BEFORE THE
PRODUCTIVITY COMMISSION OF AUSTRALIA

Harnessing Data and Digital Technology (Interim Report)

12 September 2025

The Copyright Alliance appreciates the opportunity to submit comments to the Productivity Commission ("PC") of the Australian Government in response to the Inquiry on the Harnessing Data and Digital Technology: Interim Report ("Inquiry"), specifically concerning the proposal for a copyright law exception for text and data mining (the "Proposed TDM Exception").[1]

The Copyright Alliance is a non-profit, non-partisan public interest and educational organization representing the copyright interests of over 2 million individual creators and over 15,000 organizations in the United States, across the spectrum of copyright disciplines.[2] We are dedicated to advocating policies that promote and preserve the value of copyright, and to protecting the rights of creators and innovators who rely on copyright law to protect their creativity, efforts, and investments in the creation and distribution of copyrighted works for the public to enjoy.

---

[1] PRODUCTIVITY COMMISSION, *Harnessing Data and Digital Technology: Interim Report* (Aug. 5, 2025) https://www.pc.gov.au/inquiries/current/data-digital/interim.

[2] The Copyright Alliance is the unified voice of the copyright community. A full list of Copyright Alliance organizational members is available online. *See Who We Represent*, COPYRIGHT ALLIANCE, https://copyrightalliance.org/about/who-we-represent/ (last visited Sept. 12, 2025). Positions taken here represent the views of the Copyright Alliance and may not reflect the specific views of any individual or organization.

The Copyright Alliance and our members support the responsible, ethical, and respectful development and use of AI technologies. The continuing development of artificial intelligence ("AI") systems brings many opportunities. Many in the creative industries are already using or plan to use AI-based technologies to assist in the creation of a wide range of works that benefit society. Some—like the motion picture, video game, and music industries—have been using AI-based assistive tools for many years. Others—like many independent illustrators and authors—are now incorporating AI tools into their work processes.

But the development and deployment of AI systems also bring many challenges, especially related to copyright. As Australian government policymakers decide how best to address these challenges, it is essential that copyright law and the underlying incentives to encourage investment in creative works not be cast aside in favor of new policies obligating creators to effectively subsidize AI technologies under the misguided belief that doing so is necessary to incentivize AI technologies.

It is well known that copyright fosters and drives technological innovation when it is respected, upheld, and enforced. Copyright is not an obstacle to AI innovation—in fact, ensuring continued investment in and respect for copyright works is a means of sustaining and accelerating innovation of AI system. That is why it is essential that policymakers do not take any action that devalues the rights and interests of creators and copyright owners or undermines copyright protections.

Copyright law is what empowers independent to large-scale creators and rights holders to create inspiring, innovative, and pioneering works—works that are crucial for generative AI development. These works not only benefit the public, but they also drive and benefit AI development and use, positively contributing to the economy and employment. A 2020 report showed that the copyright industries in Australia contributed $124.1 billion to the economy, or 6.8% of Australia's GDP and resulted in 1.03 million jobs.[3] The continued vibrancy of the copyright industries, the productivity stemming from these industries, and their ability to contribute to the Australian economy depend on robust copyright laws.

[3] PRICE-WATERHOUSE COOPER, *The Economic Contribution of Australia's Copyright Industries – 2006-2018* (June 2020), https://www.copyright.org.au/pwc2020.

Like certain other countries, the Australian government is considering whether legal or policy changes are necessary or appropriate to foster AI innovation. In the Inquiry, the PC is considering whether a text-and-data mining ("TDM") exception should be specified as a fair dealing exception in Australian copyright law. While fostering AI innovation is a worthy goal, obtaining that goal should not be accomplished at the expense of Australia's and other countries' creative sectors nor should it undermine the intellectual property laws that support them. To ensure that policy goals are achieved in a sustainable, win-win manner, it is vital that any revisions to Australia's law respect intellectual property—and in particular copyright.

Unfortunately, the Proposed TDM Exception fails to meet this goal and seems to myopically disregard the value and virtues of strong and effective copyright laws. Like many others, the Copyright Alliance therefore strongly urges the PC and the Australian Government to not adopt the Proposed TDM Exception. Creators and rights holders around the globe are concerned about the negative impacts the Proposed TDM Exception would have on creativity and the creative community. Such an exception would undermine the rights of creators and copyright owners and run counter to other countries' approaches to AI.

Meanwhile, the licensing markets for copyrighted works and licensing markets for AI use of copyrighted works, continue to flourish.[4] Grounded in and supported by copyright law, these free-market license agreements have enabled Generative AI ("GAI") developers to use rich, large, and quality datasets for AI training, fine-tuning, and grounding for their commercial AI products. There are more AI companies and developers reaching licensing agreements with rights holders all the time—agreements which are providing clarity and certainty for more AI companies and rights holders. AI copyright licensing has resulted in productive and valuable partnerships and synergies between the AI and creative industries. *Adopting the Proposed TDM Exception would harm these developing markets and undermine fundamental rights of creators and copyright owners around the globe.*

It is vital to ensure that human creators (i) continue to have sufficient protections for their works used in training, fine-tuning, and grounding such systems, (ii) are remunerated for their works

[4] *See* COPYRIGHT ALLIANCE, *AI Licensing for Creative Works*, https://copyrightalliance.org/artificial-intelligence-copyright/licensing/ (last visited Sep. 12, 2025).

that are exploited to build and operationalize AI models, and (iii) remain incentivized to continue creating, which in turn will lead to greater volume and diversity of works that can be licensed for use as GAI training materials. The Proposed TDM Exception fails on each of these grounds.

Copyrighted Works Drive Development of AI Technologies

We urge the PC and the Australian Government to not adopt the Proposed TDM Exception. A TDM exception would undermine both existing and developing licensing markets for the use of copyrighted works for AI training purposes, harming the continued productivity of the creative industries and their ability to contribute to Australia's economy and society.

Copyright law is not a barrier for the use and development of GAI technologies. Quite the opposite—copyright law enhances and fuels the development of GAI technologies. The reason that the creative community, from independent to large-scale creators and rights holders, is able to create quality works—and the reason they are in demand for training—is that strong copyright laws incentivize and reward creation of high-value works. And it is these high-quality, expressive works (that often require significant investment by the creator) that are ideal to train AI machines to generate high-quality output.

Many creators and rights holders, particularly publishers and image/media licensors, already license their copyrighted works for commercial AI uses and many more are on the cusp of doing so. Just a few public examples of licensing solutions, initiatives, and agreements for AI use of copyrighted works include those launched from or created by Created by Humans, Dataset Providers Alliance, Copyright Clearance Center, Elsevier, Getty Images, Bertlesmann, Kobalt, Shutterstock, Jstor, Sage Journals, Rightsify, Universal Music Group, and other major media publishers including The New York Times, News Corp, Associated Press, The Atlantic, Vox Media, Dotdash Meredith, Financial Times, Fortune, Time, Entrepreneur, The Texas Tribune, WordPress.com, and countless other rights holders.[5] Through a variety of licensing deals of all sizes and forms, copyright owners not only provide high-quality copyrighted works for better AI training and development, but also make these expressive works useful for various AI-based use

[5] *See id.*

cases, including scientific research, through semantic enrichment, metadata tagging, content normalization, and data cleanup. The Proposed TDM Exception would undermine all of that.

Not only would these robust AI licensing markets be harmed, but a GAI system trained on copyrighted works can generate outputs that also displace and dilute the markets for the very copyrighted works on which the GAI system is trained. The Proposed TDM Exception would obligate human creators and rights holders—who are rarely consulted for approval or compensated for their works—to subsidize AI developers. The proposal would thus devalue the copyrights of all creators and rights holders, hindering development of additional quality works that benefit AI developers and the public in the first place. A categorical exemption for an infringing use where the markets for the underlying, ingested creative works are being harmed by the outputs would be inherently unfair and damaging. Therefore, it is crucial that such uses not be granted legal presumptions or favor.

Existing Limitations on TDM Exceptions Do Not Resolve Rights Holders' Concerns

The PC mentions some approaches regarding TDM exceptions taken by a handful of foreign governments, including the European Union, which implements a few conditions on use of its exceptions. We briefly address at least two of these conditions to illustrate why technology-specific legal exemptions like the Proposed TDM Exception do not work:

> *Lawful Access:* One kind of condition in the few preexisting foreign TDM exceptions requires that TDM users must have "lawful access" to the copyrighted works. This is not an adequate safeguard. Under copyright law, rights holders may determine the circumstances under which their works are made available to others. Rights holders provide legal access to creative works through licensing terms that specify allowable uses of the works. Under a license, a licensee may be granted legal access for certain uses but not others. If the licensee makes use of a work that is not within the scope of a license, that is considered to be an unauthorized use. The term "lawful access" is misleading in that it suggests that lawful access confers lawful use, which is incorrect.
>
> The mere fact that copyrighted works can be *accessed* from lawful services or otherwise made publicly available on the internet does not also mean that such works can be *used*

for the training of AI models without authorization from the rights holder. By way of example, a variety of digital platforms and services make creative works accessible or available through personal and enterprise subscriptions, through memberships, and other business models. A user could claim to "lawfully access" creative works at a personal subscription level and then turn around and use those works for TDM purposes for a commercial-scale enterprise. But in most cases under copyright law, this would violate the terms of use governing the personal subscription. This undermines the workability of the lawful access requirement and the traditional flexibility and nuance offered by copyright law to rights holders to create tailored licensing models. It will not only fail to bolster rights holders' ability to manage their works, but it will also further restrict access to copyrighted works for the public by incentivizing copyright owners to sequester creative works behind paywalls.

Even where a rights holder has "opted out" (a framework that is deficient for a variety of reasons described below), a "lawful access" requirement may still cause problems. For example, when pirated copies of copyrighted works are uploaded to publicly accessible websites or social media platforms or lawfully accessed copies are illegally uploaded to publicly accessible websites or social media platforms, those copies can be and are scraped by AI companies for AI training purposes. An unintended consequence of the requirement will be that rights holders will be forced to increasingly restrict access by placing copyrighted works behind a paywall. This would considerably reduce the availability of high-quality content to the public, which is not in anyone's interest.

Opt-Out: Another kind of condition found in certain TDM regimes states that the exception would not apply if the copyright owner has expressly reserved their rights (i.e., an opt-out requirement). An opt-out condition subverts the foundational rule of copyright law that a copyright owner has the right to choose whether to authorize others to use their work. Copyright owners should not be required to take affirmative action to signal others to refrain from using their works.

As an opt-in regime, copyright law recognizes that the large universe of potential users of copyrighted works, such as AI companies, are in a far better position to determine which works they want to use and approach the copyright owners to secure authorizations to do

so, rather than require a copyright owner to monitor the universe of potential users, especially in the age of a ubiquitous internet. Copyright licensing agreements best facilitate these interactions. The marketplace should continue to properly value and incentivize creativity, and AI policy should not interfere with the right or ability of copyright owners to license, or choose not to license, their works for AI purposes. A copyright owner should be free to decide whether they want to license their work to an AI company (or anyone else).

Importantly, an opt-out regime also conditions rights holders' copyrights on effectuating the opt-out, which, if not properly executed, would strip them of their ability to enjoy and enforce their rights. This requirement would violate Article 5 of the Berne Convention, which states that copyrights are to be enjoyed and exercised by authors without being subject to any formality.[6]

In particular, independent creators and artists are at a disadvantage when it comes to using technological solutions and monitoring for theft of their works. Most do not have the resources or technical expertise to regularly monitor for theft of their works or to take technical and other steps to prevent piracy. Expecting them to now also monitor for and affirmatively prevent AI ingestion of their works would burden them even more and divert them from creating new works for the public to enjoy. This is particularly true for creative sectors where works are widely available across many different platforms and distribution channels. Even larger rights holders will struggle to ensure that opt-outs are exercised in relation to their works across all those platforms and channels.

Opt-out schemes that already exist are ineffective for various reasons. For example: (i) the copyrighted works might already have been copied and used for training at the time of opt-out; and (ii) despite opting out, copies of the copyrighted works may still be included in the datasets through other means, such as when copies are scraped from other sources such as a licensee of the copyright owner or from a third-party platform where a copy has been posted.

[6] Berne Convention for the Protection of Literary and Artistic Works, art. 5(2), Sept. 9, 1886, as revised at Stockholm July 14, 1967, 828 U.N.T.S. 221, 231-32.

The practical effects of opt-out, particularly with regards to works already used to train AI, are also negligible given that it is challenging to remove entire works at scale from an AI model that has already been trained on the works. Existing tools that might assist with opt-out also have significant limitations, especially since they are often not designed to be targeted to address scraping for GAI ingestion.

Moreover, copies of works that are available on pirate sites are outside the copyright owner's control. Allowing broad-scale web scraping means the work will end up in a training dataset even if the copyright owner has opted out. Current opt-out regimes in the EU are being exploited by AI developers, who continue to train on scraped content despite legitimate efforts from copyright owners to opt out. So, in sum, there can be no doubt that both the concept and application of opt-outs do not protect the legitimate interests of rights holders.

U.S. and International Approaches to TDM Exceptions

For many years now, lawmakers and policymakers in many countries, including the United States, have been carefully examining the intersection of copyright law and AI.[7] Most countries, including the United States, have not enacted any TDM exceptions or new exceptions to copyright law for AI purposes. And for good reason: AI licensing markets are robust and legislation may cause significant harms to such markets. In fact, the U.S. Copyright Office in a report examining AI ingestion, training, and licensing issues, concluded that no changes to U.S. copyright law were warranted regarding AI use of copyrighted works.[8] The Office further noted

[7] U.S. federal agencies that have examined the issues include the National Institute of Standards and Technology (NIST) and the U.S. Patent and Trademark Office (USPTO). *See* Study to Advance a More Productive Tech Economy, 86 Fed. Reg. 66287 (Nov. 22, 2021), https://www.regulations.gov/document/NIST-2021-0007-0001; Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation, 84 Fed. Reg. 58141 (Oct. 30, 2019), https://www.regulations.gov/document/PTO-C-2019-0038-0001. Most recently, the U.S. Copyright Office issued a report on generative AI training, ingestion, and licensing issues. *See generally* U.S. COPYRIGHT OFFICE, COPYRIGHT AND ARTIFICIAL INTELLIGENCE PART 3: GENERATIVE AI TRAINING: PRE-PUBLICATION VERSION, 107 (May 6, 2025), https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf [hereinafter "USCO Report"].

[8] *See* USCO Report *supra* note 7 at 107.

that "[e]ffective licensing options can ensure that innovation continues to advance without undermining intellectual property rights."[9]

Nor does the U.S. have a categorical TDM exception as some have claimed. U.S. copyright law requires courts to engage in a fact-specific and case-by-case analysis using an established four-factor fair use test. U.S. law forbids categorically excusing otherwise infringing uses of copyrighted works. As such, any AI use is not categorically exempted or given legal preferences and must be carefully analyzed by applying the facts in any given case to the four fair use factors.

The non-categorical nature of fair use law is apparent when examining a few of the early U.S. court decisions in the over fifty lawsuits filed on AI training. Since the filing of the initial lawsuits, novel factors continue to be considered by courts in the fair use analysis like the rapid rise of new AI applications including retrieval-augmented generation (RAG), the discovery that pirated copies of creative works from illicit websites were used to train AI, and the explosive development of the AI licensing market which will influence the most important fourth fair use factor analyzing market harm. U.S. fair use law allows for evolving precedent that enables beneficial and equitable uses of copyrighted works to move forward in ways compatible with the goals of copyright, while avoiding harmful and inequitable uses. A broad categorical copyright exception would be unable to keep up with the fact-specific and rapid development of AI use cases and the technology itself.

As the Australian Government is aware, there are jurisdictions in which copyright exceptions have been legislated with regard to TDM use, including Singapore, Japan, and the EU. It is essential to understand that, from a global perspective, a TDM exception is a minority approach and that many of these exceptions preceded the launch of generative AI models and did not contemplate many of the core issues presented by generative AI training, including the market displacement and dilution effect of generative AI output which replicates or is substantially similar to ingested copyrighted works.[10]

[9] *Id*.

[10] *See generally id*. at 76-78.

The UK has a limited exception but a proposal to further expand that TDM exception failed to gain any traction and was pulled by the UK Government in 2023 over grave concerns that it would significantly devalue copyright and severely harm rights holders.[11] In 2024, the UK Government once again raised the possibility of a TDM exception, which was met with a massive uproar from creators globally as well as prominent UK artists who rightfully questioned why their government would place their interests below those of foreign tech companies. The continued proposals from the UK Government to undermine copyright and the fair value for use of creative works have sparked incredible backlash from nearly every industry within the UK creative and news sectors.[12]

Indeed, evidence demonstrates the existence of strong and vibrant AI licensing markets, soundly refuting the need for legal changes that favor AI developers at the expense of rights holders and creators. The desire to lead technological innovation does not justify compromising the copyright foundations that aid such innovation in the first place and also have other overwhelming public and economic benefits to consumers, creators, and nations across the globe. That is why there is unified opposition over copyright exceptions and proposals that clearly undermine the creative sector.[13] Without strong copyright laws that incentivize and protect the creation and dissemination of copyrighted works, there cannot be trustworthy, reliable, and ethical AI technologies.

For a technology where the development and evolution has proven to be quicker than the ability to enact new law and regulations, a blanket legal exception that is technology-specific risks becoming quickly irrelevant. Copyright law, on the other hand, provides

[11] On 1, February 2023, George Freeman—Minister for Science, Research, and Innovation—announced that the United Kingdom government will take a step back from its original proposal put forth in the summer of 2022 for a broad exception for the text-and-data mining of copyrighted works for any purpose in the UK's copyright laws. The Minister noted that the government will further consult stakeholders in the coming months to "ensure that we do not rush precipitately into a knee-jerk move that is wrong." *See* HC Deb (Feb. 1, 2023) (727) cols. 152-68WH (UK), https://hansard.parliament.uk/commons/2023-02-01/debates/7CD1D4F9-7805-4CF0-9698-E28ECEFB7177/ArtificialIntelligenceIntellectualPropertyRights.

[12] *See* Rachel Kim, *A Global Phenomenon: The Creative Community's Viral Outrage Against AI Theft* (Mar. 6, 2025), https://copyrightalliance.org/outrage-against-ai-theft/.

[13] *See* Eileen Bramlet, *The Global Creative Community Stands Unified Against Unchecked AI Use* (Mar. 27, 2025), https://copyrightalliance.org/creative-community-against-unchecked-artificial-intelligence/.

flexibility for rights holders and AI developers to respond to new marketplace and technological developments and to incentivize the further creation of creative works and license arrangements to advance innovation.

The Proposed TDM Exception would stunt growth in the AI licensing market and copyright owners' rights. Undermining the ability of creators to develop, manage, protect, and benefit from their works in favor of a new technology would harm the Australian creative industries and the PC's goals or promoting innovation and economic growth.

For the reasons cited above, we urge the PC to withdraw the Proposed TDM Exception and the Australian Government to not adopt it.

Respectfully Submitted,

Keith Kupferschmid

CEO
Copyright Alliance
1331 F Street, NW, Suite 950
Washington, D.C., 20004