**NUJ submission to the UK government’s Copyright and Artificial Intelligence consultation.**

**February 2025**

The National Union of Journalists (NUJ) is the voice for journalism and journalists in the UK and Ireland. It was founded in 1907 and has more than 22,000 members working in broadcasting, newspapers, news agencies, magazines, book publishing, public relations, photography, videography and digital media.

The union represents among many, the reporters, photographers and analytical writers who inform the public and the sub-editors who fact-check their work and turn words and pictures into knowledge. The NUJ is not affiliated to any political party.

**Executive Summary**

* The NUJ welcomes this consultation as an important opportunity to relay the views of journalists significantly impacted by the actions of AI developers and their technologies. We recognise opportunities presented by AI but vehemently oppose the development of technologies that involve breaches to the intellectual property rights of our members, where their works are used without consent.
* We are therefore concerned that government has indicated a preferred approach that has received widespread opposition from journalists and creators, action that must have been anticipated, not least because a wider text and data mining exception was met with staunch opposition in 2023 with plans dropped by the previous government. We do not agree that there is ambiguity around the current copyright regime.
* The role of trusted, accurate journalism within our democracy cannot be overstated. Democracy is not possible unless members of the electorate have access to reliable news on which to base their voting decisions. That cannot happen without ethical, independent journalism. The NUJ Code of Conduct provides a commitment to producing ethical and accurate journalism, including visual media. Journalism by humans must remain core.
* The union believes journalistic images should not contain AI-generated elements or deletions of parts of the image which were present when the image was taken. It is vital that people can trust the accuracy and integrity of journalistic images and footage that is published or broadcast by media outlets, wires and agencies.
* Generative AI outputs should always be labelled as AI generated, ensuring audiences are informed about the content they engage with, used to inform their decision making. We need regulation that ensures this action is being undertaken as this is not being adopted voluntarily across all organisations. Government should mandate that all content including AI-generated material is clearly labelled as such.
* We do not indicate a preference for a policy proposal in this consultation due to concerns with approaches outlined. In our submission, we detail key concerns with government’s preferred option and outline measures required to ensure the copyright-protected works of journalists and creators are respected. The current rights infringements experienced by our members must not be allowed to continue; the text and data mining exception would reward 'big tech’ resulting in financial gain for AI developers at the expense of journalists.
* We stress there must be greater transparency requirements on developers to provide data that discloses the sources of their data used in the pre-training, training and fine-tuning processes and that rightsholders must be able to access this information to identify where their works may have been used.
* The NUJ supports the Trades Union Congress’ view that the UK government should ensure unions are able to introduce legal proceedings in the collective interest of their members, and the law should be modified to enable this where necessary.
* We strongly disagree with the proposal outlined for a TDM exception and do not believe there is a need for it.

**NUJ response**

Free markets are impossible unless buyers and sellers have access to truthful independent information, and journalism is the prime means of transmitting such information.

The UK’s copyright framework has sustained journalists and other creative workers for centuries, through the fundamental principle that the author of a creative work has the exclusive right to authorise use of that work, whether directly as a freelance or through a contract of employment.

Concerning the importance of the “creative industries” to the UK economy, the NUJ adopts views on their value outlined in submissions by the British Copyright Council and by the Creators’ Rights Alliance of which we are members.

Mass breach of copyright in news reporting has played a substantial part in creating the crisis that now faces the news industry. Monopolistic digital advertising provision that depends on purloined news reporting to draw eyeballs has been particularly destructive. The EU, Canada and Australia have recognised this with measures designed to force the internet corporations to negotiate licences – which those corporations continue to resist.[[1]](#footnote-2)

More recently, digital corporations – including those that now face competition law proceedings internationally for their seizure of the advertising market – have “scraped” words and images they can find to “train” their large language models (LLMs), which they market as “artificial intelligence”. They have done so without permission or recompense.

Though this consultation deals only with the copyright and authors’ rights implications of this development, the NUJ cannot comment on the subject without noting the implications of LLMs for journalistic ethics and for the proliferation of misinformation and disinformation.

There is a matter that is sometimes – as for example in the text of this consultation document – not recognised as an aspect of copyright law. Ultimately, the authenticity of a journalistic work is guaranteed by the byline on it: the named journalists take responsibility for their work. As the NUJ has often observed, the theoretical right to demand attribution (a byline) and to defend the integrity of a piece of work against distortion (for example by the distribution of altered copies) is far too hard to enforce in UK law. In practice, the experience of members of the NUJ is that these so-called “moral rights” to identification and integrity are applied only as aggravating factors causing an uplift to damages for clear breach of their economic rights. The provision against false attribution is for this reason rarely used: in order to bring a case, someone who had a work attributed falsely to them, to the detriment of their honour or reputation, would have to establish “quantum” for the case. What would be the fairly negotiated price for publication of a work that you did not create and whose publication is damaging to you? Do you have to establish the cash value of your good name to bring a civil suit for that amount?

Even enforcing the economic rights is of course difficult for the freelance journalists and other authors who typically retain copyright in their work. These are almost all sole traders, and typically the infringers of their copyright are corporations that enjoy the services of fleets of lawyers.

LLMs are machines for generating altered copies of the works on which they are trained.

LLMs are also notorious for “confabulating” responses (called by those who over-estimate their intelligence “hallucinating”). They will invent source references for untruths, often using the names of authors who are active in the field of the question put to them.[[2]](#footnote-3)

LLMs will output “works” that are recognisable as being based on the work of specific writers and photographers. They will make explicit claims that named individuals have written words that they have not: claims that are contrary to the honour or reputation of those named.

This has two effects, one of global importance and one very particular to NUJ members. It will further diminish trust in news reporting and will damage the reputations of individual journalists. These are just some reasons why the only basis on which use of works to train LLMs may be permitted is with explicit prior consent and payment.

The NUJ notes that when government launched this consultation on 17 December it was not accompanied by publication of an economic impact assessment. We are concerned government’s approach was designed to promote one response, which is a policy promoted by big tech organisations seeking to pursue profits and circumvent existing laws. Government’s press release [[3]](#footnote-4) on 13 January states “the Prime Minister is throwing the full weight of Whitehall behind this industry by agreeing to take forward all 50 recommendations set out by Matt Clifford in his game-changing AI Opportunities Action Plan.” Whilst recommendation 24 on text and data mining reform references the UK Copyright and Artificial Intelligence consultation launch, in our view, the press release inferred a decision had already been made regarding proposals, causing confusion and concern that this was government policy.

The NUJ would have considered option 1 in the consultation document had it been accompanied by the recognition of transparency measures as outlined in option 3. It is disappointing that robust measures to ensure developers are transparent about the works their models are trained on, and measures for rightsholders either individually or collectively to easily reserve their rights, were not included in the first option presented, instead linked to a text and data mining exception. We cannot in good faith express a preference for any of the four options presented in the consultation, given the above and the way the options are expressed.

We have grave concerns over the implications for journalists and creators if government’s proposed approach as outlined in option 3 is pursued. A data mining exception to copyright law would improve access to content by AI developers but be at the detriment of rightsholders facing a considerable burden in the ability to opt-out of the use of their works.

It is our view that copyright law should be strengthened to ensure the breaches that have occurred can no longer continue at scale and speed, with rightsholders instead empowered to seek redress through clear routes where rights infringements occur. Journalists who provide consent to the use of their works to train LLMs should receive fair remuneration for the use of their material. For too long, journalists and others within the creative industries have had unfavourable contract terms that seek to use the creativity of individuals without fair payment or conditions. We urge attempts to rectify this inequality in the bargaining power of our members.

There must be clear and enforceable sanctions for those breaching copyright law, and processes for rightsholders must be accessible and user-friendly. Without adequate deterrents for companies many of whom have considerable financial resources, it is journalists, creators, and freelances who bear the burden of infringements and the emotional and economic toll. Government could consider the approach adopted by the Competition and Markets Authority and issue penalties of 10% of annual income for infringements.

Existing copyright law is clear on the scope of what AI developers must adhere to, yet many have chosen not to abide by legal requirements. We do not believe the weakening and devaluing of the contribution of journalists and creators through the proposal of a data mining exception should be pursued in response.

There should be an opt in mechanism, either for individuals or as part of collective licensing, revocable by the creator at any time. An opt-out process would not promote transparency or trust. Some NUJ members receive payment for revenues from license organisations including the Authors’ Licensing and Collecting Society (ALCS) and the Design and Artists Copyright Society (DACS). Collective management organisations will likely provide useful information and thus be well-placed to inform licensing approaches.

Licensing arrangements should include representative organisations and agreements that allow journalists to receive fair payment for their work. As publishers agree licensing deals with AI companies, there must be recognition too, of the need for fair remuneration for freelance journalists whose works are included in deals for data sets to train/develop technologies.

**We insist**:

1. That there must be a scheme to compensate journalists (and other authors and performers) for unauthorised use of work to date by digital corporations.
2. Such a scheme must fairly apportion compensation between journalists – whether employed or freelance – and their publishers – except in cases where the payment is due entirely to a freelance author or photographer.
3. That the only basis on which such uses may be permitted in future is with explicit prior consent. The proposal for an “opt-out” regime is identical to the policy that Google (now Alphabet Incorporated) has pursued since its inception: namely that there shall be no controls except the ROBOTS.TXT file on a website. We recall the strenuous efforts made by Google and other internet corporations to kill news industry proposals for a comprehensive and flexible standard for machine-readable permissions. The ROBOTS.TXT mechanism clearly does not work. It does not in practice allow opt-out at the level of individual works. It does not in practice allow the owner of a website to opt out of being scraped by new “robots” since their existence is announced primarily by the site log recording the fact that they have just scraped the entire site.
4. As search engines based on deterministic indexing are increasingly replaced by LLM-based tools, opting a website out of being “crawled” to train LLMs effectively turns into opting it out of being findable. This replicates a tactic used by Google in its battle against regulation requiring it to compensate journalists (and media owners) for the contribution the work makes to its advertising revenues: if they do not agree to it making uncompensated use of the work, that work will be excluded from search and invisible to members of the public who do not already know of its existence.
5. No new exception to copyright is required or permissible.
6. Copyright rules relating to AI should apply fairly to all. The purpose of an AI model, or the size of an AI firm should not alter this.
7. It is not possible to craft an exception of the type suggested that does not fall foul of the UK’s international treaty obligations, in the shape of the “three-step test” set out in the TRIPS agreement,[[4]](#footnote-5) which we remind readers provides that exceptions are permissible *only*:
   * in certain special cases;
   * which do not conflict with a normal exploitation of the work or other subject-matter; and
   * do not unreasonably prejudice the legitimate interests of the right holders.
8. Further, it is hard to see how the proposed mechanism for “reserving rights” would not breach Article 5(2) of the Berne Convention, which holds that: “The enjoyment and the exercise of these rights shall not be subject to any formality...”[[5]](#footnote-6)
9. Any requirement to opt out would, besides breaching the two fundamental components of UK international commitments noted above, generate problems for the rule of law domestically. How could an opt-out be enforced? By an individual author suing Meta (Facebook) for the amount they would have charged for permission to train an LLM on their work, had they negotiated such a licence? Is it time to consider how a regime of statutory damages covering all abuses of the economic and moral rights in the works of an individual human author might be developed?
10. We welcome government’s recognition of the importance of transparency and stress the current imbalance between rightsholders and technology companies must be urgently addressed. There should be a “transparency” requirement on the owners of LLMs and on “AI developers” to disclose the sources of their training material.
11. That “transparency” requirement must provide highly “granular” data. It must for example provide a free interface that allows individual authors and performers to query the use made of all their works, searching either by author name or by the content of an individual work.
12. It must require detail on the URLs accessed by web crawlers deployed by AI developers or their operators, include the text and data used at stages including the pre-training, training and fine-tuning and the timeframe of data collection. These requirements were among those passed by peers when presented as amendments to the Data (Use and Access) Bill in the House of Lords in January, demonstrating the strength of feeling on the issue.
13. Transparency requirements must extend too, to synthetic data to address the practice of the manipulation of copyright works to create synthetic data sets to train models.
14. If government proceeds with an opt-out mechanism, it would be absolutely vital to the entirety of the UK “creative industries” that this mechanism relate only to text and data mining and in particular that it does not affect any of the existing exceptions that are linked to licensing schemes, as in education. The effects of the alternative are illustrated by the near-total collapse of Canada’s educational publishing industry following the passage of an ill-advised exception regime.
15. That the concept of “not-for-profit” text and data mining is of no use is illustrated by the history of image generation systems. A large part of the corpus of images on which these were trained was scraped by the LAION initiative.[[6]](#footnote-7) The owners of this claim that it gathered images for non-profit research under the German implementation of the EU exception for non-profit text and data mining. LAION was used to train the Midjourney and Stable Diffusion machine-learning image generators, among others; but both appear to be heading toward being for-profit operations. More recently OpenAI has, we understand, been moving from an open source to a for-profit model.
16. The NUJ does not believe there is any utility in repealing section 9(3) of the Copyright, Designs and Patents Act 1988 concerning “computer-generated works”. It could have had unforeseen consequences, but the only case law that we can find seems to have ruled these out and it appears to be doing no harm; repealing it could have unpredictable unforeseen consequences.

If an opt-out approach is pursued despite widespread opposition, there must be recognition of the technical limitations, financial and practical burdens with efforts made to remove these for rightsholders. An opt-out approach should not result in unfair barriers to journalistic content being found online; there must be proactive approaches to ensure an opt-out approach meets emerging challenges regarding web crawlers; varying types of material including images and text must be possible to opt-out with ease.

On key principles, there have been efforts in the EU to address issues relevant to rightsholders and AI developers, including on transparency. There are outstanding concerns regarding transparency however, as AI firms resist disclosure of training data. We note too, that under the EU’s opt-out scheme it is not possible for rightsholders to successfully opt their works out of training at a level required to protect their rights.

First-mover advantage means that the chances of a UK-owned industry succeeding to the same level in competition with the US-based corporations that have invested dollars and precious carbon emissions (and our members’ work) thus far are slim, but irrespective, should not mean the wholesale disregard of the rights held by journalists and others across the creative industries.

“Cloud” technologies don’t have a geographical location. Many of the same US corporations building LLMs are at the same time propelling users to move everything “into the cloud”. A data centre’s users can “mirror” elsewhere everything they do using the computational resources it contains.

If developers can pay for electricity and infrastructure required to ensure the smooth-running of technologies, they must not evade responsibility in ensuring fair remuneration for data where consent is granted.

Government recognises the creative industries are worth £125 billion to the UK economy. It is crucial therefore, that the voices of the sector and those representing their interests including trade unions are recognised and considered as part of this consultation process. The NUJ is not opposed to the use of AI within journalism but does recognise this must only occur with technologies used as an assistive tool and crucially, retaining human oversight. Such an approach will ensure inaccuracies can be identified, content generated is context appropriate with bias and errors identified, and that journalism remains trusted by audiences.

We are acutely aware of the lobbying efforts of big tech who have amassed large amounts of wealth and stand to gain more if they are successful in their pursuit of a broad text and data mining exception including for commercial purposes.

We urge greater consideration by government of the detrimental impact option 3 would have on rightsholders. Copyright belongs to rightsholders and must not be given away – we assert that it is not the place of government to do so.

1. <https://www.londonfreelance.org/fl/2410nz.html> [↑](#footnote-ref-2)
2. <https://www.londonfreelance.org/fl/2303ai.html> [↑](#footnote-ref-3)
3. <https://www.gov.uk/government/news/prime-minister-sets-out-blueprint-to-turbocharge-ai?utm_source=substack&utm_medium=email> [↑](#footnote-ref-4)
4. <https://www.wipo.int/wipolex/en/treaties/details/231>  [↑](#footnote-ref-5)
5. Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979) Art. 5(2) <https://www.wipo.int/wipolex/en/text/283698#P109_16834>  [↑](#footnote-ref-6)
6. <https://laion.ai/> [↑](#footnote-ref-7)