BAPLA
British Association of Picture Libraries &Agencies

BAPLA Response to the IPO AI Consultation on Copyright and Related Rights

About BAPLA

Founded in 1975, BAPLA is the UK trade association for picture libraries and agencies representing over 115 members, of a unique area of the creative industry. We have a broad and diverse membership of image rights holders and purveyors, from sole traders to major news, stock and production agencies, as well as many SMEs, archives and cultural heritage institutions.

BAPLA members are the main source of licensed images you see every day in print and digital media, and as such have contributed to the UK economy for over 45 years. Our members generate revenue for, and manage the interests of, over 120,000 professional photographers, videographers, illustrators and other image makers, supporting their ability to derive income and reinvest in their creativity. The majority of BAPLA's membership consists of small and medium enterprises (SMEs) accounting for approx. 72% of the membership, with 28% as sole traders - many operating on very tight economic margins. In addition, the UK represents the 2nd largest territory for the global still images market - a market which has the potential to grow by £1.24 billion ($1.55 billion) between 2019-2023[1].

BAPLA members have invested heavily over many years in technological innovation allowing them to move from analogue to digital, digitising millions of images of great historical and artistic value in order to offer access to digital copies through an extensive and adaptable range of licensing solutions. Significant investment is made on an ongoing basis to accurately keyword and add metadata to over 300 million images, allowing for seamless customer experiences on the one hand, and full remuneration of creators on the other. Image libraries use a range of AI-based applications to better store and separate images, as well as providing search and discovery functions that drastically improve usability. They use image recognition APIs to provide image tags, auto-generated keywords, and automatic categorisation tools based on visual categories, often across devices. Image library websites use AI image recognition tools to assist both in the upload and appropriately tagging of image content, and to provide better support to customers looking for images they intend to license. Several of our larger members have invested in new technology application systems, such as the use of artificial intelligence and machine learning techniques[2].

IPO's Artificial intelligence call for views

BAPLA welcomes the IPO's AI call for views in relation to copyright and related rights and is pleased to present our response to this consultation, and acknowledges the Government's objective to support the AI sector. We also acknowledge the ambition of the UK in becoming a global centre for AI, and its review of the current framework focusing on obstacles and incentives to achieve this objective.

Within the image industry, artificial intelligence (AI) and machine-learning (ML) is being applied with greater regularity to image search and capture, metadata and image tagging.

[1] Still Images Market Industry Analysis, 2019, Source: TechNavio https://www.technavio.com/report/still-images-market-industry-analysis

[2] Getty Images releases an artificial intelligence (AI) tool for publishers https://bit.ly/2OLGK9v; Shutterstock uses AI to help people avoid image copyright issues https://bit.ly/2ZBYSJ3 and in partnerships to support image licensing https://bit.ly/2HKwFK0

The interest by picture libraries and agencies in adopting the technology is because AI tools can be used to discover consumer preferences in a way that humans cannot, as well as speed up certain processes used by agencies handling 100s of thousands to millions of images. In addition, some picture agencies are already licensing content to be used for AI training purposes, and potential new entrants to the market are emerging offering AI created synthetic content[3].

With the photographic industry adapting at pace to AI, we strongly believe a fair and equitable system that supports both the AI industry and image industry is achievable with only minor clarifications to the current copyright regime aimed at system transparency. It is our position that the copyright licensing framework already offers flexible arrangements attractive to AI and ML developers and the AI ecosystem, as well creative content rightsholders, without the need to either revise or make new changes to UK copyright law. There is a suite of other protections in the IP portfolio that can be better adapted to any gaps in the current IP framework for AI.

Key Points:

To summarise our key points, discussed in more detail below, we set out the following:

1. AI-generated content is a derivative of the underlying content and data. As such the input of content, and content-related data, should be recognised, with owners of underlying works being acknowledged as an important part of the AI value-chain deserving protection.
2. Clear standards must be set out in relation to the use of copyright works and data by AI systems particularly around the maintenance and transparency of data used in AI development;
3. Licensing frameworks are responsive and adaptable, and are best placed to support AI innovation. We believe that licensing is more agile and better suited to adapt to the still-developing technologies than broadening any exceptions that might prove to be a short-term and overly rigid solution, and can have unintended consequences on human creativity;
4. Images are extremely valuable and important in the context of content resourced by AI programs, as images are condensed and extremely rich source of information. Careful consideration needs to be given to all uses of visual works by AI;
5. On liability and enforcement, recognising the importance of identifying original sources and ownership of underlying visual works used in AI programmes & tools, to ensure permission has been sought to use the underlying creative works. As AI is often developed by a cluster of enterprises (data companies, software designers, AI model designers and corporate AI users) liability for infringement (on a primary or a vicarious basis) should attach to all those deriving benefit from copyright infringement.
6. Copyright should apply specifically to human endeavour. Protection of AI-works should fall under other IP protections, such as patents, as it is more suited to a rapidly

[3] A few examples:

- https://icons8.com/articles/ai-generated-faces/
- https://www.rosebud.ai/tokkingheads
- https://generated.photos/

evolving technology. Providing AI-works with copyright protections would simply add an unfair advantage that would inadvertently impact original human creativity.

Questions:

The use of copyright works and data by AI systems

1. ***Do you agree with the above description of how AI may use copyright works and databases, when infringement takes place and which exceptions apply? Are there other technical and legal aspects that need to be considered?***

Yes, in general we agree with the descriptions set out by the IPO explaining the current framework for the use of copyright works and databases. There are several key principles reflected in the Copyright, Designs and Patents Act 1988 (**CDPA**) which are sufficient in providing fair and flexible requirements for AI/ML technology developers and trainers, alongside other technology and innovation developers.

First, the CDPA balances the rights of rightsholders and users by a carefully calibrated system of exclusive rights of copyright holders on the one hand, and exceptions from such rights on the other, in line with the United Kingdom's commitments under international agreements.

The most recent set of exceptions added to the CDPA in 2014 included a broad text and data mining (**TDM**) exception for non-commercial research. The Impact Assessment produced in connection with the introduction of the exception explained that the exception was aimed at improving the UK's non-commercial research activity by saving research time[4]. This is consistent with the educational orientation of copyright exceptions in general. We submit that broadening or adapting the TDM exception for the benefit of commercial AI enterprises would not only be at odds with the rationale of the copyright exceptions but will also seriously erode the fundamental right to ownership of intellectual property belonging to copyright holders. It seems inconceivable that the Government would choose to support the development of the AI sector by encroaching on well-established property rights.

In the Impact Assessment referred to above, the Government argued that the TDM exception is necessary in order to provide time savings for non-commercial research. It was considered that although wide-ranging licensing solutions exist, they might limit academic access. Outside of this limited use, the Government acknowledged that there is no general "right to mine". As such, we are firmly of the view that licensing solutions are adequate in the context of commercial use of AI technologies. It does not seem right, and would be discriminatory, that profit-making AI businesses should be given access to data and content without acknowledging the value and the cost of the investment in content generation by rightsholders. To grant an exception would be to sacrifice one business sector, at the expense of another, which appears currently more alluring. This arbitrary selection should not be the basis for policy making.

We stress that there is a positive opportunity here to use licensing solutions to resolve uses that go beyond research undertaken by non-commercial institutions. For example, when photocopying machines were first introduced as new revolutionary technology, collective

[4] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/308738/ia-exception-dataanalytics.pdf

licensing offered the flexibility required; the same can easily be applied without the need to revise the current UK exceptions.

If the Government proposes to adjust existing exceptions it would be helpful to have further legislative clarity about when such an exception would apply to AI and ML (as they are not specifically referred to), including clarification on when copyright exceptions can never apply - we would argue they should never be able to apply where potential output of an ML or AI process could be substitutable with input data (where images are used as data to create synthetic content). Visual content is also likely to contain personal data, therefore any clarification should also include considerations as to how a copyright exception would affect data protection rights of individuals whose data is mined. In all cases, a fair dealing test would need to be applied, and acknowledgement of the Berne three-step test.

If notwithstanding our contentions a new exception is created, or mining of content is undertaken by seeking to rely the existing TDM exception, determining whether the exception applies or not to particular uses may require a granular assessment of content used purportedly under the exception. This could be a challenge if the used works, and their uses, cannot be traced and made transparent.

We believe that a straightforward solution would be to have a statutory requirement for developers of AI to retain auditable records of what data has been used. Then, where input data contains copyright works, questions about whether an exception should apply, or – in a licensing scenario – whether use falls in the scope of the licence – could be determined subsequently. In each case, a legal requirement to maintain auditable and records would help instil trust in AI systems, by enabling developers and operators of those systems to demonstrate that they have used "good data", which is less likely to lead to discriminatory or biased outcomes.

Ultimately, there should be no need to rely on copyright exceptions; if AI developers are doing the right thing and keeping a log of data sources, not only would it give peace of mind to AI developers who have sought licences, it would expose bad actors who have not acquired permission.

Another helpful goal is to aim for better understanding of how algorithms and AI works in the context of copyright-protected creative content. Furthermore, if there is confusion over the use of copyright-protected works and data by those developing AI systems it would make sense for the IPO to consider working with organisations, like the British Copyright Council, to increase educational resources and bridge any gaps of knowledge or understanding within the AI industry, and vis versa.

2***. Is there a need for greater clarity about who is liable when an AI infringes copyright?***

Yes, we believe there is a very strong need for greater clarity about who is liable when copyright is infringed in relation to AI. Effective copyright enforcement for rightsholders is essential in ensuring our members are given the incentive to continue to invest and produce new artistic works such as photographs, illustrations, and video clips.

It is defined within the CDPA (Section 16.2) that copyright is infringed by a 'person', and case law has made it clear that any individual, group of associated individuals, as well as legal persons (incorporated or unincorporated) might be liable for infringement either under the primary or secondary liability regime, as well as under vicarious liability for acts of persons

within its control. AI technologies depend on an ecosystem of enterprises including data scrapping companies, software developers, AI trainers, analysts, and companies utilising AI software. Copyright liability should extend to everyone involved in the process, where copyright infringement takes place, and in particular to those within the ecosystem who store and process copyright works.

Whilst enforcement is the same whether as a human or an AI programme in a legal context, the significant challenge is for rightsholders to prove infringement has taken place, bearing in mind the vast volume and speed at which an AI programme can digest information - it is extraordinarily difficult, if not impossible, to determine. There is an issue with which works have been used and how much, as humans we cannot monitor the use in an easy or meaningful way. We would have to further invest in developing and using new AI enforcement tools to prove AI output infringements, which could prove to be costly and disincentivise the industry further. Ownership is the key to the issue of enforcement, although that should not muddy the water, courts would have to track back to the code itself - who owns the code and who owns the new output material.

3. ***Is there a need to clarify existing exceptions, to create new ones, or to promote licensing, in order to support the use of copyright works by AI systems? Please provide any evidence to justify this.***

As discussed above, we do not believe that new exceptions should be introduced. We strongly advocate licensing solutions as the best approach to maintain fair practice and opportunity for both human creative-endeavour and AI-endeavour. Promoting the licensing of copyright works used in the development and training of AI and ML systems is of paramount importance. Images are valuable and important in the context of content resourced AI programmes, as images are condensed and extremely rich, they carry a lot of valuable information. Images, together with associated metadata, provide immensely rich fuel for training AI and ML systems. Therefore, it seems only fair that human image creators should derive value from the technology, and as such licensing would offer the most appropriate framework.

Some BAPLA members have already taken the initiative of licensing images for AI trainers[5], and working with certain AI developers[6], paving an opportunity for other image libraries to follow suit. Other licensing solutions are potentially in scope, such as collective licensing, as long as there are clear parameters set to protect rightsholders, particularly those who are not members of a collective management organisation.

4. ***Is there a need to provide additional protection for copyright or database owners whose works are used by AI systems? Please provide any evidence to justify this.***

[5] For example, Getty Images has recognised this market as deserving of special attention and excludes the right to use its content for AI and ML purposes unless a customised licensing agreement is entered into. Considering the high volume of content that is often required for such uses, Getty Images is able to facilitate delivery of content using its own APIs, details of which can be found here: https://www.gettyimages.co.uk/solutions/en/api

[6] Shutterstock in partnership with IBM Watson Content Hub https://www.digitalstrategyconsulting.com/online-advertising/online-advertising-research-tips-and-news-for-marketers/shutterstock-uses-ibm-watson-ai-to-help-marketers-find-images/16406/

As set out in our response to Question 1, the current copyright regime adequately addresses the balance between rightsholders and users. Provided that the development of AI systems is supported by appropriate licensing arrangements with rightsholders, we do not believe that copyright needs strengthening.

Databases of photographic content and metadata may qualify for the database protection under Section 3A CDPA, and/or sui generis database rights under the Copyright and Rights in Databases Regulations 1997. Owners of rights in databases are protected from substantial copying under general principles of copyright, and from extractions from databases (other than those of insignificant parts of databases) under the Regulations.

Any use of our members databases for AI training would likely involve wholesale scrapping of the relevant content and thus, if conducted without a licence, would fall outside lawful copyright or database use.

While the letter of the law does not require changing, in order for the rights it creates to be enforceable, and legal protection effective, it is crucial that rightsholders, whose creative works are used as input for AI development and training, have an ability to trace the use of their content and data.

Also, as noted above, greater clarity over the parameters of the TDM and temporary copying exceptions in the specific context of AI and ML use is desirable. At the very least, in addition to the existing caveats contained within these exceptions, it would be helpful if a fair dealing test were to apply to such exceptions. Another helpful clarification would be for it to be explicitly acknowledged that data crawled from websites does not constitute "lawful access" under Section 29A(1) of the TDM exception, especially if the website includes user Terms & Conditions that prohibits such data mining or extraction.

Protecting works generated by AI

5. Should content generated by AI be eligible for protection by copyright or related rights?

In order to answer this question, it needs to be clarified what is meant by “content generated by AI”. For content generated exclusively by AI without any human creative intervention, we think the answer should be, no - such works should not be eligible for protection by copyright or related rights. This is because no economic incentive is required in these circumstances and the implications for human creative endeavour could be devastating human creators unable to compete with AI content created both at scale and at negligible cost, and being deterred from creating new content for fear of infringing rights of AI-generated content. It is vital that the UK copyright framework maintains human incentives for creation (making free and creative choices) and dissemination.

However, where AI is used as a tool to aid human creative endeavour, resulting in original works according to established copyright principles, i.e., author’s own intellectual creation reflecting their personality, then we would argue that such works could, in very limited circumstances, be afforded protection by copyright, with the threshold for originality and human creative input in this context being set at an exceptionally high level compared to other methods of creation. Guidance, or clarification, on how such tests should be interpreted, where AI is used as a tool would be useful.

Also, a pre-condition to any new, original "AI-assisted" works being afforded copyright protection should be that if any copyright works have been used as input data at any stage in the development or deployment of the AI system, then copyright in the output should not arise. In addition, if there is a possibility of joint ownership of a derivative work created with AI technology, a threshold of human creative input would need to be applied – a Related Right as protection could be applied.

This leaves the question of Section 9(3) of the CDPA regarding computer generated works and identifying who the author is. Considering that this is an outlier globally, that there has been only very limited reliance on this provision in the English courts, and that it has the potential to confuse the position with regard to works created by or with the assistance of AI technology, we queried whether it should be repealed in favour of the above suggested guidance regarding new "AI assisted" works. Our members' views diverge on this topic and arguments were made in favour of retaining Section 9(3) CDPA as is, given that it had been specifically designed to address technology developments where machines create works without human creative input, and if the provision requires further interpretation, the courts are ideally placed to give guidance to the provision.

Referring once more to the apparent intention behind the consultation namely, to explore whether the UK IP laws are fit to make the UK a global hub for AI, the question about protection of AI works becomes a question of economic incentives for technology companies. It is submitted that the IP law is only one of many such incentives and often the deciding factors are to do with digital infrastructure and financial/tax incentives rather than copyright ownership. This notwithstanding, and considering copyright as an incentive in isolation, we do not believe that additional copyright protection for AI technology is the most appropriate tool to incentivise the AI sector.

The "secret sauce" that makes an AI system competitive is the software which programmes the machine to derive a set of characteristics from data. Those characteristics are subsequently compared with the data sample for the purpose of establishing statistical patterns. This is the "learning" of which is reinforced across a different data set and so on, and in the process either reinforced of modified. These processes are possible because of the instructions of the programmer contained in the software code. The latter is subject to copyright protection. An additional incentive in the same field does not seem necessary nor justified.

If one considers whether attaching copyright to output data, for example, where millions of pictures sourced from picture library website databases are used as samples to generate synthetic images could, in fact, reduce the competitiveness of technology companies. A competitive environment, in which AI technology applications and output is tested by the market is, one would assume, necessary to create a global AI hub. Only a competitive environment will attract a range of undertakings seeking to innovate and disrupt. Granting copyright protection (lengthy and non-cancellable in contrast to other IP rights) could result in dominant technology companies monopolising AI learnings. To give an example from the image industry – 'Company A' produces synthetic images based on the learnings (pattern recognition) described in the paragraph above. If images created by 'Company A' are protected by copyright, a competitor ('Company B') whose software discerns the same patterns could be liable for copyright infringement of 'Company A'. As copying occurs in the machine's "neural networks", which are not well understood by programmers themselves, 'Company B' would face an impossible evidentiary burden in defending itself against a charge

of copying. This would lead to the monopolising of the AI market by few dominant players, with a disadvantage to creativity and society as a whole.

Our final argument against affording copyright protection to "pure" generated-AI works (those which can be said to be created without human creative intervention) is to do with issues arising out of copyright protection online. AI technology has arisen out of vast data sets made possible by the world wide web and cloud technologies. These conditions create infinitely complex issues to do with jurisdiction in which works can be said to have been created, as well as proper (let alone cost-effective) forum for enforcement. The image industry, and other creative sectors, continue to suffer significant losses of revenues as a result of online infringement[7] and the inability of national legislatures and international bodies to ensure effective protection of copyright works online. Creating a new type copyright protection without addressing the desperately needed measures to protect against digital piracy will hardly serve as an incentive for AI businesses.

6. ***If so, what form should this protection take, who should benefit from it, and how long should it last?***

Please refer to our response to Question 5, above. In general, the UK offers a wide variety of protections available within its IP portfolio that could be better applied to AI in order to provide those investing in this area, which is still adapting and changing, without a drastic need to amend the UK copyright framework. It makes sense for the IPO to undertake a mapping exercise on current AI licences by programmers, platforms, and users, in order to identify gaps in understanding where the benefits of copyright licensing solutions can support AI development and training, in order to address misconceptions that it is necessary to introduce revised or new copyright protections. Areas such as ownership of human created content used for AI input and AI-generated output should be included in the process, particularly in order to understand where policy needs to be clearer in relation to the current framework.

7. ***Do other issues need to be considered in relation to content produced by AI systems?***

As commented in Question 5, "AI assisted works" may need to be considered in relation to originality and human endeavour, and therefore depending on the level of human input may not necessarily be provided with copyright protection. Whatever the level of contribution from AI, it should not automatically trigger co- or joint authorship, as copyright protection is afforded to human endeavour, even if the human element is considerable, as the AI element still lacks human input. This is where the importance of documenting the input sources used and technology applied is critical, including from a moral rights perspective. An appropriate threshold of human creativity would also need to be determined.

Copyright protection for AI software

8. Does copyright provide adequate protection for software which implements AI?

[7] BAPLA Research into Online Copyright Infringement https://bapla.org.uk/wp-content/uploads/2019/12/BAPLA-Research-Into-Online-Copyright-Infringement-2019.pdf

Yes. Again, please refer to our response to Question 5 above. Whilst in general software is developed by human beings, as AI technology complicates this approach and human involvement reduces, it would follow that an appropriate threshold of human creativity would need to be determined.

9. Does copyright or copyright licensing create any unreasonable obstacles to the use of AI software?

No. Copyright licences have a proven track record of being suitable for streaming, video gaming, online education, and similar previously unknown forms of exploitation. There is no reason why licensing should not be appropriate for AI applications of content. From a policy perspective, the Government should ensure that there are obstacles that keep copyright-protected content from being used without authorisation. Any such obstacles in law is not unreasonable in light of established property rights and are necessary to protect human creators and incentivise them to continue to create original works.

Final Comments

In our final comments, we believe further granular assessment and greater stakeholder engagement needs to be undertaken by the IPO in order to fully outline the scope of the role AI takes in applying creative content protected by copyright (such as whether there is a greater requirement to promote licensing solutions across the AI sector) in this area before any detailed policy is outlined. It may help to start with mapping the level of use and understanding of licensing solutions, and to follow this with setting out clear standards on what AI developers and trainers should undertake when sourcing data inputs for their programmes, which should include high level principles. If we stimulate content-based AI programme development, then we must continue to incentivise support for the creation of the underlying content used as input.

[redacted]
http://www.bapla.org.uk