sourcing of training data that contains protected works. Open-source should not be equated to non-commercial research. Doing so would open up loopholes that will be ripe for abuse by bad actors. In particular, the Australian government should not facilitate the practice of data laundering through the use of text and datamining exceptions. Open-source foundation models, known to have been pre-trained on protected data, need to be held to the same level of legal compliance as closed-source foundation models. Moreso because downstream use of open-source models will be more difficult to regulate than closed-source models. Users of such generative foundation models who create generative AI tools or produce generative content need to be assured that that consent for protected training data was legally obtained.

To better understand the copyright issues related foundational models, we recommend that the paper, "Foundational Models and Fair Use" published by Stanford University on March 29, 2023 be consulted. (https://arxiv.org/pdf/2303.15715.pdf). The article focuses on the problems of relying on "fair use" as a defense to copyright infringement in the US and discusses potential techniques for risk mitigation.