



About The Copyright Licensing Agency Ltd (CLA)

The Copyright Licensing Agency Ltd (CLA) is a collective management organisation (CMO) as defined in The Collective Management of Copyright (EU Directive) Regulations 2016.

Regulated by the UK Intellectual Property Office (IPO), CLA is the recognised UK CMO for collective licensing of text and images from book, journal and magazine content (including some websites) to the corporate, education and public sectors. CLA exists to simplify copyright for content users and copyright owners, and our mission is to help users to legally access, copy and share published content, while also making sure that creators and rightsholders are fairly remunerated for the use of their works.

Established over 40 years ago, CLA has provided copyright licences to a growing user base, enhanced by a range of related products and services, evolving its offering over time to meet changing user needs and requirements, but always with the aim to simplify copyright and make it easier for users to access content.

CLA is a not-for-profit, membership owned organisation. It has four members: Authors Licensing and Collecting Society Ltd (ALCS), Design and Artists Copyright Society (DACS), PICSEL Ltd (Picture Industry Collecting Society for Effective Licensing) and Publishers' Licensing Services Ltd (PLS). CLA distributes the revenue it collects from licensing to its members, who in turn distribute to authors, publishers and visual artists.

About Voluntary Collective Licensing

Voluntary collective licensing is a cost-effective blanket licensing solution and offers a practical alternative where it is not easy to license on an individual basis for specific uses due to the volume of rightsholders or users and/or the scale of use. Since it is not possible to take account of the exact rights ownership of each extract which may be copied or used, the licence revenues are shared between all the relevant rightsholders.

Through licensing, CLA ensures that creators and rightsholders are fairly remunerated for their time and investment and continue to innovate to support the UK's creative economy. Since 1983, CLA has distributed over £1.5 billion to creators and rightsholders, and in 23/24 distributed £91.5m.

Copyright and AI

At CLA, we believe that compliance with copyright law is crucial to facilitate the ethical, safe and legal development and scalability of generative AI technologies and is essential

for the sustainability of the UK's creative industries worth £124¹ billion annually to the UK economy.

To this end, CLA has developed a set of principles for copyright and generative AI which we believe will ensure the responsible development of AI.² These principles are:

- Permission Based
- Fair Remuneration and Compensation
- Transparency and Attribution
- Authenticity

CLA's responses to the consultation questions are below.

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

CLA does not believe that option 3 is most likely to meet the Government's stated objectives of control, access and transparency.

There are several legal, practical and technical challenges to designing and implementing an effective and robust solution that would support a rights reservation, as contemplated by option 3. Any solution would need to work for, and accommodate, all types of copyright protected works, rightsholders and creators.

Instead, we believe that the focus of the Government in the immediate term should be to introduce mandatory transparency measures for AI developers to disclose publicly, in an accessible, accurate and complete way, meaningful information to allow the public, general users, creators and rightsholders, to identify what works have been used to train their systems. The transparency measures must take into account works that have already been copied by AI developers in the past, and there needs to be a mechanism in place to support retrospective licensing and compensation for use.

We do not believe that such requirements will stifle innovation, but instead encourage the safe, ethical and legal development of generative AI, and thus promote greater adoption by consumers and business in the UK, as well as create a fair playing field for rightsholder and creators by facilitating greater licensing and the development of new licensing solutions through a clear understanding of which copyright protected works have or are being used.

Question 2. Which option do you prefer and why?

CLA believes that option 1 (strengthen copyright, requiring licensing in all cases) is the preferable option put forward by Government in this Consultation.

¹ <https://lordslibrary.parliament.uk/contribution-of-the-arts-to-society-and-the-economy/>.

² <https://cla.co.uk/ai-and-copyright/principles-for-copyright-and-generative-ai/>.

Under existing copyright law in the UK, licensing **is** required for all commercial text and data mining (TDM) activity. Paragraph 5 of the Government’s consultation document acknowledges this by stating that “...*copying works to train AI models requires a licence from the relevant rights holders unless an exception applies*”. This should continue to be the case, and CLA is of the view that no change needs to be made to the law in this regard.

To effectively ‘strengthen copyright’ and to facilitate licensing as is required by law where copyright protected works are used for the training of generative AI systems, it is essential that mandatory transparency measures are introduced by Government as to what works have been used in training. As noted in CLA’s response to Question 1 above, there are societal and economic benefits for the UK in mandating generative AI developers to provide accessible, accurate and complete information around works used to train their systems. In respect of licensing, this will support the current licensing market, facilitate greater licensing and promote the evolution of new licensing solutions – either on a direct or voluntary collective basis. Rightsholders and creators will benefit from greater access to information and the ability to exert control over the use of their works by identifying which of their works have been used, by whom and how, allowing them to offer an appropriate licence if they choose to do so.

Regulators must have the powers to audit disclosures and enforce non-compliance with any transparency requirements by a generative AI developer. The IPO’s role in ensuring generative AI developers comply with their transparency obligations and support the promotion of licenses available in the market, will further support licensing for the benefit of rightsholders, creators, society and the UK economy.

CLA’s Success in Licensing

CLA is well established in offering voluntary, collective copyright licences which operate alongside exceptions under The Copyright, Designs and Patents Act 1988 (CDPA 1988). Notably, within the education sector, the fair dealing provisions of s32 of the CDPA 1988, alongside CLA’s licences, provide a convenient and affordable end-to-end access model for copyright works for educational institutions. This proven model also ensures rightsholders are fairly remunerated for the use of their works beyond the fair dealing permitted by s32.

Outside of the education sector, CLA provides collective licences covering many types of internal copying and reuse of extracts of published content by more than 10,000 businesses and public sector organisations.

CLA’s Licensing of TDM and Generative AI in the Workplace

For over forty years, CLA’s licences have adapted to enable the copyright-compliant use of content with new technologies, and generative AI is no different.

As professionals increasingly use copyright protected works as prompts in generative AI tools in the workplace, CLA has developed new **workplace AI permissions** which permit content users to carry out this activity lawfully and ensure that rightsholders and creators are fairly remunerated for the use of their works. These new workplace permissions will be followed later in the year with the launch of a premium model that includes the right to perform retrieval augmented generation (RAG). CLA is also working on developing a training licence to compliment direct licensing deals already being negotiated between the larger AI firms, to ensure the medium and long tail is represented and compensated where those rights holders wish to participate.

This evolution in CLA's blanket licences follows the development and launch of a text and data mining (TDM) extension for the corporate and public sector, permitting TDM for commercial purposes of whole works and covering a wide range of published content.

These enhancements to CLA's voluntary licensing schemes developed in consultation with CLA's members – ALCS, DACS, PICSEL and PLS – demonstrate the flexibility of licensing as a solution without the need for changes to copyright law.

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

Without a fully detailed and scoped proposal from Government as to how a rights reservation will operate, CLA currently cannot support the introduction of an exception as outlined in its consultation document. We believe that there is clarity in the existing law, and AI developers are required to comply with the law. It is worth noting that there are examples of AI developers entering into agreements with rightsholders³ and this further supports the view that a new exception to copyright law in the UK is therefore not required nor justified. Furthermore, in our view it is unlikely that introducing an exception as contemplated in option 3 of the consultation will comply with the UK's obligations under the Berne Convention, notably Art 5(2) ('the enjoyment and exercise of these rights shall not be subject to any formality') and Art 9(2) (exceptions may only be introduced in 'certain special cases').

CLA has not seen any evidence that supports the premise that a rights reservation would promote or encourage licensing of copyright works for training by AI developers in the UK. We note that Article 4 of the Directive on Copyright in the Digital Single Market (Directive (EU) 2019/790) (DSM Directive) came into force in 2019 and was required to be implemented in the EU Member States by 2021, yet there is no demonstrable evidence that this has led to increase in licensing by AI developers in any Member States.

³ <https://pressgazette.co.uk/platforms/news-publisher-ai-deals-lawsuits-openai-google/>

CLA further notes that to-date the Government has not provided any evidence or economic impact assessment to support the need for a new exception, as contemplated by option 3 in this Consultation, to be introduced to the existing law.

Given the value of the creative industries to the UK economy, the Government must have due regard to protecting the rights of creators, and ensuring the UKs 'gold-standard' copyright framework is not weakened. As the Prime Minister has stated:

*"We recognise the basic principle that publishers should have control over and seek payment for their work, including when thinking about the role of AI"*⁴

The Government's fundamental objectives in relation to AI and copyright must be that rightsholders have effective control over their work and are able to seek compensation for past use and remuneration for any future use of their works to ensure the future growth and success of the creative industries in the UK and we do not believe that option 3 will successfully achieve this.

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

To achieve the Government's objectives, first and foremost there needs to be mandatory transparency measures. Such transparency measures must incorporate:

1. Accurate and detailed record-keeping of works used, including meta-data attached to the works;
2. Publication of up-to-date records of the works and data that have been used / ingested by generative AI systems and such information to be available, in an accessible and easily understood form on AI developer's public website or equivalent in a timely manner; and
3. Labelling of AI-generated outputs.

For the text and image works repertoire currently managed by CLA, the detailed information provided by AI developers must include, at least:

- a) What works have been used by AI developers for training and fine-tuning;
- b) Any meta-data associated with the works;
- c) The source and owners of the data sets; and
- d) Work level information of those works ingested.

Requirements will be different for different types of creative works and creators, and the transparency measures will need to account for this.

By requiring transparency from AI developers, rightsholders will be able to identify whether their works have been used in training, offer a licence should they so choose to

⁴ <https://www.theguardian.com/commentisfree/2024/oct/28/keir-starmer-journalism-lifeblood-british-democracy-labour>

do so, or take appropriate action under existing copyright law if AI developers are found to have used / are using their works without a licence. Transparency will permit and encourage a fair market for licensing. As noted earlier in the response, licences may be individual licences offered directly by rightsholders, or voluntary collective licences.

Regulation of transparency may also help, ensure compliance with all relevant laws in the UK, including for example, data protection laws. Being able to enforce these mandatory transparency measures is key to the successful safe, ethical and legal development of AI: rightsholders will be able to licence their rights where they choose to do so, and AI developers will be certain they have the necessary permissions to operate their businesses lawfully.

Voluntary Collective Licensing as a solution

With adequate transparency as to what works AI developers wish to use, a CMO such as CLA is better able to understand the licence requirements and to seek the appropriate rights from creators and rightsholders to offer a collective licence, should that be the most suitable solution. Voluntary collective licensing can offer a practical solution to licensing a high volume of works, particularly where rightsholders may not be able to effectively manage their rights individually. Under voluntary collective licensing, a CMO issues licences to copy content on behalf of rightsholders and creators who have given their mandate. In CLA's case, licences are available to organisations in the corporate and public sectors, as well as educational institutions. As a regulated CMO that operates voluntary collective licensing, CLA is well placed to offer licence solutions to AI developers, but this is impossible without transparency. A model of transparency, choice and licensing will meet the Government's stated objectives of control, access and transparency.

As an example of effective and transparent collective licensing of a technology sector, CLA licenses Media Monitoring Organisations (MMOs) operating in the UK and internationally for indexing and TDM of magazine and news website content. [The CLA MMO Licence](#) is available to all MMOs. CLA has licensed MMOs for over 20 years and engages with MMOs regularly to ensure that the licence terms remain relevant to the industry.

The licence permits MMOs to carry out various restricted acts protected by copyright on relevant website content, including 'accessing and copying web pages using so-called "web-crawler", "spider" or "robot" software or other automated and/or manual review methods, processes or means in order to derive and store searchable index entries' (i.e. TDM) and permitting MMOs to 'store the indexed material in order to search, retrieve and evaluate (by automated and/or manual review methods, processes or means) [the material]'. The licence repertoire is continuously growing and currently covers more than 10,500 websites, including many thousands of news and current affairs sources relevant to MMOs and their clients.

As a condition of the licence, MMOs are required to be transparent and report to CLA, on a quarterly basis, exactly which articles the MMO has scraped, as well as which clients have received articles. Since 2021, MMOs have reported to CLA more than 89 million instances of articles being copied or shared under the licence, comprising content from more than 6,000 magazines, journals and websites.

The detailed data (URLs, ISNs and titles, as well as which client the work has been sent to) received from MMOs is used to support the distribution of revenue to publishers, authors and visual artists, as well as providing CLA with data on businesses which may need a licence for any internal use of articles received from MMOs. The provision of data ensures rightsholders are remunerated, whilst providing a solution for businesses to comply with copyright law and demonstrates that licensing is a flexible and practical solution which enables transparency as part of the lifecycle of content use.

We are supportive of an approach which will promote and support an evolving licensing market to develop and strengthen. As stated in our response to Question 2 above, before any intervention in the existing law by Government is proposed, there is an urgent requirement for transparency.

As noted above, CLA has developed new market-based licensing solutions covering TDM and the use of Generative AI tools in the workplace, without the need for legislative intervention. Greater transparency as to the use of copyright protected works for the training of AI systems, and other related uses by AI developers, would allow CLA to develop a meaningful licence proposal.

Question 5. What influence, positive or negative, would the introduction of an exception along these lines have on you or your organisation? Please provide quantitative information where possible.

CLA's view is that there is insufficient evidence available to support the assertion that the introduction of an exception with a rights reservation, as contemplated by this consultation, will promote greater licensing or enable the licensing market (both direct and voluntary collective) to develop and evolve, and therefore have a positive impact. As stated earlier in CLA's response, there is no demonstrable evidence from the EU which suggests that a rights reservation has led to an increase in licensing and the implication is therefore, that were such a proposal to be implemented in the UK, this would have a negative impact on CLA.

In the absence of enforceable, mandatory transparency measures the proposed exception with a rights reservation is likely to create an environment in which it would be futile for CLA to offer a licence for generative AI training, due to the risk of non-compliance by AI developers, and the likelihood, based on what has transpired following the implementation of Art 4 of the DSM Directive in the EU, that a rights reservation will not be complied with.

Furthermore, there is the risk that introducing a rights reservation will lead to additional burden in time and costs for creators and rightsholders which may then impact on the cost of licences, making them more expensive as the cost of managing reserved rights is likely to need to be factored in. The Government should not therefore introduce an exception that makes it prohibitively expensive for rightsholders to enforce their rights. Any rights reservation solution that utilises technology also needs to be fully accessible by all creators and rightsholders.

As the Government rightly notes in the Consultation, the requirement for a rights reservation in the EU is not always clear and there is therefore, no proven and workable rights reservation framework in existence. This remains the case four years after the first implementation of Article 4 of the DSM Directive. There is also the risk of a lack of enforcement with a rights reservation approach, the challenge of a rights reservation staying with a work as it ‘travels’ online with downstream uses, the issue of any retrospective permission where works that have already been copied for training and no permission has been sought and no remuneration paid.

Legal effects of the rights reservation

Question 6. What action should a developer take when a reservation has been applied to a copy of a work?

In accordance with the CDPA 1988, if the developer does not have permission to use a work protected by copyright, then the work should not be used. Where the rights have been reserved, as envisaged by the Government’s proposal, the developer must respect the rightsholder’s choice whether or not to offer a licence for use, and on what terms. This is a fundamental principle of copyright law.

The onus of checking whether a work has its rights reserved must lie with the AI developer, and the developer must check the rights status of all expressions of the same work. The responsibility of checking the status of a work and complying with the law must lie with the AI developers.

Question 7. What should be the legal consequences if a reservation is ignored?

Any legal consequences need to be sufficiently robust to address the current power imbalance between rightsholders and generative AI developers. Currently, the lack of transparency around the use of copyright protected works to train, as well as the imbalance of power that exists between the rightsholder and creator and developers and companies deploying generative AI systems, makes it difficult, verging on impossible, for rightsholders to be able to take action when their works have been used unlawfully.

Legal consequences must include the normal remedies for infringement under the law including injunctive relief and fines. Fines would need to be reviewed and aligned more closely with revenues of AI developers and companies deploying the systems to operate

as an effective deterrent and promote compliance, akin to those fines that can be issued by the ICO under current data protection laws in the UK. The IPO, in this context, should play an effective role, and have the powers (similar in principle to the FCA, the CMA and ICO in their respective work) for regulating compliance and enforcing the legal consequences of non-compliance.

Question 8. Do you agree that rights should be reserved in machine-readable formats? Where possible, please indicate what you anticipate the cost of introducing and/or complying with a rights reservation in machine-readable format would be.

For ease of use by creators and by generative AI developers it is preferable that, where feasible, rights are reserved in a machine-readable format. However, this is unlikely to be possible for all rightsholders and all types of copyright protected works and therefore, alternatives must also be possible. By way of example, where works of physical art are photographed, for example by a gallery or museum, the artists' rights reservation would need to be applied to the photograph itself for it to be effective. There are likely many more examples that rightsholders will be able to provide in relation to this and, if developing a rights reservation with a machine-readable format, the IPO must engage with rightsholders and creators directly on these issues.

There is no evidence that the requirement for rights reservation in Article 4 of the DSM Directive has been successfully implemented to date, and therefore the technology being used to reserve rights is not currently fit for purpose.

Finally, it is worth noting that there is often a misconception the works that are free to view on the internet are therefore free to copy. However, this is usually not the case, and the vast majority of websites will have terms and conditions that should be checked before any use is made of copyright protected content contained within a website. Given the technical proficiency of AI tools today, it should be technically possible for a machine to read websites terms and conditions and understand whether website content can be used or not. In many cases, rights are already reserved under the relevant website's terms and conditions of use however, these terms are not being complied with.

C2 Technical Standards

Question 9. Is there a need for greater standardisation of rights reservation protocols?

For rights reservation protocols to be workable, there must be standardisation. This benefits both rightsholders and generative AI developers by ensuring that any rights reservation process is not onerous, is unambiguous and is consequently respected as widely as possible.

As the Government's consultation document notes in para 95, it will be important for CMOs to be able to reserve the rights of their members effectively. As an example of effective rights management, with CLA's voluntary collective licences, CLA manages rightsholders rights on both an opt-in and opt-out basis. CLA currently holds 34 million titles in its database (32 million books and 2 million journals and magazines). Where rightsholders have opted-out their rights, this can be checked using CLA's online tool, Check Permissions (<https://cla.co.uk/resources/tools/check-permissions/>). Moreover, CLA offers an API to the Check Permissions tool, whereby content users can embed the search function within their own systems.

Question 10. How can compliance with standards be encouraged?

To ensure compliance, any standards developed will need to be regulated, with legal consequences if not complied with. The Government must demonstrate to both rightsholders and AI developers that a rights reservation can be successful, and standardisation is a necessary component for this. Without any standardisation, it is highly likely that a rights reservation system will fail. Standardisation is needed to make it straightforward, in terms of accessibility, time and cost, for rightsholders to reserve their rights, but is also needed to ensure it is straightforward for AI developers to comply with a rights reservation. How standardisation is achieved is critical and the Government must engage with the creative industries on this.

Question 11. Should the government have a role in ensuring this and, if so, what should that be?

The Government must have a pivotal role in ensuring ongoing compliance with a rights reservation by AI developers, but also have ownership, independent oversight and control over the development, testing, validation and implementation the rights reservation, ensuring due and proper consultation of any technological solution at all stages with rightsholders and creators. The Government must ensure that any rights reservation system meet a standard so that if it chooses to introduce an exception the rights and interests of rightsholder and creators are protected at all times and that it is continually reviewed against the purported objectives of the regime.

The Government must produce a clear proposal as to how a rights reservation will work, for all rightsholders, for all types of work, and who (and how they) will bear the cost. The reservation of rights involves considerable effort by rightsholders, for little reward, particularly if there is no certainty as to whether licensing is possible. Reserving rights must not become a burden and the cost of a rights reservation should not fall to rightsholders. There is a risk that smaller rightsholders will be left behind by a rights reservation system, unable to effectively assert their rights as the burden of reserving rights will be too onerous and too costly to be viable.

In developing a proposal for a rights reservation, the Government must consult with rightsholders and their representatives, including testing any proposed system, to ensure confidence in the framework. The Government must also have a role in how to address and reconcile past infringing uses of copyright protected works which no new exception or rights reservation, as proposed, will resolve retrospectively.

As per its statement in paragraph 68, the Government must also commit to not introduce any exception unless the 'robust measures' referenced are in place.

C.3 Contracts and licensing

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

Based on CLA's understanding, including insight from its member organisations who represent authors and visual artists, because so few licences have been entered into, the current practice is not meeting the needs of creators and performers. We are not aware that creators have received remuneration for the use of their works, creators have had no choice as to whether their works are used or not, and there is no transparency of what works have been used.

CLA has engaged with the IPO working groups on an AI and Copyright Code of Practice, and has sought to engage directly, via personal and industry networks, with AI developers. Despite being the UK's sole CMO for the reuse of book and journal content, CLA has been unable to engage substantially with any AI firms to discuss their licensing requirements and how any perceived barriers to copyright compliance might be overcome.

Without engagement from AI developers, it is impossible to establish any licensing that will subsequently ensure creators needs are met. As referenced in CLA's response to Question 3, there are some licensing arrangements in place between AI developers and some larger rightsholders, but these do not appear to be comprehensive enough to cover all creators. Details as to what terms have been agreed are generally not known due to the confidentiality of such agreements.

Question 13. Where possible, please indicate the revenue/cost that you or your organisation receives/pays per year for this licensing under current practice.

CLA does not currently receive any revenue for the licensing of copyright works for AI training, i.e. the scaping of websites and ingestion of copyright-protected works for training by AI developers.

Question 14. Should measures be introduced to support good licensing practice?

To support good licensing practice, CLA believes the Government needs to introduce good transparency practices, in the form of mandatory transparency measures. It is only

with a framework for transparency that good licensing practices can successfully develop.

We believe that good licensing practice is at the core of what CLA does as a CMO. As outlined in the response to Question 2, CLA has evolved its licence offer to include additional permissions to cover TDM and workplace generative AI permissions and continues to engage with its member organisation (who represent publishers, authors and visual artists) on future licence enhancements to meet user needs.

CLA is willing to engage with the IPO as to how licensing may be supported beyond the scope of the four options proposed in the consultation.

Collective licensing and data aggregation

Question 15. Should the government have a role in encouraging collective licensing and/or data aggregation services? If so, what role should it play?

In a developed licensing market, such as the one CLA currently operates in when offering licences to the education and corporate sectors, there is limited need for the Government to play a role. The IPO regulates CMOs under The Collective Management of Copyright (EU Directive) Regulations 2016 and provides information on gov.uk on each CMO. The light-touch approach of the IPO works well in this regard. In an emerging market, such as the licensing of works for training AI, there is scope and arguably a need for the IPO to have a more proactive role. This could involve signposting, recommending or endorsing licences offered by CMOs to help ensure confidence in the emerging licensing market.

Most importantly, the Government needs to introduce mandatory transparency measures, as outlined in our response to Question 4, which will facilitate licensing. Knowledge of what works have been used / are being used by AI developers, and how such works are being used will enable CLA, in consultation with its members, to develop an appropriate licence offering to ensure that AI developers can then comply with copyright law. This will then ensure that creators and rightsholders are fairly compensated for the use of the work, which supports the continued growth of the creative economy, meaning creative works continue to be produced which further supports AI developers when training, fine-tuning and improving their generative AI systems.

Use of AI in education

Question 16. Are you aware of any individuals or bodies with specific licensing needs that should be taken into account?

CLA is not currently aware of any individuals or bodies in education with specific licensing needs but is in the process of engaging in focus groups and roundtables with

representatives in the education sector to understand what AI activities are taking place and any subsequent licensing requirements.

C.4 Transparency

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

It is imperative that AI developers disclose the sources of their training material. Without this information, it is impossible to offer licences and for a functioning market for licensing to develop, and therefore impossible to remunerate rightsholders for the use of their works and provide rightsholders with choice as to whether their works are used or not. It is also impossible for rightsholders to be able to enforce their rights and take action where their works have been used unlawfully, as there is insufficient evidence available due to the lack of transparency.

Furthermore, it stands to reason that whatever works AI developers use for training, will influence the output of those systems. It is likely that AI systems trained on unauthorised and unreliable works will produce untrustworthy and unreliable outputs. The lack of adequate transparency as to the source of works creates a risk of low-quality outputs, misinformation and bias, as well as an overall lack of trust. For AI to develop and grow in the UK, as the Government envisages, there must be public trust in the AI systems that are developed. Such trust can only be realistically achieved with transparency by AI developers.

Question 18. If so, what level of granularity is sufficient and necessary for AI firms when providing transparency over the inputs to generative models?

For licensing of text works, CLA requires title level data, such as ISBNs and ISSNs. This information can then be used to identify the publisher, and author(s) of a work. We also hold URL data for websites, and additional work will need to be undertaken to ensure that all works contained within a URL can be identified. Additional granularity will be needed in relation to image works, and we refer to our DACS and PICSEL's submissions in relation to this. We have outlined our requirements for transparency in our response to Question 4.

CLA currently holds 34 million titles in its database. This data is used to enable the distribution of licence revenue to rightsholders, via our member organisations (ALCS, DACS, PICSEL and PLS) and is also used for users to be able to check whether a work is covered by a CLA licence, or not, via our Check Permissions tool.

It is therefore feasible for data to be captured on large volumes of works, as many CMOs already do.

Question 19. What transparency should be required in relation to web crawlers?

It must be clear what crawlers are being deployed by AI developers, and for what purpose they are scraping content. As with the transparency requirements for AI developers, a level of granularity is necessary for rightsholders to be able to exercise their rights. It must be possible for rightsholders to choose whether works available online are scraped or not, depending on the purpose the scraping is being undertaken for.

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

We believe that regulation is necessary to ensure that transparency requirements are met. Without regulation and enforcement, there is little incentive for AI developers to comply and the current status quo is likely to continue.

Question 21. Where possible, please indicate what you anticipate the costs of introducing transparency measures on AI developers would be.

CLA cannot anticipate what the costs of measures will be. However, the information that we are advocating be included in mandatory transparency measures should already be recorded by AI firms as a matter of good business practice. We do not anticipate that it would incur a high cost to organise and present this information in an accessible, accurate and complete manner, that can be easily understood by the public and creators and rightsholders.

Question 22. How can compliance with transparency requirements be encouraged, and does this require regulatory underpinning?

The Government, or an appropriate regulatory body, needs to be able to take action if developers are not compliant with any mandatory transparency measures. The burden should not fall on rightsholders, many of whom will not have the financial means or otherwise be able to take civil action against AI developers.

Question 23. What are your views on the EU's approach to transparency?

CLA has limited experience of the EU's approach to transparency. Anecdotal evidence suggests that, to date, it has not been successful although work continues in the EU with the progression of the General Purpose AI Code of Practice.

C.5 Wider clarification of copyright law

Question 24. 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 right holders are respected?

The growth and success of generative AI should not come at the expense of the UK's creative industries, worth £124 billion annually to the UK. There is no evidence to suggest

that amending copyright law and introducing a new exception will encourage AI developers to train their models in the UK or encourage greater licensing. As noted earlier in our response, there is no evidence that the introduction of Article 4 in the DSM Directive in the EU has not led to either increased licensing by rightsholders of their works, nor to AI developers choosing to train their models in the EU.

Question 25. 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?

No response

Question 26. Does the temporary copies exception require clarification in relation to AI training?

s28A of the CDPA 1988 (Making of temporary copies) must not be used by AI developers as justification for not requiring permission from rightsholders. There is ‘independent economic significance’ attached to the outputs of AI training, and absent a licence from the rightsholder, it does not appear to form part of a process intended to enable the otherwise lawful use of the work.

Question 27. If so, how could this be done in a way that does not undermine the intended purpose of this exception?

No response

C.6 Encouraging research and innovation

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

The existing data mining exception for non-commercial research remains fit for purpose and we understand that it is used effectively by the UK public and research sectors today.

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?

There should not be a precedent set that copyright rules are different depending on the purpose of the AI model or the size of an AI firm.

However, it is foreseeable that licensing fees, whether under a direct licence, or a voluntary collective licence, would take into account the size of an AI firm (in terms of users under the licence, for example), as well as the purpose of a model. This type of flexibility in a licensing market is one of the many reasons why licensing can benefit AI developers and users.

D.1 Computer-generated works: protection for the outputs of generative AI

Question 30. Are you in favour of maintaining current protection for computer-generated works? If yes, please explain whether and how you currently rely on this provision.

Question 31. Do you have views on how the provision should be interpreted?

No response.

Question 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?

No response.

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

No response.

Question 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.

No response.

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

No response.

Question 36. What would be the economic impact of doing this? Please provide quantitative information where possible.

No response.

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

No response.

D.4 Infringement and liability relating to AI-generated content

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

For effective of enforcement of copyright, mandatory transparency measures as discussed in this response are required.

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

First and foremost, AI providers need to ensure that the necessary permissions have been obtained for the works that have been utilised to train and fine-tune to underlying generative AI systems.

D.5 AI output labelling

Question 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?

Labelling works as AI generated should be a legal requirement. The labelling of outputs will have many benefits including ensuring that the licensing market does not become flooded with AI generated works, which risks damaging rightsholders livelihoods.

Moreover, labelling of AI outputs is needed for AI to be developed safely, ethically and legally for wider societal benefit. In addition to copyright infringement, there are serious risks associated with AI outputs, including misinformation, bias and deep-fakes which pose an unknown and massive threat to civil society. Labelling must therefore be overseen by regulation and enforced by the appropriate regulatory authorities in the UK.

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

The Government should ensure that any mandatory transparency measures include the requirement that outputs are labelled as AI generated.

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

We are broadly supportive of the proposed approach in the EU, where labelling is mandatory under the EU AI Act. However, this is not yet fully implemented in the EU, and therefore the workability of this approach has not been fully tested.

D6 Digital Replicas

Question 43. To what extent would the approach(es) outlined in the first part of this consultation, in relation to transparency and text and data mining, provide individuals with sufficient control over the use of their image and voice in AI outputs?

No response

Question 44. Could you share your experience or evidence of AI and digital replicas to date?

No response

D7. Other emerging issues

Question 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?

As stated in our responses to previous questions, the legal framework is clear: a licence is required for the copying works by AI developers to train AI models.

Question 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?

No response

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

Currently, there are not any other developments related to the UK copyright's framework that we would like to raise.

CLA supports the submissions made by each of its member organisations.

CLA is a member of the British Copyright Council (BCC) and supports its submission.