The COMPUTER & INTERNET Lawyer

Volume 40 ▲ Number 8 ▲ September 2023

Ronald L. Johnston, Arnold & Porter, Editor-in-Chief

Copyright Implications of Generative AI Systems

By Harrison Ottaway and Jonathan Feder

Generative AI systems like ChatGPT and DALL-E have been attracting media attention for their potential to cause disruption across a range of industries. In a recent report, Goldman Sachs estimated¹ that generative AI systems could impact 300 million fulltime jobs globally. In the same report, Goldman Sachs found that the same AI systems could also boost global productivity and lead to a 7% increase in annual global gross domestic product (GDP).

Generative AI systems present a number of challenges from a copyright law perspective. Two questions are particularly pressing:

- 1. Can copyright subsist in AI-generated content?
- 2. Does the use of generative AI models infringe the copyright in pre-existing works?

This article explores these questions, with a focus on text-to-image generative AI systems, and in particular, an AI model called Stable Diffusion. However, much of what is discussed in this article will be equally applicable to text-to-text generative AI models like ChatGPT.

The authors may be contacted at Harrison.Ottaway@klgates.com and jonathan.feder@klgates.com, respectively.

WHAT IS GENERATIVE AI?

Generative artificial intelligence systems are machine learning tools which can be used to create content, including text, images, videos and software code. Generally, these AI systems learn patterns from existing data, then use this knowledge to generate new outputs based on prompts from a user.

While generative AI systems have been in existence for some time, recent breakthroughs in the field have significantly advanced their capabilities, catapulting them into the global spotlight.

STABLE DIFFUSION

Stable Diffusion is a text-to-image generative AI system developed by the start-up Stability AI, in collaboration with a number of academic researchers and non-profit organizations. The model was trained on approximately 5 billion text-to-image pairings derived from a "general crawl" of the internet.

The inner workings of Stable Diffusion are quite complex. The model is trained through a process known as "diffusion." Essentially, it adds a random quantum of visual noise to an image, then teaches itself to successfully "de-noise" the image by predicting the original image and comparing its prediction to the actual image. This process typically requires a very large amount of computing power. However, Stable Diffusion solves



Figure 1. Author: Asanagi

this problem by creating highly-compressed, encoded (or latent) versions of these images and running the diffusion process in latent space. Through this process, Stable Diffusion teaches itself to create new images from random noise. The model also includes a text encoder, which enables it to transform text inputs into values that can be understood by the diffusion model and "steer" the diffusion process based on a text prompt. The end result is that Stable Diffusion is able to produce new image outputs responsive to user text prompts.²

For example, Figure 1 is an image generated by Stable Diffusion using the prompt "a photograph of an astronaut riding a horse."

CAN COPYRIGHT SUBSIST IN AI-GENERATED CONTENT?

A key question is whether copyright can subsist in content created by generative AI systems like Stable Diffusion. This question has significant ramifications for the ability of businesses and individuals to commercially exploit AI-generated content. Without legal rights of ownership in such content, the ability to license or sell it is effectively nil.

Under Australian law, copyright will only subsist in works that have been created by a human author. The

key authority for this proposition is *Telstra Corporation Limited v. Phone Directories Company Pty Ltd*,³ in which the full federal court found that Telstra's computer-generated telephone directory did not qualify for copyright protection in light of Telstra's inability to identify human authors responsible for the "material form" of those directories.⁴ This case was decided in the wake of the High Court's decision in *Ice TV Pty Ltd v Nine Network Australia Pty Ltd*,⁵ where the court found that for copyright to subsist, a work must originate with an author, from some "independent intellectual effort."⁶

Does entering a text-based prompt into an AI model like Stable Diffusion satisfy the requirements for human authorship and "intellectual effort," such that copyright will subsist in the output of the AI model? This question is yet to be tested in Australian courts (or, as far as the author is aware, courts in any other jurisdiction).

The answer is likely to depend on the particular case. In circumstances where little time and effort is spent on the input provided to the AI model – for example, just typing the text prompt "cat wearing a tie," reflected in Figure 2 - it will be difficult to argue that the generation of the resulting artwork involved some "creative spark" on the part of the human author.



Figure 2. Image created using Stable Diffusion and the prompt "cat wearing a tie."

On the other hand, it is possible to conceive of works which might require considerable creative effort on the part of the human operator: for example, works generated by an elaborate string of text prompts, or after an iterative process of trial and error. Such works may be in a stronger position to qualify for copyright protection, based on current Australian authority.

Indeed, user accounts of Stable Diffusion suggest that, at present,⁷ obtaining high quality output takes a lot of work. Typically, the user must refine the prompt until it is ultra-specific, generate a very large set of images and select the best option, refine this further using an image to image generator (another form of generative AI), and finally, add finishing touches using a program like Photoshop.⁸ Clearly, this is a labor-intensive process. However, whether the resulting artwork ultimately qualifies for copyright protection is likely to depend on the extent to which the "skill and labor" of the user was directed to the actual material form of the resulting artwork.⁹

The upshot is that the creators of AI-generated artworks presently have little certainty regarding the legal rights in their creations.

INFRINGEMENT OF THIRD PARTY COPYRIGHT

Another important question is whether the use of generative AI systems amounts to an infringement of copyright in the pre-existing content used to train the AI-model.

There are at least two ways in which the use of Stable Diffusion could potentially infringe third party copyright:

- First, the use of Stable Diffusion to create new images could infringe the copyright in works comprised in the AI model's training set; and
- Secondly, the training process itself could infringe the copyright in works comprised in the training set.

This article considers both in further detail below.

There are two principle questions that a court will consider when determining infringement of copyright.

First, whether the allegedly infringing work was derived (or "copied") from the copyright work.

Generative AI

Figure 3. A "garden in the style of Monet" created using Stable Diffusion.



Second, whether the allegedly infringing work takes a substantial part of the copyright work. Both derivation and substantiality must be established for infringement to arise.

Creation of New Images

Let us start with the use of Stable Diffusion to create new images.

It should be noted that the creation of new images using the Stable Diffusion model does not involve any direct or "literal" copying from the images in the model's training dataset. Rather, new images are created from random noise, based on patterns that the model has learned from these images during its training process.

This in itself is unlikely to be fatal to derivation. Copyright law recognizes that copying can occur in various forms and may be direct or indirect.¹⁰ The key question is really whether there is some causal connection between the allegedly infringing work and the copyright work.¹¹ This requirement is likely to be satisfied where it can be shown that the copyright work forms part of an AI model's training set.

The question of substantiality is likely to prove more challenging. It is important to remember that to establish copyright infringement in Australia, it must be shown that the allegedly infringing work reproduces a substantial part of a particular copyright work. It is not an infringement of copyright to combine a "non-substantial" part of a multitude of different works. Indeed, this is arguably how the human creative process works, Figure 4. The Artist's Garden at Giverny, Claude Monet, 1900.



because all new works of art are built to some extent on what has come before them.

The difficulty here is that AI-generated works may be derived from a very large number of sources. In most cases, it is likely to be quite difficult to demonstrate that an AI-generated artwork takes a "substantial part" of any one copyright work in its training set.

The case for infringement may be stronger for AIs trained on more specific data sets, or AIs that are instructed to produce images "in the style" of a particular artist or artwork. In such cases, it may be easier to identify the reproduction of a substantial part of a particular work. However, there is still considerable scope for legal uncertainty.

For example, Figure 3 is an artwork generated by Stable Diffusion using the prompt "a garden in the style of Monet." Figure 4 is an actual painting by Monet of his garden at Giverny.

Figure 3 is a passable imitation of Monet's style, and one can see how it might be the cause of some concern to the artist (or in this case, his estate).

Assuming that Figure 4 forms part of Stable Diffusion's training set, then derivation ought to be established. However, whether Figure 3 actually reproduces a "sub-stantial part" of Figure 4 is a more difficult question.¹² The works depict the same subject matter and share a distinct impressionistic style. But there are also notable differences in their composition and color palette. In fact, it is difficult to precisely identify any particular aspect of Figure 4 which is reproduced in Figure 3. In practice, where there is clear evidence of derivation, Courts may be more inclined to make a finding

of substantiality. The challenge here, though, is that Figure 3 may also be derived from a very large number of additional sources, potentially diluting the substantiality of any single instance of copying. This issue is compounded by the lack of transparency around Stable Diffusion's image creation process and the sources relied upon to create new output in any particular case.

This exercise illustrates the challenges presently faced by copyright owners whose works are used, and in some cases imitated, by generative AI-models.

Training

A separate question arises as to whether the process of training generative AI-systems itself amounts to an infringement of copyright.

In Stable Diffusion's case, this question is likely to turn on whether the training process, and in particular, the creation of latent versions of images in the training set, amounts to a reproduction in material form of those images. It is not contentious that the reproduction of a work in a non-visible, digital form may amount to an infringing reproduction under Australian copyright law.¹³ A key question will be whether the highly-compressed, latent images created by the Stable Diffusion system still reproduce a "substantial part" of the original image – or whether the training process involves the creation of other "non-latent" digital reproductions of the training images.

These question are at the core of U.S. court proceedings recently commenced¹⁴ by stock image database Getty Images against Stability AI in the District Court for Delaware. Getty alleges that Stability AI has copied at least 12 million copyright images from its websites, along with associated text and metadata in order to train its Stable Diffusion model. While the proceeding is still at an early stage, it is clear that Getty considers that Stability AI's training process amounts to an infringement of Getty's copyright - Getty alleges that the training process involved wide scale copying, encoding and storage of Getty's stock images and text pairings. Getty also alleges that Stable Diffusion has violated U.S. trademark and copyright law by producing images which either reproduce, or distort, or remove the watermarks on Getty's images.¹⁵ The determination of Getty's claim is likely to entail a detailed, technical consideration of the operation of the Stable Diffusion system, and how this interacts with the conventional copyright principles of derivation and substantiality.

While there are relevant differences between Australian and U.S. copyright law (most notably, the

absence in Australia of a general "fair use" defense), the proceeding is still likely to provide helpful guidance to Australian artists, businesses and intellectual property lawyers, in the absence of local case law or legislative intervention.

Notes

- 1. https://www.businessinsider.com/generative-ai-chatpgt-300-million-full-time-jobs-goldman-sachs-2023-3.
- 2. For more information about how Stable Diffusion works, see "The Illustrated Stable Diffusion," Alammar, J, 2022 – https:// jalammar.github.io/illustrated-stable-diffusion/; this article by Hugging Face: https://huggingface.co/blog/stable_diffusion; or this article by Stable Diffusion Art: https://stable-diffusion-art.com/how-stable-diffusion-work/#:~:text=Stable%20 Diffusion%20is%20a%20latent,why%20it's%20a%20lot%20 faster.
- 3. Telstra Corporation Limited v. Phone Directories Company Pty Ltd [2010] FCAFC 149.
- 4. See [90] per Keane CJ; [117] to [119] per Perram J; [169] per Yates J.
- Ice TV Pty Ltd v Nine Network Australia Pty Ltd (2009) 239 CLR 458.
- 6. See [48], [97], [98].
- 7. It is expected that obtaining quality output will require less work as the model continues to be developed.
- 8. See https://news.ycombinator.com/item?id=33902248.
- 9. The High Court in Ice TV queried the relevance of "skill and labor" per se to questions of copyright subsistence and infringement the critical question is whether skill and labor is directed to the particular form of expression see [47], [52] [54].
- 10. For example, in Fred Fisher Inc v Dillingham (1924) 298 Fed 145, the court found that copying occurred unconsciously, on the basis that the copyright work was stored somewhere in the memory of the infringer see 147-8 per Learned Hand J.
- 11. See Francis Day & Hunter Ltd v Bron [1963] Ch 587 at 614.
- 12. The copyright in Monet's works is, in fact, expired but let us assume for the purpose of this exercise that it is still current.
- 13. See Copyright Act 1968 (Cth) Section 10.
- 14. https://copyrightlately.com/pdfviewer/getty-images-v-stability-ai-complaint/?auto_viewer=true#page=&zoom=auto&pagemode=none.
- 15. Another proceeding was commenced against Stability AI (and another generative AI called MidJourney) by a collective of artists in January this year in the U.S. district court for the Northern District of California, involving similar allegations. https://stablediffusionlitigation.com/pdf/00201/1-1-stable-diffusion-complaint.pdf?_sm_pdc=1&_sm_rid=NTT-WQqD4DS8tF01Znkqt1VJkSH5SCkQq4RjPWp8.

Copyright © 2023 CCH Incorporated. All Rights Reserved. Reprinted from *The Computer & Internet Lawyer*, September 2023, Volume 40, Number 8, pages 9–13, with permission from Wolters Kluwer, New York, NY, 1-800-638-8437, www.WoltersKluwerLR.com

