MACQUARIE LAW SCHOOL
Faculty of Arts

14 February 2020

World Intellectual Property Organization

Re: Consultation on Impact of Artificial Intelligence on IP Policy

Dear Sir/Madam

We welcome the initiative taken by the World Intellectual Property Organization to seek input through the *Draft Issues Paper on Intellectual Property Policy and Artificial Intelligence*. Artificial Intelligence has been increasingly used by various industries and public sector organizations and its legal status with relation to intellectual property rights need to be clarified as soon as possible.

These submissions are intended to be made public.

These submissions deal only with a selection of particular questions posed in the Issues Paper, in particular, Copyright and Related Rights Issue 6: 'Authorship and Ownership' and Issue 7 'Infringement and exceptions'.

Sincerely

Dr Rita Matulionyte
Senior Lecturer
Macquarie Law School, Room 512
6 First Walk, NSW 2109 Australia
T: +61 2 9850 7091
E: rita.matulionyte@mq.edu.au

Prof Niloufer Selvadurai
Macquarie Law School
6 First Walk, Room 621
Macquarie University NSW 2109
E: Niloufer.selvadurai@mq.edu.au

Response to Issues Paper on Intellectual Property Policy and Artificial Intelligence

Issue 6: Authorship and Ownership

Introduction

AI technology has already generated music, works of arts, poems and even a film script. The ownership of AI-generated works has been extensively discussed in legal doctrine. Currently, the ownership question is addressed differently in different jurisdictions. For example, in Australia, a human being should provide an original contribution for a work to be protected under copyright. In particular, in *Telstra Corporation v Phone Directories Company*, the Full Federal Court of Australia clarified that the work should originate with an author or joint authors who should exert 'independent intellectual effort'.[1] In the USA, a requirement of a human involvement in the creation process has been established through the 'Monkey selfie' case where ownership of a photo made by a monkey was denied copyright protection.[2] Similar rules apply in most continental European jurisdictions.

In contrast, UK copyright law allows copyright to be awarded to the work generated by a computer. It provides that the author of a computer generated work is taken to be 'the person by whom the arrangements necessary for the creation of the work are undertaken'.[3] Similar provisions exist in Ireland, Hong Kong, India and New Zealand.[4] Although this provision has been little tested in courts,[5] it could arguably apply to AI-generated works.

[1] *Telstra Corporation Ltd v Phone Directories Company Pty Ltd* (2010) 194 FCR 142, [82].
[2] *Naruto v. Slater*, No. 16-15469 (9th Cir. 2018)
[3] *Copyright, Designs and Patents Act 1988* (UK) s 9(3)
[4] See, eg, in New Zealand "the person by whom the arrangements necessary for the creation of the work are undertaken": *Copyright Act 1994* (NZ) s 5(2)(a); in India "the person who causes the work to be created": *Copyright Act 1957* (India) (2)(d)(vi); in Hong Kong "the person by whom the arrangements necessary for the creation of the works are undertaken": *Copyright Ordinance* (Hong Kong) cap 528, s 11(3); in Ireland "the person by whom the arrangements necessary for the creation of the work are undertaken": *Copyright and Related Rights Act 2000* (Ireland) s 21(f).
[5] See, e.g *Nova Productions Ltd v Mazooma Games Ltd* [2006] EWHC 24; [2006] RPC 379.

Most recently, the issue of ownership of AI-generated works has been tested in Chinese courts. In *Feilin v Baidu* case[6] in 2019, Beijing Internet Court ruled that reports generated by AI were not protected by copyright. However, even so, the court suggested that they do not fall into public domain and cannot be freely reused by anyone.[7] In early 2020, in *Shenzhen Tencent v Yinxun* case, Nanshan District Court of Guangzhou Province decided that a press article generated by AI is protected by copyright. [8]

These development shows that the issue of ownership of AI-generated works is of increasing relevance in practice and has no single solution so far.

Question (i) - Should copyright be attributed to original literary and artistic works that are autonomously generated by AI or should a human creator be required?

AI systems apply machine learning algorithms to process pre-fed inputs and data and generate a creative work.[9] The nature and degree of human intervention and oversight involved in such a creation can vastly vary, ranging from being confined to merely data selection and labelling to encompassing all stages of creative life cycle including the selection and presentation of the work. As intimated by the *WIPO Conversation on Intellectual Property and Artificial Intelligence* report,[10] this issue requires a consideration of the purpose of copyright law.[11] If the purpose of copyright law is to reward and incentivise creativity and innovation by human beings, eligibility for copyright protection should require a substantial element of human action. In

[6] Full decision in Chinese available here: https://www.bjinternetcourt.gov.cn/cac/zw/1556272978673.html
[7] Kan He, "Feilin v. Baidu: Beijing Internet Court tackles protection of AI/software-generated work and holds that copyright only vests in works by human authors" (9 November 2019) http://ipkitten.blogspot.com/2019/11/feilin-v-baidu-beijing-internet-court.html
[8] http://ipkitten.blogspot.com/2020/01/another-decision-on-ai-generated-work.html
[9] Burrell, J, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 3(1) *Big Data and Society* 10.
[10] WIPO, WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI), Second Session, 13 December 2019, p. 5.
[11] Kaminski, ME, 'Authorship, Disrupted: AI Authors in Copyright and First Amendment Law' (2017-2018) *UC Davis Law Review* 589; *See further* Ruipérez, C, Gutiérrez, E, Puente, C & Olivas, J, 'New Challenges of Copyright Authorship in AI,' 2017, *International Conference on Artificial Intelligence*, 291-296 at p. 294.

comparison, if the purpose is to support the efficient allocation of resources in a society, such a position would more readily support the recognition of machine creation, irrespective of the absence of any direct or indirect contribution by a human creator.

In addressing the issue it is also relevant to consider the way in which the law should respond to technological disruption. In this regard, it has been suggested that laws to address technological change should, where possible, seek to build and extend existing laws rather than create new law.[12] This is because the latter approach can lead to a patch-work quilt of discrete and potentially inconsistent provisions. The absence of an overarching holistic approach can also lead to double-regulation and gaps in regulation. In developing copyright laws to govern AI, it is hence submitted that such an approach should be adopted and that the laws governing AI works should be consistent, as far as feasible, with existing copyright principles on authorship and originality.

Accordingly, a work created using AI should only be considered sufficiently original to attract copyright protection if its creation has involved substantial human input. Such a position is consistent with Australian law and we believe properly calibrates competing interests. The Australian *Copyright Act* 1968 (Cth) stipulates that copyright can only subsist in a literary, dramatic, musical or artistic work if it was created by an author. Considering the issue of originality and machine creation, the Australian High Court held in *IceTV Pty Limited v Nine Network Australia Pty Limited* [2009] HCA 14 that only material that has originated from an author who has expended "independent intellectual effort" will be original for the purposes of the *Copyright Act*. The Australian Federal Court in *Telstra Corporation Limited v Phone Directories Co Pty Ltd* (2010) 194 FCR 142 further stipulated that "[a]uthorship is central to the determination of whether copyright subsists, noting that "to suggest that copyright does not require the identification of authors where a work is sufficiently original

[12] *See further* Selvadurai, Niloufer, 'Not just a face in the crowd: Addressing the intrusive potential of the online application of face recognition technologies' (2015) 23(3) *International Journal of Law and Information Technology* 187-218; and 'Designing copyright licensing laws to support the converged digital economy' (2015) 37(5) *European Intellectual Property Review* 285-295.

(howsoever that question of originality is to be answered) puts the cart before the horse. It ignores the fact that it is the original work of an author or authors who contribute to the particular form of expression of the work and reduce the work to a material form that is the act giving rise to the statutory protection of copyright".[13] Significantly, the court noted that an entity who is a mere "scribe" cannot be the subject of copyright protection.[14] Applying this principle to the above scenario, it could be argued that the AI system contributing to the creation of the work would be analogous to a scribe carrying out the instructions of a human author. In such a case, copyright would subsist in the human author. Such a position would be consistent with established case law on authorship and originality, such as *Cummins v Bond*[15] *and Walter v Lane*[16]as applied in more recent cases such as *IceTV* and *Telstr*a.

If, however, the final work does not have sufficient human input to qualify for copyright protection as a literary work, the underlying algorithm could be sufficiently original to qualify for protection as a literary work. This is because the definition of "literary work encompass "computer programs". In such a case, the author would be the human creator of the algorithm.

Finally, if the algorithm does not qualify for protection as a computer program that is a literary work, an alternative basis of potential protection could be found under patents legislation if the algorithm satisfies the criteria for an invention.

Question (ii) - In the event copyright can be attributed to AI-generated works, in whom should the copyright vest? Should consideration be given to according a legal personality to an AI application where it creates original works autonomously, so that the copyright would vest in the personality and the personality could be governed and sold in a manner similar to a corporation?

[13] *Telstra Corporation Limited v Phone Directories Co Pty Ltd* (2010) 194 FCR 142; See discussion in Alpin, T, 'When are compilations original? Telstra Corporation Ltd v Desktop Marketing Systems Pty Ltd' (2001) 3(1) *Digital Technology Law Journal* 1.
[14] *Telstra Corporation Limited v Phone Directories Co Pty Ltd* (2010) 194 FCR 142 at [59].
[15] [1927] 1 Ch 167.
[16] [1990] AC 539.

In order to address the issue of ownership, it is first necessary to delineate the various potential forms and degrees of interaction between humans and machines in AI works, and address each such scenario separately. To do so, it is further necessary is identify the various stages of the creation of an AI work and the differing forms and levels of human input at each stage. The AIPPI working paper provides useful guidance in this respect:[17]

Stage 1 – Selection and classification of data

Three potential levels of human intellectual effort:

(a) Human intellectual effort is involved in selection and classification of the input data used to be used to train the AI system; or

(b) Human intellectual effort is provided to select data but the classification of such chosen data is done by the AI system; or

(c) No human intellectual effort is used in the selection and classification of data.

Stage 2 - Creation of AI work

Two potential levels of human intellectual effort:

(a) Human intellectual effort is involved in providing instruction to the AI system to direct its creation of the work; or

(b) The AI system creates a work without human intellectual effort.

Stage 3 – Checking and delivery of final AI work

Two potential levels of human intellectual effort:

(a) Human intellectual effort is involved in checking and delivering the final work to an end-user; or

[17] Based on delineations outlined by the AIPPI, *Copyright in artificially generated works, Study Guidelines*, 2019, pp.8-9.

(b) The AI system delivers the final work to end users without human checking or intervention.

Recommendations

It is useful to consider each of the above scenarios separately in order determine the appropriate level of copyright protection that should be afforded. It is suggested that where human intervention is limited to data selection and classification, the resulting work cannot be considered original under copyright law because it has not involved the independent intellectual effort of a human creator. Such a scenario would be analogous to the circumstances of *Telstra Corporation Limited v Phone Directories Co Pty Ltd*. In *Telstra* the central legal issue to be determined was whether the applicants' telephone directories attracted copyright protection. The creation of this work involved an automated system that used a set of pre-determined rules to conduct automatic checks of data and compile data into telephone directories. The court found that copyright did not subsist in the telephone directories, with Justice Gordon stating that "there are substantial parts of the directories that do not have human authors (for example, many of the service order listings), are automated to the extent that human involvement is minor" (at [335]). Significant to the present discussion, the court noted that the persons who monitored the operation of the system did not exercise "discretion" but merely "applied the Rules".

In contrast, where independent human intellectual effort is provided to guide the AI system in the process of creation in stage 2, it would suffice to establish originality. Moreover, where human intellectual effort is also directed checking and delivering the final work, it would clearly attract copyright protection.

As far as affording of legal personality to AI is concerned, we do not need an imminent need for such a solution at this stage. We suggest copyright protected should be afforded only to works where a human being has substantially contributed to the

creation of the work. In this case that human being(s) or their successor (e.g. employer, rights transferee etc) will be the owners of the copyright in the work.[18]

Question (iii) Should a separate sui generis system of protection (for example, one offering a reduced term of protection and other limitations, or one treating AI-generated works as performances) be envisaged for original literary and artistic works autonomously generated by AI?

It has been suggested above that in order to address technological change law makers should, where possible, seek to build and extend existing laws rather than create new law (see reasons listed above). Our general position is that AI-generated works should be awarded copyright only if there is a substantial human contribution. This complies with originality and authorship requirements under many jurisdictions. If such position is upheld, current copyright law system is able, with certain modifications[19], to adequately protect AI-generated works and no additional sui generis system is needed.

Issue 7: Infringement and Exceptions

13. Introduction – *does the use of copyright-protected material in the training of AI currently constitute an infringement of copyright law?*

Machine learning includes feeding AI systems with large amount of data. When copies of entire or significant parts of works – music pieces, press articles, artistic works etc – are made and used to "train" AI, the question of copyright infringement arises: do these uses fall in the scope of exclusive rights and whether any copyright exception applies. Currently, the answer to these questions is likely to vary from jurisdiction to jurisdiction.

[18] For a detailed analysis of ownership issue from the perspective of Australian law see White, Courtney and Matulionyte, Rita, 'Artificial Intelligence Painting The Bigger Picture For Copyright Ownership' (December 5, 2019). Available at SSRN: https://ssrn.com/abstract=3498673or http://dx.doi.org/10.2139/ssrn.3498673

[19] Eg human contribution might need to be better defined or redefined

In Australia, if an entire work or a significant part of it has been reproduced in whatever form or for whatever duration, there will be an infringement of reproduction right[20]. Thus, if during machine learning process copies of protected works are made for training purposes, this use will fall in the scope of exclusive reproduction right. Similarly, EU and US copyright laws have very broad reproduction rights that include temporary and technical reproductions[21].

As far as exceptions are concerned, answers will be different from jurisdiction to jurisdiction. In Australia, fair dealing exceptions for research and study, criticism and review, parody and satire, reporting news or for use in judicial proceedings[22] are unlikely to apply. Among specific exceptions, temporary reproduction exceptions[23] could be considered as most relevant. In particular, according to s 43B, "the copyright in a work is not infringed by the making of a temporary reproduction of the work if the reproduction is incidentally made as a necessary part of a technical process of using a copy of the work". One could argue that for the purposes of AI training, reproduction that is taking place is "temporary" (until AI training process is finished), and their use is a "necessary part of a technical process". However, it is unclear whether copies fed in AI system are deleted after the training is finished. Also, it would be difficult to argue that the reproduction is "incidental" and the process is not merely directed at "the use of work" (as in case of RAM copies) since it serves a broader purpose of training software. Generally, this exception is unlikely to apply since it was meant to cover very specific technical reproduction, such as RAM copies, and Australian courts tend to interpret exceptions by narrowing down rather than expanding their scope[24].

Under the EU law, no exceptions are likely to apply to the situation either. The technical reproduction exception[25] has a very narrow scope[26]. Use of copies of works

[20] s 31 Copyright Act 1968 (Australia)
[21] See art 2, Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (Information Society Directive), *Official Journal L 167 , 22/06/2001 P. 0010 – 0019;* 17 U.S. Code §106
[22] ss 40-43 Copyright Act 1968 (Australia)
[23] S 43A and s43B Copyright Act 1968 (Australia)
[24] See eg *National Rugby League Investments Pty Limited v Singtel Optus Pty Ltd* [2012] FCAFC 59
[25] Art 5(1) EU information Society Directive 2001/29/EC
[26] It excludes only reproductions which "are transient or incidental [and] an integral and essential part of a technological process and whose sole purpose is to enable:(a) a transmission in a network between third parties by an intermediary, or (b) a lawful use -- of a work or other subject-matter to be made, and which have no independent economic significance"

in AI training is unlikely to meet 'transient', 'incidental' and most other elements of this provision. Similarly, the newly introduced TDM exception[27] would not apply to most cases since the exception covers only the reproduction acts of research organizations and cultural heritage organizations. Also, it is doubtful whether the use of copyrighted content in AI training would fall under the definition of 'text and data mining'.[28]

In contrast, in the US which is one of the leaders in AI technology, the use of copyrighted content in machine learning is in many cases likely to be covered by fair use.[29] To adjudicate fair use, judges apply a four-factor standard that evaluates: (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used; (4) the effect of the use upon the potential market for or value of the copyrighted work. [30] The first and the most important criteria favours uses that are transformative. Based on this criteria Google Books and Google Images services that use unlicensed content for transformative purposes where allowed under fair use doctrine. The use of protected content in machine learning is likely to satisfy the transformativeness requirement since the output created by AI system is supposed to be very different from input data. Overall, a detailed analysis of fair use test shows that the use of protected content in machine learning is likely to be covered by fair use, in most cases at least.[31]

Question (ii) If the use of the data subsisting in copyright works without authorization for machine learning is considered to constitute an infringement of copyright, what would be the impact on the development of AI and on the free flow of data to improve innovation in AI?

We argue that the current legal situation with relation to AI and copyright infringement is problematic.

[27] Art 3 Eu Copyright in digital Single market directive
[28] See a narrow definition of TDM under Directive recital 8 (" Text and data mining makes the processing of large amounts of information with a view to gaining new knowledge and discovering new trends possible.")
[29] For a detailed analysis of how fair use applies to machine learning see Benjamin L. W. Sobel, "Artificial Intelligence's Fair Use Crisis", 41 Colum. J.L. & Arts 45 (2017)
[30] 17 U.S.C. § 107 (2017)
[31] For discussion which particular uses might fall outside the scope of fair use see Benjamin L. W. Sobel, "Artificial Intelligence's Fair Use Crisis", 41 Colum. J.L. & Arts 45 (2017)

In countries where the use of copyrighted material in machine learning *constitutes a copyright infringement*, such legal situation is likely to impede the development of AI technology. AI developers face legal risks of being sued for copyright infringement and this may disincentivize them from developing AI tools.

If the development of AI technologies is hampered, this might negatively affect different industries, country's economic and social development. Copyrighted content is needed when developing AI tools, such as face recognition and text recognition technologies, which are used in a range of industries, as well as in security, educational and health sectors. Legal barriers in developing AI tools for these industries and public sectors would put them in a competitive disadvantage with industries in countries where use of copyrighted content in machine training is legal, eg in the US.

If the use of content in AI training needs licensing, only powerful and large corporations may afford this; this would entrench them as the only innovators in the field.[32] Those who do not afford paying are likely to use readily available or freely licensed data. However, this data might codify pernicious biases present in that data.[33] At last but not least, unfavourable legal regime might drive AI innovators to jurisdictions with more favourable legal regime (e.g. US).

On the other hand, in countries where use of copyrighted material in machine learning *is legal* under copyright law, the problem arises that authors and right holders whose content is used in machine learning and in developing AI tools are not compensated and do not participate in revenue sharing produced by these new technologies. Instead, large companies that develop AI tools by using creative content to train AI reap all benefits and maintain IP rights over the software they develop. The gap between large tech companies and individual authors or small creative business is likely to increase.

[32] Benjamin L. W. Sobel, "Artificial Intelligence's Fair Use Crisis", 41 Colum. J.L. & Arts 45 (2017), p 18
[33] Benjamin L. W. Sobel, "Artificial Intelligence's Fair Use Crisis", 41 Colum. J.L. & Arts 45 (2017), p 18

This issue becomes especially relevant as new technologies such as Blockchain creates opportunities for right holders to track use of their works and manage micro-payments. If Blockchain-based licensing becomes readily available and micro-licensing issue is solved, it might be less reasonable to allow tech companies use creative content in developing new technologies but at the same time not share revenues with the creators of the content.

More generally, copyright laws that permit unauthorised use of content for AI development purposes would contribute to AI innovation. However, AI is likely to bring not only benefits but also risks to the society. AI threaten to exacerbate inequality, shift income from labour to capital, threaten livelihoods of millions of people, etc. While copyright law is not sufficient to address these challenges, more restrictive copyright laws may slow progress in AI industry and thus slow down the negative impacts and provide time to deal with them appropriately.

Question (i) Should the use of the data subsisting in copyright works without authorization for machine learning constitute an infringement of copyright? If not, should an explicit exception be made under copyright law or other relevant laws for the use of such data to train AI applications? (i), (iii) – Is a new exception needed? What exception?

Question (iii) If the use of the data subsisting in copyright works without authorization for machine learning is considered to constitute an infringement of copyright, should an exception be made for at least certain acts for limited purposes, such as the use in non-commercial user-generated works or the use for research?

For the reasons outline above, it is difficult to give a straightforward reply whether countries in which the use of copyrighted material in machine learning constitutes an infringement need a copyright exception to cover such uses.

Introducing an exception that permits the unauthorised use of creative content for machine learning purposes would mean providing legal certainty and security for technology companies and incentivizing development of AI technology that has high

potential in a range of industries and public sector. It also means putting AI developers on the same competitive edge as their US counterparts.

At the same time, introducing a copyright exception would mean that authors and other right holders are taken away an opportunity to receive remuneration for the use of their content in this fast-growing industry with high future potential.

The best way forward would be to meet the demand of all sides and develop a compromised-based approach. AI companies require an easy access to content that they need for AI training purposes without an obligation to clear rights with each and every potential right holder. Individual licensing becomes virtually impossible due to large amounts of content used in machine learning. At the same time, right holders want to be paid for the use of their content. With decreasing revenue of individual authors[34] this becomes especially important.

As a compromise, countries could consider introducing a remunerated exception where AI developers would be able to use content without the authorisation from the right holders but would have to pay right holders (especially, authors) a fee for it. In order to solve problems caused by individual licensing, mandatory (and possibly extended) collective licensing could be introduced, whereby AI developer would need to approach a single collective society (or several ones specializing in different types of works) to acquire a license for all content that AI developer needs.

Such exception would cover both commercial and non-commercial uses of any content for machine learning. Limiting the exception to the use of non-commercial user-generated works or for research purposes would mean introducing a too narrow exception that is not able to achieve its goal, eg ensure that innovation in the field of AI is not unreasonably impeded.

[34] Recent studies around the world show that revenues of authors gradually decrease, see eg Macquarie University, Australian Authors. Industry Brief No. 3: Authors' Income (October 2015), http://www.businessandeconomics.mq.edu.au/our_departments/Economics/econ_research/reach_network/book_project/authors/3_Authors_Income.pdf; 'Most UK authors' annual incomes still well below minimum wage, survey shows', https://www.theguardian.com/books/2016/oct/19/uk-authors-annual-incomes-below-minimum-wage-survey-average-earnings; 'Income for US authors falls below federal poverty line – survey', https://www.theguardian.com/books/2015/sep/15/income-for-us-authors-falls-below-federal-poverty-line-survey.

The proposed exception is likely to comply with the TRIPS three-step-test.[35] Namely, the exception would be (i) limited to certain special cases, i.e. for the use in machine learning or similar technologies. (ii) Such use would not conflict with a normal exploitation of the work since the use in machine learning is non-consumptive and do not have an effect on an ordinary exploitation of work. Finally, (iii) use of content in AI training would not unreasonably prejudice the legitimate interests of the rights holder since right holders would receive remuneration for such uses.

The proposed exception presents a compromise approach where the interests of AI developers and creators are taken into account. At the same time, if it was introduced in a specific country, AI developers in that country would still be disadvantaged if compared to their counterparts in countries with fair use or broader exceptions without remuneration requirement. This might especially disadvantage small AI developers eg who carry out experimental art projects. On the other hand, if well designed, remunerated exception could benefit creators and help address the low revenue issue in most creative industries and the increasing gap between small right holders and large tech companies.

It is worth noting that the proposed solution would have a few problems that would need to be addressed. First, while in some creative industries collective societies have quite extensive catalogues (eg music), in other sectors collective management is underdeveloped (eg visual arts). This could be remedied by introducing additional solutions, such as an extended collective licensing or an indemnification of AI developers after they sought for all available collective licenses. Second, collective rights societies normally can license rights on national basis only, not internationally. AI developers might need a license for international use if AI-training and input of data is taking place internationally. Possible solutions to this international collective licensing problem could be discussed in international IP fora, such as WIPO.

Question (iv)If the use of the data subsisting of copyright works without authorization for machine learning is considered to constitute an

[35] art 13 WTO TRIPS

infringement of copyright, how would existing exceptions for text and data mining interact with such infringement?

As mentioned at (i),(iii) above, TDM exceptions available under EU law are unlikely to apply to machine learning scenario.

Question (v) – Would any policy intervention be necessary to facilitate licensing if the unauthorized use of data subsisting in copyright works for machine learning were to be considered an infringement of copyright?

We argue at (i),(iii) above that licensing, and especially compulsory collective licensing and extended collective licensing schemes, could provide a compromised-based solution to the problem where both the interests of AI developers and right holders are addressed. Introducing such schemes might be of challenge in countries where compulsory licensing is being increasingly challenged (eg Australia) but could serve as an inspiration when searching for a balanced solution to the problem.

Question (vi) How would the unauthorized use of data subsisting in copyright works for machine learning be detected and enforced, in particular when a large number of copyright works are created by AI?

A number of AI ethics guidelines impose transparency requirement on AI development[36]. One way to facilitate the detection of unauthorised use of protection material is to extend this transparency requirement to copyright field. This could mean a requirement that AI developers should indicate what sources were used to collect data for training purposes and what licensing arrangements apply with relation to the use of that data.

In addition, attribution right generally requires indicate the author or right holder of the work used. As AI developers use huge amounts of works in AI training, it would be unreasonable to ask them to comply with such a specific attribution requirement.

[36] See eg OECD Principles on AI https://www.oecd.org/going-digital/ai/principles/; Australian Ethics Framework on Artificial Intelligence https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles

However, this right could be used as an inspiration to require more transparency as to what material has been used in the training of AI.