



INSTITUTE FOR
Ethics in AI



Accelerator
Fellowship
Programme

THE OXFORD CONSULTATION ON COPYRIGHT AND AI

October 2026

THE OXFORD CONSULTATION ON COPYRIGHT AND AI



Second AI and Creativity Workshop, House of Lords, May 2025



AI and Creativity Workshop, Oxford, April 2025

Table of contents

PREFACE	2
EXECUTIVE SUMMARY	3
INTRODUCTION	7
THE CONTEXT	8
PILLAR 1 – OWNERSHIP AND CONSENT	12
PILLAR 2 – TRANSPARENCY, TRACEABILITY AND DISCLOSURE	15
PILLAR 3 – FAIR LICENSING AND MARKET MECHANISM	19
WHAT GOVERNMENT MUST DO	24
THE OPPORTUNITY FOR THE UK	28
CONTRIBUTORS	29
KEY DEFINITIONS	30



Baroness Beeban Kidron OBE
Advisor, Institute for Ethics in AI



Dr Caroline Emmer De Albuquerque Green
Director of Research, Head of Public Engagement, Institute for Ethics in AI

PREFACE

In 2025, a very public battle emerged over the use of creative copyright by artificial intelligence (AI) developers. It gave prominence to a debate that had been circulating for some time in the AI ethics and legal community: when is ‘learning from’ or ‘scraping’ other people’s work just a part of the inspiration that we all take from others, and when is it stealing?

In broad terms, some in the AI world felt this was a new dawn of creativity in which AI – rather than humans – would be the creators of the future. Some were more pragmatic, understanding that property and work must be paid for, but pointing out how difficult it is to gather all copyright holders’ permission, and insisting that there is so much data in the largest models that only very rarely does an individual piece have value. On the other side of the argument, creators pointed out that the models cannot be made without data; companies pay for the other raw materials, the compute, and the skills, so why not the data? And a significant minority cited their right to their own creative and intellectual output – suggesting that any kind of inclusion in an AI model was their decision alone.

It was in this climate that the University of Oxford’s Institute for Ethics in AI brought together a group of interested parties across academia, industry, policymaking and the creative industries. This meeting gave birth

to a wider project of work, including a series of meetings, discussions and interviews with a broad set of stakeholders across these sectors. This report is the outcome of that work. It marks a significant step forward in demonstrating that a consensus can be found between the many different interests that have populated the debate.

Despite a heated public and parliamentary disagreement last year, the UK Government has yet to put forward a plan for how it intends to find a solution to the differing positions on use of copyrighted work by AI developers. This report, and the consensus that underpins it, shows there is a clear route forward.

Our thanks go to Dr Lily Rodel and Dr Simeon Goldstraw for conducting interviews, synthesising conversations, and drafting the recommendations. Whilst the issue was contentious at times, coming to agreement on how to move forward was an exemplary experience in consensus building and pragmatic policy development. It was a privilege and pleasure to work alongside all the participants who gave their time and expertise to shape the final form.

We would also like to thank the Institute for Ethics in AI’s staff and the wider community for their interest and support during our deliberations.

EXECUTIVE SUMMARY

In recent years, there has been widespread concern about the unauthorised use of copyright works to train artificial intelligence (AI). Publishers, creative workers and other intellectual property (IP) holders have launched legal action against AI companies across the globe. The UK Government has faced huge pressure, inside and outside parliament, including from high-profile artists such as Elton John and Dua Lipa, to do more to protect creatives from AI companies using their work without permission. But, despite consultations and promises, the UK Government has not acted yet, meaning that the unauthorised use of copyright work by tech companies remains widespread.

Many creatives – alongside news media, publishers, and other rights holders – fear their profession will soon become unsustainable if they are not remunerated for the use of their work in training these models.

This paper focuses on the threats facing the UK's creative sector and the policy solutions required to address them. Its recommendations also have wider relevance for the UK's IP-based economy, including financial services, medicine and pharmaceuticals, defence, and hi-tech manufacturing, all of which depend on rights that establish ownership and control.

Tech companies require high quality 'data' to boost the accuracy and quality of their models. Some technology companies are becoming more cautious about the threat of legal action from rights holders for using their work without permission, resulting in a growing licensing market for the use of copyright work. However, without regulations which support transparency and the wider background conditions for licensing, the impasse will remain and the full potential of licensing the UK's intellectual property will not be realised.

UK rights holders own some of the world's most valuable data for AI, not least in its creative industries. There are opportunities for the UK to become a global leader in:

- A new licensing market for copyright material backed by government regulations and incentives, so that the UK can become the hub of purchasing and selling licensed material for use by AI companies worldwide.
- Sectors which support the licensing market, including technology, mature regulatory systems and law.

- Establishing a level playing field across major markets and reducing incentives for jurisdiction shopping, bolstering soft power in turn.
- Trusted and ethical AI, which polling repeatedly shows is a priority for UK citizens as AI is integrated at all levels of government, business and society.
- Cyber-security, as copyright transparency feeds directly into regulations promoting data and system security.

Global competition among states to capitalise on these opportunities is growing rapidly. Whilst the UK has a natural advantage due to the strength of its creative, technological and legal sectors, the Government must move immediately to harness this advantage.

The key recommendations of this paper are presented as three pillars:

PILLAR 1 – OWNERSHIP AND CONSENT

- Recognise existing ownership and consent rights of rights holders over their work within a new legally binding licensing regime. This licensing regime should enable rights holders to express permissions and preferences, and refuse consent.
- Explicitly rule out text and data mining exceptions, automatic opt-out, and commercial research exceptions which do not uphold these standards.
- Expand primary and secondary copyright infringement in AI contexts.
- Products, models or services imported into or made available in the UK must have all relevant licences and consent to use UK copyright work in the UK market, wherever it was scraped.
- Retain current exceptions to copyright restrictions in UK law, such as those secured under principles of 'fair dealing' or 'lawful access'.
- Legally protect the use of a person's name, image, and likeness, which includes a person's face, voice, body, or artistic performances.

PILLAR 2 – TRANSPARENCY, TRACEABILITY AND DISCLOSURE

- Mandate traceability of training data and source material through minimum, robust standards for AI provenance systems and licensing information, and expressly mandate that it may not be removed by a third party. Licensing data must be embedded as standard.
- Mandate ingredient transparency, including the detailed disclosure of training data and source materials, with full public transparency as the default for general-purpose systems and a confidential process for specialised systems trained on sensitive datasets or where there are national security concerns.

- Mandate minimum interoperability standards, rather than a single Government-dictated system, to allow rights holders, regulators, and other interested parties to scan and compare the metadata of different models and products.
- Mandate bot usage transparency and disclosure with a public bot register which lists all web crawlers, their stated purpose, and the responsible entity.
- Mandate labelling of AI-generated works to ensure people know when they are interacting with AI-generated content.
- Empower rights holders to seamlessly access information on how their works have been used in a way that is simple and easy to understand to ensure proper control and attribution.
- Allow limited transparency exemptions through confidential processes for national security, specialised systems trained on sensitive datasets and for commercial sensitivity in exceptional circumstances, while ensuring full regulatory access.

PILLAR 3 – FAIR LICENSING AND MARKET MECHANISMS

- Allow the organic emergence of a market for lawful, fair and efficient licensing of the use of copyright material by AI developers, by implementing the Consultation's recommendations for establishing ownership, transparency and supportive governance.
- The Government should establish and enforce the conditions for a competitive licensing market, rather than operate the market itself. It should rule out a centrally administered levy system.
- Facilitate the development of three complementary models in which creative content is licensed which reflect the different preferences of creators and the differing values of their work: (A) data banks and combined datasets; (B) collective licensing; and (C) direct negotiation.
- Establish a UK Government-run pilot system to test whether structured, licensed, machine-readable access to public datasets can support transparency and compliance obligations for AI developers.

What government must do

- Introduce new legislation to support the enforcement of the three pillars that protects rights holders, mandates provenance, transparency, and bot identity standards, and establishes meaningful penalties for non-compliance.
- Designate a competent regulator as primarily responsible for regulating and enforcing the legal standards of ownership and consent, and transparency, traceability and disclosure.
- Include workers/creatives/unions and copyright owners in policy decisions about AI and creative copyright.
- Incentivise compliance through whitelisting mechanisms that reward compliant firms with access to Government procurement processes, data, grants and innovation funding, tax incentives, and participation in regulatory sandboxes.
- Ensure that government procurement, development and use of AI models across every department centrally and locally comply with these principles.
- Support a dual enforcement regime for infringing copyright law, in which traditional private enforcement through the courts is maintained, while the regulator retains the ability to take further enforcement against repeat offenders.
- Move swiftly to implement the measures in this report, with cross-Government, coordinated action across DCMS, the Cabinet Office and involving the cabinet Minister for AI.

INTRODUCTION

Since generative-AI models became commercially available, there has been widespread controversy about the data used by developers to create them. Most of the data used by AI developers is third party data, rather than their own. They acquire it by scraping websites, accessing internet-connected databases, copying data from physical materials, or licensing it. As a result, a substantial proportion of material that is absorbed is confidential personal data, or is under copyright. This material is a fundamental element of building proprietary commercial AI products – often outside of the UK and without the authorisation, participation, or control of UK rights holders.

Neither rights holders nor tech companies believe this is a sustainable arrangement. It poses an existential risk to the livelihoods and creative purpose of rights holders,¹ whilst tech companies are increasingly at legal risk and encountering barriers to the high-quality data they need to create good content.

In February 2025, a group of computer scientists, technology developers and members of the creative industries convened in Oxford under the auspices of the University of Oxford's Institute for Ethics in AI to confront the crisis and opportunities surrounding AI, copyright, and the creative industries in the UK. This meeting catalysed the Oxford Consultation on Copyright and AI – a cross-sector effort to develop a sustainable solution.

This report synthesises the voices of members of the Oxford Consultation, in dialogue with a range of other creatives and AI companies, with the goal of presenting a fair and competitive framework for AI innovation in the UK. The contributors are a diverse group with different perspectives, and the report represents a broad consensus arrived at over many months.

We urge the UK government to move quickly to implement the recommendations contained in this report.

¹ Independent Society of Musicians (ISM), Society of Authors (SoA), Equity, Association of Illustrators (AOI) and Association of Photographers (AOP) (2026) *Brave New World? Justice for Creators in the Age of GenAI*. Available at: <https://www.the-aop.org/uploads/brave-new-world-justice-for-creators-report-29-01-2026.pdf>

THE CONTEXT

Why do rights matter?

The creation, operation, and improvement of AI models requires vast datasets across the spectrum of human knowledge. This includes using copyright material which is owned by creators or businesses without their permission.

Upholding copyright is essential to the livelihoods of UK creatives and the future of their sector. They are key to the UK's economy, contributing £145.8bn in gross value added (GVA) in 2024.² This is equivalent to 5.6% of the GVA of the whole UK. Protecting the livelihoods of those in the creative industries is therefore highly important to sustaining the sector's economic contribution and supporting the UK's wider prosperity. But currently, rights holders have no direct way of knowing if their work has been used without their consent.

How is data used by AI?

There is an ongoing need for AI developers to access data including copyrighted work.

One way AI developers gather data is by deploying 'web crawlers' or 'bots' that 'scrape' material at scale. During model training, an AI model is exposed to large amounts of data. The system repeatedly adjusts its internal parameters to *learn* the patterns, correlations and structures within the data used during training. When the AI is prompted by a command or question, it uses these patterns to calculate its output or answer. There is evidence that this process can lead to models producing extensive verbatim passages of encoded training data.³

All AI systems access and create new data to remain relevant. Many new AI models, including large language models (LLMs) and text to image or audio models, are still being developed. Smaller or specialist models also require specialist training data for 'fine-tuning'. These specialist models build on an existing model but often require specialist data and instructions aimed at specific tasks – such as those designed for educational purposes, healthcare, news or film production. Emerging models, such as world and spatial models, also require sensory audiovisual data for their training.⁴

Many AI systems draw on sources *in real time* – often referred to as 'grounding'. This real time information is essential to ensure that outputs and answers are up to date, more accurate and to situate them in broader context.

This means that there can be use of copyright material during the initial and ongoing training of an AI model and as the model is being used. Currently, much of this process happens without the permission of rights holders, but it is possible to use data legitimately when a copyright holder has approved use of their material.

² Department for Culture, Media and Sport (2026) *DCMS Sectors Economic Estimates: Gross Value Added 2024 (provisional)*. Available at: <https://www.gov.uk/government/statistics/dcms-economic-estimates-gva-2024-provisional>

³ Buick, Adam. 'Copyright and AI training data—transparency to the rescue?', *Journal of Intellectual Property Law & Practice*, Volume 20, Issue 3, March 2025, Pages 182–192, <https://doi.org/10.1093/jiplp/jpae102>; Carlini, Nick et al, 'Quantifying Memorization Across Neural Language Models' (2022) *arXiv* preprint arXiv:2202.07646

⁴ Zhang, D., Wald, R., Adeli, E., Cryst, E., Ho, D.E., Meinhardt, C., Wu, J., Zegart, A. and Li, F.-F. (2026) *The World Model and Spatial Intelligence Era: Governing AI Beyond Language*. Stanford Institute for Human-Centered Artificial Intelligence. Available at: <https://hai.stanford.edu/assets/files/hai-issue-brief-the-world-model-and-spatial-intelligence-era.pdf>

Good AI needs good data

The AI sector uses talent, compute, natural resources, and data. Data is a central component of the AI ecosystem; it is the 'lifeblood' of AI and an 'essential' part of its infrastructure.⁵ This is particularly evident in AI, which requires data for training and development. If the data that is used is of low quality (for example, it is incorrect or homogenous), then the outputs generated by the AI will be of a lower standard. Experts have already warned of a looming 'AI data scarcity crisis' as the production of high-quality data becomes less abundant or more difficult to access.⁶ Creative material is an important source of high-quality data – its use is valuable to AI companies – and a system that protects its continued production is good for creators and the AI sector.

Creatives are worried

Copyright is an integral part of the creative value chain, often acting as income, maternity, holiday pay, and a pension, in a sector heavily populated with freelancers. If rights holders are not paid for their work, creative careers become unsustainable. Some creators also face the challenge of competing for business with AI models which have been trained – without remuneration – on their work.⁷ Many UK artists also have concerns about the integrity of their work: they wish to maintain the moral right to their creative and intellectual output, which is established in a range of international conventions.⁸ For these artists, their work is their individual original cultural contribution first and foremost.

Emerging licensing market

Many rights holders, particularly larger production, music, publishing and news outlets, have begun to privately agree their own licensing deals.⁹ This growing licensing market shows a willingness amongst both creators and tech companies to find solutions. Creators recognise it is a key part of the future of their sector,¹⁰ and tech companies have paid for access to good data with legal certainty.

These licensing deals are limited by the present context. Currently, only the most powerful rights holders have sway over the most powerful AI companies, worsening the asymmetry of power in both sectors, with individual creators and smaller businesses – which possess less bargaining

⁵ UK Government. (2025) 'AI Opportunities Action Plan'. CP 1241, January 13, 2025. GOV.UK. <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan>; Joint Committee on the National Security Strategy. (2025) *Oral Evidence: The National Security Strategy*. Oral evidence transcript no. 16738, November 17, 2025. UK Parliament. <https://committees.parliament.uk/oralevidence/16738/pdf/>

⁶ *The Independent* (2025) 'Experts warn of "crisis point" ahead for AI data'. Available at: <https://www.independent.co.uk/bulletin/news/ai-peak-data-training-b2838959.html>

⁷ Independent Society of Musicians (ISM), Society of Authors (SoA), Equity, Association of Illustrators (AOI) and Association of Photographers (AOP) (2026) *Brave New World? Justice for Creators in the Age of GenAI*. Available at: <https://www.the-aop.org/uploads/brave-new-world-justice-for-creators-report-29-01-2026.pdf>

⁸ In a letter demanding the UK Government protect creators' copyright, creatives and creative organisations cite the International Covenant on Economic, Social and Cultural Rights (1996), the European Convention on Human Rights and Fundamental Freedoms (1950), and the Berne Convention for the Protection of Literary and Artistic Works (1886). National Union of Journalists (2025) 'UK creators demand prime minister recognises creators' human rights and protects copyright'. Available at: <https://www.nuj.org.uk/resource/uk-creators-demand-uk-prime-minister-recognises-creators-human-rights-and-protects-copyright.html>

⁹ E.g. *PitchFork* (2025), 'Major Labels Sign Licensing Deals With AI Music Company Klay'. Available at: <https://pitchfork.com/news/warner-music-group-signs-licensing-deal-with-ai-music-company-klay/>

¹⁰ 77% of members of the British Phonographic Industry are open to licensing their music for ethical use by AI, and the same number agree that licensing is key to driving business growth in the creative industries. WPI Economics (2026) *Driving UK growth: The role of licensing music in the age of AI*. Available at: <https://wpieconomics.com/wp-content/uploads/2026/05/Driving-UK-growth-the-role-of-licensing-FINAL.pdf>

power – being left behind. One contributor also reported that they are seeing trends of US creators being remunerated at a higher level than their British counterparts, despite work being of the same standard. Unabated, this trend will repeat the concentration of value in the US market, based on companies' size and location, that we have seen across the tech sector. This is of particular concern to a mid-size country with creative industries that are central to its economy, employment and soft power.

Most problematically, rights holders and individuals still lack the automatic means to know if their work or likeness has been used. This means that, whilst it comes with the risk of legal uncertainty, AI companies are not incentivised to seek licensing agreements. Again, this favours larger companies who have deep pockets to fight legal challenge, while smaller and mid-size companies (including virtually all UK AI companies) are in a more vulnerable position, exacerbating the asymmetry in which a UK start-up pays whilst a frontier company – capitalised to trillions of dollars – does not.

Meanwhile, a shadow market is emerging in which bots belonging to third parties illicitly scrape data, then sell the data to AI companies.¹¹ These intermediaries act as market makers and effectively de-risk data access for AI developers who can remain at arm's length, paying for data, with the money going to entities that stole the work, rather than the work's creator. This added layer makes it harder for creators to identify the legally correct party to enforce their rights against.

Government policy

The UK Government has repeatedly stated that copyright law must be upheld, but has stood in the way of steps that would provide copyright holders with the means to uphold the law, through transparency and labelling.¹²

In December 2024, the government opened a consultation with a 'preferred option' of a text and data mining exception, which would routinely apply if a rights holder had not expressly 'opted-out'. This was abandoned in March 2026, following significant pushback from creators and parliamentarians. Ultimately, only 3% of consultation respondents supported the government's preferred proposal, but the government did not act on the overwhelming preferences of the consultation to strengthen transparency and uphold copyright.¹³ Instead it announced that it would focus on developing policy in four areas: 1) digital replicas and likeness rights; 2) labelling AI-generated content; 3) creator control and transparency; and 4) ensuring a solution which

¹¹ Pachal, Pete. (2026) 'AI scraping has become its own media business'. *Fast Company*. Available at: <https://www.fastcompany.com/91539092/ai-scraping-become-own-media-business>

¹² See amendments tabled to the Data (Use and Access) Bill at the Committee, Report, and "ping pong" stages in the House of Lords, and corresponding debates. Committee amendment nos. 204, 205 and 206 (debated 18 December 2024); Report nos. 61, 62, 63, 64, and 65 (debated 28 January 2025); Ping pong debates on 19 May 2025, 2 June 2025, 4 June 2025, 11 June 2025

¹³ Department for Science, Innovation and Technology, Department for Culture, Media and Sport and Intellectual Property Office (2026) *Report on Copyright and Artificial Intelligence*. Available at: <https://www.gov.uk/government/publications/report-and-impact-assessment-on-copyright-and-artificial-intelligence/report-on-copyright-and-artificial-intelligence>; Department for Science, Innovation and Technology, Department for Culture, Media and Sport and Intellectual Property Office (2025) *Copyright and Artificial Intelligence: Statement of Progress under Section 137 Data (Use and Access) Act*. Available at: <https://www.gov.uk/government/publications/copyright-and-artificial-intelligence-progress-report/copyright-and-artificial-intelligence-statement-of-progress-under-section-137-data-use-and-access-act>

suits independent creatives as well as larger creative companies and publishers.¹⁴ Alongside this, it committed to parliamentary engagement and a 'Creative Content Exchange' (CCE) – a government-backed online marketplace for exchanging creative assets from galleries, libraries, archives and museums.

At the time of writing in September 2026, there has been one meeting with parliamentarians in each house, but no further action.

Other jurisdictions are beginning to act

In 2025, California passed laws requiring AI companies to publish high level summaries of their training data sets on their websites, and label AI-generated content.¹⁵ In June 2026, the European Union published its Code of Practice on transparency of AI-generated content, which set out rules for marking, detection and labelling of AI-generated content.¹⁶ In July 2026, the Australian Prime Minister announced the intention to move from voluntary AI standards to a new set of mandatory Australian AI Standards, which include policy commitments on copyright protections that would protect the right of Australian creators to retain control and ownership of their works.¹⁷ Whilst none of these initiatives establish the full range of transparency requirements needed for creatives to enforce their copyright, they demonstrate that other jurisdictions are willing to legislate to protect their economy and creators, and they offer the UK Government an incentive to act swiftly if it is to capitalise on the opportunity to create a global market in licensing with the support of other like-minded jurisdictions.

The UK's AI and copyright policy is really vital, not just for the UK but for the world. It has the potential to become the global standard in this area that other countries can then adopt, leading to achieving consensus on global standards... A clear, innovation-friendly legal framework will incentivise AI companies to train models in the UK with trust and confidence.

Zahra Shah, UKAI

¹⁴ Department for Science, Innovation and Technology (2026) *Copyright and AI Progress*. Written statement HCWS1416, 18 March 2026. Available at: <https://questions-statements.parliament.uk/written-statements/detail/2026-03-18/hcws1416>

¹⁵ State of California (2025) California Civil Code § 3111: *Artificial Intelligence Training Data Transparency*. Available at: <https://law.justia.com/codes/california/code-civ/division-3/part-4/title-15-2/section-3111/>; California State Legislature (2024) *SB-942 California AI Transparency Act*. Available at: https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB942

¹⁶ European Commission (2026) *Code of Practice on Transparency of AI-generated Content*. Available at: <https://digital-strategy.ec.europa.eu/en/policies/code-practice-ai-generated-content>

¹⁷ Albanese, Anthony. (2026) 'AI in Australia's interests'. Prime Minister of Australia, 15 July 2026. Available at: <https://www.pm.gov.au/media/ai-australias-interests-0>

PILLAR 1

OWNERSHIP AND CONSENT

Key principle

The UK must set a clear standard: no unauthorised use of copyright creative content in AI, whether for training datasets or for grounding models in live information. Subject to existing exceptions for Creative Commons and open-access content, all access to British copyright content must be licensed, traceable, and fairly compensated.

Rationale

Neither the scale of modern AI systems nor their purpose provides a justification for the unlicensed use of copyright work. To the contrary, the data is what makes an AI system useful, and it is that founding utility that enables all downstream value creation in AI. Clear and enforceable rights of ownership and consent are foundational because they protect creators and provide legal certainty for AI developers. The same legal protections and commercial enablement would support the UK's other intellectual and data industries, including medical and pharmaceuticals, defence, aerospace and hi-tech manufacturing, and all associated academic research and technology transfer.

1. **Publicly recognise ownership and consent.** Copyright law in the UK recognises that the creator or authorised rights holder holds the exclusive right to authorise the use of their work. The same principle must apply in the context of AI systems: any use of copyright content for training or grounding must be explicitly authorised by the applicable rights holder unless they have chosen to make their work available on open access or Creative Commons terms, or an existing legal exception applies.
2. **Expand primary and secondary copyright infringement in AI contexts.** Under UK law, primary copyright infringement occurs when copyright material is copied or reproduced without authorisation. Secondary infringement occurs when unauthorised material is distributed, shared, or commercialised. These provisions should be reviewed and adapted to ensure that they adequately address the ways in which copyright works are used in the development of AI models and products.

Strengthen secondary copyright through infringing beneficiaries

In *Getty vs Stability AI*, the High Court held that Stable Diffusion was not an 'infringing copy' for the purposes of secondary copyright infringement, because the model did not store or reproduce copyright works, despite Stability AI acknowledging that at least some images from the Getty Images Website were used during the training of Stable Diffusion.¹⁸ This highlights a potential gap in the existing framework.

One stakeholder suggested that the UK's secondary copyright regime could be strengthened by expressly extending its scope to products and services that, regardless of whether they are themselves copies, derive commercial value from acts committed during their development that were copyright infringement in the UK or would have constituted an infringement had they taken place in the UK.

¹⁸ *Getty Images (US) Inc and others v Stability AI Ltd* [2025] EWHC 2863 (Ch), Case No. IL-2023-000007, High Court of Justice, Business and Property Courts of England and Wales, Intellectual Property List. Courts and Tribunals Judiciary. Available at: <https://www.judiciary.uk/wp-content/uploads/2025/01/Getty-Images-and-others-v-Stability-AI-14.01.25.pdf>

- 3. Extraterritorial protection.** AI models that are built using unlicensed UK copyright content constitute a breach of UK copyright, even if the act of copying (for example, in training) took place outside of the UK.

Are we willing to regulate in the national interest to allow British creatives to exploit the content that they create?

Sebastian Cuttill, News Media Association

- 4. Retain current exceptions.** Existing UK law includes exceptions to copyright restrictions for purposes including parody, caricature and pastiche, non-commercial research, and reporting current events. Such exceptions are secured under principles of ‘fair dealing’, ‘lawful access’ and others, which are assessed case-by-case.¹⁹ These exceptions are widely acknowledged and are essential to creatives, journalists and researchers amongst others and must be retained without dilution.
- 5. Extend rights to protect name, image, and likeness.** There have been high-profile cases where AI has emulated the identities of creative workers.²⁰ These cases sit within the wider proliferation of deepfakes and AI-generated impersonations of citizens, politicians and other high-profile figures which can present a security threat, and cause distress for those whose identities are mimicked.²¹ Copyright must be extended to apply to the use of name, image, and likeness,²² which includes a person’s face, voice, body, or artistic performances.

Some stakeholders supported greater protection but raised concerns around the complexities of implementing such a regime. Rights to name, image and likeness exist in a range of European jurisdictions, including Germany, Italy and Denmark. The foundations for a right to style can be found in the US CREATOR Act. These examples provide models for the UK Government to consider.

There are also concerns about how AI outputs may emulate the ‘style’ of a particular artist rather than any individual work (for example, by creating an image in the style of Salvador Dali or L.S. Lowry). While this raises different considerations from the protection of name, image and likeness, creators should also have appropriate protection against the unauthorised commercial exploitation of their distinctive style.

¹⁹ Intellectual Property Office. (2014) Exceptions to copyright. GOV.UK. Published June 12, 2014; last updated January 4, 2021. <https://www.gov.uk/guidance/exceptions-to-copyright>

²⁰ *The Guardian* (2024) ‘ChatGPT suspends Scarlett Johansson-like voice as actor speaks out against OpenAI’. Available at: <https://www.theguardian.com/technology/article/2024/may/20/chatgpt-scarlett-johansson-voice>

²¹ Several respondents raised their concern about the harm of deepfakes to victims and individuals, and the potential security threat. Whilst there are other steps the government must take to address these beyond immediate AI copyright concerns, the right to name, image and likeness is a very important first step. (see Office of Communications (Ofcom). (2024) *A deep dive into deepfakes that demean, defraud and disinform*. Ofcom. <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/deepfakes-demean-defraud-disinform>; U.S. Department of Homeland Security. 2024. *Increasing Threats of Deepfake Identities*. Department of Homeland Security. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf)

²² Some respondents raise concern about the challenges of such a regime, but rights to name, image and likeness exist in a range of European jurisdictions, including Germany, Italy and (more recently) Denmark, so it can be done. The foundations for a right to style can be found in the US CREATOR Act.

Why exceptions and opt-out regimes are not enough

Governments across the world have looked at text and data mining exceptions and automatic opt-out regimes, whilst recent academic articles have suggested a commercial research exception.

None of these recognise ownership and consent:

- Text and data mining exceptions exempt some companies from seeking consent.
- Automatic opt-out regimes are too unreliable and demanding on creators, who would be forced to identify every use of their work by an AI company and opt-out separately.
- Commercial research exceptions exempt some companies from seeking consent if their use of work is for research purposes. (This classification is also open to abuse by large companies investing in research for commercial purposes).

Instead, **authorisation must be secured through a legally-binding licensing regime, which enables rights holders to express permissions and preferences, and refuse use of their work if they wish.**

PILLAR 2

TRANSPARENCY, TRACEABILITY AND DISCLOSURE

Key principle

Transparency, traceability and disclosure are essential components of any trusted system because they allow third parties to verify the claims made by those subject to the law and allow those subjects to verify their compliance. This principle develops how transparency, traceability, and disclosure should be established.

If none of the material [used to train an AI model] is in any way tagged in shape or form — it's almost impossible to track it.

Guy Gadney, Charismatic.AI

Rationale

Without transparency of data sources, rights holders cannot know how their work has been used and AI developers cannot prove that their models comply with the law. In a survey of British creatives that asked whether they thought their work had been used to train AI systems, 33% answered they did not have enough information to know and another 30% suspected it had been used but did not know for certain.²³ External projects like The Atlantic's AI Watchdog, which allows users to search whether a work is included in a dataset used by AI companies for training, are a start, but cannot be comprehensive and have no legal authority.²⁴

As good AI systems need good data, both in the training process and continuously, transparency measures must apply to training datasets and retrieval pipelines, ensuring that sources used in grounding are attributable and auditable. Implemented together, seven measures will create a transparent and ultimately trusted data-use environment for AI:

- 1. Mandate traceability through minimum standards for data provenance and licensing information.** Require minimum standards for provenance systems that track training datasets and source material that include provisions to prevent tampering or removal of provenance data by a third party.²⁵ Provenance standards must ensure licence data can be embedded as standard.²⁶ This approach would complement existing UK copyright law, which makes the deliberate removal or alteration of electronic rights management information unlawful.²⁷

²³ Katell, Michael et al. (2025) *Creative Industries and GenAI: Good Work Impacts on a Sector in Transition (CREAATIF Good Work Report)*. Available at: <https://doi.org/10.5281/zenodo.15535400>

²⁴ The Atlantic's AI Watchdog. Available at: <https://www.theatlantic.com/category/ai-watchdog/>

²⁵ Werder, Karl et al. 2022. 'Establishing Data Provenance for Responsible Artificial Intelligence Systems'. *ACM Trans. Manage. Inf. Syst.* 13, 2, Article 22 (June 2022), 23 pages. <https://doi.org/10.1145/3503488>

²⁶ Similar recommendations have been made in the Co-Star ACCT report: see CoSTAR National Lab, DECaDE, and Sheridans. (2025) *Time to ACCCT: An AI Copyright Framework for UK Creative Industries*. CoSTAR Network, May 21, 2025. <https://cms.costarnetwork.co.uk/api/media/file/costar-national-lab-ai-copyright-framework-report-may-2025.pdf>

²⁷ Copyright, Designs and Patents Act 1988 (UK), ss. 296ZG–296ZH.

What is provenance?

Provenance in AI-generated content refers to a record that describes the origins and processing of data or a piece of digital media

2. **Mandate ingredient transparency.** Transparency regarding training inputs is essential to enabling accountability, consent enforcement, and fair remuneration. Company records must include detailed disclosure of data used to train the model, and ongoing transparency of source materials that contribute to AI-generated content.²⁸ Records need not include weights and would exclude personal prompts and data.²⁹ These records must be retained and available for scrutiny. Unless special exemptions apply, they will be fully publicly available (see point 7).
3. **Mandate minimum interoperability standards.** Different companies encode metadata in incompatible ways, impeding interoperability. Minimum interoperability standards allow rights holders, regulators, and other interested parties, to more easily scan and compare the metadata of different models and companies' products. Mandating interoperability standards, rather than a single (government-dictated) technical system, will allow innovation, while ensuring a consistent baseline for transparency.

What is interoperability?

Interoperability means that different systems and data formats can work together without requiring custom code for every integration. UK competition law already recognises that interoperability can be necessary in contexts where proprietary technical choices undermine fair competition or transparency. The Digital Markets, Competition and Consumers Act 2024 enables the Competition and Markets Authority (CMA) to impose conduct requirements on firms designated with Strategic Market Status. These requirements are enforceable, with the CMA able to impose significant financial penalties for non-compliance. This provides a clear precedent for mandating minimum interoperability standards without prescribing a single technical solution.

4. **Mandate bot transparency and disclosure.** Establish a public bot register which lists all web crawlers, their stated purpose, and the responsible entity. Bots must also be required to hold cryptographic proof of identity which easily enables web domains to distinguish between human and AI-related traffic. This would aid enforcement against bots which scrape data anonymously to sell it on a black market and strengthen the confidence of tech companies that licensing agreements give them privileged access to the high quality data they need.

²⁸ The Competition and Market Authority's 2026 'Publisher Conduct Requirement' for Google's general search services includes requirements for Google to publish information for users on how search content is used by its generative-AI. This establishes a precedent that transparency measures for source materials are possible and expected, which the government must build upon. Competition and Markets Authority (2026) *Publisher Conduct Requirement: Google's general search services*. Available at: https://assets.publishing.service.gov.uk/media/6a1f0098b95db968c8f3b9db9/Publisher_CR_final_decision.pdf

²⁹ The recent court case between OpenAI and the *New York Times* has elicited privacy concerns that retaining source data may require AI companies to publish private interactions that users have had. That does not need to be the case to establish sufficient levels of transparency for rights holders to exercise their rights and to create the foundations of a licensing market.

When an entity scrapes a website without disclosing that it is automated, and deliberately impersonates a human user using tools such as headless browsers and residential IP addresses to evade detection, that crosses the line into a cybersecurity violation in my view. The intent is not simply to access publicly available information, but to circumvent standard technical controls that publishers rely on to manage automated access.

Regulators should draw a clear boundary that deceptive evasion of access controls is unacceptable. Addressing this kind of conduct is unlikely to harm the broader internet economy or legitimate AI research, but it would provide important clarity and protection for publishers and online services.

Lucky Gunasekara, Miso.AI

5. **Mandate labelling of AI-generated works.** Companies must ensure people know when they are interacting with AI-generated content, including images, audio, and text. There is recent precedent: the EU's AI Act states that synthetic text, images, video, and audio designed to resemble existing persons, objects, places, entities, or events must be visibly marked as AI-generated and contain a digital EU icon watermark to show the origins of the content.³⁰
6. **Empower rights holders to attain information on how their works have been used in AI training.** Based on points 1-4, rights holders and any user should be able to see how copyright work or name, image and likeness rights have been used in AI training and/or grounding in a way that is simple and easy to understand.
7. **Allow limited transparency exemptions for national security, specialised systems trained on sensitive datasets, and for genuine commercial sensitivity, while ensuring full regulatory access.** Stakeholders highlighted that transparency obligations must distinguish between systems for which records should be fully publicly available, and systems trained on sensitive datasets.

The first group would comprise most AI systems, including 'general-purpose AI models'³¹ and specialist models which do not use sensitive data. This would enable rights holders and individuals with personal image rights to easily identify if their work or image has been used and a data rights market to function effectively and quickly.

³⁰ European Commission (2026) 'EU Icons for labelling AI-generated content'. Available at: <https://digital-strategy.ec.europa.eu/en/policies/eu-icons-labelling-ai-generated-content>

³¹ General purpose AI models are defined in the EU AI act as 'an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications.' European Union. "Regulation (EU) 2024/1689 of the European Parliament and of the Council 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)." *Official Journal of the European Union* L 2024/1689 (July 12, 2024): 1-144.

The second group would consist of a subset of specialised systems trained on sensitive datasets such as medical datasets. These should not be subject to identical public disclosure obligations because this may breach confidentiality obligations or create security risks. Companies could apply for a 'specialised systems status' to a regulator, which would then also be responsible for assessing the system's ingredient transparency.

How could confidential regulation work?

A comparable model exists in UK pharmaceutical regulation. Companies seeking marketing authorisation must submit data to the Medicines and Healthcare products Regulatory Agency (MHRA) in the form of an electronic common technical document (eCTD), a standardised format for clinical evidence. This includes confidential clinical trial data and quality control data. The regulator has full access to this information, which must be attributable to the individual or system that created it, but it is not made publicly available.

Commercial sensitivities have traditionally been a barrier to transparency. Where an emphatic case can be made for commercial sensitivities, the regulator may have access to relevant information that the public will not – equivalent to specialised system status. The threshold for claiming commercial sensitivity should be high, and the provisions for ensuring regulator access should be equal to other sensitive datasets.

Specialised systems status should only affect the level of public disclosure, not the record-keeping obligations themselves. Confidentiality can risk reducing accountability because it limits who can see the underlying evidence behind regulatory decisions.³² Therefore, the regulator's independence, capacity, and integrity is paramount.

³² Gotham, Dzintars and Barber, Melissa (2024) *Pharmaceutical Regulation in the UK: A Case Study*, Ada Lovelace Institute, October 31, 2024, Available at: <https://www.adalovelaceinstitute.org/resource/pharmaceutical-regulation-uk/>

PILLAR 3

FAIR LICENSING AND MARKET MECHANISM

Key principle

The UK economy needs a framework that lays the foundations for an organic data licensing market which enables lawful, fair, and efficient exchange between rights holders and AI developers. This principle works in tandem with the previous two pillars.

Just as AI companies pay for their own talent and compute, they must pay for the talent and intellectual property of creators on which their models rely. An active market could ignite new forms of trade-supporting businesses, such as in data preparation and aggregation, helping data producers become data traders, while simultaneously adding value to raw data by bringing utility to data users.

Rationale

Stakeholders highlighted that the current licensing system in the UK for AI-related content is not “fit for purpose”, particularly in audiovisual content, due to the fragmentation of rights across multiple parties. For example, rights in a single television production may be split across multiple rights holders for writing, original production, broadcast and distribution. Music is similarly complicated. For AI developers, this creates significant complexity because establishing licensed access may require negotiating across multiple layers of ownership. Reducing this friction would incentivise the creation of a dynamic licensing market in the UK.

Many in the AI and creative sectors expressed concern that the government is seeking to create a marketplace for selling, buying, and licensing content with its Creative Content Exchange proposal, rather than set the regulatory conditions for a commercial market. A government-run market risks creating a UK-only silo and is unlikely to command the confidence of either sector. Both the creative and the AI sectors were clear that the Government’s role is not to operate or build a market but to establish and enforce the conditions and principles – as set out elsewhere in this framework - which support an independent market.

This paper outlines a forward-looking framework that focuses on establishing conditions for future use, and the following complementary models could co-exist within a regulated ecosystem. Each relies on opt-in participation by rights holders, transparency obligations for AI companies, and interoperability standards overseen by the regulator. **It is not for an AI company to dictate which of the three different models they are responding to, but for the copyright owner to determine their desired route and the value of their work.**

SCENARIO

A

DATA BANKS AND COMBINED DATASETS

Some material may not be valuable on its own but still provide value in combination with other material.³³ This data may be owned in small batches by individuals or in vast data sets by entities such as a library or broadcaster, and only of value in aggregate, if readily available with little interaction, and at low cost per data point. It makes sense to pool such data.

Accredited databanks could allow rights holders to contribute works voluntarily under simple, machine-readable licenses, and would be subject to a micro-payment system based on usage. Certain conditions could be built into the accredited database criteria, for example, a right to withdraw from future uses, freedom to license elsewhere, and a bar on anyone making membership a condition of work.

This could enable large-scale and low-friction licensed access to data for AI developers. Participation should remain voluntary.

SCENARIO

B

COLLECTIVE LICENSING

Collective copyright licensing is a familiar and well-tested mechanism. Properly designed, opt-in collective licensing agreements would enable the ethical and lawful use by AI developers of large bodies of creative work, while ensuring creators are compensated for use of their intellectual property and retain meaningful control.

Nicola Solomon, Independent legal consultant, former Chief Executive of the Society of Authors and former Chair of the Creators' Rights Alliance

³³ For example, videos of people talking to cameras in which the content is irrelevant (but which can be used to help mirror human reactions in an avatar), a social media feed, or historic notes and memos toward a thesis.

Existing collection societies such as Directors UK, the Educational Recording Agency (Era.org.uk), the Performing Rights Society (PRS), and the Design & Artists Copyright Society (DACS), plus collective management organisations (CMOs) such as Authors' Licensing and Collecting Society (ALCS) and Publisher's Licensing Services (PLS), represent proven structures for aggregated rights management.

Expanding collective licensing for AI training could address fragmentation and friction by enabling aggregated clearance of rights across performers, writers, and contributors whose rights may otherwise be negotiated separately.

The collective licensing approach recognises opt-in at the point of contract: creators contractually give an organisation permission to manage certain rights, and that organisation then licenses those rights on behalf of its members. This approach directly addresses concerns raised by stakeholders that individual opt-in mechanisms are impractical, while also responding to arguments that licensing should be voluntary, since consent is embedded at the contractual stage.

The long-established collective licensing model in publishing offers a scalable, voluntary, market-based solution alongside direct licensing, enabling publishers of all sizes to benefit from the use of their content in generative AI while helping to ensure that content is used responsibly.

Will Crook, Publishers Licensing Services

In the context of AI systems, a collective licensing regime could evolve into automated or agent-based collective licensing systems, where price and usage terms are negotiated dynamically between AI agents within agreed parameters.

SCENARIO

C

DIRECT NEGOTIATION

Property which is in high demand, or for which control is particularly important to a rights holder could be withheld for direct negotiation. This would not rely upon a broader market setting, as it would require the AI company or purchaser to go directly to the rights holder or their representative, but it still demands the background legal and regulatory regime that delivers provenance and transparency.³⁴

³⁴ An example of this is the deal signed by Michael Caine with ElevenLabs to create virtual replications of his voice. See 'Announcing Partnership with Sir Michael Caine to Newly Launched Iconic Marketplace', ElevenLabs Blog, February 18, 2026, Available at: <https://www.elevenlabs.io/blog/announcing-partnership-with-sir-michael-caine-to-newly-launched-iconic-marketplace>

Licensing not levies

In Section C of the Government's report on Copyright and Artificial Intelligence, the report proposed a 'levy system', in which AI developers are permitted to use copyright material through a copyright exception, on the condition that they pay into a pot administered by government in which royalties are then distributed to rights holders.³⁵ A similar arrangement has been proposed in India. However, a system like this does not uphold the principle or rights of rights holders that they have power over how their work is used by others, and on what terms. It would force rights holders to comply, and it would prevent them from assigning a value to their own work. In addition, the UK would miss out on the opportunity to become a licensing hub for data, including the growth stimulated both directly by the licensing market, and in the emergence of new areas of technological and legal expertise that would emerge alongside it. The Government must avoid a levy system which disrupts the market or is tantamount to micromanaging how value is allocated.

Proposal for a Government-led pilot

Through the Government's own position as a rights holder and also a developer, it could test a fair licensing and market mechanism to show it works. [...] We could test it out with Government data sets.

Founder, Tech company

Stakeholders expressed broad support for a Government-led pilot using public and Government-owned data before mandatory rules are introduced. The objective of the pilot is to demonstrate technical and commercial feasibility, rather than to intervene in the licensing market. The pilot would aim to use structured, licensed, machine-readable access to public datasets to support transparency and compliance obligations, while simplifying licensed access to data for AI developers. It would help to establish best practices for how to manage access to datasets, supplying provenance data, reporting, and payment systems.

The pilot would organise and structure access to selected datasets that are already publicly available. Stakeholders suggested piloting a model that enables both human-led retrieval of licensed data and controlled machine access via an agent data layer.³⁶ This is one system with two access modes: a human-facing portal and authenticated API access. Both modes sit within a single controlled framework with shared licensing terms.

³⁵ Department for Science, Innovation and Technology (2026) *Report on Copyright and Artificial Intelligence*. March 18. Available at: <https://www.gov.uk/government/publications/report-and-impact-assessment-on-copyright-and-artificial-intelligence/report-on-copyright-and-artificial-intelligence>

³⁶ Krishnan, Naveen. (2025). 'AI Agents: Evolution, Architecture, and Real-World Applications' (arXiv:2503.12687). arXiv. <https://doi.org/10.48550/arXiv.2503.12687>; Zu, Bin. (2026). 'AI Agent Systems: Architectures, Applications, and Evaluation'. arXiv. <https://doi.org/10.48550/arXiv.2601.01743>

How would an agent data layer work in the pilot?

An AI agent is a system or program capable of performing tasks on behalf of a user or another system in pursuit of a defined objective. Rather than replace the underlying system, an AI agent layer sits on top of an existing system, such as the pilot database of Government-owned data, and provides a structured interface for machine access.

In the context of the pilot, an AI agent layer would allow selected AI agents to access the designated data through a secure interface. It would embed licensing information into that access, log and record usage, and prevent unlicensed scraping.

The layer therefore will test whether structured, machine-readable, licensed access to public data can operate in alignment with transparency and compliance principles set out in this paper.

Participation will be conditional. Developers who access data through this pilot will be required to comply with transparency requirements and copyright law. Failure to comply with transparency or record-keeping agreements will result in the termination of access.

The pilot will be confined to publicly available datasets and will exclude highly valuable, sensitive, or commercially strategic datasets, e.g. NHS and BBC data. The Government will remain the data owner, and the pilot will not create a marketplace or broker relationships with external data owners. It would be a time-limited pilot, expected to run for an initial period of six to 12 months. A transparent communication and evaluation strategy will be built into the pilot from the outset. Given the technical nature of the pilot, communication plans must include clear and accessible language to explain its operation and objectives to a non-technical audience.

WHAT GOVERNMENT MUST DO

The Government must lead on putting the three pillars into practice and making them a success. UK policy on AI and copyright should be developed in close consultation with creators, creative organisations, unions, and other rights holders. Meaningful engagement will be crucial in ensuring policy works effectively in practice and in building consensus and support for the Government's approach.

There is a world where you could have all this provenance technology in place, but if there's no policy support, then no one is really obliged to use it. When it comes to protecting works from being included in datasets, this is an area where we need legislation to get anywhere.

Bence Szilágyi, Garandor

RECOMMENDATIONS

1. Introduce new legislation which:

- Amends laws on 'secondary infringement of copyright' so that mental elements such as knowledge or intent are not required for an AI model to be in breach of secondary infringement.
- Expands primary and secondary copyright infringement protections in AI contexts. This should ensure that existing infringement provisions adequately address the use of copyrighted works in the development and commercialisation of AI products.
- Extends copyright to name, image and likeness.
- Establishes offences for:
 - Ignoring provenance standards or removing provenance tags;
 - Making an AI system available for commercial use without publicly available ingredient transparency;
 - Operating bots without proof of identity.
- Enables the imposition of meaningful fines on companies which breach copyright in the training or operation of AI systems.
- Creates statutory penalties for copyright breaches by AI without requiring a copyright holder to pursue litigation.

2. Designate a regulator with primary responsibility for regulating and enforcing the legal standards of the principles across ownership, consent and transparency, traceability and disclosure, fair licensing and market mechanisms.

The regulator does not need to be new, but it must have the confidence of the creative industries and the tech sector, and have robust powers and a record of using them where necessary.

Previous regulatory approaches have struggled because the bodies involved were not equipped with the necessary mandate, enforcement powers, or resources. A future regulator must have the authority and independence to mandate compliance with technical standards and impose meaningful statutory penalties. Government must decide whether this will be best achieved by strengthening an existing regulator, or appointing a new regulator.³⁷

The regulator will be responsible for:

- Establishing and publishing:
 - Minimum provenance standards.
 - Source data transparency requirements.
 - Minimum interoperability standards.
 - Ingredient transparency standards and expectations.
- Registering crawlers with a legal ID.
- Creating and maintaining public registries of compliant AI developers.
- Ensuring rights holders can access information about how their work has been used, and setting standards for what information is available.
- Enforcing copyright infringements pertaining to AI, including by imposing statutory penalties.
- Penalising failures to uphold other legal standards.
- Categorisation of sensitive AI models and overseeing their compliance confidentially.
- Operating a whistleblower hotline for the anonymous reporting of copyright breaches or failures to uphold standards.

3. Incentivise compliance

Tech developers are worried about the risk of future litigation. A credible and balanced licensing framework would give developers what the current environment does not: clarity and predictability in their legal obligations.

Uptake can be reinforced through whitelisting mechanisms, including:

- Whitelisting compliant firms for eligibility in Government procurement processes.
- Whitelisting for access to the National Data Library and other public data sources.
- Whitelisting as a condition for eligibility to Government grants and innovation funding.
- Whitelisting for access to tax incentives or participation in regulatory sandboxes.

Other mechanisms at the Government's disposal include using the Sovereign AI Fund to invest in start-ups developing new technology which will aid technological solutions to transparency and the licensing market, and using a government scheme to certify 'responsible AI use' that complies with the three principles.

Developers and creatives are also frustrated by proposals for different regimes in different jurisdictions – a lack of regulatory alignment is burdensome and costly. A clear UK framework

³⁷ The Intellectual Property Office does not have the confidence of the creative industries after previously proposing an opt-out mechanism as a resolution, and the ICO has historically been too weak in its enforcement processes.

with reciprocity agreements with other jurisdictions will enable parties from both groups to trade their data in a broader global market that is more aligned. By moving first and defining a credible model, the UK can anchor international reciprocity agreements to its framework. This will help create a level playing field across major markets, reduce incentives for jurisdiction shopping, and ultimately boost the UK's soft power.

The most meaningful incentive is a fair, efficient and robust licensing model: AI tools, rights holders and creators could earn revenue from licensing intellectual property safely. If tech can commercialise content using IP legally, then they will participate.

Benjamin Woollams, TrueRights

4. Ensure enforcement

The most effective and innovative copyright regime will be one where AI developers and rights holders willingly enter into deals. However, setting the scene for a strong and trusted copyright regime requires robust enforcement. Enforcement and incentives shouldn't be seen as opposites, but two complementary levers working in tandem to deliver a trusted regime.

The current mechanism for enforcing copyright breaches of individual lawsuits is slow, expensive and benefits those breaking the law rather than the victims. Each individual breach may be low value, and an individual or SME company may not be able to afford to challenge, but the scale of the breaches transfer huge wealth to large AI companies as a whole.

There should be two routes for enforcing copyright:

- **Private enforcement:** rights holders must retain the ability to pursue legal action directly. The UK could build a fast-track dispute resolution process to relieve creators of the burden and enable both parties to move on quickly. The regulator could play a role in this, for example, as the regulator in a final offer arbitration process.³⁸
- **Regulatory enforcement:** The regulator could act as a backstop for enforcing copyright law in cases where companies are repeat offenders. In these cases, the regulator will be responsible for enforcing fines for unlicensed use regardless of whether a claim has been brought forward. Given historic challenges to regulatory decisions, fines ought to be based on the principle of penalty per breach rather than fines for non-compliance in general.

The regulator will also issue fines for failure to comply with standards as set out.

A robustly enforced regime will foster trust and collaboration amongst consumers, creatives, AI companies, and the wider tech sector.

³⁸ Final offer arbitration processes tend to work effectively in incentivising settlements – in practice, the regulator would be unlikely to be required to act very often.

5. Lead from the front

Led by the Cabinet Office, the Secretary of State for Digital, Media, Culture and Sport, and the new cabinet Minister for AI, the government must take a leading role and make responsible, trusted and universally beneficial AI a strategic priority.

This will require the swift actioning of the above: passing legislative measures through Parliament, establishing the duties of the regulator, and initiating the wider incentives. It will also demand moving quickly with the pilot scheme.

To increase the size of the market and potentially ground a large export market for technical and legal expertise, the government must promote its copyright regime by entering into equivalence or recognition agreements with other countries which have similarly fair compliance regimes, bringing them into alignment wherever possible.

Finally, it must work closely with the CMA to prevent companies from using data and model market dominance to create walled gardens which stifle competition.

6. Retain a role for the courts

The courts will continue to be used for individual litigation (potentially alongside the regulator in any final offer arbitration process) and apply the law to specific fact patterns, and will be responsible for adjudicating any historical claims.

THE OPPORTUNITY FOR THE UK

The reward for the UK of implementing this report's recommendations can be seen in at least five domains:

Creating a robust copyright market for the age of AI. The UK's body of creative work is a unique and highly valuable data asset. A globally recognised rights market will protect and reward creatives, whilst giving AI developers legal certainty, access to high quality data, and streamlined routes to market.

Leading in supporting areas around the copyright market. A copyright market will present spillover opportunities in related sectors, including data preparation and provision, bidding, authorisation and payment processing, and legal support. The UK's strengths in tech and law, combined with the global reach of the English language, present an opportunity to build markets and expertise with global demand.

International influence. The UK could embed international reciprocity agreements in its framework, helping establish a level playing field across major markets, reducing incentives for jurisdiction shopping, and bolstering economic influence and soft power.

The commercial opportunity in trusted and ethical AI. 'Trust' and 'legal certainty' are increasingly commercial necessities in sectors including security, aerospace, medicine, and law.³⁹ A trusted regime for AI and copyright could become a global selling point for the UK, for its broader public-facing AI market and for business-to-business use.

Strengthening the cyber security sector. The UK is a world leader in cyber security and AI regulation with significant research capacity and technical expertise, supported by the AI Security Institute. Advancements in copyright licensing for AI will expand the technology and market for provenance, metadata interoperability, and human and agent verification.

Now is the time to capitalise on these opportunities. High-quality data is increasingly financially unviable for its creators, and support for upholding copyright is widespread amongst the British public.⁴⁰

³⁹ 91% of the UK public feel it is important that AI systems are developed and used in ways that treat people fairly; three quarters agree that social and ethical considerations are good grounds for restricting AI models. Polo, Nuala and Modhvadia, Roshni (2025) *Great (public) expectations*, Ada Lovelace Institute (December 2025). Available at: <https://www.adalovelaceinstitute.org/wp-content/uploads/pdfs/32877/great-public-expectations.pdf?v=1782122600>

⁴⁰ One poll finds the protection of copyright to be the single AI policy position with the broadest support. Stone, Daniel. (2026) *The UK AI Compass: What the UK thinks, feels and wants from AI Regulation*. Available at: <https://diffusion.au/articles/uk-compass-report>

CONTRIBUTORS

Tim Carter, IP, Strategy and global partnerships specialist

Clementine Collett, BRAID Research Fellow, University of Cambridge

Will Crook, Head of Policy and Communications, Publishers' Licensing Services Ltd

Sebastian Cuttill, Head of Parliamentary and Legislative Engagement, News Media Association

Tim Flagg, Chief Executive, UKAI

Guy Gadney, Founder and CEO, Charismatic.ai

Anna Ganley, Chief Executive, Society of Authors

Lucky Gunasekara, Co-Founder and CEO, Miso.AI

Ali Hossaini, Professor, Digital Media and Culture, SOAS, Co-Director, Centre for Philosophy & Art, King's College London

Henrik Kugelberg, Associate Professor of Moral and Political Philosophy, University of Warwick

Joshua Lamerton, Co-Founder, PropTexx

Graham Lovelace, Journalist and generative AI and copyright commentator

Dan Neely, Co-Founder and CEO, Vermillio

Mathilde Pavis, Legal counsel of AI and digital replicas and Agency Director, Replique

Sam Radclyffe, Co-Founder, Strand

Ian Roper, Founder and CEO, Fribbler

Nigel Shadbolt, Distinguished Senior Scientist, Institute for Ethics in AI, Professor of Computer Science, University of Oxford, and Chair of AI@Oxford Research

Zahra Shah, Responsible AI and Frontier Tech Expert, Chair and Founding Member, UKAI Women in AI Working Group

Nicola Solomon OBE, Independent Legal Consultant, Past Chief Executive of the Society of Authors and Past Chair of the Creators' Rights Alliance

Bence Szilágyi, Co-Founder, Garandor

Alexandru Voica, Head of Corporate Affairs and Policy, Synthesia and Advisor, MBZUAI

Benjamin Woollams, Founder and CEO, TrueRights

Further contributors from the creative, media and tech industries offered their input anonymously.

KEY DEFINITIONS

- **AI:** Artificial intelligence refers to computer systems that can perform complex tasks normally done by human reasoning, decision making, and/or creating. The document concerns all forms of AI because all AI requires data. This includes, but is not limited to, Generative AI systems, which can create text, images, and/or other forms of media. Generative AI systems use enormous datasets to identify and reproduce patterns, enabling them to generate new outputs that resemble the data they were trained on.⁴¹
- **Bots:** Automated software applications that are programmed to perform tasks online.⁴² Bots can be used for legitimate functions, such as indexing websites for search engines, or for unauthorised scraping, where large volumes of data are copied without permission or licensing. In this document, “bots” refers primarily to automated agents that collect, process, or generate data for AI systems.⁴³
- **Crawlers:** A specific type of bot that scans and copies online content from websites, which are then used to build datasets, often for search or AI training. For example, Google has a search crawler (Googlebot) and an AI crawler (Google-Extended), which they use to inform their Google Search and AI Overview.⁴⁴ Crawlers can be legitimate when used under license, but they can also be unauthorised and bypass access controls or ignore “robots.txt” directives.
- **Copyright creative content:** Any material protected under the Copyright, Designs and Patents Act 1988, including literary, artistic, musical, dramatic, film, photographic, journalistic, or performance works.

⁴¹ Nasa, (2020) ‘What is artificial intelligence?’ Available at: <https://www.nasa.gov/what-is-artificial-intelligence/>

⁴² Monaco, Nick and Woolley, Samuel C. (2022) Bots. Digital Media and Society Series. Polity Press.

⁴³ Kwon, Diana. (2025) ‘Web-Scraping AI Bots Cause Disruption for Scientific Databases and Journals’. Nature 642 (8067): 281–82. <https://doi.org/10.1038/d41586-025-01661-4>.

⁴⁴ Davies, Jessica (2025) *Cloudflare CEO Matthew Prince on Why Google Must ‘Play by the Same Rules’ as Other AI Companies*, Digiday, October 27. Available at: <https://digiday.com/media/cloudflare-ceo-matthew-prince-on-why-google-must-play-by-the-same-rules-as-other-ai-companies/>

This document has been made possible thanks
to the Accelerator Fellowship Programme of the
Institute for Ethics in AI, University of Oxford

www.oxford-aiethics.ox.ac.uk/

The Stephen A. Schwarzman Centre for the Humanities

Radcliffe Observatory Quarter

Woodstock Road

Oxford

OX2 6GG



INSTITUTE FOR
Ethics in AI



Accelerator
Fellowship
Programme