

Artificial Intelligence and Intellectual Property

Copyright and Patents

This response has been authored by the Publishers Association to inform the IPO's call for views on AI and IP covering: copyright in works made by AI; text and data mining using copyright protected material; and patents for inventions devised by AI.

The options for computer generated works, text and data mining and patent inventorship are summarised in the following tables.

Computer generated works	
Option 0	Make no legal change
Option 1	Remove protection for computer-generated works
Option 2	Replace the current protection with a new right of reduced scope/duration

Text and Data Mining (TDM)	
Option 0	Make no legal change
Option 1	Improve licensing environment for the purposes of TDM
Option 2	Extend the existing TDM exception to cover commercial research and databases
Option 3	Adopt a TDM exception for any use, with a rights holder opt-out
Option 4	Adopt a TDM exception for any use, which does not allow rights holders to opt out

Patent Inventorship	
Option 0	Make no legal change
Option 1	"Inventor" expanded to include humans responsible for an AI system which devises inventions
Option 2	Allow patent applications to identify AI as inventor
Option 3	Protect AI-devised inventions through a new type of protection

Section A – questions arising from this consultation

Copyright – computer generated works (CGW)

1. Do you currently rely on the computer-generated works provision? If so, please provide details of the types of works, the value of any rights you license and how the provision benefits your business. What approach do you take in territories that do not offer copyright protection for computer-generated works?

Our members are in the early stages of exploring the potential uses of this technology and have a keen interest in the functionality of the CGW provision. As a notable example, Springer Nature published ‘Lithium-Ion Batteries: A machine-generated summary of current research’ in 2019. The book, which took 1.5 years to develop and produce, demonstrates how content generation by AI could develop in the coming years.

This being said, the use of AI technology to develop academic and literary content without any human input is still nascent. Evidence outlined in the 2020 Frontier Economics report “People Plus Machines: The Role of Artificial Intelligence in Publishing” clearly found that it will take longer to develop market-ready AI-generated content for mass consumption. While there are lauded examples of AI-produced content, there is industry-wide consensus that we are a long way from CGWs approaching the standard of human creativity. With this context in mind, we do not advocate for a change in legislation at this time and would instead advocate for legal certainty and stability to allow the technology to develop further.

In event that the IPO does move ahead with new legislation, it should be aware of the following business practices and the associated risks of changing the existing CGW provision:

Currently, publishers more commonly use AI to support a human author’s endeavours. While this is beyond the intended scope of any proposed change in legislation, we would emphasise that AI-supported titles require substantial publishing intervention and thus attract significant investment and creativity from publishers. It is therefore crucial that any new legislation clearly defines what is meant by a “computer-generated work” and an “AI-assisted work”, including the extent to which these involve at least some human input, interaction or investment.

Indeed, similar consideration and clarity must be applied with regards to original literary works. The IPO has already made some assurances on this topic (i.e. p. 18 of the impact assessment), but we would urge the careful transposition of these assurances in the event new legislation is brought forward. By way of an example, although publishers are exploring the role of AI technologies in generating audiobooks, they would not want the transposition process to undermine the protections afforded to the foundational literary text, and in turn the rationale for investment in that original content.

Similarly, any legislative change should be accompanied by provisions and/ or guidance that a CGW protection cannot be used as a justification to bypass other forms of copyright – i.e. feeding a machine with copyright protected content in breach of licence terms.

Finally, consideration must be given as to how any new CGW provision may work alongside a newly expanded TDM exception. In the event that the IPO also decides to expand the existing TDM exception, there could be a significant increase in the number of CGWs acting as derivatives to original copyright-protected works. If these derivatives are all given protection, it could be difficult to protect the original works against new derivatives that have CGW protection and were generated from the use of TDM.

2. Please rank these options in order of preference (most to least preferred) and explain why.

Our most preferred option is Option 0, make no legal change. Given that these technologies are still in their infancy, and given the evolving dependencies between human and AI authors, we would urge caution in making alterations to the legislation at this stage. The impact assessment does not provide a clear evidence base for making the change, nor does it make a thorough, quantitative assessment of what harms or benefits might result from government intervention.

Indeed, we would again direct the IPO to the 2020 report “People Plus Machines: AI and the Future of Publishing”, which confirmed that protections for AI-generated content and legal responsibility for AI-generated works are highly relevant to publishers’ investment decisions. The report makes the case for legal certainty and stability regarding UK IP law, so that the investment cases both for the development of AI products and the UK content that underpins quality outputs is preserved.

While we do not advocate for any change to the existing system, our second preferred option is Option 2, replace the current protection with a new alternative. This would at least continue to offer some incentive for the production of CGWs.

Our least preferred option is Option 1, the removal of all protection for computer-generated works. We believe this would undermine investment in CGWs when the market is still in its early stages – thus risking the UK’s leading position in developing such works.

3. If we introduce a related right for computer-generated works, as per option 2, what scope and term of protection do you think it should have? Please explain how you think this scope and term is justified in terms of encouraging investment in AI-generated works and technology.

There is no evidential reason to remove or reduce the current period of protection. As outlined above, it would be premature to make changes to copyright law at this stage, and we suggest that stability and consistency in the UK's legal framework should be prioritised.

4. What are your views of the implications of the policy options and of AI technology for the designs system?

N/A

5. For each option, what are your views on the risk that AI generated works may be falsely attributed to a person?

We would urge the IPO to ensure human authors can continue to receive protection for AI-assisted works, without accusation of “falsely attributing” a CGW to a person.

We would note that copyright legislation is not necessarily the best and only vehicle for protecting against false attribution, irrespective of whether the author is human or computer. We would suggest there is an intersection with fraud and consumer laws that we anticipate could be utilised.

Finally, we would note that this issue perhaps intersects with the question of legal responsibility. This matter is particularly important where there is scope for AI algorithms to introduce bias or inaccurate representation when summarising content (which could also infringe an author's moral rights) or making editorial recommendations. The government should give these risks its due consideration if it chooses to legislate.

Copyright – text and data mining (TDM)

6. If you license works for TDM, or purchase such licences, can you provide information on the costs and benefits of these? For example, availability, price-point, whether additional services are included or available, number and types of works covered by the licence etc.

The Publishers Association does not offer or license works for TDM or purchase such licences directly. However, the licensing of works in both directions is common practice across our membership, particularly in the education and academic sector, to permit and undertake text and data mining. Our members are at the forefront of digital innovation and routinely manage TDM licensing, tech partnerships, and increasingly invest in AI technology themselves at all points in the supply chain.

The publishing industry firmly asserts that licensing solutions remain the best tool for facilitating AI development. Licensing arrangements provide flexibility, and also provide rights holders and data users certainty with respect to their rights and obligations. Experience has shown that licensing enables copyright to adapt to meet new markets and changes in technology, and the submission made by the Publishers' Licensing Services (PLS) offers more evidence in this regard.

To be absolutely clear, where an AI provider is not able to access content, because it is not publicly available or covered by the non-commercial TDM exception, then licensing models are always available by way of open and transparent commercial negotiations. As far as we are aware, licences have never been refused for TDM access.

We would point to Elsevier's Scopus as a case study of how licensing can be used effectively by data acquisition companies to collect copyrighted material. Scopus is a world-leading citation database of peer-reviewed literature and hosts over 1.7 billion cited references from nearly 12,000 publishers. The foundational data was obtained through commercial negotiations with publishers, and by obtaining licences for the data. The same model is also used by similar platforms like Web of Science owned by Clarivate. These examples highlight (i) that commercial access negotiations are already possible on a grand scale, (ii) that IP protected data can be licensed effectively and (iii) the UK's current copyright framework is not an impediment to AI development, negating any need for new exceptions.

Moreover, as the Scopus example further demonstrates, we would caution against a false dichotomy of "rights holders versus AI developers". IP and AI are often positioned in contrast, but this is a mischaracterisation they are fundamentally complementary. Publishers are themselves developers and users of AI technologies, and themselves utilise licensing solutions to do business.

We can therefore say with some certainty that licensing agreements are wholly suitable for granting access to copyright-protected content with the aim of furthering AI developments. These agreements continually evolve to deal with digital technology, while also fairly rewarding rights holders.

A further conversation on this topic with the IPO and wider stakeholders is of course very welcome, to ensure that companies of all sizes and specialisms can access licensing arrangements and benefit accordingly.

7. Is there a specific approach the government should adopt in relation to licensing?

It is imperative that the government continues to champion the role of licensing in facilitating AI development. Licensing remains the most flexible tool through which AI training can be promoted, while also recognising rights holders and incentivising investments in high-quality data sets.

The licensing marketplace is fast-moving, meaning that negotiations can respond to incremental advancements in AI technologies in real-time. Meanwhile, bespoke contractual agreements offer rights holders and users of data greater stability and certainty with respect to their rights and obligations i.e. the scope of use, authorship, ownership of data sets, and ownership of AI outputs. This clarity and agility is critical for investment confidence, and for accelerating AI developments in line with the government's ambitions. As noted in response to Q6, the evidence from members of the Publishers Association is that such agreements are readily accessible and can support major initiatives.

It is also appropriate that licensing ensures adequate compensation for rights holders, in the event that TDM is deployed for commercial purposes. Given the downstream commercial applications to which a commercial data miner might apply the output of the mining, and the associated benefits that will accrue for them, it is reasonable for the original rights holder to be compensated for use of their content specifically for TDM, AI or some other technological use, as with any other use. The alternative solution, whereby copyright exceptions allow unfettered free access to UK content for AI purposes, would merely result in a wholesale transfer of value from the UK content sector (including authors and journalists) to predominantly large technology firms and, more importantly, would not support greater investment or higher-quality outputs.

Indeed, it is imperative that the UK's IP framework continues to incentivise investment in robust and reliable data sets, as well as the accompanying infrastructure to access it. It is commonly accepted that reliable data is essential to successful AI training and development, as data biases and data hygiene issues can quickly lead an AI output astray. This is particularly important for disciplines that are likely to have the biggest impact on

people's lives, such as healthcare. However, rights holders must have the confidence to invest in producing, managing and curating this foundational information. Licensing solutions can offer the necessary confidence to deliver high-quality TDM services, whereas broad exceptions would likely lead to a reduction in standards, not simply for AI outputs but more generally for levels of investment in the data collections.

Similarly, rights holders must have means to invest in effective and efficient technological access methods to content data (i.e. publisher-provided API tools and data downloads). Investment in the development and maintenance of such tools is dependent on the commercial model that underpins a publisher's viability. Efficient access to data by miners is likely to be negatively impacted by the absence of any commercial model supporting it.

By the same token, we also need to incentivise investment in quality author led text, which could be used to train AI. As above, broad exceptions would undermine such investment and be severely detrimental to authors, journalists and the wider publishing ecosystem.

Finally, we would acknowledge that more might be done to ease licensing frictions and facilitate further access to licences for UK businesses of all sizes and specialisms. Whilst we are confident in the UK's existing licensing framework, improvements can and should continue to be made. Publishers would be delighted to work with the IPO in this regard. We hope that this consultation will fully surface where there is room for improvement in the licensing framework, and we invite the IPO to communicate these to ensure there is shared consideration of how market-based developments, including the option of licensing by collective management organisations, might resolve the issue.

8. Please rank the options in order of preference (most to least preferred) and explain why.

Our preferred option is best articulated by Option 1, Improve licensing environment for the purposes of TDM. We believe that current copyright exceptions in the UK are fair and balanced (more below), and as evidenced in earlier responses, the UK is already home to a functioning licensing market. That being said, we are keen to maintain an open and constructive dialogue with the IPO to further develop licensing best practice. As a first step, we would invite clear and meaningful data regarding the existing gaps and frictions in the market for content for TDM use. We would then welcome the opportunity to work to further improve the licensing environment on a technical – as opposed to legislative – basis. Any further exploration of model licences or collective licensing should be delivered as part of a voluntary, transparent and consultative process, with input from publishers as well as other stakeholders.

Our second preference would be **Option 0, Make no legal change**. As expressed throughout this response, we believe the existing framework is fair and balanced. Most importantly, it is already working to facilitate commercial activity with no evidential need for further intervention. When asked, our members have (to the best of their knowledge) never turned down a request from customers to carry out non-commercial TDM on lawfully accessed content. Requests to carry out commercial TDM have also always been granted on commercial terms.

Publishers have reservations about all the options that would create a new exception. **Option 3, Adopt a TDM exception for any use, with a rights holder opt-out**, is perhaps regarded as the least risky, provided it be framed in a similar manner to Article 4 of the EU Copyright Directive, and with the expectation that this route be combined with the incentivisation and championing of licensing. However, we would still flag a series of major concerns about pursuing this option: (1) as expressed in further detail below, Article 4 is yet to be transposed into law in a number EU Member States, meaning it largely remains untested; (2) this option is more feasible for multinational rightsholders to implement, but could be prohibitively expensive for smaller rightsholders leading to market distortion; and (3) this route risks morphing into Option 4, which, as explained below, would be unworkable for the industry.

We would fundamentally reject **Option 2, Extend the existing TDM exception to cover commercial research and databases**, and **Option 4, Adopt a TDM exception for any use, which does not allow rights holders to opt out**. These two options are disproportionate, and their introduction could be catastrophic for UK rights holders. Not only would they ultimately benefit large, often overseas-based, non-UK tax paying, technology firms, but there would be an inevitable destruction of value for UK-based content providers. There is also a real risk that these options could lead to non-compliance on the part of the UK with the provisions of the Berne Convention “three-step test” (which confines exceptions and limitations to certain special cases which do not conflict with the normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the right holder).

With regards to **Option 2**, it should be understood that database and reference resources are extremely expensive to create and, given the increasing societal concern about veracity and the potential for “fake news” to have a corrosive effect on the democratic function, such resources are now more important than ever in providing accurate and authoritative information throughout society. Any steps to broaden an exception in this way would undermine the revenue streams necessary to sustain these resources in the future, particularly as consumption of those resources becomes increasingly machine-driven. Publishers also note the increased privacy and security risks that may be brought about by a rise of data laundering (i.e. illegally acquiring data and making it seem authentic).

Meanwhile, a broadened TDM exception (such as **Option 4**) would have a significant impact on publishers' and rightsholders' ability to monetise their content in fast-growing and innovative fields such as AI. It would prevent the development of business models and new services, which are continually evolving, to support content discovery, licensing and delivery. In short, this legislation would be a blunt instrument that would destroy the market from a rights holders perspective.

We would reiterate our belief that copyright is compatible with innovation. Indeed, AI innovation is dependent on a strong and stable copyright framework which encourages investment in the underlying data and original content. We would remind the government that a balance must be struck, so that there are opportunities to build on – rather than undermine – existing investments. Publishers also have a significant stake in maintaining accuracy and trust in the scientific record, and market conditions must continue to facilitate this responsibility.

9. If you have experience of the EU exception with opt out for rights holders, how has this affected you?

As a number of EU Member States are yet to implement Article 4, with some transpositions not expected until 2023, it is currently too early to know with any certainty how it is likely to work in practice within publishing.

However, we would note that the 'opt-out' language in the Copyright Directive is unclear, and the idea of a machine-readable opt-out conflates two separate processes: accessing content for TDM (which can be done in a variety of ways, one of which is crawling the web and/or publisher platforms) and the activity of TDM itself (which typically takes place offline).

It also perhaps would have been more helpful for the Directive to have provided for further stakeholder dialogues to refine and further develop what a 'machine readable' opt-out actually means in practice.

10. How would any of the exception options positively or negatively affect you? Please quantify this if possible.

We would reiterate that the UK's IP framework has demonstrated ongoing adaptability in the face of new technologies. We therefore see no reason as to why the existing copyright system cannot successfully support

the continued development of AI in the UK, and would urge extreme caution against creating any new or broadened exception.

Our response to Q.8 outlines many of the likely impacts expected by the different exception options included in the consultation document. We would add here that it is extremely difficult to quantify the future impact of such hypothetical exceptions, given this is an emerging market. Publisher investment depends on the security and stability of the UK's copyright framework, and it is impossible to accurately predict the services that will not come to fruition and the investments that will be redirected if copyright is eroded in this way.

We would also emphasise that the impact risks extending beyond lost licensing revenue, and could threaten the very foundations of the publishing enterprise. The IPO must seriously consider the relationship between the end product of TDM and the original content mined. There are very real risks that the output of TDM might be substitutable for, or otherwise damaging to, the original content. Certain non-database content types are especially vulnerable to substitution by mined outputs – commercial or non-commercial – where source data and content is achieved only through long-term academic endeavour and investment.

This affects copyright works across both the sciences and humanities but there is significant risk of damage in relation to the latter, as publication of the content in question may be the only or main way in which this scholarly output is disseminated by researchers and Universities in the UK. The output does not have to be the same as the input material for this to be an issue; of equal concern is that the output may remove a large part of the use case for the original content.

The other output-related TDM risk concerns the ability of a third party to use TDM/AI to produce new works 'in the style of' the original content – leveraging and spring-boarding off the creative efforts of authors in a way that could potentially undermine their work as well as being reputationally damaging.

With this in mind, some publishers have advised that over time **a majority of their revenues would be put at risk** by a new TDM exception, as consumption patterns of their content shift to a TDM-enabled machine-driven approach, replacing human usage. This in turn would undermine the sustainability of the publishing enterprise, and fundamentally weaken the data and text inputs required for AI development.

Instead, we would urge the IPO to encourage a licensing model, which ensures rightsholders have the flexibility to take substitutability into account when negotiating usage.

General

11. What role does the IP system play in the decision of firms to invest in AI?

We would again direct the IPO to the 2020 report “People Plus Machines: AI and the Future of Publishing”, which found that protections for AI-generated content and the legal responsibility for AI-generated works are highly relevant to publishers’ investment decisions. For this reason, we would be concerned that having a broad exception could actually place unintended limitations on the development of the market.

Without a gold-standard intellectual property regime, publishers are less likely to invest in AI – particularly as the technology is still in its early stages – owing to the costs involved. If publishers do not know that they can recoup their monies, they are less likely to invest. This is especially the case with smaller publishers who do not have the resource nor monies to do so. Some level of protection is thus necessary to support and incentivise investment in AI technology.

The AI sector is no different from any other market when it comes to price negotiation, particularly with regard to commercial use of copyright-protected content. The fact that some users are unwilling to pay market value for content does not necessitate the creation of new exceptions to support the use of copyright works by AI systems. Uncompensated or undercompensated commercial access to content that requires skill and investment to create, manage and curate, raises a very real risk to sustainability and quality of both content creation and AI development in the UK.

We would also again reiterate that the government should not place a disproportionate focus on quantity over quality. Good and effective AI applications rely not only on volume, but on high-quality training data and the ongoing support of curated reference data and text content. Copyright underpins and is a mark of the value information created and curated in published works.

12. Does the first mover advantage and winner-take-all effect prevail in industries adopting AI? How would this affect the impact of the policy options proposed on innovation and competition?

As outlined above, we would emphasise that rights holders must be able to receive adequate compensation in the event that TDM is deployed for commercial purposes. This is imperative to maintain an innovative and competitive market for data services.

Given the downstream commercial applications to which a commercial data miner might apply the output of the mining, and the associated benefits that will accrue for them, it is reasonable for the original rights holder

to be compensated for use of their content specifically for TDM, AI or some other technological use, as with any other use.

The alternative solution, whereby copyright exceptions allow unfettered free access to UK content for AI purposes, would merely result in a wholesale transfer of value from the UK content sector to predominantly large, often overseas-based, technology firms and, more importantly, would not support greater investment or higher-quality outputs.

13. How does AI adoption by firms affect the economy? Does the use of AI in R&D lead to a higher productivity?

As reiterated throughout this response, publishers are themselves developers and users of AI technologies. These technologies can be applied in a wide range of ways throughout the publishing supply chain to increase the absorptive capacity, and thus enhance productivity, of both publishers and our customers.

Publishers directly facilitate absorption of research outputs through the academic infrastructure provided to businesses and institutions across the country. We curate the ever-growing canon of research and ensure its usefulness, and we offer bespoke tools for R&D professional and corporate users – reducing the time and effort spent finding relevant information for further innovation. Where possible, publishers are now increasingly utilising AI to deliver new and improved tools for our users, with benefits including enhanced quality, additional insights, improved operation efficiencies, cost savings and better customer experience.

These initial publishing benefits should generate subsequent wider benefits for the economy as a whole, including companies' improved ability to compete thanks to innovation and efficiency gains, increased academic attainment levels, more effective and efficient medical break-throughs, and the wellbeing benefits associated with improved customer satisfaction.

We would note though that these benefits are not contingent on a weakened copyright framework arising from the introduction of any further exceptions. Instead, we would stress that a lack of AI-related skills and difficulties applying AI solutions with existing IT infrastructure are the most common AI investment barriers faced by large AI-active publishers.

Lack of awareness of the potential benefits of AI also appears to be a significant barrier for large AI-active publishers. The evidence gathered by Frontier Economics suggests that overcoming this awareness barrier

could also unlock several other issues identified by some stakeholders, such as organisational barriers (e.g. willingness to invest in AI solutions that cut across organisational silos, requiring buy-in from multiple decision makers). There is a clear role for government to highlight the benefits of AI, as is being done through the likes of the National AI Strategy and wider work of the Office for AI.

14. Do the proposed policy options have an impact on civil society organisations? If so, what types of impacts?

We would note that civil society organisations may already rely on existing TDM exceptions in the UK's IP legislation. However, there is still a negative societal impact that could be brought about by broadening the existing exceptions further.

It is imperative that the UK's IP framework properly supports high-quality data sets. As outlined above, reliable inputs are critical to successful AI training and development. This is particularly important in disciplines that are likely to have the biggest impact on people's lives, such as healthcare.

Rightsholders must therefore have the confidence to invest in producing, managing and curating the foundational content. It should be understood that database and reference resources are extremely expensive to create – and such resources are more important than ever in providing accurate and authoritative information.

We would reiterate that licensing solutions can offer the necessary confidence to deliver high-quality TDM services, whereas broad exceptions for commercial use would likely lead to a reduction in standards. Any steps to broaden the existing exception in this way would undermine the revenue streams necessary to sustain these resources in the future, particularly as consumption of those resources becomes increasingly machine-driven.

This would ultimately be to the detriment of civil society organisations who rely on high-quality data and high-quality AI outputs.

Section B – Information about your organisation

A: Please give your name (name of individual, business or organisation).

The Publishers Association

B: Are you responding as an individual, business or on behalf of an organisation?

Organisation – The Publishers Association

C: If you are responding on behalf of an organisation, please give a summary of who you represent.

The Publishers Association is the member organisation for UK publishing, representing companies of all sizes and specialisms. Our members produce digital and print books, research journals and educational resources across genres and subjects.

D: If you are an individual, are you?

N/a

E: If you are responding on behalf of an organisation, are you?

- 1) An academic institution
- 2) An industry body
- 3) A licensing body
- 4) A rights holder organisation
- 5) Any other type of organisation - please specify

F: If you are responding on behalf of a business or organisation, in which sector(s) do you operate? (

- 1) Agriculture, forestry and fishing
- 2) Mining and quarrying
- 3) Manufacturing – Pharmaceutical products
- 4) Manufacturing – Computer, electronic and optical products
- 5) Manufacturing – Electrical equipment
- 6) Manufacturing – Transport equipment
- 7) Other manufacturing
- 8) Construction
- 9) Wholesale and retail trade; repair of motor vehicles and motorcycles
- 10) Transportation and storage
- 11) Information and communication – Publishing, audio-visual and broadcasting
- 12) Information and communication – Telecommunication

13) Information and communication – IT and another Information Services

14) Financial and insurance activities

15) Real estate activities

16) Scientific and technical activities

17) Legal activities

18) Administrative and support service activities

19) Public administration and defence

20) Education

21) Human health and social work activities

22) Arts, entertainment and recreation

23) Other activities – please specify

G: How many people work for your business or organisation across the UK as a whole?

1) Fewer than 10 people

2) 10–49

3) 50–249

4) 250–999

5) 1,000 or more

H: The Intellectual Property Office may wish to contact you to discuss your response. Would you be happy to be contacted to discuss your response?

Yes.

I: If you are happy to be contacted by the Intellectual Property Office, please provide a contact email address.

J: Would you like an acknowledgement of receipt of your response?

Yes.