Written evidence submitted by Dr Nicola Searle[1], Goldsmiths, University of London

[1] PhD. Economics (St Andrews), MSc. International Strategy and Economics (St Andrews), BSFS. International Economics (Georgetown)

Response to consultation by the Intellectual Property Office (IPO), the Department for Science, Innovation and Technology (DSIT) and the Department for Culture, Media and Sport (DCMS)

Consultation: Copyright and Artificial Intelligence

I am Reader (Associate Professor) at the Institute of Creative & Cultural Entrepreneurship at Goldsmiths, University of London where I research innovation and specialise in business models and copyright in the digital media industries. I have worked on this speciality since 2010. In 2010 and 2011, I held an academic placement fellowship at the UK Intellectual Property Office where I developed evidence to inform the Hargreaves Review of Intellectual Property. I provided evidence to the 2022 inquiry by the House of Lords' Communications and Digital Committee titled 'At risk: Our creative future'.

This submission stems from 15 years of research and my role as an academic in informing and influencing policymaking related to copyright. This submission draws from my analysis of the digital creative media industries experiences following the advent of digital technologies at the turn of the century, policy debates in 2010-2014, and the parallels with current copyright and AI debates.

Executive summary

1. The current situation has undeniable parallels to the copyright policy debates around digital media platforms in the 2010s. **Where digital media technologies reduce the cost of *copying* content, AI technologies reduce the cost of *creating* content.** These reduced costs will largely benefit technology firms.

2. In the 2010s, new technology firms replaced more traditional retail outlets and created a more concentrated market, upending long-established value chains and leading to rightsholders' bargaining power dropping. This scenario is likely to repeat, with the AI technology firms enjoying significant growth and digital media losing market share to the competition these firms pose.

3. While option 3 is a considered response, it will not overcome the market forces that will dictate future licensing structures.

4. Attempts at requiring transparency of the contents of training data may be blocked by claims of trade secrets. The government will need to take a strong stance to ensure transparency.

Question 1: Do you agree that option 3 is most likely to meet the objectives set out above?

Option 1 is untenable as it would create far too much uncertainty for both rightsholders and AI developers.

Option 2 would unduly constrain the growth of the creative and media sectors.

Unfortunately, I do not think Option 3 will meet the objectives as related to control and access, since it will not enable a balance between the two. However, it may represent the least-worst option.

My research indicates Option 3 will fail to produce a market that adequately remunerates rightsholders, as it will not overcome the likely market failure of copyright licensing. While legal control may be preserved, control over remuneration will be minimal.

While enabling rightsholders to opt-out of data mining exceptions ("reserve their rights") seems fair on principle, in practice this will lead to a market failure in the licensing market. Option 3 requires firms developing AI technologies negotiate ex-ante licensing agreements with rightsholders, and this involves prohibitively high transaction costs. It will result in the same market failure that persisted in the development of digital media platforms in the 2010s, where **it was not copyright that solved the licensing market failure but rightsholders' loss of market and bargaining power with respect to platforms.** The same dynamic will likely play out between rightsholders and the AI industries.

A core assumption behind Option 3 is that it will meet the objective of enabling licensing between AI developers and rightsholders[2]. However, history does not suggest that a market in licensing will develop in a timely manner. As discussed in my research[3], the 2011 Hargreaves Review of Intellectual Property[4] recognised there was a market failure in copyright licensing as the costs of licensing were prohibitive at the time for the development of consumer-facing digital media outlets such as YouTube, Spotify and Netflix. Post-Hargreaves copyright reforms had very little impact on the ease of licensing; rightsholders instead spent many years attempting to negotiate, individually, licensing contracts with consumer-facing digital media outlets. Those outlets were sometimes uninterested in directly licensing and/or unwilling to offer licensing terms that were favourable to rightsholders.

---

[2] "AI developers are able to train on large volumes of web-based material without risk of infringement. Importantly right holders are also able to control the use of their works using effective and accessible technologies and seek payment through licensing agreements." (66). "Create licensing opportunities for right holders and reduce barriers to remuneration" (70)

[3] Searle, N. (2017). Business Models, Intellectual Property and the Creative Industries: A Meta-analysis. CREATe Working Paper Series, 201709.

[4] Hargreaves, I. (2011). Digital opportunity: A review of intellectual property and growth: An independent report. UK Intellectual Property Office. http://dera.ioe.ac.uk/16295/7/ipreview-finalreport_Redacted.pdf

The likely progression of the licensing of copyrights will be increasingly less favourable outcomes for rightsholders, as my analysis of the evolution of music licensing agreements from the late 1990s to 2020 indicates. The copyright policy debates began in the late 00s, but it was not until the late 10s that the market failure in licensing was resolved (e.g. Spotify was only able to license a full catalogue in 2017 via direct licensing with individual music record labels). In the interim, there was a combination of a) ambiguous or minimal licensing (e.g. pre-Content ID YouTube); b) the development of technology firms as competitors to the creative firms as these firms generated new content (e.g. Netflix originals); and c) the use of existing rightsholders' content to inform that generation of new content and the development of the technology firms' offering to consumer (e.g. YouTube's distribution of copyrighted content to gain market share).

This history of the progression of licensing terms in the 00s and 10s is important because the parallels in the present situation are undeniable: licensing for AI development is hotly debated (a); the development of technology firms as competitors as they create novel content – which essentially describes how AI works (b); and the use of existing rightsholders' content to inform the generation of content and develop the technology firms' offerings (e.g. TikTok's AI generation of content) (c). In both the 'digital' debates and the current AI debates, the ambiguity of copyright policies and the fast development of technologies has made it difficult for rightsholders to adapt and for technology firms to operate within copyright policies.

The **market failure** arising from points a, b, and c above for the 00s and 10s digital copyright debates **was addressed not because copyright policy changes influenced transaction costs or licensing was otherwise encouraged, but because the market fundamentally changed** – new technology firms replaced more traditional retail outlets and created a more concentrated market, upended long-established value chains and rightsholders' bargaining power dropped. By the mid to late 10s, the traditional licensors (traditional retail outlets) of the music industries were no longer viable and music rightsholders had no other options but to license to technology firms, who had more bargaining power as this firms had vast amounts of users and there was less competition from other retailers.

AI technologies change the role of copyright licensing. **Content used in the development of AI technologies is substitutable in a way that the previous digital technologies were not, and this will have impact on the incentives for licensing.** In the 2000s and 2010s, technology firms were interested in licensing specific content, with low substitutability (e.g. consumers wanted a service that offered a specific music track, not a generic substitute.) The content required for the development of AI technologies is instead substitutable, as the content can be more generic (e.g.

consumers want a service offering a style, not a specific music track.) The incentives for AI to license specific copyrights are weaker when there are more substitutes.

History is likely to repeat itself. It will take some time for a new licensing regime to develop, and the market failure of licensing will persist. In the meantime, the AI technology firms will enjoy significant growth, and digital media will lose market share to the competition these firms pose. The technology firms may find substitutes to the content of existing rightsholders.

Given the speed at which AI technologies and businesses are developing, the government should consider that the likely repeat of the lag in licensing practices between a policy recognition of a licensing market failure and a market solution will involve years of market failure. Option 3 does not overcome this. However, Options 1 and 2 are even less likely to do so in a way that supports the creative industries.

Question 3: Do you support the introduction of an exception along the lines outlined above?

The rights reservation is akin to the distinction between opt-out and opt-in licensing regimes we see in some digital media consumer-facing business models. **Opt-out models (YouTube, social media) generally have lower levels of remuneration for rightsholders and involve costly copyright infringement monitoring.** Rights reservation, as essentially an opt-out licensing approach, will significantly shift the burden of copyright enforcement to rightsholders.

I have developed a table of the progression of estimated licensing terms from the late 1990s to 2020 in the music industries as music retailers (licensors) changed (see Appendix 1, Table 1)[5]. Over this period, there was minimal legal changes to the copyrights of the music labels in their legal right to be compensated for their copyrighted content, but market changes still resulted in markedly different licensing terms and revenue outcomes. The traditional licensing contract is associated with traditional retail licensors, in which case the license is signed before the content is distributed (ex-ante) and rightsholders have agreed for the copyrights to be licensed (opt-in). These practices are associated with the highest revenues for rightsholder. In contrast, licensing contracts signed after (ex-post) the copyrighted content is distributed, require the rightsholder to identify, monitor and enforce their copyright (opt-out) and these contracts are associated with lower revenues for the rightsholders.

[5] Searle, N., Baden-Fuller, C., (2023). Appropriability and Competition between Complementors: Copyright and Music Business Models. In Academy of Management Proceedings (Vol. 2023, No. 1, p. 18801). Briarcliff Manor, NY 10510: Academy of Management.
https://journals.aom.org/doi/10.5465/AMPROC.2023.18801abstract

Please note that the information presented in Table 1 is an estimate based on publicly available data. While I cannot say these are fully accurate for the specifics of otherwise confidential licensing contracts, I am confident that the overall trend of deteriorating conditions for rightsholders is accurate.

The proposed approach will significantly shift the burden of copyright enforcement to rightsholders. In addition, as the content used to train AI systems is hidden, detection of copyright infringement will be extremely difficult and costly. As I observe with current opt-out approaches in technology firms (e.g. social media platforms like Facebook and Tik-Tok, and user-upload platforms such as YouTube), this will weaken the bargaining power of rightsholders. In licensing negotiations, rightsholders will have less information on demand and consumption of their copyrights and face higher enforcement costs relative to previous iterations of licensing.

Question 4: If so, what aspects do you consider to be the most important? If not, what other approach do you propose and how would that achieve the intended balance of objectives?

The consultation does not consider **copyright levies**, which is an established policy related to copyright in some jurisdictions and already existing in the UK related to gambling and a proposed levy for stadiums. Levies could be attached to AI technologies to be paid into a collective fund for the UK media and creative industries. There are some downsides as it raises the cost of AI technologies compared to other countries and assumes all AI development involves unlicensed content. It may also pose a barrier to the development of business models based on AI technologies, such as open source. To mitigate the potential negative impacts on innovation incentives, the levy could be attached to revenues arising from copyrighted training data rather than when it is used in training. For example, the levy could come out of advertising revenues (when AI-generated content is supported by advertising), subscriptions or retail prices.

A major advantage of this approach is its relative simplicity. It involves minimal transaction costs and overcomes the technical and practical challenges to monitoring the copyright of inputs. Additionally, the revenues described above are already subject to taxation and therefore there are existing mechanism to collect the levy.

Question 12: Does current practice relating to the licensing of copyright works for AI training meet the needs of creators and performers?

As discussed above, as technology firms have gained market share and evolved their business models, rightsholders have progressively had lower bargaining power and lower revenues. As traditional retail outlets failed, rightsholders had very little option except to contract with the emerging technology firms. The same will likely happen

again, but the situation for rightsholders will be exacerbated as **technology firms will have even weaker incentives to license copyrighted content** as AI technologies can be built with alternatives to rightsholders' content.

As AI reduces the cost of content creation, technology firms will move further up the value chain. A firm using or developing AI could pursue vertical integration by performing all aspects of the value chain and creating, distributing and selling content (see Figure 3 below). These firms may not need to license. Additionally, existing rightsholders will be licensing to their soon-to-be competitors, putting the former at a great disadvantage.

The emergence of the technology firms as competitors, not just retailers, to content creators will exacerbate existing concerns about creators, performers and other rightsholders' ability to negotiate fair contracts. As copyright involves the fundamental right to contract and this is a market in flux, setting licensing or remuneration terms in law might both violate that contract and be ill-informed. The development of an industry standard would seem advisable albeit difficult to develop.

Visualisation of the evolving creative content to consumer value chain

The process of how AI technology will enable firms currently not involved in content creation to become competitors, with a significant cost advantage, to current content creators is visualised below in Figures 1-3. Licensing occurs between actors.

Figure 1: Actors in the traditional, largely pre-digital, value chain of licensing (pre-2000s)

Figure 2: Actors in the initial digital value chain of licensing (2000-present)

Figure 3: A vertically integrated single actor in the AI value chain (emerging)

Firm using or developing AI: content creation, distribution and sale

Question 17: Do you agree that AI developers should disclose the sources of their training material?

Yes, transparency is necessary for fair bargaining conditions. The data collected by firms who aggregate content (e.g. the user data Prime Video collects) grants those firms a significant data capital, which can be used to create further revenues. Current content creators are unable to amass that data capital, despite the fact the sale of their content directly contributes to data capital. Transparency on what copyrighted content is used is one of the few ways in which this power imbalance can be adjusted

Question 20: What is a proportionate approach to ensuring appropriate transparency?

In addition to my work on copyright and business models, I am also an expert on innovation aspects of trade secrets. My research indicates that firms can and do use trade secrets to shield their information and data from oversight – both from regulations (e.g. disclosure requirements) and other firms (e.g. audits from collaborators or licensees.) By claiming its information is a trade secret, a firm may be able to reduce transparency significantly**. Transparency is essential to fair bargaining conditions, and if this is impeded by claims to trade secrets, rightsholders will be further disadvantaged in their ability to monitor whether their content has been used.** The government should nonetheless consider how regulations can enable transparency and disclosure.

A further challenge is that firms often overclaim how much of their information qualifies as a trade secret. Assessing whether the information is actually a trade secret is usually done via the court system, which will mean significant cost to UK taxpayers.

Appendix: Table 1: Progression of **estimated** licensing terms from in the music industries (late 1990s-2020) based on the business model of the licensor (retailers including technology firms)

Licensor Business model	Total revenue per unit of consumption[6]	Estimated licensor revenue share	Contract timing and structure
Traditional retail (physical sales)	Highest Per copy	~20%[7]	Ex-ante (opt-in); Direct sale of end product
Peer-to-peer infringement (e.g. Napster)	Lowest Per copy	100% although often no revenues	Ex-post (opt-out) although licensing non-existent[8]
Download digital retail (pay-per-download, e.g. iTunes)	Medium Per copy	~30%[9]	Ex-ante (opt-in); Direct licensing
Subscription streaming (e.g. paid-for Spotify)[10]	Low Per stream	~30%[11]	Ex-ante (opt-in); Direct licensing, equity and kickbacks[12]
Ad-funded, user-uploaded streaming (e.g. YouTube)[13]	Low Per stream	~45-55%[14]	Ex-post (opt-out); Direct licensing

[6] This comparison is difficult to do consistently as there are differences between what a sale consists of between types of licensors (e.g. sale of file, sale of one stream etc.) See (Towse, 2020) for a discussion.

[7] Donovan, N. (2013, August 26). If CDs cost £8 where does the money go? BBC News. https://www.bbc.com/news/magazine-23840744

[8] Some labels signed contracts with peer-to-peer services, but they had only limited success and are excluded from analysis.

[9]Sholomove, I. (2019, September 4). Make Your Music Make Money: Understanding Royalties. Attack Magazine. https://www.attackmagazine.com/reviews/books-dvds/make-your-music-make-money-understanding-royalties/

[10] Excludes services free to the user (e.g. ad-funded.)

[11] Generally speaking, Spotify retains 30% of revenues although this can vary. Weatherbed, Jess (28/01/2025), "Spotify says that streaming has made the world 'value music'" The Verge https://www.theverge.com/news/597155/spotify-streaming-adds-value-music-payouts Accessed 24/02/2025.

[12] Details for labels' contracts with streaming services are subject to confidentiality agreements, therefore precise details are unknown. This lack of transparency has been criticised by the industry, particularly by artists. For Spotify, equity and revenue shares with big labels are well-known. These contracts also prohibit Spotify from 'turning into a label' by owning rights (see Sisario, B. (2018, September 6). A New Spotify Initiative Makes the Big Record Labels Nervous. The New York Times. https://www.nytimes.com/2018/09/06/business/media/spotify-music-industry-record-labels.html)

[13] Excludes paid-for services.

[14] It is difficult to pin down the revenue split on YouTube as it is complex. A 2018 BASCA publication notes, "After a 20% cut from Google Ads YouTube takes another 45% of all advertising income on its platform." BASCA. (2018, January 1). BASCA delivers letter to the Bulgarian Ambassador to ask for support in the protection of copyright. BASCA. https://basca.org.uk/2018/09/13/basca-along-with-fac-mmf-mpg-and-mu-launches-council-of-music-makers/ More recent publicly available information suggests 45-55% of ad-revenues are shared between creators (uploaders of content) and music copyright holders. Google, (24/02/2025), YouTube partner earnings overview, https://support.google.com/youtube/answer/72902?hl=en#zippy=%2Cwhats-my-revenue-share, accessed 24/02/2025. This also depends on the type of content, with revenue splits between different between long and short-form content. Khalid, Amrita (28/03/2024), A quarter of YouTube's paid creators are earning money with Shorts, The Verve, https://www.theverge.com/2024/3/28/24114031/youtube-shorts-partner-program-ad-sharing-revenue accessed 24/02/2025