Copyright and generative artificial intelligence

By David Wilson, Principal, Stratocumulus Legal

David Wilson is the founder and principal of Stratocumulus Legal, a legal consultancy with a particular focus on intellectual property, technology and related fields.

Introduction

Generative artificial intelligence has captured the public imagination recently, particularly since the release of OpenAI's ChatGPT in November 2022. Its capability to seemingly interact in natural language has been lauded, but it has also induced fears of disruptively displacing humans from a wide range of activities. AI systems have also reached the capability of being able to produce quite advanced images generated from descriptive text-based inputs from users.

At first glance, copyright law seems ill equipped to deal with AI technology. The Copyright Act 1968 (Cth) was conceived in an analogue world, and only over more recent times it has been retrofitted by bolting on new provisions to address digital technology. Existing copyright law, usually lagging behind new technologies, has not developed with AI in mind.

Although it may seem unintuitive, current Australian copyright law results in perhaps nobody being the owner of content generated by an AI system due to lack of an identifiable human author, and moreover, using copyright content to train a machine learning system without the consent of copyright owners risks infringement. Without reform, this could greatly stifle the development of AI technology, although this status quo does retain a measure of protection for human creators concerned about the pendulum swinging too far in favour of robots.

After all, copyright law has always been about balancing the interests of creators and users, while still being vexed by new technologies. Where the balance should rest in the case of generative AI is something policymakers need to determine. Should anyone own a robot's creation? If so, who?

Historical context – the cricket team portrait

Photography was also, at first, a disruptive technology, and its initial protection by copyright in the Fine Arts Copyright Act 1862 (UK) caused controversy given its mechanical and scientific nature of production compared to traditional art forms. Photography even started to replace the artist's manual labour. These themes no doubt sound familiar in the context of contemporary emerging generative AI.

In June 1882, an Australian cricket team visiting England posed for a photograph at Kennington Oval. In Nottage v Jackson, the two proprietors of the photographic studio claimed to be authors of the resulting photograph, and registered their copyright under that Act accordingly. They had not been personally present, rather leaving the actual taking of the photograph to their employees. In rejecting their infringement claim against pirates of the photograph, Brett MR lamented the statutory use of the term 'author' in the context of photography:

But now we have the 'author' of a photograph. I should like to know whether the person who drew this Act of Parliament was clear in his mind as to who can be the author of a photograph.

To assess authorship, Brett MR noted that using the cameras of the day, the person taking the photograph:

... had to arrange the group, to put them in the right position and the right focus. But he does not do it all; because I suppose there is another man who gets the plate ready; and there is another man who, when the thing is ready, takes the cap off. It is difficult to say who is the author of the photograph. Neither of them [the proprietors] made the picture because, after all, that is done by the sun.

Instead, the author was the person who "superintended the arrangement", who was the "effective cause of the picture", and not the employee who came up with the idea for the photograph. Indeed, the "great photographers of London" ought to "superintend the work themselves", or given that the duration of copyright was for the life of the author, when choosing an artist, "consider not only his skill but his state of health"!

Misstating the author in the photograph's registration rendered the copyright unenforceable, an early example demonstrating how authorship has always been the touchstone for copyright ownership.

Back to the 21st century – copyright in generative AI output

The Copyright Act provides that copyright subsists in an original literary, dramatic, musical or artistic work of which the author was a "qualified person", being an Australian citizen or resident. At face value, this requires that an author must be human.

More substantively, originality requires that the work in question have originated with the author, and that the creation of the work required some independent intellectual effort. Although literary or artistic merit are not required, this requires a qualitative assessment of what the author did to bring the work into being, more than just being industrious. Authorship and originality are thereby correlative.

Just as Brett MR, with disdain, grappled with the word 'author' being used in respect of a photograph, who might be the 'author' of material generated by a computer? Computerisation of our daily lives has brought increasing automation of tasks that were once undertaken manually. It seems uncontroversial that a person using a computer as a tool (for example, using word processing software to write this article) still constitutes authorship of the work, but other situations require more analysis. In Telstra v Phone Directories, Perram J explained that:

Software comes in a variety of forms and the tasks performed by it range from the trivial to the substantial. So long as the person controlling the program can be seen as directing or fashioning the material form of the work there is no particular danger in viewing that person as the work's author. But there will be cases where the person operating a program is not controlling the nature of the material form produced by it and in those cases that person will not contribute sufficient independent intellectual effort or sufficient effort of a literary nature to the creation of that form to constitute that person as its author.

Telstra v Phone Directories examined whether copyright subsisted in White Pages and Yellow Pages telephone directories as databases or compilations (types of literary work). The directories were listings of telephone subscribers, and the evidence did not show any particular authors who had compiled them, but rather that such work had been undertaken by a computer program. Human contribution might have existed in individual data entries, but this did not amount to the making of the overall directories as compilations. Therefore, copyright did not subsist.

Acohs v Ucorp determined that a process contained in source code to automatically generate documents (being material safety data sheets for hazardous substances), compiled from a central database when users inputted relevant information, did not make the authors of the source code the authors of the generated material safety data sheets. Their work was antecedent, and there was no human author of the automatically generated output.

This is perhaps where the analogy to the above photography example diverges from AI creations. It is not the programmer who has coded the AI system who owns copyright in its individual outputs, if that output is the automated result of code being executed. Nor does the user necessarily contribute authorially to the output. It seems likely in many instances that nobody will be considered an author.

By contrast, a different approach is found in the United Kingdom, where 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". The only time this provision has been considered was in litigation concerning a computer game. The software developer was deemed to be the author of an on-screen work generated by the program that he coded, on the basis that his programming constituted the arrangements necessary for the creation of the work. By contrast, the input of a person playing the game was "not artistic in nature" and "contributed no skill or labour of an artistic kind", therefore did not meet the requirement.

For AI-generated works, this analysis may differ considerably, depending on how the specific system operates. For example, AI algorithms operate differently to the coded instructions of traditional software, and software developers therefore have less control over how a work is generated by an AI system. Further, in addition to the software developer, there may also be others who may arguably undertake arrangements, such as data providers. Identifying the deemed author of AI generated works may not be straightforward.

Australia could also have gone down this path, with the Copyright Law Reform Committee recommending in 1995 the protection of "computer generated material" for a period of 25 years, again the author deemed to be the person by whom arrangements necessary for the creation of the material are undertaken. This was never enacted.

This leads to the inevitable policy question

whether AI-generated material should be protected. Arguments flow in opposing directions. In very broad terms, in favour are arguments for incentivising the development of AI and investment in emerging technologies , while against are arguments that only actual human intellectual creativity is worthy of protection. Despite rapid advancement of AI technology, this conversation still has some way to go.

Stifling AI – machine learning as copyright infringement

Another looming copyright issue has the potential to either stifle the development of generative AI systems, or make them fairer for human creators, depending on your point of view.

AI systems are trained via a machine learning process using a vast volume of content. Inevitably, much of that input material will itself be protected by copyright. It is reported that the image generating AI system called Stable Diffusion, operated by Stability AI Ltd, was trained on 2.3 billion images scraped, or data mined, from across the Internet.

The principal copyright subsisting in respect of a literary, artistic, dramatic or musical work is the right to reproduce the work as a material form.

Since the definition of 'material form' was amended in 2004 , even temporarily loading a copyright work into the read-access memory (RAM) of a computer without consent of the

copyright owner constitutes infringement. A recent example of this in operation was QAD v Shepparton Partners Collective Operations , in which a computer program previously installed under licence, but later used beyond licence terms, during which it was loaded into the RAM of the computer to operate, constituted infringement.

This means that using content to train a machine learning system – even if not storing that content in permanent storage system – will likely infringe copyright if there is no consent from the owner. Of course, this applies to copyright works, and not data or information alone.

At time of writing, proceedings have recently been commenced in both the United Kingdom and the United States against Stability AI Ltd, alleging exactly this. A proceeding in the United Kingdom High Court brought by stock photo provider Getty Images is reported to allege that Stability AI infringed copyright by processing millions of its images. A like proceeding has also been commenced in the US District Court.

Separately, a class action brought in the US District Court by three artists who claim their art was used without their consent, makes similar allegations, a motivation being to make "AI fair & ethical for everyone" . A further US class action alleges that AI-powered coding assistant "GitHub Copilot" was trained on open-source code scraped from the Internet published with copyright licences that require

anyone reusing the code to credit its creators.

Such litigation has the potential to heavily impact the development of generative AI systems, if the cost of training AI systems will require royalties or licence fees to copyright owners.

The United Kingdom currently permits data mining, making it an exception to copyright infringement to make a copy of a lawfully accessed work, but only for carrying out a computational analysis solely for non-commercial research purposes.

There is no equivalent exception for copyright infringement in the Australian Copyright Act. Calls have been made to introduce an exception to infringement that would allow text and data mining , so that the development of an emerging technology with "huge potential" should not be "unnecessarily impeded".

The Australian Law Reform Commission's report on Copyright and the Digital Economy suggested that text and data mining be considered under a general fair use exception to copyright infringement recommended elsewhere in the report. However, this was on the basis that text and data mining is a "non-expressive use" of copyright material, which does not appear to anticipate contemporary generative AI.

In response to proposed changes to permit data mining generally, the House of Lords Communications & Digital Committee recently labelled those proposals as "misguided", noting they "take insufficient account of the potential harm to the creative industries", and that "Developing AI is important, but it should not be pursued at all costs". The Committee recommended that proposed changes to the text and data mining regime be paused immediately.

Whether AI-generated material could itself infringe copyright in works on which it is trained will depend on whether each specific item of output constitutes a substantial reproduction of an identifiable original work. This is a qualitative assessment in each case, and perhaps a daunting task given the huge volume of content on which AI is trained. A more fruitful route for copyright owners may be to focus on training material, rather than AI systems' output.

Conclusion – an AI poem

Just as Nottage v Jackson was required to resolve copyright issues in respect of the new technology of the 19th century, particularly in respect of authorship, again courts and policymakers are challenged with similar issues for the emerging technology of the 21st.

Generative AI system ChatGPT was instructed by this author to "Please write a rhyming poem about intellectual property and robots", and the prescient response, despite minor hiccups in its expression, is a fitting summary of the ground covered in this article:

In a world of robots so advanced,
Where minds and machines have fused and danced,
The question of IP comes to mind,
Who owns the creations of robot kind?

Do we treat them as mere machines,
Or recognize their novel schemes?
If they create, can they own,
Or does that right to us alone?

As they learn and evolve each day,
Can we claim what they make in every way?
Or should we give them their fair share,
And treat them as creators, fair and square?

Perhaps the answer lies ahead,
As robots and humans forge ahead,
Collaborating in new ways,
To build a future bright and blaze.

But until then, we must debate,
And determine the laws and fate,
Of intellectual property and robots,
As they become our creative co-pilots.

■

End Notes

1 Bowrey, Kathy, "'The World Daguerreotyped – What a Spectacle!' Copyright Law, Photography and the Commodification Project of Empire" (1 May 2012). UNSW Law Research Paper No. 2012-18, accessible at SSRN: https://ssrn.com/abstract=2060803

2 Ibid at page 3

3 Copyright is no longer a registered right. The Fine Arts Copyright Act 1862 (UK) afforded copyright to "the author … of every original painting, drawing and photograph", which subsisted for the life of the author.

4 Nottage v Jackson [1883] 11 QBD 627 at 631

5 Ibid at 632

6 Ibid at 632

7 Ibid at 633

8 Likely because of these issues, 'author' is now only defined in respect of photographs (and not for other types of work) as "the person who took the photograph". See Copyright Act 1968 (Cth), s10

9 Copyright Act 1968 (Cth), s32

10 IceTV Pty Ltd v Nine Network Australia Pty Ltd (2009) CLR 458 at [23] per French CJ, Crennan & Kiefel JJ

11 Telstra Corporation Ltd v Phone Directories Company Pty Ltd (2011) 90 IPR 1

12 Ibid at [118] per Perram J

13 Acohs Pty Ltd v Ucorp Pty Ltd [2012] FCAFC 16

14 Ibid at [76] – [87]

15 Copyright, Designs and Patents Act 1988 (UK), s 178 defines a work that is "computer-generated" to mean that "the work is generated by computer in circumstances such that there is no human author of the work"

16 Copyright, Designs and Patents Act 1988 (UK), s 9(3). Other jurisdictions with like provisions include New Zealand, Hong Kong, India, Ireland.

17 Nova Productions Limited v Mazooma Games Limited & Ors [2006] EWHC 24 (Ch)

18 Ibid at [105]

19 Ibid at [106]

20 See the excellent discussion of these issues in Matulionyte, Rita and Lee, Jyh-An, "Copyright in AI-generated works: Lessons from recent developments in patent law" (2022) 19:1 SCRIPTed 5 accessible at https://script-ed.org/?p=4036, section 2.1.1

21 Copyright Law Review Committee, Computer Software Protection, Office of Legal Information and Publishing, Attorney General's Department, Canberra. See recommendations 2.42 (a) to (c) in Chapter 13, pp 247 - 248

22 For just one view, see Abbott, Ryan "The Reasonable Robot: Artificial Intelligence and the Law", Cambridge University Press, June 2020, Chapter 4 ("Artificial Inventors"), arguing for a legal framework that incentivises the generation of intellectual property by AI systems. See also Matulionyte & Lee, refer note 19 above, arguing that the owner of the AI-system ought to be owner of any AI-generated content.

23 Again, for just one view, see Zurth, Patrick, "Artificial Creativity? A Case Against Copyright Protection for AI Generated Works" (25 March 2021) UCLA Journal of Law & Technology, published at https://uclajolt.com, arguing that artificial intelligence does not equate to "artificial creativity" and copyright protection is the wrong approach to reward AI technology.

24 Baio, Andy, "Exploring 12 Million of the 2.3 Billion Images Used to Train Stable Diffusion's Image Generator", Waxy, 30 August 2022, accessible at https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator

25 Copyright Act 1968 (Cth), s31(1)(a)

26 US Free Trade Agreement Implementation Act 2004 (Cth), s186 amended the definition of "material form" in the Copyright Act 1968 (Cth), s10

27 QAD Inc v Shepparton Partners Collective Operations Pty Ltd [2021] FCA 615

28 Getty Images Statement, 17 January 2023, accessible at https://newsroom.gettyimages.com/en/getty-images/getty-images-statement

29 Vincent, James, "Getty Images sues AI art generator Stable Diffusion in the US for copyright infringement", The Verge, 7 February 2023, accessible at https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion

30 Vincent, James, "AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit", The Verge, 16 January 2023, accessible at https://www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart

31 Vincent, James, "The lawsuit that could rewrite the rules of AI copyright", The Verge, 22 May 2022, accessible at https://www.theverge.com/2022/11/8/23446821/microsoft-openai-github-copilot-class-action-lawsuit-ai-copyright-violation-training-data

32 Copyright, Designs and Patents Act 1988 (UK), s29A

33 One example is Matulionyte, Rita, "Australian Copyright Law Impedes the Development of Artificial Intelligence: What Are the Options?", accessible at SSRN: https://ssrn.com/abstract=3720289

34 Ibid at page 22

35 ALRC Report 122, "Copyright and the Digital Economy", November 2013

36 Ibid at para 11.74

37 Ibid, Recommendation 4-1

38 Ibid at para 11.60. See generally discussion in Chapter 11, paras 11.57 to 11.84

39 House of Lords Communications and Digital Committee "At risk: our creative future", 2nd Report of Session 2022-23, published 17 January 2023, HL Paper 125, paras 26 to 35

40 Ibid at para 35

41 At time of writing, accessible at https://chat.openai.com