THE UNIVERSITY OF MELBOURNE

Centre for Artificial Intelligence and Digital Ethics

CAIDE AI Policy Forums

FORUM #1 ISSUES PAPER: Demystifying Generative AI, Intellectual Property and the Creative Industries

About the Author

Calvin Collins is a Research Assistant at the Centre for Artificial Intelligence and Digital Ethics (CAIDE) at the University of Melbourne. Calvin completed his Bachelor of Arts and his Juris Doctor at Melbourne Law School, before becoming a paralegal, and eventually a solicitor, at leading global law firm Herbert Smith Freehills. At CAIDE, Calvin pairs his legal knowledge and experience with an interest in emerging technology regulation. Calvin would like to acknowledge the contributions of Professor Jeannie Marie Paterson to the editing of this paper.

What is CAIDE?

The Centre for Artificial Intelligence and Digital Ethics (CAIDE) is a cross-disciplinary research centre at the University of Melbourne. CAIDE facilitates cross-disciplinary research, teaching and leadership on the ethical, technical, regulatory and legal issues relating to AI and digital technologies. CAIDE is directed by Professor Jeannie Marie Paterson from Melbourne Law School. For more information about CAIDE, see our website: https://www.unimelb.edu.au/caide.

Acknowledgements and Sponsorships

The AI Policy forums are supported by:

- The Centre for AI and Digital Ethics, funded by the Faculty of Engineering and Information Technology and Melbourne Law School at the University of Melbourne
- The Ninian Stephen Law Program powered by the Menzies Foundation, as part of the project *New Legal Thinking for Emerging Technologies*
- Microsoft, Atlassian and the Tech Council of Australia as part of the project *Demystifying Generative AI*.

CAIDE would also like to specially acknowledge Terri Janke and Company Lawyers and Consultants Neane Carter and Claire McKenzie for help in preparing the material on Indigenous cultural and intellectual property.

CAIDE AI Policy Forums

This issues paper was prepared as reading for the CAIDE AI Policy Forum #1, held at the University of Melbourne in 2024. It should be read in conjunction with the discussion paper on Generative AI, Intellectual Property and the Creative Industries, available here. The aim of the AI Policy Forums is to provide an opportunity for discussing policy and law issues raised by the emergence of AI that go beyond the headlines. This event considered the impact of AI, particularly Generative AI, on intellectual property and the law.

Contact CAIDE

Email: uom-caide@unimelb.edu.au.
Website: https://www.unimelb.edu.au/caide.
LinkedIn: https://au.linkedin.com/company/caide-unimelb.

AI Forum #1- Demystifying Generative AI & IP

The advent of multimodal large language models ('LLMs') — which can be used in applications spanning healthcare, material science, accessibility and sustainability as well as offering an expedient avenue for producing content — has raised unprecedented questions about the meaning of original creation.[1] AI-generated content has already been utilised across a range of creative industries, including by Marvel Studios,[2] by the BBC,[3] in art competitions,[4] and by musicians Drake,[5] Grimes[6] and Paul McCartney.[7] Yet LLMs have also become a focus point for concerns about data protection,[8] control over creative output,[9] the misuse of Indigenous Cultural and Intellectual Property ('ICIP'),[10] and the future of work in the creative industries.[11]

Those in the creative industries have been particularly outspoken about the risks of copyright infringement posed by generative AI, as well as the wider social and moral consequences of abdicating creativity to AI. For example, the Writers Guild of America and the Screen Actors Guild-American Federation of Television and Radio Artists participated in a months-long strike in 2023, seeking (among other things) to restrict the use of AI in reproducing actors' likenesses and generating scripts.[12] The discovery that thousands of Australian books were allegedly pirated in the Book3 dataset used to train Meta, Bloomberg and EleutherAI's LLMs prompted outrage from publishers and authors alike, with author Richard Flanagan describing it as 'the biggest act of copyright theft in history'.[13] A recent survey conducted by the Australian organisation Media, Entertainment & Arts Alliance ('MEAA') suggests that up to 75% of creative professionals are concerned about the potential for theft of intellectual or creative work using AI technologies.[14]

1. The challenge: AI's generative process

Currently, LLMs' generative process is of such complexity that it is essentially impossible to unpick or to pinpoint a given piece of copyrighted data that a model has relied upon to produce a given output.[15] Nor does the mere existence of copyrighted data in a model's training data indicate that the model has somehow 'borrowed' from the data. In fact, LLMs generally do not have access to their training data at the time of generation. Rather, as part of the training process, models use data to identify the commonalities in the expression of language, as opposed to deriving the meanings or intentions contained in the data. Data is broken down into a multitude of 'tokens', tiny fragments of language attached to biases and indicators of statistical significance that allow the model to predict when tokens should be paired together. At the time of generation, generative AI models only have access to these tokens and the associated probabilities.

Accordingly, even if you had full access to a model's training data and a given output, you could not draw a definitive causal link between the two. Those seeking to protect their copyrighted works or style (or to seek redress for a perceived infringement) must target either:

1. the input stage, namely by restricting the use of copyrighted data in the material used to train generative AI models or by alleging that an AI entity did not have the necessary access rights to use the copyrighted data; or
2. at the output levels, by suggesting that a given output infringes upon existing copyright on the basis that it:
    a. seemingly 'regurgitates' a copyrighted work; or
    b. appears to 'borrow' copyrighted styles or iconography.

2A. The issues: The use of copyrighted training data

It has become a truism that data is the 'new oil', though in many ways it outpaces oil in its economic potential. Data is an excludable, non-rivalrous good: 'excludable' in the sense that those who control datasets can readily mediate access to them; and 'non rivalrous', as an infinite number of people can simultaneously benefit from the same data pool. The potential value of data has skyrocketed with the take-off of generative AI. As noted by Fred Havemeyer of the financial research firm Macquarie, 'A.I. begins and ends with data.'[16] LLMs require huge volumes of training data, frequently 'scraped' from across the internet. Already, technology companies are raising the alarm that the internet may contain insufficient material to continue training their AI models.[17]

The issues presented by the compilation and use of datasets to train LLMs are manifold and, at times, contradictory. As ever-data hungry models continue to grow, it is inevitable that copyrighted material will increasingly be sought to train models. Meta has gone as far as to propose purchasing publishing giant Simon & Schuster to obtain swathes of older, written material for training its AI model.[18] There has been understandable backlash from creatives that this data mining

appears to be a commercial enterprise utilising creative output without due compensation for the creators of the output. There are also fears that the use of copyrighted data in training sets increases the likelihood for generative AI output to mimic existing works or to otherwise borrow liberally from existing styles or iconography (discussed further below). Conversely, many in the technology sector see data mining as an essential cog in AI innovation that, in the absence of laws insulating the practice from legal challenge, may be stymied by copyright law.

It remains an open question as to whether the use copyright-protected content in training data infringes copyright and the answer will vary between jurisdictions.[19] There has been widespread concerns about the risks of copyright infringement posed by generative AI. Yet on one view, the use of training data does not involve copying and therefore does not infringe copyright.[20] Alternatively, it may fall within the exception to copyright protections provided by 'fair use' doctrines, at least in some jurisdictions.[21] For example, America's 'fair use' protections may encompass training datasets, allowing them to be used for training LLMs.[22] The EU and the UK both offer exemptions from copyright infringement for data mining, or automated text and data analysis, for the purpose of non-commercial research.[23] In both these jurisdictions, however, web scraping is subject to data protection laws and the individual rights that such laws are intended to safeguard.[24] Australia's protections for data usage under 'fair dealing' are narrow and must fall within one of a few legally specified categories: temporary reproductions, research or study, criticism or review, parody or satire, or news reporting.[25]

2B. The issues: Output replicating existing work

Generative AI is already infamous for a phenomenon dubbed 'regurgitation': generative AI output that appears to simply replicate its training data, often producing output that would appear to contain copyrighted material. This is also known as the 'Italian plumber problem', a label drawn from models' propensity to generate an image of Nintendo's Mario when asked to generate an image of an Italian plumber.

Some of the most prominent debates about generative AI and IP concern cases of seeming replication of copyrighted work by generative AI. In the US, this issue is the subject of a New York Times lawsuit against Microsoft and OpenAI for copyright infringement.[26] The New York Times' evidence included an exhibit containing one hundred examples of material generated by OpenAI's GPT-4 that contains near-identical text to New York Times articles, and that allegedly permits users to effectively bypass the newspaper's paywall.[27] On the other hand, Open AI describes the examples cited by the New York Times as a rare bug that resulted in inadvertent regurgitation of 'years-old articles that have proliferated on multiple third-party websites' and involves misuse that is 'not typical or allowed user activity'.[28] In the UK, Getty Images has claimed in a lawsuit against Stability AI that the presence of similar watermarks to their own in AI-generated images is evidence of copyright infringement.[29] There have been no such law suits filed in Australia to date.[30]

2C. The issues: Output mimicking existing styles or iconography

Even where generative AI outputs do not substantially replicate or regurgitate existing work, there has been widespread concern about copying likeness or style. Some claim that the AI generation process merely mirrors the creative process that humans undertake: producing output that may be inspired by any number of previous works but is nonetheless 'original'. Others have warned that 'it may be difficult to draw a sufficient causal link between any given output and a specific input' given the vast data pool that LLMs are trained on.[31] Yet the generation of outputs by generative AI that resembles existing work continues to raise at least normative if not legal concerns.

2D. The issues: Indigenous Cultural and Intellectual Property

Generative AI's potential to facilitate the ready proliferation of derivative works poses an acute threat to Indigenous Cultural and Intellectual Property (**ICIP**). The notion of ICIP has been crystalised in art 31(1) of the United Nations Declaration on the Rights of Indigenous People:[3233]

> *Indigenous peoples have the right to maintain, control, protect and develop their cultural heritage, traditional knowledge and traditional cultural expressions, as well as the manifestations of their sciences, technologies and cultures, including human and genetic resources, seeds, medicines, knowledge of the properties of fauna and flora, oral traditions, literatures, designs, sports and traditional games and visual and performing arts. They also*

have the right to maintain, control, protect and develop their intellectual property over such cultural heritage, traditional knowledge, and traditional cultural expressions.

The ambit of ICIP is vast and may include traditional knowledge (eg knowledge of land and stars), cultural property (eg visual arts and crafts, weapons), practice of cultural expressions (eg cultural burning practices) ancestral remains (including DNA and genomes), language (both oral and written) and literary or artistic works. While the rights arising from ICIP are not fixed, Terri Janke and Company's *Our Culture: Out Future - Report on Australian Indigenous Cultural and Intellectual Property Rights* suggests that these rights may or should include the:

1. Requirement for free, prior and informed consent for the access or use of ICIP material, in addition to the ability to refuse to offer such consent.
2. Recognition of Indigenous peoples and communities as the primary guardians and interpreters of their culture.
3. Right to benefit commercially from authorised use of ICIP material, which may include monetary or non-monetary benefits, as dictated by the relevant ICIP holders.
4. Prevention of derogatory use of ICIP material.
5. Ability to maintain secrecy in relation to ICIP material.

There are gaps in the protections Western copyright law offers ICIP. In particular, copyright law differs from ICIP in its lack of recognition of communal rights, its time-limited protections, material expression and attribution to a known author.

While generative AI remains nascent, its potential to abrogate ICIP rights is already clear. The ability for multimodal models to facilitate the production of inauthentic, 'Indigenous-style' work has already been realised. It does not appear that any guardrails have been developed to reign in the appropriation of protected styles and storytelling, allowing bad-faith actors to circumvent Indigenous communities and profiteer off ICIP (a practice already rife in the art industry without the assistance of technology[34]). Nor can generative AI appreciate the sensitivities and diversity of Indigenous communities, and the production of works off broad, uninformed prompts risks producing works capable of inflicting harm. The misuse of ICIP by generative AI can cause such harms as distress to community and risk breaching the fiduciary duties and customary law of ICIP holders if certain knowledge or designs are incorporated.

Advocates have raised a need for 'Indigenous data sovereignty', defined by the Maiam Nayri Wingara Indigenous Data Sovereignty Collective as the right of Indigenous people to exercise ownership over Indigenous Data.[35] Emma Fitch, Clare McKenzie, Terri Janke and Adam Shul have argued that while AI systems can 'help maintain, contribute to, and preserve cultural knowledge', AI raises considerable risks for ICIP. They argue that AI could negatively impact ICIP rights by:

- giving no attribution to the Traditional Owners who hold the relevant knowledge and language;
- providing false or misleading information;
- disregarding cultural protocols that affect how art styles are used, or who can share certain information (for example, who is permitted to tell or illustrate a story, using gender-specific artwork and designs);
- creating works that misappropriate ICIP without undergoing consultation; and
- leaving ICIP created using AI with no legal protection in jurisdictions where AI-generated works are not covered under copyright law requirements.[36]

Responding to these concerns requires as a baseline commitment insertion of "Indigenous sovereignty and self-determination practices" at the core of data practices.[37] Inclusive design practices that aid Indigenous people in the development of AI systems is also important in preventing the misuse of ICIP. [38] See further Angie Abdilla et al, *Out of the Black Box: Indigenous Protocols for AI* [39]

3A. Possible solutions: Existing copyright mechanisms

The traditional recourse for artists and creators seeking to protect their works and styles is copyright law. While there have already been claims that copyright law is ill-equipped to grapple with the complexities presented by generative AI, it is arguably too early to determine whether the jurisprudence can adequately adapt. The court cases in the US will impact on other jurisdictions notwithstanding their statutory differences. Reform is not always a necessary response to the rise of a new technological medium. Copyright law dates back to the Copyright Act 1790 (also known as the 'Statute

of Anne') and has weathered the arrival of the printing press and the internet alike. Copyright law is and should remain technology neutral; technological advancement will always outpace legislative reform so bespoke drafting risks legislation becoming ineffective and even stifling.

Nonetheless, whether existing mechanisms will be sufficient to combat derivative or replicative works remains untested. To the extent that amendments to the current regime are required, any proposals are likely to be subject to intense scrutiny. Previous attempts have seen particularly fierce backlash from the creative and technology sectors, both of which are comparatively well resourced in Australia. As a result, copyright reform remains frozen, despite perennial pushes to modernise the regime. For example, in December 2013, the ALRC published its *Final Report into Copyright and the Digital Economy*,[40] which included recommendations relating to exemptions for data mining. To date, these recommendations remain unactioned.

A recent survey by the Cth Attorney General's Copyright and AI Reference Group Expert Group found survey participants had different views on the need for reform.[41]

3B. Possible solutions: Data access and protection

Data rights remain underdeveloped in Australia. Unlike in the United States, data does not constitute property in Australia, complicating its excludability and individuals' right to seek redress against unauthorised use of their data.

The manner in which data protection and access regimes are to be regulated will involve delicate balancing acts:

1. To invest in Australian digital industries and artforms, Australia may need to adopt some form of data mining exemption that inoculates generative AI from copyright challenges. Conversely, guardrails are required that will ensure creatives are duly compensated for the use of their works. For example, in *Who Owns the* Future, Jaron Lanier envisages a reworked internet where there is a stronger economic link between end-market products and the data underlying it, essentially creating an economy of micropayments that compensates people for original material they post to the web.
2. A similar tension arises with respect to transparency and privacy. We want technology companies to be transparent with their use of data, to enable to us to unpick concerns relating to data theft, copyright and model biases. However, depending on the nature of the data, we want stringent privacy regulations to ensure the protection of sensitive information and public confidence in the use of AI, particularly if generative AI proliferates in sectors such as health.

Already, explicitly direct and protective options for protecting training data have been proposed, and alternately lauded for protecting artists[42] and criticised for stifling innovation.[43] The Productivity Commission has pointed out that even existing uncertainty creates costs for those seeking to use the technology.[44] In the EU, the 2024 AI Act mandates that AI providers maintain copyright policies that identify and comply with opt-out decisions, and make publicly available summaries of content used for training.[45] In the US, the Content Origin Protection and Integrity from Edited and Deepfaked Media Bill (or 'COPIED Bill') would, if passed, require the provision and maintenance of provenance information alongside synthetic (i.e. AI-generated) content.[46] AI developers such as Open AI provide tools and controls for rightsholders who make their content publicly available online to opt-out of their content being used for AI training.[47] Other protections over how data can be used by LLMs include compensation or licensing fees paid to copyright owners. Pasquale and Sun argue that 'We thus also propose a levy on AI providers, to be distributed to the copyright owners whose work they use without a license'. [48]

3C. Possible solutions: Novel protections in the age of AI

Academics, policy makers and workers in creative industries have been involved in workshopping possible solutions to the complications raised by the intersection of generative AI, creative output and intellectual property rights. Hacohen and Elkin-Koen have suggested combining legal formulations familiar from trademark or copyright law with new technical solutions, like using generative AI systems themselves to quantitatively measure the genericity of any given piece of intellectual property.[49] They argue that 'GenAI may facilitate the development of new and more accurate measures to assess the originality of these works'.[50] The 'capacity to measure genericity and quantify originality against

the knowledge captured by a model could enable copyright law to more accurately distinguish between original works and those in the public domain. It may then insert more nuance into copyright analysis in deciding copyright scope.'[51]

Other solutions seek to dramatically expand the reach of existing intellectual property regulation, including through the introduction of new protections and property rights for voice and likeness. This approach has gained particular traction in the US, where it was passed into law as the ELVIS Act in Tennessee,[52] and is currently being debated as the federal No AI FRAUD Act[53], backed by major creative industry lobbying groups including the Human Artistry Campaign.[54] It is not without controversy, with its detractors raising concerns about its potential to encroach on parody and satire, and allow transferable property rights to be controlled by industry bodies rather than individual artists.[55] Here, a concept of moral rights and powers of control or veto may be closer to the normative concerns raised by copying likeness or voice. As noted above, disclosure of synthetic content will also be important as a general public protection.

The Australian Senate has established a Select Committee on Adopting Artificial Intelligence to inquire into and report on issues including copyright concerns and the impacts of generative AI on creative industries.[56] In December 2023, Commonwealth Attorney-General Mark Dreyfus announced the establishment of a copyright and AI reference group, specifically to consider 'material used to train AI models, transparency of inputs and outputs, the use of AI to create imitative works, and whether and when AI-generated works should receive copyright protection.' This group published its first report in late 2024, noting a need to investigate the full range of issues including opportunities for innovation, protections for artists and creators' moral and performers' rights, and indigenous cultural and intellectual property rights in the age of AI. [57]

4. Conclusion

There are no easy answers to these conundrums. They will only be found (if at all) at the intersection of technology, law and policy. The policy choice involves not only 'weighing the advantages of reducing the cost of content creation and the value of expertise against the potential risk to various careers and sectors of the economy',[58] but also considering the intangible societal benefits from the vast variety of AI models that rely on copyrighted content in their training data. Any resolution favoured on policy grounds must also be technically feasible. Furthermore, any decision on these matters requires a reflection on the value we place on the work of artists and those in the creative industries, along with respect for culturally significant work.

---

[1] For a comprehensive discussion of generative AI and copyright issues in the US see Katherine Lee and A. Feder Cooper and James Grimmelmann, 'Talkin' 'Bout AI Generation: Copyright and the Generative-AI Supply Chain' (2024) <https://arxiv.org/abs/2309.08133>.

[2] Adrian Horton, 'Marvel Faces Backlash Over AI-Generated Opening Credits', *The Guardian* (online, 22 June 2023) <https://www.theguardian.com/tv-and-radio/2023/jun/21/marvel-ai-generated-credits-backlash>.

[3] Max Goldbart, 'BBC Will Stop Using AI For "Doctor Who" Promotion After Receiving Complaints', *Deadline* (online, 25 March 2024) <https://deadline.com/2024/03/bbc-doctor-who-ai-complaints-1235867333/>.

[4] Kevin Roose, 'An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy', *The New York Times* (online, 2 September 2022) <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>.

[5] Maya Chung, 'Why Drake Had to Take Down His Song That Featured AI-Tupac Vocals', *Time* (online, 26 April 2024) <https://time.com/6971720/drake-tupac-ai/>.

[6] Jazz Monroe, 'Grimes Unveils Software to Mimic Her Voice, Offering 50-50 Royalties for Commercial Use', *Pitchfork* (online, 2 May 2022) <https://pitchfork.com/news/grimes-unveils-software-to-mimic-her-voice-and-announces-2-new-songs/>.

[7] Michael Sun, 'Paul McCartney Says There's Nothing Artificial in New Beatles Song Made Using AI', *The Guardian* (online, 23 June 2023) <https://www.theguardian.com/music/2023/jun/23/paul-mccartney-says-theres-nothing-artificial-in-new-beatles-song-made-using-ai>.

[8] See eg, European Data Protection Supervisor, *First EDPS Orientations for EUIs using Generative AI* (Guideline, 3 June 2024) <https://www.edps.europa.eu/data-protection/our-work/publications/guidelines/2024-06-03-first-edps-orientations-euis-using-generative-ai_en>.

[9] Sam Buckingham Jones, 'AI "Final Nail in Coffin" of Australia's Creative Sector', *Australian Financial Review* (online, 16 July 2024) <https://www.afr.com/companies/media-and-marketing/ai-final-nail-in-coffin-of-australia-s-creative-sector-20240716-p5ju1u>.

[10] Emma Fitch et al, 'The New Frontier: Artificial Intelligence, Copyright and Indigenous Culture', *Terri Janke and Company Lawyers & Consultants* (Blog Post, 30 November 2023). <https://www.terrijanke.com.au/post/the-new-frontier-artificial-intelligence-copyright-and-indigenous-culture>.

[11] Productivity Commission, *Making the most of the AI opportunity: AI raises the stakes for data policy*, (Research paper, no. 3, Canberra, 2024), 12.

[12] Gene Maddaus, 'AI Companies Praise SAG-AFTRA and WGA Contracts: "They've Done a Great Job"', *Variety* (online, 29 November 2023) <https://variety.com/2023/biz/news/ai-companies-balanced-approach-wga-sag-aftra-1235813309/> .

[13] Kelly Burke, '"Biggest Act of Copyright Theft in History"': Thousands of Australian Books Allegedly Used to Train AI Model', *The Guardian* (online, 29 September 2023) <https://www.theguardian.com/australia-news/2023/sep/28/australian-books-training-ai-books3-stolen-pirated>.

[14] 'Stop AI Theft: Media, Creative and Arts Workers Demand Action from Government' (Media Release, Media, Entertainment & Arts Alliance ('MEAA'), 16 July 2024) <https://www.meaa.org/mediaroom/stop-ai-theft-media-creative-and-arts-workers-demand-action-from-government/>.

[15] Anthropic is currently working on whether the generation process can be reverse engineered to produce explanations for a given output. See Steven Lvy, 'AI Is a Black Box. Anthropic Figured Out a Way to Look Inside', *Wired* (online, 3 October 2024) <https://www.wired.com/story/anthropic-black-box-ai-research-neurons-features>.

[16] J Edward Moreno, 'Boom in A.I. Prompts a Test of Copyright Law', *The New York Times* (online, 30 December 2023) <https://www.nytimes.com/2023/12/30/business/media/copyright-law-ai-media.html>.

[17] Deepa Seetharaman , 'For Data-Guzzling AI Companies, the Internet Is Too Small', *The Wall Street Journal* (online, 1 April 2024) <https://www.wsj.com/tech/ai/ai-training-data-synthetic-openai-anthropic-9230f8d8>.

[18] Ella Creamer, 'Meta "Discussed Buying Publisher Simon & Schuster to Train AI"', *The Guardian* (online, 10 April 2024) <https://www.theguardian.com/books/2024/apr/09/meta-discussed-buying-publisher-simon-schuster-to-train-ai>.

[19] See generally, See generally, Luke Hawthorne et al, 'It Copies, Right? Generative AI & Copyright Law', *Insight – King & Wood Mallesons* (Blog Post, 11 May 2023) <https://www.kwm.com/au/en/insights/latest-thinking/generative-ai-and-copyright-law.html>.

[20] See Mark A Lemley, 'How Generative AI Turns Copyright Law Upside Down' (2024) 25(2) *The Columbia Science and Technology Law Review* 190, 191, 205.

[21] Mark A Lemley and Bryan Casey, 'Fair Learning' (2021) 99(4) *Texas Law Review* 743, 763–701.

[22] Ibid 770.

[23] *Copyright, Designs and Patents Act 1988* (UK) s 29A; *Parliament and Council Directive (EU) 2019/790 of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC* [2019] OJ L 130/92, art 4(1).

[24] See, eg, European Data Protection Supervisor, n 7.

[25] *Copyright Act 1968* (Cth) ss 40–42, 43A and 43B.

[26] *The New York Times Co v Microsoft Corporation* (SD NY, No 1:23-cv-11195, 2024).

[27] *The New York Times Co v Microsoft Corporation* (SD NY, No 1:23-cv-11195, 2024), Exhibit J, Doc 1-68 <https://nytco-assets.nytimes.com/2023/12/Lawsuit-Document-dkt-1-68-Ex-J.pdf>.

[28] Open AI, 'OpenAI and journalism', (Blog Post, 8 January 2024) <https://openai.com/index/openai-and-journalism/>.

[29] James Vincent, 'Getty Images is Suing the Creators of AI Art Tool Stable Diffusion for Scraping its Content', *The Verge* (online, 17 January 2023) <https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit>.

[30] See, eg, Hawthorne et al, n 21.

[31] Aaron Hayward et al, 'The IP in AI: Can IP Rights Protect AI Systems?', *Insight — Herbert Smith Freehills* (Blog Post, 23 March 2023) <https://www.herbertsmithfreehills.com/insights/2023-03/the-ip-in-ai-can-ip-rights-protect-ai-systems>; Kevin Roose, 'A.I.'s Black Boxes Just Got a Little Less Mysterious', *The New York Times* (online, 21 May 2024) <https://www.nytimes.com/2024/05/21/technology/ai-language-models-anthropic.html>.

[32] United Nations, *United Nations Declaration on the Rights of Indigenous People* (online, 3 October 2024) <https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf>.

[33] https://www.un.org/development/desa/indigenouspeoples/wp-content/uploads/sites/19/2018/11/UNDRIP_E_web.pdf

[34] Emilia Terzon, 'The Missing Royalties', *ABC News* (online, 27 March 2024) <https://www.abc.net.au/news/2024-03-27/auction-houses-sell-aboriginal-art-no-royalties/103542626>.

[35] Cam Wilson, 'AI is Producing 'Fake' Indigenous Art Trained on Real Artists' Work Without Permission', *Crikey* (online, 19 January 2024) <https://www.crikey.com.au/2024/01/19/artificial-intelligence-fake-indigenous-art-stock-images/ >; 'Indigenous Data Sovereignty Everywhere... All the Time', *Maiam Nayri Wingara* (Web Page) <https://www.maiamnayriwingara.org/>.

[36] Fitch et al, n 9.

[37] Emma Fitch et al, 'The new frontier: Artificial Intelligence, copyright and Indigenous Culture', *Terri Janke and Company* (online, 30 November 2023). <https://www.terrijanke.com.au/post/the-new-frontier-artificial-intelligence-copyright-and-indigenous-culture>.

[38] Emma Fitch et al, 'The new frontier: Artificial Intelligence, copyright and Indigenous Culture', *Terri Janke and Company* (online, 30 November 2023). <https://www.terrijanke.com.au/post/the-new-frontier-artificial-intelligence-copyright-and-indigenous-culture>; Bronwyn Carlson and Peita Richards, 'Indigenous Knowledges Informing "Machine Learning" Could Prevent Stolen Art and Other Culturally Unsafe AI Practices', *The Conversation* (8 September 2023) <https://theconversation.com/indigenous-knowledges-informing-machine-learning-could-prevent-stolen-art-and-other-culturally-unsafe-ai-practices-210625>.

[39] < https://www.anat.org.au/wp-content/uploads/2021/11/Out-of-the-Black-Box_Indigenous-protocols-for-AI.pdf/>.

[40] Australian Law Reform Commission, *Copyright and the Digital Economy* (Final Report, November 2013) <www.alrc.gov.au/wp-content/uploads/2019/08/final_report_alrc_122_2nd_december_2013_.pdf>.

[41] Cth Attorney General Copyright and AI Reference Group Expert Group, *Copyright Materials as AI Input* (2024) < https://www.ag.gov.au/rights-and-protections/copyright/copyright-and-artificial-intelligence-reference-group-cairg>.

[42] Frank Pasquale & Haochen Sun, 'Consent and Compensation: Resolving Generative AI's Copyright Crisis' (2024) 110 *University of Virginia Law Review Online*.

[43] Productivity Commission, *Making the most of the AI opportunity: AI raises the stakes for data policy*, (Research paper, no. 3, Canberra, 2024).

[44] Ibid, 10.

[45] *Parliament and Council Regulation 2024/1689 of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)* [2019] OJ L, art 53(1).

[46] Cathy Virginia, 'The AI-Focused COPIED Act Would Make Removing Digital Watermarks Illegal', *The Verge* (online, 12 July 2024) <https://www.theverge.com/2024/7/11/24196769/copied-act-cantwell-blackburn-heinrich-ai-journalists-artists>.

[47] Open AI, 'OpenAI and journalism', (Blog Post, 8 January 2024) <https://openai.com/index/openai-and-journalism/>. See further 'Overview of OpenAI Crawlers', <https://platform.openai.com/docs/bots>.

[48] Pasquale and Sun, n 28.

[49] Uri Y Hacohen & Niva Elkin-Koren, 'Copyright Regenerated: Harnessing GenAI to Measure Originality and Copyright Scope' (2024) 37(2) *Harvard Journal of Law & Technology*.

[50] Ibid 9.

[51] Ibid 10.

[52] *Ensuring Likeness Image and Voice Security (ELVIS) Act*, Tenn Code Ann §§ 47-25-1101 – 47-25-1107(c), 39-14-152(g) (2024). <https://publications.tnsosfiles.com/acts/113/pub/pc0588.pdf>.

[53] *No Artificial Intelligence Fake Replicas and Unauthorized Duplications Act*, HR 6943, 118th Congress (2024) <https://www.congress.gov/bill/118th-congress/house-bill/6943/text?s=1&r=9>.

[54]The Office of Congresswoman María Elvira Salazar, Representative for Florida's 27th District, 'Hundreds of Artists, Actors, and Creators Back Salazar's No AI FRAUD Act' (Media Release, 2 February 2024) <https://salazar.house.gov/media/press-releases/hundreds-artists-actors-and-creators-back-salazars-no-ai-fraud-act>.

[55] See *Artificial Intelligence and Intellectual Property: Part II – Identity in the Age of AI, Hearing before the Subcomm. on Courts, Intellectual Property, and the Internet of the H. Comm. On the Judiciary*, 118th Congress 27-45 (2024) (statement of Jennifer E Rothman, University of Pennsylvania Law School) <https://www.congress.gov/118/chrg/CHRG-118hhrg54845/CHRG-118hhrg54845.pdf>; Joseph Fishman, 'Could Elvis Impersonators and Tribute Bands be Sued in Tennessee? Possibly Under AI Bill', *The Tennessean* (online, 1 March 2024) <https://www.tennessean.com/story/opinion/contributors/2024/03/01/tennessee-artificial-intelligence-bill-too-broad-elvis-impersonators-tribute-bands/72773778007/>.

[56] 'Select Committee on Adopting Artificial Intelligence — Terms of Reference', *Parliament of Australia* (Web Page) <https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Adopting_Artificial_Intelligence_AI/AdoptingAI/Terms_of_Reference>. To date, the Select Committee has received submissions from a significant number of representative bodies across the creative industries, including the MEAA, the National Aboriginal and Torres Strait Islander Music Office, the Australian Guild of Screen Composers, the Australian Production Design Guild, the Australian Copyright Council, the Australasian Performing Right Association, the Australian Society of Authors, the Australian Publishers Association, the National Association for the Visual Arts and the Australian Association Of Voice Actors.

[57] Cth Attorney General Copyright and AI Reference Group Expert Group, *Copyright Materials as AI Input* (2024) < https://www.ag.gov.au/rights-and-protections/copyright/copyright-and-artificial-intelligence-reference-group-cairg>.

[58] Simon Chesterman, 'Good Models Borrow, Great Models Steal: Intellectual Property Rights and Generative AI' (2024) *Policy and Society* 1, 1.

THE UNIVERSITY OF
MELBOURNE

Centre for AI and Digital Ethics

Level 8
Melbourne Connect
700 Swanston Street
Carlton 3053

unimelb.edu.au/caide