



## King's Research Portal

### *Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

### *Citation for published version (APA):*

Lee, H.-K., Bertolini, J., Terui, T., & Kawashima, N. (2024). *Copyright and generative AI*.

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

## **Sustainable Cultural Futures**

'Digitalisation of Culture' | Case Study 2

# **Copyright and Generative AI**

**September 2024**

Josepha Bertolini (King's College London), Hye-Kyung Lee (King's College London), Takao Terui (Xi'an Jiaotong Liverpool University, China) and Nobuko Kawashima (Doshisha University, Kyoto)



# Welcome to Sustainable Cultural Futures

Led by King's College London (UK) and Doshisha University (Japan), **Sustainable Cultural Futures: COVID-19 and Resetting Cultural Policy (SCF)** takes a mid-to long-term perspective to reconsider pre-pandemic assumptions, explore new frontiers for cultural policy studies and build a more sustainable future for the arts and culture. To achieve this, we focus on three themes: 1) Cultural value and public engagement; 2) Culture work; and 3) Digitalisation of culture. The research activities in the UK are supported by the Economic and Social Research Council via its Fund for International Collaboration [Grant Ref: ES/W011891/1]. Those in Japan are funded by the Japan Society for the Promotion of Science [Grant No. JPJSJRP 20211707].

This is our second AI report (Theme 3). You can access our first AI report here in [English](#) and [Japanese](#).



# JSPS

# Contents

<b>1. Introduction</b>	<b>3</b>
<b>2. Emerging concerns</b>	<b>4</b>
2.1. Copying existing creative works	
2.2. Affecting future market demand	
2.3. Monopolistically accumulating human knowledge and creative expressions	
2.4. Reducing cultural creation to data production	
<b>3. International perspectives</b>	<b>7</b>
3.1. Japan	
3.2. United Kingdom	
3.3. Singapore	
3.4. Australia	
3.5. Canada	
<b>4. Reconciling AI with copyright</b>	<b>11</b>
4.1. The 4Cs: consent, control, credit and compensation	
4.2. Licensing agreements and beyond	
4.3. Critical relevance of open access culture	
4.4. The wider ecology	
<b>5. Appendices</b>	<b>15</b>
Appendix 1 Sampling of generative AI-related lawsuits	
Appendix 2 Sampling of generative AI-related licensing agreements	

# Copyright and Generative AI

Josepha Bertolini (King’s College London), Hye-Kyung Lee (King’s College London), Takao Terui (Xi’an Jiaotong Liverpool University, China) and Nobuko Kawashima (Doshisha University, Kyoto)[1]

## 1. Introduction

Developing generative AI requires vast datasets of human knowledge and creative expressions.[2] With data quality also being a critical factor,[3] copyrighted materials have become a desirable target for AI data mining and training, for as the global tech leader, OpenAI, admits: “it would be impossible to train today’s leading AI models without using copyrighted materials.”[4] Current generative AI models disrupt copyright dynamics in terms of their inputs (data mining and training) and outputs (generative text, images, music, and videos).[5] [6] [7] Furthermore, this disruption can be more broadly considered within the context of rising “**creative precarity,**” that is “**the increasing uncertainty in terms of cultural workers’ creative roles, rights and identity, and audience’s perception of their creativity and labour.**”[8] Resultantly, artists and cultural industry leaders have raised various concerns, sometimes likening the data mining of original works to creative “theft.”[9] [10]

---

[1] Josepha Bertolini (Research Assistant) reviewed English-language sources, researched legal cases and licensing agreements and drafted this report with the guidance of Hye-Kyung Lee (PI) who structured the report and provided the conceptual framework. Takao Terui (the project’s critical friend) reviewed Japanese and further English-language sources, contributing to the report. Nobuko Kawashima (the project’s Japanese PI) advised on overarching arguments of the report.

[2] McKinsey & Company (2024) ‘What is generative AI?’, *McKinsey & Company*, 2 Apr. Available at: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai> [Accessed 16 May 2024].

[3] Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., Naumann, F. and Harmouch, H. (2022) *The Effects of Data Quality on Machine Learning Performance*. Available at: <https://arxiv.org/abs/2207.14529> [Accessed 16 May 2024].

[4] OpenAI, quoted in Popular Science (2024) ‘OpenAI argues it is ‘impossible’ to train ChatGPT without copyrighted work’, *Popular Science*, 8 Jan. Available at: <https://www.popsci.com/technology/openai-copyright-fair-use/> [Accessed 14 May 2024].

[5] Henderson, P., Li, X., Jurafsky, D., Hashimoto, T., Lemley, M.A. and Liang, P. (2023) *Foundation Models and Fair Use*. Stanford Law and Economics Olin Working Paper No. 584. Available at: <https://dx.doi.org/10.2139/ssrn.4404340> [Accessed 4 Jul. 2024].

[6] Guadamuz, A. (2024) A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs, *GRUR International*, Volume 73, Issue 2, February 2024, Pages 111–127, <https://doi.org/10.1093/grurint/ikad140> [Accessed 1 Aug. 2024].

[7] Pasquale, F. and Sun, H. (2024) *Consent and Compensation: Resolving Generative AI’s Copyright Crisis*. Cornell Legal Studies Research Paper Forthcoming, University of Hong Kong Faculty of Law Research Paper No. 2024/07. Available at SSRN: <https://ssrn.com/abstract=4826695> [Accessed 5 Aug. 2024].

[8] Lee, H.-K. (2024) Reflecting on cultural labour in the time of AI. *Media, Culture & Society*. <https://journals.sagepub.com/doi/full/10.1177/01634437241254320> [Accessed 19 May 2024].

[9] The Economist (2024) ‘Generative AI is a marvel. Is it also built on theft?’ *The Economist*, 14 Apr. Available at: <https://www.economist.com/business/2024/04/14/generative-ai-is-a-marvel-is-it-also-built-on-theft> [Accessed 14 Jun. 2024].

[10] All-Party Parliamentary Group on Music and UK Music (2024) *Artificial Intelligence and the Music Industry - Master or Servant?* Available at: <https://www.ukmusic.org/wp-content/uploads/2024/04/APPG-AI-Report-Low-res.pdf> [Accessed 14 Jun. 2024].

Although some commentators analogise data mining to artists being inspired by existing cultural expressions,[11] their voices are yet to gain traction in the current AI copyright debate. As such, the introduction of generative AI has unsettled the copyright landscape, witnessing a rising tide of legal action against tech firms.

To better understand the implications of this new, unsettled terrain, we will examine key aspects of the copyright infringement claims made by artists and cultural companies against AI companies, review policy responses in five different countries and, lastly, consider how generative AI might be reconciled within a copyright framework.

## 2. Emerging concerns

Based on the ongoing court cases and debates, we can identify the following four problematic aspects of AI's data mining of copyrighted works and its generative outputs (also see Appendix 1).

### 2.1. Copying existing creative works

As evident through the growing number of generative AI-related lawsuits,[12] artists and cultural companies are worried about the dual copyright threat that consists of the data mining of their original works without consent, credit or remuneration, as well as AI's generation of outputs that resemble those original works. First, of the generative-AI-related lawsuits we reviewed, one consistent claim is that of copyright infringement by tech firms who used plaintiffs' original works to train their models. [13] [14] [15] [16] [17] [18]

---

[11] Creative Commons (2023) *Open Letter: Artists Using Generative AI Demand Seat at Table from US Congress*. Available at: <https://creativecommons.org/about/policy-advocacy-copyright-reform/open-letter-artists-%20using-generative-ai-demand-seat-at-table-from-us-congress/> [Accessed 14 Jun. 2024].

[12] There are currently estimated to be about 20 such lawsuits. Aplin, T. (2024) *Steering through technological and market disruption: the copyright perspective on Generative AI*. Sustainable Cultural Futures International Workshop on Digitalisation of Culture. King's College London, 6 Jun., London.

[13] *Getty Images vs Stability AI* (2023) Case 1:99-mc-09999. Available at: <https://aboutblaw.com/6DW> [Accessed 16 May 2024], pp 23-34.

[14] *Andersen, McKernan, Ortiz, Southworth, Rutkowski, Manchess, Brom, Zhang, Kaye, and Ellis, vs Stability AI, Midjourney, Runway & DeviantArt* (2023) Case 3:23-cv-00201-WHO. Available at: <https://fingfx.thomsonreuters.com/gfx/legaldocs/znpnzrgyzpl/AI%20COPYRIGHT%20LAWSUIT%20amended.pdf> [Accessed 16 May 2024], pp 8-10.

[15] *Authors Guilds vs Open AI* (2023) Case 1:23-cv-08292 Available at: <https://fingfx.thomsonreuters.com/gfx/legaldocs/xmjlbqbnvr/AUTHORS%20GUILD%20OPENAI%20LAWSUIT.pdf> [Accessed 16 May 2024], pp 44-46.

[16] *Concord Music Group vs Anthropic* (2023) Case 3:23-cv-01092. Available at: <https://regmedia.co.uk/2023/10/20/anthropic-lyrics-complaint.pdf> [Accessed 16 May 2024], pp 49-57.

[17] *New York Times vs Open AI* (2023) Case 1:23-cv-1119. Available at: [https://nytco-assets.nytimes.com/2023/12/NYT\\_Complaint\\_Dec2023.pdf](https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf) [Accessed 16 May 2024], pp 60-67.

[18] *Universal Music Group vs Suno* (2024) Case 1:24-cv-11611. Available at: <https://storage.courtlistener.com/recap/gov.uscourts.mad.272063/gov.uscourts.mad.272063.1.0.pdf> [Accessed 7 Aug. 2024].

Currently, AI companies argue that such data mining occurs within the domain of fair use.[19] [20] However, their use of copyrighted materials reaches far beyond the traditional scope of copyright exceptions for fair use or fair dealing in terms of quantity and speed. Second, there are claims that some generative outputs are “substantially similar” to the original works used to build models.[21] As such, in multiple court filings, some plaintiffs aim to demonstrate the capacity of AI models to mimic or near-replicate plaintiffs’ creative works through generative outputs.[22]

Resultantly, commentators and interested parties are debating whether such outputs can be treated as derivatives of the copyrighted works used for training generative models and thus potential copyright infringements or if outputs can be considered a transformative use of the original work[23] and thus exempt from copyright under the aegis fair use.[24] [25] [26] [27] While whether AI inputs (data mining and training) and outputs (generative works) consist of copyright infringement is a complex legal, technological and artistic question, it is clear that artists and companies alike are increasingly anxious about the loss of control over their copyrighted works.

## 2.2. Affecting future market demand

Current lawsuits are also concerned with the potential impact of AI outputs on the broader cultural market. For example, it is noted that “the most important factor by far is the effect of the use on the potential market for or value of the

---

[19] The Economist (2024) ‘AI firms will soon exhaust most of the internet’s data’ *The Economist*, 23 Jul. Available at: <https://www.economist.com/schools-brief/2024/07/23/ai-firms-will-soon-exhaust-most-of-the-internets-data> [Accessed 31 Jul. 2024].

[20] Wolfson, S. (2023) ‘Fair Use: Training Generative AI’, *Creative Commons*, 17 Feb. Available at: <https://creativecommons.org/2023/02/17/fair-use-training-generative-ai/> [Accessed 16 May 2024].

[21] *Andersen et al. vs Stability AI et al.* (2023).

[22] See *Andersen et al. vs Stability AI et al.* (2023); *Concord Music Group vs Anthropic* (2023); *New York Times vs Open AI* (2023).

[23] In the 90s, “transformativeness” was added by the US Supreme Court to fair use understanding: “If someone can show that their secondary use transforms the original in some way, it is much more likely to be fair use than otherwise.” (Wolfson, 2023).

[24] Schick, N. and Ajder, H. (2023) ‘AI on trial: Can everything generated by AI be subject to copyright?’, *The Era of Generative AI*, 23 Jan. Available at: <https://ninaschick.substack.com/p/ai-on-trial-caneverything-generated> [Accessed 28 Jun. 2023].

[25] Appel, G., Neelbaur, J. and Schweidel, D.A. (2023) ‘Generative AI Has an Intellectual Property Problem’, *Harvard Business Review*, 7 Apr. Available at: <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem> [Accessed 7 Aug. 2024].

[26] Helms, S. and Krieser, J. of McDermott, Will and Emery (2023) ‘Copyright Chaos: Legal Implications of Generative AI’, *Bloomberg Law*, 23 Mar. Available at: <https://www.bloomberglaw.com/external/document/XDDQ1PNK000000/copyrights-professional-perspective-copyright-chaos-legal-implic> [Accessed 1 Aug. 2024].

[27] Wolfson (2023).

copyrighted work.”[28] Thus, the possibility that AI outputs may compete with the existing copyrighted works that were used for AI training or displace their markets, further complicates the ongoing debate.[29] Such outputs are argued to create unfair competition for cultural producers wherein they must compete against automated versions trained on their own works.[30] As generative AI models are accessible and notably productive, it has been suggested that their outputs might offer an enticing alternative to the time and expense incurred in paying for human-based services.[31]

### **2.3. Monopolistically accumulating human knowledge and creative expressions**

In more general terms, we should also note global tech firms’ accumulation of human knowledge and creative expressions at a vast scale. As so much data is necessary and costly to train and process,[32] [33] only a small number of firms have managed to train what are known as “foundation models”[34] such as the Large Language Model used for ChatGPT.[35] This potentially causes a near monopoly of data spanning the depth and breadth of human knowledge and creativity, distilling it into a few sources that are seeking to capitalise on their investments. Within this dynamic, investigations led by the US Federal Trade Commission (FTC) have commenced following mounting concerns over “monopolistic behavior in the rapidly advancing industry of chatbots and other generative AI products.”[36] Furthermore, we can more broadly contextualise this against the backdrop of creative precarity by noting how human creativity, knowledge and skills, which have thus far been “embodied” in human artists and creators, are now separable from their labour and could function as if they are capital assets of those tech firms.[37]

---

[28] Carey, T. of Sunstein LLP (2024) ‘The New York Times v. OpenAI: The Biggest IP Case Ever’, *JD Supra*, 11 Jan. Available at: <https://www.jdsupra.com/legalnews/the-new-york-times-v-openai-the-biggest-5149037/> [Accessed 16 May 2024].

[29] Henderson et al. (2023).

[30] *New York Times vs Open AI* (2023), pp 65-66.

[31] Downing, quoted in Observer (2023) ‘Will A.I. Replace Artists? Some Art Insiders Think So’, *Observer*, 21 Jun. Available at: <https://observer.com/2023/06/will-a-i-replace-artists-some-art-insiders-think-so/> [Accessed 18 Apr. 2024].

[32] McKinsey & Company (2024).

[33] The Atlantic, (2024) ‘Silicon Valley’s Trillion-dollar Leap of Faith’, *The Atlantic*, 29 Jun. Available at: <https://www.theatlantic.com/technology/archive/2024/07/ai-companies-unprofitable/679278/> [Accessed 5 Aug. 2024].

[34] “Foundation models are machine learning models trained on broad data (typically scraped from the internet) generally using self-supervision at scale (Bommasani et al., 2021). Most foundation models are not trained to accomplish specific tasks but rather to capture useful general information in the data...These models can then be tuned to align more with human preferences (Ouyang et al., 2022) or be adapted for specific tasks. Foundation models can be used for generating content.” Cited by Henderson et al. (2023), p 3.

[35] Toner, H. (2023) *What are Generative AI, Large Language Models, and Foundation Models?* Center for Security & Emerging Technology, Georgetown University. Available at: <https://cset.georgetown.edu/article/what-are-generative-ai-large-language-models-and-foundation-models/> [Accessed 27 Aug. 2024].

[36] Associated Press (2024) ‘US antitrust enforcers will investigate leading AI companies Microsoft, Nvidia and OpenAI’, *AP News*, 6 Jun. Available at: <https://apnews.com/article/nvidia-openai-microsoft-ai-antitrust-investigation-ftc-doj-0adc9a4a30d4b581a4f07894473ba548> [Accessed 17 Jun. 2024].

[37] Lee (2024).



## 2.4. Reducing cultural creation to data production

Such an AI-driven cultural ecology may result in a reductionist view of human creativity and cultural production, wherein artistic works and creative expressions are equated with data.[38] Markets for cultural content have always existed. However, the mass accumulation of human creative capacity and cultural expressions by tech companies has, until this point, been unprecedented. This “datafication of culture”[39] may lead to a protectionist dynamic within the cultural landscape. For example, creatives looking to safeguard their work from data mining may choose not to share their creations publicly,[40] whilst major tech firms may continue to strike licensing agreements to further empower their generative models.[41] [42] Furthermore, it has also been argued that the tech industry’s current narrative of open access to creative works will shift once they have accumulated their necessary data and sufficiently trained their models, later pivoting to stricter IP enforcement.[43]

## 3. International perspectives

As the current AI-copyright debate is focused on litigation in the US district courts and therefore remains US-centric, more diverse perspectives are warranted.

### 3.1. Japan

Currently, Japanese copyright law allows broad access to use copyrighted content for data mining and training AI models without the authorisation of copyright holders. This dynamic has resulted in Japan being dubbed a “Machine Learning Paradise.”[44] Noticing the shifting technological and economic conditions surrounding copyright holders due to generative AI, the Agency for Cultural Affairs (ACA) recently published an official report regarding its approach

---

[38] Coyle, D. and Manley, A. (2022) *What is the value of data? A review of empirical methods*. Bennett Institute for Public Policy Cambridge. Available at: [https://www.bennettinstitute.cam.ac.uk/wp-content/uploads/2022/07/policy-brief\\_what-is-the-value-of-data.pdf](https://www.bennettinstitute.cam.ac.uk/wp-content/uploads/2022/07/policy-brief_what-is-the-value-of-data.pdf) [Accessed 17 May 2024].

[39] Lee (2024).

[40] Pasquale and Sun (2024), pp 19-20.

[41] Financial Times (2024) ‘The Financial Times and OpenAI strike content licensing deal’, *Financial Times*, 29 Apr. Available at: <https://www.ft.com/content/33328743-ba3b-470f-a2e3-f41c3a366613> [Accessed 17 May 2024].

[42] SAG-AFTRA (2024) *SAG-AFTRA and Replica Studios Introduce Groundbreaking AI Voice Agreement at CES*. Available at: <https://www.sagaftra.org/sag-aftra-and-replica-studios-introduce-groundbreaking-ai-voice-agreement-ces> [Accessed 17 Apr. 2024].

[43] Regulation Panel (2023) *Balancing Act: AI, Intellectual Property, and Legal Landscapes*. Protecting Creativity in the Age of AI. King’s College London, 14 Dec., London.

[44] In 2019, new copyright law was legislated, including the new Article 30-4 claiming that: “It is permissible to exploit a work, (...) in any other case in which it is not a person’s purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work.” Copyright Research and Information Center (n.d.) *Copyright Law of Japan*. Available at: <https://www.cric.or.jp/english/clj/cl2.html#chapter2sect3> [Accessed 5 Aug. 2024].

to AI and copyright.[45] In preparing the report, the ACA consulted with creative industry workers and representative organisations to gain insight into their experiences and opinions of AI's impact.[46] It also opened an office tasked with recording artists' concerns about generative AI, such as imitation of artistic style or unique characters.[47] Despite such consultation, the ACA has maintained that there is not an adequate level of infringement and thus no need to amend existing copyright law.[48] Ultimately, its report reiterates existing AI and copyright regulation, whilst clarifying key definitions and the legal scope for the use of copyrighted works for data mining. The report argues that data mining should not be considered a copyright violation, with such practices remaining permissive without the consent of copyright holders.

### 3.2. United Kingdom

The UK government's approach to AI and copyright remains influx. With its wish to make the UK “an AI superpower” and to encourage development in domestic AI, [49] the government initially planned to amend the Copyright, Designs and Patents Act 1988 to permit data mining for “any purpose,” [50] [51] rather than for strictly non-commercial ends.[52] However, opposition from cultural sector representatives and policymakers – such as MPs serving on the Culture, Media & Sport Committee – has voiced concerns regarding the government’s proposal to exempt data mining from copyright protection.[53] Aiming to “balance between innovation and creator rights,”[54] the UK government’s approach has remained hesitant to regulate, anticipating tech companies and creative industry

---

[45] ACA (2024) *Approach to AI and Copyright: draft*. Available at: [https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05\\_07/pdf/94024201\\_01.pdf](https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_07/pdf/94024201_01.pdf) and abbreviated English version: [https://www.bunka.go.jp/english/policy/copyright/pdf/94055801\\_01.pdf](https://www.bunka.go.jp/english/policy/copyright/pdf/94055801_01.pdf) [Accessed 5 Aug. 2024].

[46] ACA (2023) *Representative Opinions of Creators and Copyright Holders on the Generative AI*. Available at: [https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05\\_05/pdf/93980701\\_03.pdf](https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_05/pdf/93980701_03.pdf) [Accessed 5 Aug. 2024].

[47] Yomiuri Shimbun (2024) ‘The ACA started to collect generative AI copyright infringement case’, *Yomiuri Shimbun*, 13 Mar. Available at: <https://www.yomiuri.co.jp/culture/20240312-OYT1T50242/> [Accessed 5 Aug. 2024].

[48] ACA (2024) *General Understandings on AI and Copyright in Japan*. Available at: [https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/bunkakai/69/pdf/94022801\\_01.pdf](https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/bunkakai/69/pdf/94022801_01.pdf) [Accessed 5 Aug. 2024].

[49] DSIT (2023) *A pro-innovation approach to AI regulation*. Available at: <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper> [Accessed 5 Aug. 2024].

[50] IPO (2022) *Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation*. Available at: <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation> [Accessed 27 Aug. 2024], para 86.

[51] Smith, J., Thornton, P. and Shaw, A. of Hogan Lovells (2024) ‘Government AI and copyright strategy: a change in artistic license?’ *Practical Law*, 29 Feb. Available at: [https://uk.practicallaw.thomsonreuters.com/w-042-4259?transitionType=Default&contextData=\(sc.Default\)&firstPage=true](https://uk.practicallaw.thomsonreuters.com/w-042-4259?transitionType=Default&contextData=(sc.Default)&firstPage=true) [Accessed 4 Jul. 2024].

[52] Ibid.

[53] House of Commons, Culture, Media & Sport Committee (2024) *Connected tech: AI and creative technology: Government Response to the Committee’s Eleventh Report of Session 2022–23*. Available at: <https://committees.parliament.uk/publications/42766/documents/212749/default/> [Accessed 4 Jul. 2024].

[54] UK Music (2023) *UK Music Chief Welcomes MPs’ Report and Calls for Strong Copyright Protections To Be At Heart Of UK’s Approach To Artificial Intelligence*. Available at: <https://www.ukmusic.org/news/uk-music-chief-welcomes-mps-report-and-calls-for-strong-copyright-protections-to-be-at-heart-of-uks-approach-to-artificial-intelligence/> [Accessed 4 Jul. 2023].

organisations to take the lead on a voluntary code of practice to be facilitated by the IPO.[55] Talks, however, stalled in February 2024 – with no code of conduct produced.[56] [57] As a result, the government aims to take a more active role in shaping AI copyright policy through talks "with the AI and rights holder sectors, seeking to ensure the workability and effectiveness of an approach that allows the AI and creative sectors to grow together in partnership." [58] Depending on the progress of various talks and consultations involving the government, the introduction of new policy appears to remain a potential outcome.[59] [60]

### 3.3. Singapore

The Singaporean government has responded to new practices surrounding digital content use and access with the Copyright Act of 2021. The act was precipitated by controversies surrounding copyright infringement and the legal premise of personal use becoming problematised by data mining practices. Most significantly, this law permits the use of copyrighted works for computational data analysis that include: "(a) using a computer program to identify, extract and the analysis of information or data from the work or recording; and (b) using the work or recording as an example of a type of information or data to improve the function of a computer program about that type of information or data". [61] This provision allows tech firms to use accessible data to develop generative AI services. The Singaporean government's effort to encourage AI innovation and boost domestic AI resonates with this approach.

### 3.4. Australia

In Australia, artists and cultural companies have raised concerns over the risks associated with copyright exemptions for data mining and called for better regulation. In August of 2023, the Media, Entertainment and Arts Alliance (MEAA) called on the government to cease copyright law exemptions in training

---

[55] HM Government (2023) *HM Government Response to Sir Patrick Vallance's Pro-Innovation Regulation of Technologies Review Digital Technologies*. Available at: [https://assets.publishing.service.gov.uk/media/6410aa2ce90e076cc6e370ef/HMG\\_response\\_to\\_SPV\\_Digital\\_Tech\\_final.pdf](https://assets.publishing.service.gov.uk/media/6410aa2ce90e076cc6e370ef/HMG_response_to_SPV_Digital_Tech_final.pdf) [Accessed 4 Jul. 2024], p 5.

[56] Aplin (2024).

[57] Berry, K. of Linklaters (2024). 'UK fails to agree AI/copyright code of practice', *Linklaters Law Blog*, 8 Feb. Available at: <https://www.linklaters.com/en/insights/blogs/digilinks/2024/february/uk-fails-to-agree-ai---copyright-code-of-practice> [Accessed 4 Jul. 2024].

[58] DSIT (2024) *Consultation Outcome: A pro-innovation approach to AI regulation: government response*. UK Government. Available at: <https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response#a-regulatory-framework-to-keep-pace-with-a-rapidly-advancing-technology> [Accessed 4 Jul. 2024], Para 30.

[59] Berry (2024).

[60] Smith, Thornton and Shaw (2024).

[61] Infinity Legal LLC. (2023) *Copyright Law: A Balancing Act: Singapore's Computational Data Analysis Exception*. Available at: <https://infinitylegal.com.sg/wp-content/uploads/2024/01/FINAL-PDF-CDA-Exception-Article.pdf> [Accessed 5 Aug. 2024], p 2.

generative AI.[62] The Australian Publishers Association (APA) also issued a statement asking for a similar cessation, as well as a call to safeguard publishers' rights.[63] The AWG, Australian Writers Guild, had joined in these concerns, also opposing underregulated AI.[64] The Australian government began to directly address AI copyright concerns through a Ministerial Roundtable that convened in 2023.[65] As a result of this roundtable, the Copyright and AI Reference Group (CAIRG) was established within the Attorney General's office, tasked with considering the copyright implications of AI in consultation with stakeholders.[66] Recently, the Australian government released a policy proposal laying out a vision for mandatory AI guardrails.[67] Of note is guardrail three, which addresses copyright and AI, calling for data transparency mandates and that CAIRG will be "considering the intersection of this proposed mandatory guardrail and copyright laws." [68] Momentum appears to be building in the Australian copyright sphere, and time will tell what outcomes CAIRG and the government will deliver following their consultations.

### 3.5. Canada

Canada has taken a more proactive consideration to copyright concerns spurred by the introduction of AI. The government has implemented a series of measures to secure the rights of artists and cultural practitioners, with the halting of unconditional AI data mining of artistic works being of note.[69] With the goal of taking a public-informed policy approach, a consultation committee tasked with collecting the opinions of creative practitioners has also been launched. Such a methodology suggests a policy dynamic in which the voices of cultural workers are actively shaping the AI-copyright discussion. As such, the outcomes and implications of this public-informed, creator-centric practice are likely to interest policymakers, creatives and the tech industries alike.

---

[62] IFJ (2023) *Australia: Media union outlines dangers of AI on creative industry*. Available at: <https://www.ifj.org/media-centre/news/detail/category/press-releases/article/australia-media-union-outlines-dangers-of-ai-on-creative-industry> [Accessed 4 Jul. 2024].

[63] APA (2023) *Artificial intelligence and publishing*. Available at: <https://publishers.asn.au/Web/Our-Work/Advocacy-Policy/Artificial-Intelligence-and-publishing.aspx?hkey=0a381ff9-4e92-4b54-bfc8-78e396e0f3f1> [Accessed 4 Jul. 2024].

[64] AWG (2023) *Artificial intelligence in performance and interactive writing*. Available at: <https://awg.com.au/Media/Industrial/AWG%20AI%20Position%20Paper%20September23.pdf> [Accessed 5 Aug. 2024].

[65] AGD (2023) *Ministerial Roundtable on Copyright*. Australian Government. Available at: <https://www.ag.gov.au/rights-and-protections/copyright/ministerial-roundtable-copyright> [Accessed 5 Sep. 2024].

[66] AGD (2023) *Copyright and AI Reference Group*. Australian Government. Available at: <https://www.ag.gov.au/rights-and-protections/copyright/copyright-and-artificial-intelligence-reference-group-cairg> [Accessed 5 Sep. 2024].

[67] DISR (2024) *Safe and Responsible AI in Australia*. Australian Government. Available at: [https://storage.googleapis.com/converlens-au-industry/industry/p/prj2f6f02ebfe6a8190c7bdc/page/proposals\\_paper\\_for\\_introducing\\_mandatory\\_guardrails\\_for\\_ai\\_in\\_high\\_risk\\_settings.pdf](https://storage.googleapis.com/converlens-au-industry/industry/p/prj2f6f02ebfe6a8190c7bdc/page/proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf) [Accessed 5 Sep. 2024].

[68] *Ibid*, p 37.

[69] ISED (2023) *Consultation on copyright in the age of generative artificial intelligence*. Government of Canada. Available at: <https://ised-isde.canada.ca/site/strategic-policy-sector/en/marketplace-framework-policy/consultation-paper-consultation-copyright-age-generative-artificial-intelligence> [Accessed 4 Jul. 2024].

## 4. Reconciling AI with copyright

An ethical reconciliation between AI and copyright is critical to ensuring suitable protection for cultural producers and the healthy and equitable development of generative AI. As such, policymakers should forge pathways to “ensure that creators have proper mechanisms to enforce their consent and receive fair compensation for use of their work by AI developers.”[70] As such, we propose:

### 4.1. The 4Cs: consent, control, credit and compensation

A new generative AI ecology is warranted to pave a way forward for the cultural and tech sectors alike. We have characterised such an ecology by the 4Cs: consent, control, credit and compensation:

- Artists and copyright holders should have more control over the use of their works and performances in AI data mining.[71] [72] [73] This could be secured through use of “opt-in” or “opt-out” mechanisms, via which AI developers must respect *consent agreements* and implement artistic oversight of works. At the same time, Creative Commons’ suggestion of the use of “preference signals” is worth our consideration: i.e., creators indicating how they want their creations to be used, including AI data mining, in their Creative Commons license.[74]
- Relevant to this, effective *filtering* in AI models may help restore artistic control, as well as trust between the cultural and tech sectors. [75] To this end, we propose a) filtering out copyrighted, non-consensual works when training a model and b) filtering out text prompts and generative outputs that are likely to violate copyright or the permissible use of creative works as defined by artists.[76]

---

[70] UK House of Commons: Culture, Media & Sport Committee (2024) *Creator remuneration: Fifth report of sessions 2023-24*. HC 156, UK Parliament. Available at: <https://committees.parliament.uk/publications/44143/documents/219382/default/> [Accessed 22 May 2024], p 20.

[71] Replica Studios (2024) *The New Contract Between SAG AFTRA and Replica*. Available at: <https://www.replicastudios.com/blog/how-it-works---contract-between-sag-aftra-and-replica-studios> [Accessed 23 May 2024].

[72] “The control part is as important as the compensation.” Rasenberger, quoted in The Hollywood Reporter (2024) ‘Authors Guild Exploring Blanket License For Artificial Intelligence Companies’, *The Hollywood Reporter*, 11 Jan. Available at: <https://www.hollywoodreporter.com/business/business-news/authors-guild-exploring-blanket-license-artificial-intelligence-companies-1235785941/> [Accessed 23 May 2024].

[73] “Creative control is pivotal, reflecting the need to preserve the essence of human-driven creativity from mechanistic outputs.” Lim, D. (2023), Generative AI and copyright: principles, priorities and practicalities, *Journal of Intellectual Property Law & Practice*, Volume 18, Issue 12, December 2023, Pages 841–842. Available at: <https://academic.oup.com/jiplp/article/18/12/841/7331468?login=false> [Accessed 7 Aug. 2024], p 841.

[74] Tumadóttir, A. (2024) ‘Questions for Consideration on AI and the Commons’, *Creative Commons*, 24 Jul. Available at: <https://creativecommons.org/2024/07/24/preferencesignals/> [Accessed 5 Sep. 2024].

[75] Henderson et al. (2023).

[76] “output filtering techniques will need to go beyond simple surface-level matching... a more fair-use-aligned output filtering approach would focus on detecting transformations unlikely to be fair use, such as direct translations and abridgements or depictions of copyrightable characters associated with unique text descriptions (Sag, 2018). It would ideally also take into account situations where reproduction of content is permitted, including parodies, or factual content.” Ibid, p 28.

- Furthermore, there is a general need for better transparency within generative AI, especially regarding original works. Data *transparency mandates* that require AI developers to reveal what data they have used in training their models may act as a regulatory push factor.[77] As a result, tech firms may become more proactive in seeking artists' consent and pursuing more collaborative rather than extractive relationships with the cultural sector.
- Despite generative AI being an innovation milestone, the technology is still in need of fine-tuning to better align it with societal and cultural sector norms [78] – such as crediting artists for their work. Therefore, we back proposals to advance the development of feasible *attribution mechanisms* within the field of AI. [79] Attribution can be understood as the capacity of a model to attribute an output to its source data and provide provenance information.[80] Resultantly, such developments could help to credit the artists behind generative outputs, ultimately recentring human creativity within AI dynamics. [81]
- Lastly, related to attribution, we must advance technical mechanisms that facilitate the fair compensation of artists and copyright holders whose works are used in AI data mining. In addition to licensing agreements (see 4.2), which offer one pathway toward remuneration, further research and development into attribution may someday allow for compensation in the form of royalties, potentially promoting more mutually beneficial partnerships between the tech and cultural sectors.[82]

## 4.2. Licensing agreements and beyond

With the quantity and quality of data at high value and court rulings imminent, content owners and AI firms are looking to strike licensing agreements, potentially offering further avenues of reconciliation within copyright.[83] Our review of such agreements (see Appendix 2) finds that those which align with “the 4Cs” provide the most viable roadmap for copyright reconciliation through licensing. Although licensing could perpetuate the datafication of culture and is bound up in corporate power dynamics, the agency and creative oversight

---

[77] *The EU Artificial Intelligence Act*. European Parliament. Available at: <https://artificialintelligenceact.eu/the-act/> [Accessed 1 May 2024].

[78] Schick, N. and Mosque, E. (2023) *Emad Mostaque: Generative AI as infrastructure for humanity | PIONEERS #1*. Available at: <https://www.youtube.com/watch?v=aNYPCQBqNTY> [Accessed 28 Jun. 2023].

[79] “*Instance attribution* refers to methods that assign attribution scores to training examples to understand the contribution of individual examples (or group of examples) to (test-time) model predictions (Koh and Liang, 2017; Ghorbani and Zou, 2019; Jia et al., 2019; Pezeshkpour et al., 2021; Ilyas et al., 2022)... One application of instance attribution is in determining the source of a generated output...[it] can also address the credit assignment problem by providing a clear attribution page that lists all works which contributed to the output, along with licensing information, to comply with creative commons license attribution guidelines.” cited by Henderson et al. (2023), p 29, their emphasis.

[80] Henderson et al. (2023).

[81] Pasquale and Sun (2024).

[82] Ibid.

[83] Gilbert, A. (2024) ‘Google-Reddit AI Deal Heralds New Era in Social Media Licensing’, *Bloomberg Law*, 7 Mar. Available at: <https://news.bloomberglaw.com/ip-law/google-reddit-ai-deal-just-the-start-for-social-media-licensing> [Accessed 23 May 2024].

inherent in consent and control, when combined with credit and compensation, may help to address some elements of creative precarity surrounding human cultural labour.[84]

Consequently, we must also consider the spectrum of power presented in licensing agreements. Smaller AI firms require the funding necessary to strike meaningful deals, whilst individual artists lack collective bargaining power, most especially in the absence of meaningful union representation. On the other side of the spectrum, larger cultural corporations, established institutions and high-profile artists are in better positions to negotiate fairer compensation and exert an element of creative control. To this end, we call for a more nuanced, collaborative relationship between the cultural and tech sectors, beyond court rooms and firm-level licensing deals. To this end, the establishment of an industry association tasked with representing creatives and ensuring their equitable remuneration and collection of royalties – working with both large and small AI firms – might be needed to ensure a more level playing field for all actors.

### **4.3. Critical relevance of open access culture**

Amidst the heated copyright debates and increasing licensing deals, we must not also lose sight of the Internet’s founding principle of open access and its rich history of sharing and contributions to cultural commons.[85] [86] Thus, we must aim to strengthen the public domain and facilitate fair uses of copyrighted works for individual, non-commercial and research purposes, whilst buttressing the rights of artists and copyright holders from monopolistic data mining practices for commercial gains and exclusive access.

### **4.4. The wider ecology**

Finally, we should consider the wider ecosystem within which generative AI and IP exist.[87] Observations should include research into the impact and value of robust privacy and personality rights in relation to generative AI and how such

---

[84] Pasquale and Sun (2024).

[85] Vézina, B. and Benedict, C. (2024) *Don't Be a Dinosaur; or, the Benefits of Open Culture*. Creative Commons. Available at: <https://creativecommons.org/wp-content/uploads/2024/04/Dont-Be-a-Dinosaur-or-The-Benefits-of-Open-Culture.pdf> [Accessed 5 Jul. 2024].

[86] Vézina, B., Benedict, C. and Miyara, J. (2024) Recap & Recording: ‘Open Culture in the Age of AI: Concerns, Hopes and Opportunities’, *Creative Commons*, 5 Jun. Available at: <https://creativecommons.org/2024/06/05/recap-recording-open-culture-in-the-age-of-ai-concerns-hopes-and-opportunities/> [Accessed 4 Jul. 2024].

[87] Pasquale and Sun (2024).

potential legal mechanisms might interact with and augment IP protections.[88] Such an ecosystem may also include proactive measures, such as the delivery of an AI toolkit for artists and creatives by an industry association or arms-length body.[89] An AI toolkit could offer access to anti-data-mining and model-destabilising software to protect new original works, [90] [91] as well as the provision of legal guidance, machine-learning training, meta-creative upskilling and more. It could be part of a larger national machine-learning programme that would support the creation of proprietary generative AI models by and for artists themselves. Such a programme could help return control over creative identity and expression to artists, whilst addressing the dominance of US-based tech firms in the field of AI. [92] [93] [94]

---

[88] It should be noted that US Copyright Office (USCO) has recently published the first in a series of AI reports, in which they support the introduction of a new federal law to protect all citizens from unauthorized digital replicas. “We recommend that Congress establish a federal right that protects all individuals during their lifetimes from the knowing distribution of unauthorized digital replicas. The right should be licensable, subject to guardrails, but not assignable, with effective remedies including monetary damages and injunctive relief.” USCO (2024) *Copyright and Artificial Intelligence, Part 1: Digital Replicas*. Available at: <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf> [Accessed 9 Aug. 2024], p 57.

[89] Two US Senators recently proposed the “Small Business Artificial Intelligence Training and Toolkit Act of 2024 that would authorize the U.S. Department of Commerce (DOC) to work with the Small Business Administration (SBA) to create and distribute artificial intelligence (AI) training resources and tools to help small businesses leverage AI in their operations.” An artistically aligned toolkit offered to creative practitioners could help protect and propel the UK’s cultural and creative industries: US Senate Committee on Commerce, Science, & Transportation (2024) *Cantwell, Moran Introduce Bill to Help Small Business Leverage AI Tools*. Available at: <https://www.commerce.senate.gov/2024/6/cantwell-moran-introduce-bill-to-help-small-business-leverage-ai-tools> [Accessed 19 Jun. 2024].

[90] MIT Technology Review (2023) ‘This new data poisoning tool lets artists fight back against generative AI’, *MIT Technology Review*, 23 Oct. Available at: <https://www.technologyreview.com/2023/10/23/1082189/data-poisoning-artists-fight-generative-ai/> [Accessed 19 Jun. 2024].

[91] The Glaze Project (2022) *Our Mission and Vision*. Available at: <https://nightshade.cs.uchicago.edu/aboutus.html> [Accessed 29 Apr. 2024].

[92] Emad Mostaque, former CEO of Stability AI, proposes a national AI framework for each country that is home-developed to represent the local culture and values of different peoples and counteract the power and cultural monopoly of predominantly Western ‘big tech’ firms: Schick and Mostaque (2023).

[93] Frinny Lee, musician and CEO of AV Mapping, proposed the idea of personal machine learning models so that individuals could have ownership and agency over their creative and personal expression: Lee, F. (2023) conversation with Josepha Bertolini, present at Protecting Creativity in the Age of AI. King’s College London, 14 Dec., London.

[94] Ploin et al. (2022) discuss the building process and use of proprietary models by artists in creative practice. Ploin, A., Eynon, R., Hjorth, I. and Osborne, M. (2022) *How Machine Learning Is Changing Artistic Work AI and the Arts*. Oxford Internet Institute. Available at: [https://www.oii.ox.ac.uk/wp-content/uploads/2022/03/040222-AI-and-the-Arts\\_FINAL.pdf](https://www.oii.ox.ac.uk/wp-content/uploads/2022/03/040222-AI-and-the-Arts_FINAL.pdf) [Accessed 24 Apr. 2024] pp 23-29.



## 5. Appendices

### Appendix 1. Sampling of generative AI-related lawsuits

Plaintiff(s)	Defendant(s)	Model(s)	Claim(s)	Country
Getty Images	Stability AI	Stable Diffusion	Copyright infringement, trademark infringement and dilution, unfair competition, and more [95]	USA, (UK)
			"Stability AI's brazen infringement of Getty Images' intellectual property on a staggering scale"[96]	
10 Artists (Class Action)	Stability AI, Midjourney, Runway & DeviantArt	Stable Diffusion, Midjourney, Runway	Direct and vicarious copyright infringement, Digital Millennium Copyright Act (DCMA) violations, Lanham Act violations, and more [97]	USA
			"Though the Defendants claim to be selling access to AI image products, what they're really selling is copyright infringement as a service. The scale of this misappropriation is staggering and unprecedented, with violations of law happening at every phase: the gathering and copying of the dataset, the training and deployment of the model, and the output images."[98]	
Authors Guild	OpenAI	ChatGPT	Direct, vicarious and contributory copyright infringement [99]	USA
			"Flagrant and harmful infringements of Plaintiffs' registered copyrights in written works of fiction." [100]	
Universal Music Group (UMG)	Suno Inc.	Suno	Direct copyright infringement of post-1972 copyrighted recordings, direct copyright infringement of pre-1972 copyrighted recordings [101]	USA
			"The fact that Suno's product generates digital music files that mimic readily identifiable features of the Copyrighted Recordings supports the conclusion that Suno is using the Copyrighted Recordings in training its AI model. To be clear, Plaintiffs are not presently alleging that these outputs themselves infringe the Copyrighted Recordings unless discovery reveals that they directly or indirectly recapture portions of the Copyrighted Recordings."[102]	
Concord Music Group, Capital CMG, Universal Music (and subsidiaries)	Anthropic	Claude	Direct, vicarious and contributory copyright infringement, and removal or alteration of copyright management information [103]	USA
			"Systemic and widespread infringement of their copyrighted song lyrics"[104]	
NY Times	OpenAI & Microsoft	ChatGPT	Copyright infringement, vicarious and contributory copyright infringement, alteration of copyright management information, unfair competition by misappropriation (DMCA), and trademark dilution [105]	USA
			"Defendants' unlawful use of The Times's work to create artificial intelligence products that compete with it threatens The Times's ability to provide that service"[106]	

[95] *Getty Images Inc (USA) vs Stability AI, Inc* (2023) Case 1:99-mc-09999. Available at: <https://aboutblaw.com/6DW> [Accessed 16 May 2024], pp 23-34.

[96] *Ibid*, p 1.

[97] *Andersen, McKernan, Ortiz, Southworth, Rutkowski, Manchess, Brom, Zhang, Kaye, and Ellis, vs Stability AI, Midjourney, Runway & DeviantArt* (2023) Case 3:23-cv-00201-WHO. Available at: <https://fingfx.thomsonreuters.com/gfx/legaldocs/zpnzrgrzpl/AI%20COPYRIGHT%20LAWSUIT%20amended.pdf> [Accessed 16 May 2024], pp 8-10.

[98] *Ibid*, p 3.

[99] *Authors Guilds vs Open AI* (2023) Case 1:23-cv-08292 Available at: <https://fingfx.thomsonreuters.com/gfx/legaldocs/xmvjlbqbnvr/AUTHORS%20GUILD%20OPENAI%20LAWSUIT.pdf> [Accessed 16 May 2024], pp 44-46.

[100] *Ibid*, p 2.

[101] *Universal Music Group vs Suno* (2024) Case 1:24-cv-11611. Available at: <https://storage.courtlistener.com/recap/gov.uscourts.mad.272063/gov.uscourts.mad.272063.1.0.pdf> [Accessed 7 Aug. 2024], pp 29-31

[102] *Ibid*, p 15.

[103] *Concord Music Group vs Anthropic* (2023) Case 3:23-cv-01092. Available at: <https://regmedia.co.uk/2023/10/20/anthropic-lyrics-complaint.pdf> [Accessed 16 May 2024], pp 49-57.

[104] *Ibid*, p 3.

[105] *New York Times vs Microsoft & Open AI* (2023) Case 1:23-cv-1119. Available at: [https://nytimes.com/2023/12/NYT\\_Complaint\\_Dec2023.pdf](https://nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf) [Accessed 16 May 2024], pp 60-67.

[106] *Ibid*, p 2.

## Appendix 2. Sampling of generative AI-related licensing agreements

Content Owner(s)	Tech Firm(s)	Year	Terms of Agreement
Authors Guild of America	N/A	2024	AGA is considering a “blanket licensing agreement”, wherein: <ul style="list-style-type: none"> <li>• “The control part is as important as the compensation”[107]</li> <li>• “fee for the ingestion of works and another for outputs that reference content”</li> <li>• A platform “would distribute fees for licenses”</li> <li>• “A board would be installed, alongside a new organization, to oversee the project”</li> <li>• Potential “restrictions on prompting the chatbots to produce material ‘in the style of’ authors, using characters from other works and producing summaries of books” [108]</li> </ul>
Reddit	Google	2024	Agreement struck: <ul style="list-style-type: none"> <li>• \$60million per annum</li> <li>• Access to Reddit’s rich user-generated content [109]</li> </ul>
Axel Springer SE (Politico, Business Insider, Bild, and Welt)	OpenAI	2023	Reported to: <ul style="list-style-type: none"> <li>• Be worth “tens of millions of euros for the right to use the media giant’s news articles and content to build its artificial intelligence systems...[as] part of a three-year deal” [110]</li> <li>• Include summaries from Axel Springer brand articles as well as citations, and access to full articles [111]</li> </ul>
SAG-AFTRA	Replica Studios	2024	Agreement struck: <ul style="list-style-type: none"> <li>• “Consent, compensation and control... the ability to deny their voice being used in perpetuity without their consent and the ability to opt out of the use of their voice in new works”</li> <li>• Only authorised use of voice actor data that “safeguards the privacy and intellectual property rights of voice actors”</li> <li>• “Requesting and honoring data removal on Replica’s platform”</li> <li>• “Post-project reporting and transparency”; how an actor’s voice was used on an approved project</li> <li>• “Remuneration per line for production”[112]</li> </ul>
Shutterstock	Multiple: Apple, OpenAI, Google Meta, and more	2023-24	Reported to <ul style="list-style-type: none"> <li>• Range between \$20-\$50million per firm, offering access to millions of images for training purposes [113]</li> <li>• In the case of OpenAI, the agreement contains a six-year extension to access not only images, but videos, music, and meta-data [114] [115]</li> </ul>

[107] Rasenberger, quoted in The Hollywood Reporter (2024) ‘Authors Guild Exploring Blanket License For Artificial Intelligence Companies’, *The Hollywood Reporter*, 11 Jan. Available at: <https://www.hollywoodreporter.com/business/business-news/authors-guild-exploring-blanket-license-artificial-intelligence-companies-1235785941/> [Accessed 23 May 2024].

[108] Ibid.

[109] Gilbert (2024).

[110] Cullen, A. and Davalos, J. (2023) ‘OpenAI to Pay Axel Springer Tens of Millions to Use News Content’, *Bloomberg Law*, 13 Dec. Available at: <https://news.bloomberglaw.com/tech-and-telecom-law/openai-to-pay-axel-springer-tens-of-millions-to-use-news-content> [Accessed 23 May 2024].

[111] Reuters (2023) ‘Global news publisher Axel Springer partners with OpenAI in landmark deal’, *Reuters*, 13 Dec. Available at: <https://www.reuters.com/business/media-telecom/global-news-publisher-axel-springer-partners-with-openai-landmark-deal-2023-12-13/> [Accessed 23 May 2024].

[112] Replica Studios (2024) *The New Contract Between SAG AFTRA and Replica*. Available at: <https://www.replicastudios.com/blog/how-it-works---contract-between-sag-aftra-and-replica-studios> [Accessed 23 May 2024].

[113] Nuñez, M. (2024) ‘Apple’s \$25-50 million Shutterstock deal highlights fierce competition for AI training data’, *VentureBeat*, 8 Apr. Available at: <https://venturebeat.com/ai/apples-25-50-million-shutterstock-deal-highlights-fierce-competition-for-ai-training-data/> [Accessed 23 May 2024].

[114] Roth, E. (2023) ‘OpenAI’s DALL-E will train on Shutterstock’s library for six more years’, *The Verge*, 11 Jul. Available at: <https://www.theverge.com/2023/7/11/23791528/openai-shutterstock-images-partnership> [Accessed 23 May 2024].

[115] Shutterstock Inc (2023) *Shutterstock Expands Partnership with OpenAI, Signs New Six-Year Agreement to Provide High-Quality Training Data | Shutterstock, Inc.* Available at: <https://investor.shutterstock.com/news-releases/news-release-details/shutterstock-expands-partnership-openai-signs-new-six-year> [Accessed 23 May 2024].

**If you have any questions**, need further information or would like to learn more about Sustainable Cultural Futures, feel free to contact us on the details provided below.

**Contact Details:**

**Prof Hye-Kyung Lee**

hk.lee@kcl.ac.uk

**Dr Takao Terui**

Takao.Terui@xjtlu.edu.cn

**Josepha Bertolini**

josepha.bertolini@kcl.ac.uk

**Prof Nobuko Kawashima**

nkawashi@mail.doshisha.ac.jp

**Visit us at:**

[www.sustainableculturalfutures.weebly.com](http://www.sustainableculturalfutures.weebly.com)

