Introduction

Adobe appreciates the opportunity to submit comments in response to the *Safe and Responsible AI in Australia: Discussion Paper* (Discussion Paper) prepared by the Australian Government's Department of Industry, Science, and Resources. We recognise the need to build robust AI assurance mechanisms to ensure artificial intelligence (AI) systems are designed, developed, and deployed in a trustworthy and responsible manner. We look forward to engaging with the Australian Government on this important topic.

Our submission focuses on three main points:

- **Collaborative approach with industry.** A collaborative approach that leverages industry leadership and initiataives on AI governance, especially at this early stage of AI research and development, can be very effective. The Australian Government should facilitate and recognise industry AI governance initiatives. In addition, Adobe urges the Government to promote transparency in digital content by implementing Content Credentials—open-sourced technology enabling creators to attach digital provenance information to digital content so that consumers can see the origins and edit content history online. The 1,500+ member Content Authenticity Initiative, and its underlying standards body, the Coalition for Content Provenance and Authenticity, are working to drive adoption and implementation of this free, open-source technology across industries, tools, and platforms. Adobe strongly supports the Australian Government leveraging Content Credentials to:
  - support greater transparency and trust in our online ecosystem by ensuring creator attribution for digital content and helping users understand the origins and edits of the content they are consuming
  - indicate that generative AI was used in the creation of the content;
  - encourage industry adoption of a Do Not Train standard for AI systems; and
  - endorse this industry-developed standard for use by other organisations and governments.
- **Risk-based approach.** Adobe supports a risk-based approach to AI in line with the US National Institute of Standards and Technology's *AI Risk Management Framework* and the EU's *Artificial Intelligence Act*, as well as others around the world. As the approach to regulating AI develops, we encourage Australia to remain technology neutral by focusing on the socio-technical context of AI rather than the technology itself.
- **Responses suitable for Australia.** Adobe suggests a distinct role for the Government in monitoring and assessing AI, supporting implementation, assisting innovators, and promoting education and awareness. An integral aspect of this recommendation is the Government's publication of standardised datasets to facilitate addressing bias in AI systems.

Adobe welcomes the opportunity to submit these comments to the Department of Industry, Science, and Resources. We are available at any time to discuss further as the Government finalises its strategy and approach to Safe and Responsible AI in Australia.

Sincerely,

Jennifer Mulveny
Government Relations and Public Policy Director, Asia Pacific, Adobe

Submission in response to the *Safe and Responsible AI in Australia Discussion Paper*

26 July 2023

<u>**Background**</u>

At Adobe, our mission is to empower everyone to deliver the world's best digital experiences. Since our founding in December 1982, we have continued to pioneer transformative technologies that allow our customers—who range from emerging artists to global brands—to channel their imaginations, unleash their creativity, and power their businesses.

Adobe's business consists of three cloud-based solutions: Creative Cloud, Document Cloud, and Experience Cloud. In Adobe Creative Cloud, our tools like Photoshop, Illustrator, and Premiere Pro empower millions of creators around the world to create compelling assets like photos, videos, and illustrations, while our online and mobile design app Adobe Express allows people of all skills to easily create outstanding social graphics, videos, and web pages. In Document Cloud, our products provide individuals, small and medium businesses, and large enterprises with tools to support all their digital document needs. Lastly, our Experience Cloud business supports businesses with their websites and e-commerce experiences.

Across all our products, Adobe is building on a decade-long legacy of AI innovation by leveraging the power of AI to deliver hundreds of intelligent capabilities. In Creative Cloud, AI powers many creative functions, including advanced image editing features in Photoshop, making it easier for everyone to tell their story with simpler and more intuitive tools. We recently launched Adobe Firefly, our new family of creative generative AI models designed to be both creator-focused and designed to be safe for commercial use.[1] With Firefly, users can channel their creativity in ways they never imagined possible by simply typing in a prompt and generating images in seconds.

In Document Cloud, AI powers features like Liquid Mode, which automatically converts a PDF layout to different formats to fit different screens so it can be easily read on tablets, mobile devices, and more. As part of our Digital Experience offerings, Adobe's customers can use AI-driven features to deliver relevant and meaningful insights and personalised digital experiences to the millions of visitors on their websites.

We believe that AI done right will amplify human creativity and capabilities to new levels with deeper insights, accelerated task performance, and improved decision-making ability. As we continue to harness the power of AI, we are committed to developing and deploying AI in line with our AI Ethics principles of accountability, responsibility, and transparency while considering its broader impact on society.[2]

---

[1] TechCrunch, "Adobe's thoughts on the ethics of AI-generated images (and paying its contributors for them)," Frederic Lardinois (March 22, 2023).

[2] Adobe Blog, "Seeing is believing: It's time to restore trust in online media," Dana Rao (May 15, 2023).

We have been active in this space and in addition to making submissions to several AI-related policy processes in Australia, we have provided our perspective to various governments around the world including the US government through different consultations and the UK's pro-innovation approach policy paper outlining support for the flexible and principled-based approach. We have also testified before the European Parliament's Special Committee on AI in the Digital Age (AIDA) and have engaged with the White House Office of Science and Technology Policy as they developed the AI Bill of Rights.

Globally, governments are establishing rules and standards to promote AI assurance. When standards are targeted, contextual, and harmonised across jurisdictions, they can be valuable in advancing responsible AI and reducing harm related to AI systems.

A. Adobe Favours Collaborative, Regulatory Approaches with Industry to AI Assurance

Adobe agrees with the Discussion Paper's general approach to governing harmful outcomes of AI systems, not AI technology itself. At this early stage in the development, commercialisation and adoption of AI technologies, principles for effective governance can be highly effective at promoting safe and responsible AI while enabling human innovation and creativity.

It is crucial not to overregulate so early in AI development, but to enable businesses to invest and create responsibly. Adobe generally agrees with the definitions outlined in Figure 1 of the Discussion Paper, though cautions against over-reach with the inclusion of rules-based automated decision-making systems in processes intended to govern semi- or fully autonomous computer systems.[3] The goal of AI governance is to build safe, responsible and trustworthy AI systems.

In Adobe's experience, companies, both individually and in collaboration with others, are best placed to implement these responsible AI mechanisms into business processes.

i. Adobe's Proactive Internal Responsible AI Mechanisms

As early as 2019, Adobe proactively and voluntarily developed a layered, multi-disciplinary process for responsible AI. Our process has proven to offer effective checks and balances on some of our most cutting-edge features and products before they are released publicly (Figure 1 demonstrates Adobe's AI assurance mechanisms).

First, we established AI Ethics principles of accountability, responsibility, and transparency to guide the development of our AI-powered solutions.[4]

- **Accountability** means we take ownership of the outcomes of our AI-assisted tools. We have processes and resources dedicated to receiving and responding to concerns about our AI products and taking corrective action as appropriate. Accountability also requires testing for and anticipating potential harms, taking pre-emptive steps to mitigate such harms, and maintaining systems to respond to unanticipated harmful outcomes.

- **Responsibility** means we approach designing and maintaining our AI technology with thoughtful evaluation and careful consideration of the impact and consequences of its deployment. We ensure that we design for inclusiveness and assess the impact of potentially unfair,

---

[3] Discussion Paper at q. 1.
[4] https://www.adobe.com/about-adobe/aiethics.html

discriminatory, or inaccurate results, which might perpetuate harmful biases and stereotypes. We understand that special care must be taken to address bias if a product or service will have a significant impact on an individual's life, such as with employment, housing, credit, and health.

- **Transparency** means we are open about, and explain, our use of AI to our customers so they have a clear understanding of our AI systems and their applications. We want our customers to understand how Adobe uses AI, the value AI-assisted tools bring to them, and what controls and preferences they have available when they engage with and utilize Adobe's AI-enhanced tools and services.

Guided by these principles, we have created standardised processes from design to deployment, including training, testing, and a review process overseen by a diverse AI Ethics Review Board.

Second, we have set up an AI Ethics Committee and an AI Ethics Review Board to action ethical concerns with new features and technologies. The Committee and Board are supported by an AI Ethics team that serves two main functions: (a) advises our product and development teams on ethical considerations for our AI-powered solutions; and (b) thinks expansively about ethics across various types of tools and technologies, both developed by Adobe for our customers and used internally by our teams.

- Our **AI Ethics Committee** is a diverse, cross-functional team of Adobe employees that is responsible for helping to ensure the AI Ethics principles are understood and incorporated across our development teams.

- Our **AI Ethics Review Board** is tasked with the review of potentially high-impact AI-powered features and products before their release. Like our AI Ethics Committee, the AI Ethics Review Board is a cross-functional group of Adobe employees with diverse professional, gender, and ethnic backgrounds, as we believe diversity is critical to identifying potential issues that a non-diverse team might not see.

Third, we have embedded an **AI Ethics Assessment** in our processes to help us identify the potential ethical impact of a new AI system or feature. The AI Ethics Assessment is a multi-part review designed to detect products and features that can potentially perpetuate harmful biases or stereotypes. The feature is approved if an initial assessment shows no major ethical risk and the feature meets our ethical standards. AI features with a higher potential ethical impact go through a more rigorous testing process, including a review by the AI Ethics Review Board. The AI Ethics Review Board will either approve the release, give a qualified approval with instructions to test or fix a minimal feature set or issues, or not approve due to significant testing or remediation required. Where approval is denied, the product team is required to take on additional work and re-present to the AI Ethics Review Board.

These internal responsible AI mechanisms have been effective in upholding our AI Ethics principles. For example, our AI Ethics Review Board worked with Adobe development teams on a new suite of features in Photoshop called Neural Filters which let users add non-destructive, generative filters to create things that were not previously in images using AI. Neural Filters allow for easy adjustment of someone's age or expression or can turn a black-and-white photo to colour by simply adjusting a slider. The Neural Filters development team used our AI Ethics Assessment to assess the potential ethical impact of the filters to avoid perpetuating negative biases.

When evaluating Neural Filters, one Review Board member flagged that one of the filters did not always properly model the hairstyle of a particular group. Based on this feedback, the Neural Filters development team updated the dataset to enable retraining of the AI model to address the issue before the feature went live. Another example is when a member of the Review Board flagged early in the development process that a skin-smoothing feature in Neural Filters was removing tattoos from people's faces. Tattoos are culturally significant for certain communities. By removing a person's tattoo, the AI could have been unintentionally changing someone's identity and conforming them to certain beauty norms that they may not ascribe to. Since our Review Board is comprised of a diverse group of employees all reviewing the AI from their own personal perspectives, we identified this issue early and shared this important feedback with our product and engineering teams. The teams made improvements to the model and updated the AI dataset before Neural Filters was released.

Finally, we promote **community feedback from our users**. For example, we launched our Firefly tool as a beta to seek input from our employees, customers, and our broader creative community. Firefly has a built-in feedback mechanism so users can easily report if a feature produces a result they perceive as biased or inaccurate. This feedback data is monitored to identify undesirable outcomes and allows our product and engineering teams to address them appropriately. In our view, this constant feedback loop with our user community is the best way to help ensure our tools minimise bias and uphold our values as a company.

Figure 1: Adobe's Internal AI Assurance Mechanisms
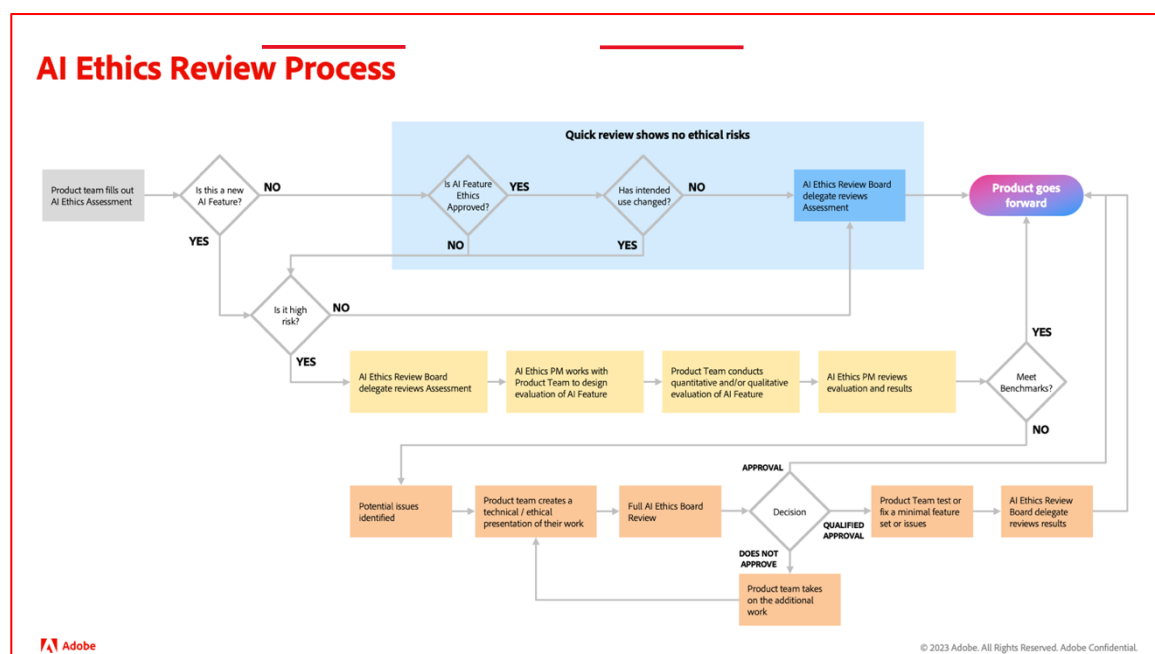
The approaches described above can be used to embed the AI Principles described in the discussion paper into existing business processes. Adobe recommends that both the public and private sectors uphold an equal level of rigour in governance regarding the use of AI technologies[5].

---

[5] Discussion Paper at q. 6.

ii.   Industry-led Transparency Mechanisms for AI Systems

Adobe believes that transparency mechanisms are most effective when made available to end-users of AI systems to improve public trust and confidence in AI.[6] Adobe agrees that "there are many examples and concerns around AI being used for potentially harmful purposes, such as: generating deepfakes…[and], creating misinformation",[7] and has worked extensively with interested parties to develop practical and effective transparency processes such as the Content Authenticity Initiative.

Adobe leads the Content Authenticity Initiative (CAI) – a global coalition working to increase transparency in digital content through open industry standards. Founded in 2019, the CAI now counts over 1,500 members from across industries including media outlets, camera manufacturers, generative AI developers, and more. The CAI promotes the use of technology to allow creators to attach metadata, called Content Credentials, to a piece of content, such as name, date, and what tools were used to create it. Content Credentials remain associated with the content wherever it is used, published, or stored and can help content viewers make more informed decisions about whether to trust it. This level of transparency enables proper attribution for creators and helps fight misinformation.

In generative AI, Content Credentials can indicate whether a piece of content was human-created, AI-edited or AI-generated. Adobe automatically attaches Content Credentials to images generated in our generative AI tool Firefly to indicate that AI was used.

Finally, Content Credentials allow creators to securely attach a "Do Not Train" tag in the metadata of their work to give creators a way to opt out of training AI datasets if they choose. Adobe supports industry adoption of this assertion to help prevent companies everywhere from using content with "Do Not Train" Content Credentials to train AI systems and solutions.

Content Credentials are a free, open-source technology based on an open standard created by the Coalition for Content Provenance and Authenticity (C2PA), for which Adobe is a founding member. Anyone can incorporate it into their tools and products, and Adobe is committed to working with our cross-industry partners to drive widespread implementation of the C2PA standard.

The CAI and C2PA showcase how industry collaborations can promote the safety, reliability, trust, accountability, and transparency of AI outputs without requiring the intervention of any third party. They are examples of creative, stakeholder-led initiatives that help to protect public trust in digital content.

Adobe strongly supports Government leveraging the C2PA standard to:

- support greater transparency and trust in our online ecosystem by ensuring creator attribution for digital content and helping users understand the origins and edits of the content they are consuming;
- enable automatic indication of AI-generated content;
- encourage industry promotion, adoption, and retention of a Do Not Train standard, and endorse this industry-developed standard for other organisations and for use by Government to leverage

---

[6] Discussion Paper at q. 9a.
[7] Discussion Paper at pg. 7.

AI responsibly

Adobe also urges the Government to incorporate content provenance and authenticity in its guidance on responsible AI as effective tools for protecting consumers and creators.[8]

B. Adobe Supports a Harmonised, Risk-based Approach Applied to the Lifecycle of the AI System

In 2021, over 130 AI-focused bills were proposed or passed in the U.S. alone.[9] Globally, the European Union, China, Brazil, Canada, and Singapore, including multilateral organizations such as the Organization for Economic Cooperation and Development (OECD), are standing up policy or laws on AI. Too much fragmentation will make it nearly impossible for global companies like Adobe to compete and comply. Adobe favours aligning AI standards globally and harmonising leading AI governance frameworks. Alignment starts with a shared lexicon.

Adobe supports a risk-based approach to AI[10] in line with emerging guidance and AI laws, including US NIST's *AI Risk Management Framework* (AI RMF). A risk-based approach empowers organisations to establish oversight over AI systems in an adaptable, fit-for-purpose manner that spurs innovation and promotes trustworthiness.[11]

Adobe believes that responsible AI mechanisms should be scoped to cover all stages of the AI lifecycle, from ideation through design, development, deployment, and ongoing monitoring. These responsible AI mechanisms, such as those outlined in U.S. NIST's AI RMF, can be used to effectively deal with systemic and/or collective risks of harm. Similar to Adobe's implementation of a risk-based approach via the responsible AI mechanisms outlined in section A of this submission, we advocate for an approach that leverages industry initiatives[12] that enables early detection and mitigation of AI-related harms by those with the most relevant expertise (i.e., the designers and deployers of the AI systems).[13] This framework should apply to public and private organisations, ensuring a consistent approach across sectors.

Adobe supports the proposed elements outlined in Attachment C[14] and has already implemented them within our business processes, as detailed in section A of this submission. Conducting impact assessments, providing training, ensuring sufficient notice, providing explanations, and ensuring adequate human oversight are all crucial steps that Adobe believes should be integrated into organisational business processes.

C. Adobe Provides Concrete Actions the Australian Government Can Take to Promote Responsible AI

Adobe acknowledges the Discussion Paper's specific request for feedback on initiatives the Government can take to build public trust in AI deployment and to encourage more people to use AI.[15] We propose three concrete actions for the Australian Government to take to promote trust in AI:

---

[8] Discussion paper at q. 9b.
[9] Center for Data Innovation, U.S. AI Policy Report Card, July 27, 2022.
[10] Discussion Paper at q. 14.
[11] Discussion Paper at q. 15.
[12] Discussion Paper at q. 20a.
[13] Discussion Paper at q. 20b.
[14] Discussion Paper at q. 17.
[15] Discussion Paper at q. 11.

First, we see that many of the draft regulations around AI development aim to mitigate harmful bias. Mitigating harmful bias is important to us at Adobe and it is the reason we established our comprehensive AI Ethics program that includes testing, training, and a review with our AI Ethics Review Board.

Curating large, globally representative test datasets to robustly evaluate an AI model is challenging but is also a key strategy to mitigating harmful bias in AI. Having Government curate and manage these test datasets could encourage more diverse contributions while ensuring transparent and accountable data governance. Government-based standardised test datasets can enable companies to run their AI models against these datasets to make sure their models meet certain performance and ethical standards. If a company's models perform within an agreed-upon "normal range" in a distribution of outcomes, that company could get a certificate of compliance. The standardised datasets would need to work across various verticals, industries, markets, and use cases. There is an opportunity for the private sector to partner with governments to establish large, globally representative and privacy compliant datasets. We think this initiative would help create an environment where companies do not have to make a tradeoff between innovation and ethics.

Second, education and awareness-building are critical for responsible AI development and for AI adoption. We support activities aimed at providing guidance and best practices for responsible AI. Adobe believes transparency is key to responsible AI, including educating users when interacting with an AI product. We encourage efforts to leverage existing industry-led initiatives such as the CAI, and its underlying standards body, the C2PA.

Third, we believe that industry must work closely with regulators and agencies to resolve discrepancies and ensure a focus on innovation. Adobe emphasises providing clear guidance and focusing on a pro-innovation approach to governing and overseeing AI. Adobe supports the establishment of a feedback loop between industry, regulators, and agencies to measure the proportionality and effectiveness of proposed guidance, oversight, and legislation.

Conclusion

Adobe appreciates the opportunity to comment on the Australian Government's Discussion Paper. We understand that the Government has taken important steps to consult with interested stakeholders to leverage existing responsible AI mechanisms, minimise barriers to AI development and adoption, and identify opportunities to harmonise such standards across Australia and the globe. We encourage using industry-developed initiatives and standards such as the CAI and C2PA as examples of best practices to maintain trust while mitigating harms caused by some AI solutions.

We look forward to continuing the discussion with the Department of Industry, Science, and Resources, acting as a hub for the broader community of interest, and are eager to engage collaboratively on these important issues.