2024 Will Be A Busy Year For Generative AI And IP Issues

By **Nina O'Sullivan, Anne Rose and Peter Nunn** (January 1, 2024)

Nina O'Sullivan

Anne Rose

Peter Nunn

With November 2023 being the first anniversary of the release of ChatGPT, 2024 looks to be another whirlwind year for generative artificial intelligence, or GenAI.

In particular, legislators, intellectual property offices and courts are all closely analyzing the complex issues involved in relation to the impact on protection and enforcement of IP. The outcomes of these assessments will have significant implications for the developers of AI tools, users of such tools and rights holders.

Clearly, there is a need for clarity on certain fundamental principles, but it remains to be seen whether this will come via rulings in the swathe of disputes currently being litigated before the courts — mainly in the U.S., but also in jurisdictions such as the U.K. and Germany — or via legislative interventions.

The IP Issues

GenAI tools generate or create outputs, e.g., text, artwork, film, voices and code, based on user instructions or prompts. In order to be able to generate these outputs, the tools have been trained on vast amounts of data from a broad range of sources.

While not all the sources of data have been disclosed, it is apparent that, in many cases, they include material that is protected by IP rights — most relevantly, copyright and database rights. The IP issues arising from GenAI fall into three main categories:

- The IP protection issue: Are AI-generated works protected by IP — in particular, by copyright — and, if so, who would be the author of the copyright in such works?
- The inputs issue: Is there infringement of IP rights when materials and data are used to train an AI tool?
- The outputs issue: Is there infringement of IP rights in relation to the content produced using a GenAI tool?

The IP Protection Issue

The approach as to whether copyright can protect AI-generated works differs from country to country. In the U.K., while the Copyright, Designs and Patents Act 1988 specifically recognizes copyright protection for computer-generated works, the relevant provision is untested in the context of content created using GenAI.

It seems very likely that the U.K. courts will be required at some point to determine whether an AI-generated work meets the criteria for copyright protection. This will require,

in particular, a complex assessment as to whether the work meets the originality requirement — an issue that in any event raises various analytical difficulties — and to identify the author of the relevant work.

Meanwhile, the U.S. Copyright Office has taken a hard line and has denied copyright registration to various AI-generated works in accordance with its statement of practice that was published March[1] to the effect that only human created parts of a GenAI work will be protected by copyright.

While the U.S. Copyright Office does continue to explore the broader policy debates around copyrightability, there may also be further developments on this issue in the U.S. courts. For example, Stephen Thaler has filed a notice of appeal[2] to the U.S. Court of Appeals for the District of Columbia Circuit against a decision[3] to reject his application for judicial review of the U.S. Copyright Office's refusal to register the AI-generated image "A Recent Entrant to Paradise."

The Inputs and Outputs Issues

The issues concerning potential infringement arising from training GenAI tools and in relation to the outputs they produce are, to some extent, linked. However, to date, all the cases before the courts are against developers of the AI tools, rather than against users, though it is possible — perhaps depending on how those cases progress — that a separate stream of litigation against users may develop.

Indeed, several AI companies now offer copyright infringement indemnities for their users, but the terms of these need to be assessed carefully in relation to:

- Who will have the benefit of the indemnity, as they are generally limited to certain categories of users;
- The indemnity's scope, i.e., what it covers and whether these is a limit to the amount it will cover; and
- The conditions that must be complied with in order to rely upon the indemnity.

The infringement questions currently before the courts are broad ranging. They include whether the collection, storage and training process of GenAI tools involves an infringement of any of the exclusive rights of the copyright owner and, if so, whether there are any potential defenses.

The assessment of whether there is infringement in the training of the models will depend upon several factors, including consideration of each of the steps involved in the training process, and where those acts take place. It may also involve complex evidentiary assessments, including expert evidence on exactly how the AI tools were trained and operate.

As for potential defenses, the U.S. courts will need to address questions around the applicability of the fair use defense, whereas in other jurisdictions, potentially relevant defenses include those relating to text and data mining, temporary copies and exceptions for parody or pastiche.

Litigation Relating to GenAI and IP

U.S. Litigation

The vast majority of the GenAI and copyright litigation before the courts is currently, unsurprisingly, in the U.S. Given the fact that all the major AI companies are based in the U.S., and will need to comply with the laws of their home country, the outcomes of these cases are likely to dictate to a large extent how they operate in future globally.

A broad variety of rights holders have brought actions, including on a class action basis, relating to most of the significant GenAI tools. This includes:

- Claims by artists and image libraries, such as Getty Images, against image generators like Stability AI Ltd. and Midjourney;[4]

- Claims by comedians, and fiction and nonfiction authors that allege that their books were copied in the training process, such as the lawsuit filed in September by the U.S. Authors Guild against OpenAI in the U.S. District Court for the Southern District of New York, to which Microsoft Corp. was added as a defendant in December;[5]

- Claims relating to Meta Platforms Inc.'s LLaMa product;[6]

- A claim brought in October in the U.S. District Court for the Middle District of Tennessee by various leading music publishers about song lyrics relating to Anthropic PBC's "Claude" model;[7]

- A claim brought by former Presidential candidate Mike Huckabee and others against Bloomberg in October in the U.S. District Court for the Southern District of New York;[8] and

- A claim brought in November 2022 in the U.S. District Court for the Northern District of California against Microsoft Corp. and GitHub Inc.[9] relating to its "Copilot" coding tool.

Further claims will no doubt be issued into 2024 and beyond.

In several of these U.S. cases, AI companies have gone on the offensive by filing motions to dismiss aspects of the claim, which has led in some instances to the issues being trimmed down or dismissed subject to amendment. For example, decisions have been considered by the U.S. District Court for the Northern District of California in Andersen v. Stability AI on

Oct. 30, and Kadrey v. Meta on Nov. 20.

In particular, the claims that the tool itself is an infringing work and that all content generated by the relevant tool will infringe copyright on the content used to train the tool were dismissed with leave to amend. For instance, on Nov. 29, an amended complaint was filed in Andersen v. Stability AI, with a number of new plaintiffs joining the complaint.[10]

However, the central issue of whether the act of training the tool involves an infringement of IP rights and, if so, whether a fair use defense arises, remains very much in play before the U.S. courts. This part of the various claims has not, yet, been the focus of a motion to dismiss.

Ultimately, however, given the multifaceted fair use arguments, these questions may take the U.S. court system many years to resolve and, indeed, an answer may only be forthcoming in relation to the specific scenario before the particular court, as opposed to one of general application. As such, in November, a representative of Meta told the U.K. House of Lords Communication and Digital Committee that they anticipated a decade of litigation to settle AI copyright issues.[11]

U.K. and EU Litigation

Turning to the U.K., there is one piece of ongoing litigation between Getty Images and Stability AI[12] that involves claims relating to copyright, database right and trademark infringement or passing off.

Getty Images' argument is that the training and development of Stability AI's Stable Diffusion models infringed almost 12 million of its copyright works. Stability AI applied for summary judgment or strike out in respect of various, but not all, issues in the case. This included the training and development of its model, i.e., that Getty Images had no real prospect of succeeding on those claims or issues and there was no other compelling reason why they should be disposed of at trial.

However, on Dec. 1, the High Court of Justice of England and Wales rejected Stability AI's application and allowed Getty Images to amend its claim to bring in a new claim in relation to an image-to-image feature of Stable Diffusion, which Getty Images argues allows users to make "essentially identical copies of copyright works."[13]

The key issue before the court on the summary judgment application was whether there was any prima facie evidence that the training and development of Stable Diffusion took place in the U.K. Getty Images argues that it did, based on an inference that — during development and training of Stable Diffusion — its copyright works would have been downloaded on servers or computers in the U.K.

While the judge noted that Stability AI's evidence on its face suggested that no development or training had taken place in the U.K., she was not satisfied that there was no reasonable prospect of Getty Images being able at trial to refute that evidence. The case will therefore continue to trial, with the next stage being for Stability AI to file its defense.

Proceedings are also on foot in other jurisdictions, including copyright infringement proceedings in the Hamburg Regional Court in Germany against LAION eV in relation to its widely used LAION-5B dataset. A hearing date has been set for April 25, 2024, in these proceedings.[14]

Policy and Legislative Developments

As noted above, it will likely take some years for the various issues to be fully ventilated and determined before the courts in the various jurisdictions. As a result, policymakers around the world are also urgently considering the various issues, recognizing the overarching need to ensure a fair and equitable balance between protecting valuable IP and encouraging innovation in AI, in particular, in their relevant territory.

In the U.K., for example, the government announced plans in 2022[15] to introduce a very broad text and data mining exception with the aim of promoting investment in AI, with no opportunity for rights holders to opt out. This proposal was pulled in early 2023, following significant opposition by rights holders.

However, in March, Government Chief Scientific Adviser Sir Patrick Vallance published the report, "Pro-innovation Regulation of Technologies Review,"[16] in which he called for an enabling environment for text and data mining.

Following this, the government announced in March that the U.K. Intellectual Property Office would work with users and rights holders on a code of practice on copyright and AI, with the stated aim of making licenses for data mining more readily available for rights owners.

However, the government also said that if the code could not be agreed, or was not adopted, it would consider introducing legislation. The code of practice was due to be published during the summer. However, since there is no sign of it, this suggests that it is proving difficult to agree on it with the various stakeholders, and it is looking unlikely that it will be possible for agreement to be reached on a voluntary basis.[17]

Clearly, the issues involved traverse some significant points of principle, as highlighted in the recent evidence hearings before the U.K. House of Lords Communication and Digital Committee. It now appears very unlikely that it will be possible for agreement to be reached on a voluntary basis.

Meanwhile, at the European Union level, the EU AI Act has completed its passage through the trilogue negotiations process, with a provisional agreement being reached between the EU institutions on Dec. 8, although the final text is awaited. Included within this debate was the question of copyright, with the final compromise text set to include transparency requirements around the training, design and development of models, e.g., requiring sufficiently detailed summaries of content used for training to be published.[18]

Concluding Comments

2024 is shaping up to be a year of fast-moving development in the world of GenAI technology, and this can only be expected to put into ever sharper focus the challenges these developments continue to present for protection and enforcement of IP rights.

While the courts will, in time, assess the extent to which these issues can be effectively determined under existing legal frameworks, and individual governments and regulators will no doubt propose and adopt legislation or regulations that attempt to navigate a path through the various competing interests, the underlying challenge remains that this is a global issue requiring a global solution.

Nina O'Sullivan is a partner, Anne Rose is a managing associate and Peter Nunn is a partner at Mishcon de Reya LLP.

[1] Copyright Registration Guidance for Works Containing AI-Generated Materials, 16 March 2023: https://www.govinfo.gov/content/pkg/FR-2023-03-16/pdf/2023-05321.pdf.

[2] Notice of Appeal in Stephen Thaler v Shira Perlmutter, Register of Copyrights and Director of the United States Copyright Office, et al, 1:22-cv-01564-BAH, 11 October 2023.

[3] Memorandum Opinion in Stephen Thaler v Shirla Perlmutter, Register of Copyrights and Director of the United States Copyright Office, et al, 1:22-cv-01564-BAH, 18 August 2023: https://www.copyright.gov/ai/docs/district-court-decision-affirming-refusal-of-registration.pdf.

[4] Getty Images (US), Inc v Stability AI & anr, 1:23-cv-00135-GBW; Sarah Andersen & ors v Stability AI & ors 3:23-cv-00201-WHO; Getty Images (US), Inc & ors v Stability AI Ltd, IL-2023-000007.

[5] Julian Sancton v OpenAI & ors, 1:23-cv-10211; Authors Guild & ors v OpenAI & ors, 1:23-cv-8292; Michael Chabon & ors v OpenAI & ors, 3:23-cv-04625; Sarah Silverman & ors v OpenAI & ors, 3:23-cv-03416; Paul Tremblay v OpenAI & ors, 3:23-cv-03223.

[6] Michael Chabon & ors v Meta Platforms, Inc, 4:23-cv-04633; Richard Kadrey & ors v Meta Platforms, Inc, 3:23-cv-03417.

[7] Concord Music Group & ors v Anthropic PBC, 3:23-cv-01092.

[8] Mike Huckabee & ors v Meta Platforms, Inc & ors, 1:23-cv-09152.

[9] J.Doe 1 & anr v Github, Inc & ors, 3:22-cv-06823.

[10] Amended Complaint in Andersen & ors v Stability AI Ltd & ors filed 29 November 2023.

[11] Oral evidence of Rob Sherman, Vice President and Deputy Chief Privacy Officer for Policy at Meta before House of Lords Communications and Digital Committee, 14 November 2023: https://committees.parliament.uk/oralevidence/13806/pdf/.

[12] Getty Images (US), Inc & ors v Stability AI Ltd, IL-2023-000007.

[13] Getty Images (US), Inc & ors v Stability AI Ltd [2023] EWHC 3090 (Ch).

[14] Robert Kneschke v Laion e.V.

[15] Government response to consultation on Artificial Intelligence and Intellectual Property: copyright and patents, 28 June 2022: https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-

and-patents-government-response-to-consultation.

[16] Pro-innovation Regulation of Technologies Review: Digital Technologies, 15 March 2023: https://www.gov.uk/government/publications/pro-innovation-regulation-of-technologies-review-digital-technologies.

[17] Government response to Pro-innovation Regulation of Technologies Review, March 2023 : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1142798/HMG_response_to_SPV_Digital_Tech_final.pdf.

[18] Press Release of the Council of the EU, 9 December 2023: https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/.