



UK Government Copyright and AI Consultation

<https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence/copyright-and-artificial-intelligence>

Submission From OpenUK, 8 Coldbath Square, London EC1R 5 HL

Contact: <mailto:admin@openuk.uk>

Date: 25.02.2025

OpenUK

OpenUK is the UK's organisation for the business of open technology, being open source software, open hardware, open data, open standards and AI openness. Its purpose is UK leadership and global collaboration in open technology. It is staffed by a small team supported by the contributions of many hundreds of volunteers.

<https://openuk.uk/>

Answers to the Consultation Questions

B.4. Policy Options

1. Q1 Does option 3 satisfy the objectives of the consultation?

OpenUK does not believe that Option 3 satisfies the government's objectives of the consultation - control, access and transparency.

Why: Option 3 does not in satisfy the government's objectives as it creates:

1. Uncertainty: The opt out creates additional compliance costs, and friction would be introduced into the AI ecosystem and AI supply chain.
2. Impracticality: It is practical only for companies of scale to comply with such a regime of opt out and effectively this enables a form of regulatory capture.
3. Inhibiting Innovation: The additional compliance costs, and friction introduced by the opt-out, would disproportionately affect small AI developers, startups, and open source communities. As a consequence it would reduce their ability to innovate.
4. Uncompetitive: As well as making individuals, communities and small developers less competitive, the UK will not be on a level playing field with many other innovative countries including the US and Japan which may dissuade foreign companies from coming to the UK.
5. Talent flight: The lack of competitive market and inhibition of innovation for those based in the UK may also encourage talent flight from the UK and stop UK companies scaling in the UK.
6. Incoming talent: organisations will not come to the UK to innovate if we are not as pro-competitive in our innovation as other countries.
7. International Discrepancy: Fails to Address the issue that those based in non-UK-jurisdictions would be able to use UK content despite any

international scope envisaged in this proposal whilst UK-based developers would not be able to create an uneven playing field and making the UK less competitive.

8. Open source: open source development is rapidly being established as the pre-eminent form of AI development and normalising, as the meaning of this catches up with reality. Many if not most AI providers are opening weights, looking to open more code as we see from DeekSeek this week and even OpenAI founder, Sam Altman has described OpenAI as being on the “wrong side of history” when referring to “weights” and open weights in GenAI . Successful open source development has relied for 30 years on a free flow that enables anyone to use the open source licensed code for any purpose. A reservation to the TDM exception caused by the opt out would introduce friction to the freeflow with the effect of disrupting it and reducing the value of open sourcing. This will have a significant detrimental impact to open source in the AI context in the UK and will be out of step with countries of significance and their approach to this.
9. Data Access: In short open development requires freely available and legally accessed data to train and fine tune models and this is explored in more detail later.
10. Disadvantaging open source: A licensing requirement, even through an opt-out mechanism, favours large corporations with legal teams which can manage the additional compliance burden and creates challenges for collaborative or open source projects which do not have the same level of compliance resource as a company. Implementing machine readable rights reservations could place a significant burden on AI developers, necessitating costly infrastructure to ensure compliance at scale.
11. Increased litigation: The necessity to check and comply with opt outs across extensive datasets would increase legal uncertainty and potential for litigation, especially given the varying conditions that different rightsholders may impose. This could discourage AI innovation within the UK.
12. Out of Step: The UK risks being at a disadvantage compared to jurisdictions such as the US (which permits TDM under fair use) and Singapore (which has a broad TDM exception). Adopting a more restrictive approach could lead to the UK lagging in AI innovation and investment.
13. Delays and Uncertainty: The need for greater certainty was called out in [The Hargreaves Review of Intellectual Property](#) in 2011, and the delays to date are unacceptable. This must be driven forward in a pragmatic manner at pace to stop those based in the UK being at a disadvantage in innovation.

Q2. Which option do you prefer and why?

OpenUK prefers, Option 2, A broad data mining exception

Why:

1. **Certainty:** clarification that there is no need to undertake compliance checks on publicly available data correctly used for AI training will enable certainty for innovators. A broad TDM exception is aligned with the intended scope of copyright. TDM extracts facts and patterns rather than expressive content protected by copyright. Extending copyright to cover facts, data, or knowledge extraction would violate international obligations, including TRIPS Article 9(2).
2. **Practicality:** all companies are able to comply due to this being simple and practical.
3. **Promoting Innovation:** Access to data in the UK for training AI will promote innovation.
4. **Economic Growth:** Option 2 will give rise to the highest level of economic growth, by reducing barriers to AI adoption and development. Making the UK a favourable place for AI development. [Copyright & AI: The Case for a Pro-Growth Approach - UK Day One](#)
5. **Increasing Innovation:** avoiding the unnecessary compliance costs of managing opt outs will encourage more innovation particularly in smes, individual developers and open source communities from the UK - little tech which should be on a level playing field.
6. **Stem talent flight:** being on a level playing field with other countries will encourage innovators to stay in the UK.
7. **Incoming talent:** more innovators will come to the UK if we are not out of step with other countries.
8. **Level Playing Field:** both those based in the UK and in key international countries will be able to innovate on the same basis with UK data. Countries such as Singapore, Japan and the USA already have copyright regimes that are in line with option 2.
9. **Open source:** recognising the importance of open source and a free flow in software and AI openness, this will be promoted and increased in companies, individuals and communities based in the UK.
10. **Data Access:** certainty of increased access coupled with transparency requirements will enable more innovation.
11. **Supporting open source:** the UK has long been recognised as Europe's number one in open source software and number five globally (according to GitHub and other sources). In "open source AI" the UK has been seen to fall behind France in Europe (according to Tortoise Index). This UK leadership will be promoted by better access to data and the removal of barriers which will lead to better future innovation.
12. **Removing unnecessary litigation risk:** this will be achieved by simplicity from a clarification of the ability to use data without an opt out/ reservation.
13. **In-step:** UK will be in-step with the regulation and practices in key competitive regimes.

14. Removing delays: finally dealing with the Hargreaves report recommendations to promote innovation.
15. Multi-sectoral support: AI is not only about innovators but our traditional sectors being enhanced by AI, TDM supports AI innovation across multiple sectors. It is crucial for AI training and vital research in fields such as drug discovery, where analysing medical literature can lead to new treatments, financial markets for detecting fraud or economic trends, and scientific research areas like climate modelling and materials science.
16. Revolutionise AI: a broad exception can provide legal certainty for AI developers engaged in data mining. By removing ambiguity regarding the legality of training AI models, a broad TDM exception would prevent fragmented and complex rights management systems that may hinder smaller developers. This would also enable open source AI projects, which often lack the resources to negotiate individual licenses with numerous rightsholders, fostering transparency, collaboration, and public benefit.

C.1. Proposed Approach: Exception with Rights Reserved

Q3. Do you support the introduction of an exception along the lines outlined above?

A: No, we do not support the introduction of an exception, as this includes rights reserved and is not a broad exception.

Q4. 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?

We support Option 2, creating a broad data mining exception without rights reserved.

Managing the reservation of rights simply will not work - see above and throughout. Responsibility must sit with rightsholders to only make works publicly available which they are happy to see used in this way and not where they wish to reserve rights. They now make work publicly available on the internet with full knowledge of how it will be used. AI tools are helping rightsholders identify and remove AI content that is made available online without permission

Q5. What influence, positive or negative, would the introduction of an exception along these lines have on your organisation? Please provide quantitative information where possible.

The exception would have limited impact on OpenUK per se, but it would have a significant impact on the open source AI sector as set out herein.

Q6. What action should a developer take when a reservation has been applied to a copy of a work?

OpenUK does not support the reservation of rights via a technical measures approach.

We believe that this is difficult if not impossible for “Little tech” and individual developers to manage.

If the government insists on this route, we would suggest that the only possible way for this to be managed would be through software tooling which should be open source although at this point it is unclear whether this is feasible.

Q7. What should be the legal consequences if a reservation is ignored?

OpenUK believes that the use of such a reservation is burdensome and inhibitive to innovation so would not support any action in these circumstances.

If the decision is made to enable a reservation then it ought not be considered a copyright infringement when a work is computationally analysed, suggesting otherwise creates a barrier to innovation. Creating legal consequences for when an opt out is ignored exposes AI developers to a significant risk of copyright infringement that would be difficult to mitigate and would deter AI development.

Q8. Do you agree that rights should be reserved in machine-readable formats?

No, rights should not be reserved.

C.2 Technical Standards

Q9. Is there a need for greater standardisation of rights reservation protocols?

As we do not support reservation protocols we would not encourage any standards in the context of rights reservation protocols.

However, to the extent reservation of rights protocols are considered, OpenUK supports a general need for increased standardisation through open and/or de facto standards in innovation, digital, and AI. Coupled with an increase of open source tooling allowing management of such standards.

The recent launch of the [ROOST](#) open source tooling at the Paris AI Safety Summit, is a great example of this in action. However, closed standards and associated Standard Essential Patents are prohibitive to innovation and do not work with open source ecosystems as the need to pay royalties and comply with additional licensing requirements introduces friction into the open source free flow.

Should the government decide to use rights reservation protocols then we would encourage on a more general basis that the use of any standards to be associated

with this being open and ensuring that access to the creation process for these standards is made freely available to broad representation of the open source communities. Also, the existing UK Open Standards Board is in process of being replaced and this and the skills of government in open standards should be strengthened at a central level with better skills training provided broadly across the core public sector team and users to enable understanding across the public sector users and for enterprise.

Looking to specific standards, should a reservation protocol approach be used, then we believe a robots.txt exclusion is not sufficiently granular either in terms of the precise rights reserved/permissions granted, or in terms of the individual works covered (as opposed to all works at the domain), to be a complete solution. Other approaches are available which may provide partial solutions (e.g., TDMRep from the W3C, C2PA, ISCC), but these may not be appropriate to all cases.

Q10. How can compliance with standards be encouraged?

Standards should not be tantamount to “lawmaking through the back door”.

Standards are created by standards bodies and often those bodies are accessed and participated in by large, wealthy companies only. This is the case as those who can afford the standards bodies’ fees, are able to employ experts and pay for the associated travel to attend meetings are those companies and civil society and open source communities are de facto excluded. If standards are to have a critical place in the UK’s innovation policy, then the standards themselves and their creation process must be designed with inclusion and access at their heart, enabling access to the process of their design, usage and management.

In short, any applicable standards must be simple, accessible and open which will encourage adoption.

Offering access to tools as opposed to rules (see comments on ROOST above) would encourage “standards” being achieved using the language of innovation and development today. Any such tools must be open source software to enable their free usage by all. In this way adoption can normalise and the open source software can be used to create de facto standards, which unlike traditional standards can adapt at pace. Also these are readily understandable and used by developers in a RegOps - like GitOps etc- way. Like many challenges in innovation and technology which appear to be legal, regulatory or policy-based, the solution may be technical.

Q11. Should the government have a role in ensuring this and, if so, what should that be?

Government ought - to the extent standards are necessary - to have a role in the selection, implementation and management of standards and we would advocate for those to be open standards.

This role should recognise the costs involved in the standards creation and management system, that standards bodies are for-profit companies and are not not-for-profit bodies, i.e. that they have economic drivers.

Government is well placed to support business models enabling the creation and management of open standards, accessible to all at no cost and importantly provided without the burden of Standard Essential Patents to enable a free flow that aligns with open source.

Funding for this and any appropriate or associated tooling should be reserved as a priority. Such funding should not only cover any necessary innovation, but maintenance and also dissemination, training and building a contributing and healthy ecosystem around those tools.

C.3 Contracts and Licensing:

Q12 Does current practice relating to the licensing of copyright works for AI training meet the needs of creators and performers?

We do not believe that licensing of copyright should be required for AI training. However, neither do current practices meet the apparent wishes of some creators and performers. The use of rights reservation or opt out to the proposed exception will not in our view fix this challenge for creators or performers. Any usage of this will be tantamount to a “sticking plaster” covering a bigger fundamental problem and nothing more than a short term solution.

The creative sector and content providers are a critical part of society and we wish to see them well supported and provision made for their future.

However, the model to do this appropriately will in our view, be a new and creative one not dependent on royalties, or licensing for training. The current mode of funding is out of date and rapidly becoming unfit for purpose today. This needs to see a shift away from the opt out or reserved rights model and a bigger and challenging conversation to be had on how the future of human content providers not just in the creative sector will be enabled.

Q13 Where possible, please indicate the revenue/ cost that you or your organisation receives/ pays per year for this licensing under current practice.

N/A

Q14 Should measures be introduced to support good licensing practice?

We do not, however, support this in the context of a reservation to the exception. Contrary to the understanding of some, licensing and governance are fundamental to the success of open source. However this must be appropriate to the circumstances.

We do not believe that licensing is required for AI training. However, content providers would in some cases benefit from support in understanding of sharing their content publicly via an online medium.

Q15 Should the government have a role in encouraging collective licensing and/ or data aggregation services? If so, what role should it play?

As we do not believe in an opt out or reservation to an exception we would not encourage their use in managing an opt out.

Q16. Are you aware of any individuals or bodies with specific licensing needs that should be taken into account?

N/A

C.4 Transparency

Q17. Do you agree that AI developers should disclose the sources of their training material?

Yes, AI developers should disclose the sources of the input data. It would be reasonable for a data mining exception for AI model training to be conditional upon full transparency as to the data sources used. Note that such disclosures may be difficult or impossible to verify and this may be a use case where software tooling can be developed and evolved over time.

Q18. If so, what level of granularity is sufficient and necessary for AI firms when providing transparency over the inputs to generative models?

If rightsholders are given an ability to reserve their rights then transparency needs to be given at the most granular level at which a rightsholder could opt out and both ought to be in sync. We suspect this is likely to be at the file level.

Whilst a file (e.g., web page, pdf, mp4, epub, etc.) may in reality contain multiple works, if a rightsholder is not technically able to reserve their rights at a more granular level then a more detailed level of transparency would appear to be unnecessary.

Systems and tools implemented to manage this process would likely dictate the practical level at which governance might apply.

If no reservation of rights is permitted, as OpenUK would recommend using option 2, then more flexibility could be allowed and an appropriate balance would need to be struck between:

1. The ability to demonstrate data provenance (from both a quality and copyright perspective);
2. The ability for rightsholders to understand when and by whom their content has been accessed for TDM purposes; and
3. The technical feasibility of transparency reporting at a very granular level.

The flexibility in this latter scenario is part of our justification for supporting Option 2.

Q19. What transparency should be required in relation to web crawlers?

Similar principles as those described in answer to Q18 apply to web crawlers and scraping.

Q20. What is a proportionate approach to ensuring appropriate transparency?

Rightsholders should be able to determine whether their work has been included, either because the work has been identified directly or as part of a collection. It may not be technically feasible for a model to identify which works resulted in a given output (i.e., explainability).

Q21. Where possible, please indicate what you anticipate the costs of introducing transparency measures on AI developers would be.

The costs of introducing transparency measures depend on the nature of the AI systems being developed and the use case for which they are being developed. We believe the costs of this transparency would be minimal. The processing code needs to "copy" the content in order to use it for training, at which point it could log e.g. the URLs or collections being consumed.

Q22. How can compliance with transparency requirements be encouraged, and does this require regulatory underpinning?

The challenge is one of trust in what the AI vendors say. It may be difficult or impossible to verify the information provided, unless the model memorises and reproduces copyright content verbatim. Whilst our view is that regulation of AI is generally unhelpful, compliance with transparency requirements may be one of the few areas which need to be supported by regulation.

This is one of the few areas where regulation in AI would be helpful across the board and specifically required. Ultimately, encouragement by itself is not enough and is unlikely to incentivise the desired behaviours. Enforcement powers are likely to be required to drive the right behaviours.

Q23 What are your views on the EU approach to transparency?

We presume this refers to the information required under the EU's AI Act, which is more than just the origin of the materials used for training and is expansive, requiring virtually every aspect of the design, development and deployment of AI systems. We would view this as cumbersome with respect to TDM, and the purpose of the data being gathered as referred to in the EU AI Act is broader than the transparency required for the purposes of this consultation (i.e., it is gathered to perform an effective assessment of these systems for broader risk purposes).

The EU AI Act requirement calls for a different approach to collecting the information needed to satisfy the transparency requirements because traditional manual methods will soon be insufficient to meet these needs as the pace of development, deployment and update accelerates further. We would encourage supporting the evolution of open source tooling to support this. We have commented on granularity of transparency measures for these purposes at Q18 above.

C5. Wider Clarification of Copyright

Q24. What steps can the government take to encourage AI developers to train their models in the UK and in accordance with UK law to ensure that the rights of rightsholders are respected?

We would encourage the government to enable a wide TDM exception without reservations on par with the exceptions in Japan.

Q25. To what extent does the copyright status of AI models trained outside the UK require clarification to ensure fairness for AI developers and right holders?

If the UK wishes to enable a globally level playing field this requires copyright status of models trained outside the UK to be clarified. Without this level playing field it is difficult to envision how the UK's goals for AI can be achieved.

We consider it important that AI developers be held to the same standards irrespective of where the model was trained, but equally important not to discourage AI developers from making their models available in the UK and to encourage their working and innovating from the UK.

Q26. Does the temporary copies exception require clarification in relation to AI training?

Yes

Q27. If so, how could this be done in a way that does not undermine the intended purpose of this exception?

We believe it would be helpful to clarify whether or not creation of an AI model has independent economic significance in the context of the temporary copies exception.

C6. Encouraging Research and Innovation

Question 28. Does the existing data mining exception for non-commercial research remain fit for purpose?

No, it is neither adequate nor fit for purpose and there is an existing level of uncertainty that must be resolved.

Question 29. Should copyright rules relating to AI consider factors such as the purpose of an AI model, or the size of an AI firm?

No, copyright rules should not relate to factors such as use of the model, they should be broader than this and non-discriminatory. Other requirements such as transparency might however.

Any copyright regime which means that a work is licensed for one purpose and not another, and requires downstream licensing conditions on the model developer would be incompatible with open source licensing, as a condition on who can use and for what purpose would be introduced.

This would sit in contradiction to the free flow of open source and the long established principles of the [Open Source Definition](#) which at points 5 and 6 enable anyone to use outputs for any purpose. The free flow of open source is enabled by this and any conditions are problematic and introduce friction.

Whilst not considering the purpose of the model or size of the developer, we do think it appropriate to consider the behaviour of the developer in the context of responsible use of the mined materials. In the same way as open source software is subject to laws over its licensing we would expect legal requirements to ensure appropriate behaviour, as we have seen for many years across technology. It is important to remember here that like any other digital tool, AI is a tool to support human behaviour.

Of the economic actors in the ecosystem, those that are training AI models are best placed to put any legally required safeguards in place in a scenario where the AI model end user is innocent of wrongdoing and we would encourage any governance or requirements to be placed at this level.

D2. Policy Options

Q30. Are you in favour of maintaining current protection for computer-generated works?

Yes, we support maintaining protection for computer generated works and extending this to AI for the sake of clarity.

If yes, please explain whether and how you currently rely on this provision.

As an organisation we do not rely on this, however the software development industry in the UK currently operates on the basis that copyright can subsist in code generated by a machine, and we've observed little in the way of governance within developer organisations to distinguish between human authored code, AI-assisted code, and AI-generated code.

Q31. Do you have views on how this provision should be interpreted?

We have provided commentary addressing this point below.

D2. Policy Options

Q 32 Would computer-generated works legislation benefit from greater legal clarity, for example to clarify the originality requirement? If so, how should it be clarified?

Yes computer-generated works legislation would benefit from greater clarity as there is substantial confusion with regard to:

- (1) the word "author" in the drafting; and
- (2) What is meant by "no human author".

In s.178 a computer-generated work is one that "*is generated by computer in circumstances such that there is no human author of the work*". However section 9(3) refers to the existence of an author (which may be a natural or legal person), which concept is then used in relation to first ownership of copyright.

A key question is what human involvement is required in the entire lifecycle of development (e.g., creation of AI, training/fine-tuning of AI, configuring AI settings/parameters, providing instructions to AI), and whether there is a threshold as to the level of detail in the instructions given to an AI, in order for copyright to subsist in the work ultimately produced by an AI tool. Under the current law, we understand that no human involvement is required in creation in order for copyright to subsist, but the dividing line between computer-generated and human authored works (and the different terms of protection arising) is currently unclear. See our response to Q35 for further explanation.

Clarity could be added by adjusting the definition of computer-generated at s.178 to refer to the degree of human involvement in creation, as opposed to there being "*no human author*". This needs to be considered in the context of our response to Q35 below.

Q 33 Should other changes be made to the scope of computer-generated protection?

Other than as described in our response to Q35 below, we don't believe other changes are required to the scope of protection.

Q 34. Would reforming the computer-generated works provision have an impact on you or your organisation? If so, how? Please provide quantitative information where possible.

It would not affect our organisation directly, but would have a very material impact on the software development sector and UK open source communities.

Software development best practice is for the provenance of computer code to be recorded and passed on as part of onward licensing and in the case of open source software in the downstream.

This is typically done through 'attribution' provisions in the licensing, which are common across all forms of open source software licensing. Where the contributor of the relevant software component is named in the downstream licence (or associated documentation) and attributed for their contribution. Attribution is a common feature of open source software licences, even in very permissive licences when no other rights are reserved.

The current copyright protection regime allows for this important recognition of a developer's contribution. For human-authored works moral rights can be asserted, including the right to be named as author. However, when a work is computer-generated, moral rights are unavailable and therefore to support attribution the subsistence of copyright is necessary. This is because the attribution component will be covered as a copyright licence condition, rather than an assertion of moral rights.

Other open source licences rely on the protections afforded by copyright to allow developers to require that the materials that they make available on an open basis remain treated as open by all downstream users, under copyleft licensing. This is an important feature of the software development landscape, and is an important safeguard against works shared on an open source copyleft basis being co-opted and made 'closed' by those wishing to take from the community but unwilling to contribute back. In effect the use of copyleft protects the user from themselves.

Q 35. Are you in favour of removing copyright protection for computer-generated works without a human author?

No, we support computer-generated works without a human author receiving the benefit of copyright. However we believe a modification should be made to the term of protection, and detailed consideration must be given to distinguishing this scenario from one where there is greater human involvement.

We see there being a number of distinct scenarios that need to be addressed:

- (1) A purely AI-generated work with no human involvement at any stage of the creative process;
- (2) a purely AI-generated work with minimal human direction as to the subject, purpose, style, or other creative choices;
- (3) AI assistance as a tool supporting human creation (i.e., with substantial human involvement in the creative process).

Whilst we do not believe the involvement of human endeavour in the development or training of the AI should be considered involvement in the creative process (and therefore an AI autonomously creating works should fall into category 1), we consider the creation of a work in response to an instruction given by a human should fall into category 2. The distinction between categories 2 and 3 is difficult to quantify, and may come down to a determination as to the degree of originality and human creativity in the instructions given to a generative model.

Our view is that copyright ought to subsist in all 3 scenarios. However, we consider the term of protection for scenario 1 (purely AI generated work) ought to be substantially shorter reflecting the ease with which such works can be created with only minimal investment.

A different treatment between scenarios 2 and 3 should only be considered if a clear and well-defined distinction can be drawn between scenarios to give clarity as to the degree of protection afforded.

Q 37. Would the removal of the current CGW provision affect you or your organisation? Please provide quantitative information where possible.

It would not affect our organisation directly, but could have a very material impact on the software development sector and open source software communities, if doing so resulted in the removal of the current CGW provision which might lead to uncertainty as to whether copyright arises upon the AI-assisted creation of works. As mentioned elsewhere AI is rapidly becoming a normal tool to support this process.

We believe a better approach would be to clarify what is CGW, see our answer to Q35.

D.3 Computer-generated works interaction with designs

Q 38. Does the current approach to liability in AI-generated outputs allow effective enforcement of copyright?

A properly trained AI ought not to deliver a “substantial part” of a protected work but a poorly trained model might do so (e.g., when there is a bug or overfitting occurs).

Overfitting is where a model learns a specific detail of the training data and reproduces these rather than the idea of it. Ideas and concepts as such are not protected, whilst individual expressions of creativity are. Outputs ought not to include the creative expression in those works. Where AI models output individual copyright works, or substantial parts of the creative expression, it might today be considered to be a copyright infringement. The user and provider of a model may both be liable for infringement of copyright today.

We believe that any broader TDM exception should be conditional upon the use of the work not unreasonably prejudicing the legitimate interests of the rightsholder (in alignment with the three step test under the Berne Convention, Article 13 of TRIPs, and the approach taken in Japan), thereby incentivising developers to take reasonable measures to avoid infringement by their models (in the absence of malfeasance on the part of the user), and resulting in primary liability for infringement should their failure result in an unfair prejudice to the rightsholder’s legitimate interests.

The end user’s use (and any malfeasance by the end user) should always be the end user’s responsibility.

Q 39. What steps should AI providers take to avoid copyright infringing outputs?

AI providers should be able to rely on a broad training exception and with this in place should ensure that they meet transparency requirements and appropriately manage the detail of what data their AI is trained upon and keep appropriate records of such to enable trust through transparency.

D.5 AI output labelling

Q 40. Do you agree that generative AI outputs should be labelled as AI generated? If so, what is a proportionate approach, and is regulation required?

Yes, AI outputs ought ideally to be labelled as such, but any requirements should not place obligations on the developer to waterfall into any form of responsibility for downstream compliance.

Q 41. How can government support development of emerging tools and standards, reflecting the technical challenges associated with labelling tools?

The technical challenges around labelling systems relate to the quality and accuracy of the labelling. These challenges have existed for some time in relation to manual labelling of data for supervised training.

Whereas quality assurance (QA) for labelling of images or video data can be done quite easily, and metrics for these are well-known, QA for contextual analysis and labelling of text is more nuanced; and no standards are currently available. In this context, individual industry verticals (and associated trade bodies/regulators) may wish to devise individualised standards or tooling that reflects the nuance and semantic sensitivity of their individual needs.

In any event we recommend that software tools that are open sourced are used wherever possible to manage these needs enabling these to in effect become de facto standards. To the extent that formal standards are utilised these must not be closed and must be open standards with an open process of design and management as explained above.

Question 42. What are your views on the EU's approach to AI output labelling?

Assuming this refers to the requirements for the labelling of AI outputs from GenAI systems as set out in the EU's AI Act, this requirement forms only a small part of the much broader raft of requirements for AI providers under the EU AI Act, especially in relation to information to be provided to users. The approach adopted by the EU AI Act is proportionate as it reflects the likely risk profile of the use case of the outputs, rather than the nature of the AI systems themselves in the context of the Act.

However we would not support a regulatory approach to labelling in all cases as this will likely be over burdensome. We would in the alternate support an open standards approach which we would recommend like the AI Act requires the (ideally machine-readable) labelling of synthetic output that could mislead the public (e.g. journalism, political ads, etc.) without requiring e.g. an employee to label their work internal to a company. A PR written by AI and reviewed by a human for example might well not meet the requirements or need to be labelled.

D.6 Digital replicas and other issues

We do not believe that personality rights are relevant to the issues under consideration and so have not answered 43 - 44.

D.7 Other emerging issues

Q 45. Is the legal framework that applies to AI products that interact with copyright works at the point of inference clear? If it is not, what could the government do to make it clearer?

The temporary copies exception is relevant to how AI products interact with copyright works at the point of inference, however this exception is limited and applies under certain conditions. Implementing a broad TDM exception would provide clarity that all acts of inferencing are permitted.

Q 46. What are the implications of the use of synthetic data to train AI models and how could this develop over time, and how should the government respond?

Synthetic data can be generated in a variety of ways and is a useful technique to enable model training. It also allows for privacy to be respected. It enables more data to be available for AI training, which is a significant challenge for open source model development.

Q 47. What other developments are driving emerging questions for the UK's copyright framework, and how should the government respond to them?

Outdated concepts in 2025:

Fundamentally there is an inevitable concern as to whether copyright is fit for purpose in 2025's innovation landscape and this must be considered at a holistic level. This sees the need for big picture thinking around business models not only for content creators but for standards providers, to enable innovation at pace and keep the UK competitive both in its AI products and the support of AI for the UK's traditional industries.

Downstream risk: AI can be trained on code under an open source licence, but the connection between the original work and the downstream must not create liability on the upstream for a downstream behaviour.

Approach to be taken - tools not rules: An approach of tools made available as open source software to enable developers to comply with de facto standards as opposed to rules is generally recommended. In the absence of such we strongly support open standards before closed standards or regulation. This "RegOps" type approach speaks to GitOps etc and how developers understand management of their development practices and avoids burdensome and inhibitive regulations. Opening these tools offers access and the ability of all to contribute and use.

The Creative Bias and its impact on AI:

Whilst the tech sector may be considered more than capable of giving a good account of itself, the reality is that in the UK the creative industries have been much more powerful and have been able to have a greater influence on this process.

In submitting this response we recognise the power of the creative industries to do what they do best...to tell a story. The stories they paint in pictures, sing in songs



and write stories, plays and television are accessible to all. The creative industries have historically brought a huge amount of revenue to the UK. The tech, and in particular the AI sector is not made up of story tellers. The tech and AI sectors' work is narrow and complex, and unlike the creative sector's stories it is inaccessible to many. Whilst tech more generally has generated considerable income for the UK, the AI sector has the potential to generate tens of billions of future revenue.

Choices today will impact the UK's digital future, our technological success and the UK's future economy and global position and must be made carefully, with consideration of those difficult to understand nuances of technology that will impact the UK's future AI leadership.

OpenUK supports the UK and global creative industries and hopes to see them flourish, whilst recognising a need for access to Input Data - data, images, sounds etc that are publicly available. The rights in those are broad and the consultation is focused only on one category of the potential suite of applicable Intellectual Property (IP) rights, copyright and does not touch upon contractual rights, data privacy rights or other potential applicable rights. There is also an assumption in our response that content considered is legitimately and legally placed online and that its public availability is not as the consequence of any breach of any rights.

In considering our answers to these questions, we recognised that creators include developers, who also own copyright in works created, as the authors of those works.

End