LONDON

INSIGHT SERIES

23 / 04 / 2023

AI & Intellectual Property: COPYRIGHT and the fruit of the inventive tree.

INSIGHT NUMBER 12
2023

INSIGHT NOTES ARE PERIODIC COMMENTARIES ON SUBJECTS OF LAW WHICH ARE OF INTEREST TO OUR PARTNERS OR WHICH AFFECT THE AREAS WE PRACTICE

THEY ARE NOT DESIGNED TO BE SUBSTITUTES FOR LEGAL ADVICE OR AS ACCURATE SUMMARIES OF THE LAW AND NO LIABILITY IS ACCEPTED FOR THE CONTENTS OR ANY RELIANCE UPON THE CONTENTS OR OPINIONS EXPRESSED THEREIN.

INSIGHTS ARE ONLY ACCURATE AT THE DATE OF PUBLICATION.

CERTAIN OPINIONS EXPRESSED IN INSIGHTS MAY BE SUBJECT TO LEGAL DEBATE OR UNCERTAINTY. IN SOME CASES, THAT UNCERTAINTY MAY BE RESOLVED BY SUBSEQUENT CASELAW.

TECHNOLOGY

AI and Intellectual Property:
COPYRIGHT and the fruit of the inventive tree.

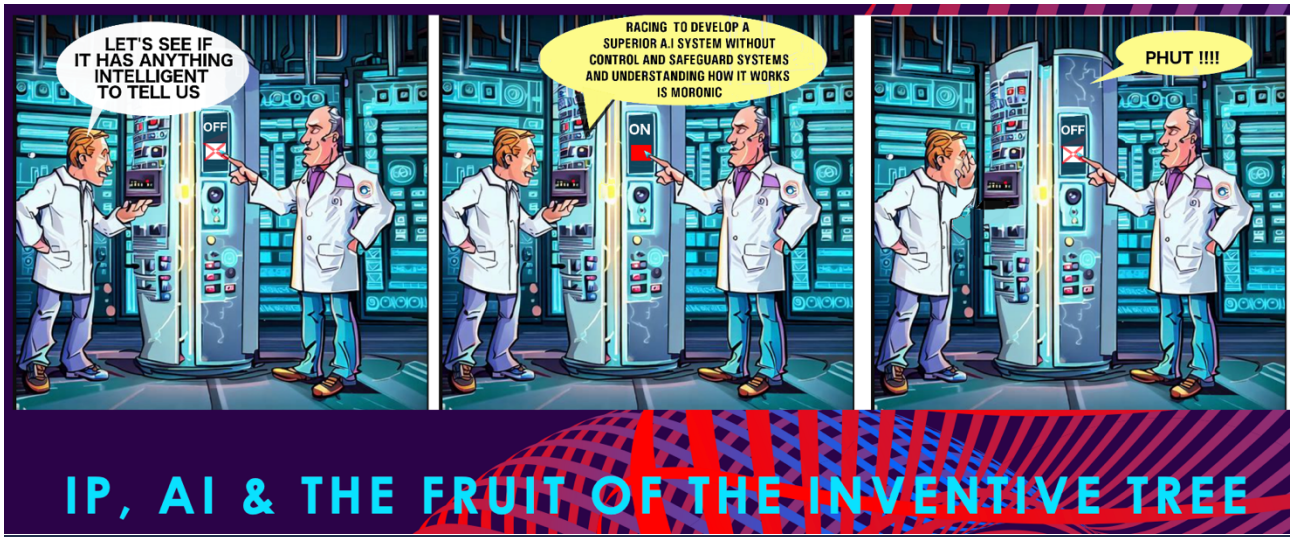
Before We Start

In the UK, the predominant reaction to intellectual property infringement is through the Civil Courts, and outside the scope of large wholesale piracy and counterfeiting systems, including piracy reselling of television and movies. As a law enforcement command of the Home Office, Border Force is charged with protecting the UK from security threats at the border, which includes IP crime. Border Force are responsible for intervening in and processing suspected IP infringing material at the border on behalf of Her Majesty's Revenue and Customs (HMRC). The IPO works closely with Border Force to develop evidence and action in this area. Two key agencies that take the lead in tackling IP infringement on the ground are Trading Standards and the Police; however whilst trading standards regularly seize small-scale counterfeits, police actions are largely limited to large-scale operations or sting operations on particular area, or actions in conjunction with other agencies such as HMRC or Border Force. The main role of the police lies with the Police Intellectual Property Crime Unit (PIPCU), a dedicated IP crime unit, run by the City of London Police and funded by the UK Intellectual Property Office (IPO), the intellectual property registration and regulatory authority. PIPCU coordinates the activities of industry, government, and law enforcement agencies to help counter IP crime, working closely with others such as Trading Standards, Border Force, and the UK Visas and Immigration division of the Home Office.

Although technically an IP case can be proceed without a complainant, in practice, all criminal prosecutions require a robust and compliant, co-operative complainant and many complainants do not see the cost-benefit analysis of a prosecution because of the poor integration of compensation orders for the damage suffered.

Thus, for 99.9% of criminal intellectual property cases, it is the State, through the Crown Prosecution Service (CPS), that controls the prosecution, with very few private criminal prosecutions because of the very high costs involved and it is also the reason why criminal intellectual property cases are dropping as a criminal intellectual property case takes the same resource requirements and costs very similar to a murder or corporate manslaughter case. These CPS cases, outside the Border Force actions, tend to concentrate on counterfeiting, drugs, cigarettes, alcohol and online piracy.

In practice, most piracy and infringmen t actions, together with all trade-secret actions, are taken in the civil courts and this is reflected in the truism that if you're not going to get over the civil burden of proof (greater than 50% likelihood), you're not going to reach the criminal burden (beyond reasonable doubt).



1. What is AI?

When we talk about AI, we need to first determine what we mean as the term encompasses everything from Siri to Chat-GPT4 and one of the problems with dealing in this field is determining what we mean by AI. Even the regulators get this wrong as has been seen by the extensive criticism of the terminology used in the EU AI strategy.

Things largely changed in 2015, a landmark year for artificial intelligence, with the number of neural net projects utilising some form of AI within Google passing 3,000, a massive change from the handful in use just three years later. This change was attributable to affordable neural networks and a rise in cloud computing infrastructure with resultant drop in costs and significant increases in research tools and datasets.

These datasets led to narrow AIs (ANIs) which do narrow tasks far in excess of human capability such as Deepmind's Alpha-Go and in 2020 Alpha-Fold which solved the protein folding problem[1] whilst developers also worked on trying to create Artificial General Intelligence (AGIs) which can do most tasks at a near human level and matching human intelligence (the "singularity") which in 2018, it was widely expected that the singularity was expected at about 2065.

Since 2015 and particularly in the last 5 years however, the AI landscape has changed with the emergence of large language datasets – vast data collections, often curated, that Artificial intelligence is trained on. These Large Language Models treat whatever the problem is as a language problem, whether it is DNA – merely a biological language, Art – an artistic language, Sound or Music, Finance – a mathematical and statistical language etc. and the LLMs had a radical effect on AI. Suddenly, everyone was using these large language models with the effect that improvements in one area of AI was applicable to all AI areas making development gains almost exponential.

These General open Large Language Models (GoLLMS[2]) such as Chat-GPT4 are doing things that their own designers didn't train then for and are making decisions and changes that were entirely unforeseen and which the designers don't understand why and can't interpret the algorithmic output.

So AI is really a neural net system which has the capability of carrying out tasks in broadly the manner that a human would do so, albeit a rather stupid human.

2. How does AI impact on Copyright?

Under the Berne convention, there is a requirement for an author to be a national of a signatory state, in other words, a human or legal person, and it is widely accepted that the lack of definition of this is due to a common understanding amongst member states, that authors are those humans who create "original works of expression"[3]. This requirement to be a human or legal person is also reflected in the universal declaration of human rights and supported by the current caselaw. The obligations in Berne and TRIPS are however differently implemented.

In the UK there have just been a number of reviews by the IPO, the Intellectual Property Office looking at copyright, patents, design rights and the impact of AI which looked at

- How do we protect AI systems
- How do we protect the creations and outputs from AI
- Should IP be linked to human endeavour and where does ownership sit?
- ML, RL and text and data mining (TDM)
- Where does liability arise for infringement etc.

[1] https://www.future-science.com/doi/10.2144/btn-2022-0007

[2] Term coined by Aza Raskin and Tristan Harris CHT

[3] Berne convention Art 3

The review was necessary following concerns about AI's impact and following the Thaler line of patent cases where courts looked at whether AI could own a patent and the Monkey Selfie case that suggested that AI could not create copyright.

The UK's CDPA which is 30 years old specifies that an author is the person who creates a work and fortunately some foresighted draftsman thought to add s9(3) which says that *"In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken".*

The legislation clearly states that AI cannot own copyright, but because of s9(3), AI generated works are usually protected in the UK, unlike in other jurisdictions such as the USA where AI generated works are incapable of copyright protection according to the ruling in Zarya of the Dawn[4] (which follows the Monkey Selfie Case[5]) by the Copyright Office[6] . Zarya ruled that although Ms. Kashtanova was the author of the text under the law, and the comic book's selection, coordination, and arrangement of the Work's written and visual elements, the image was a Midjourney image, not the product of human authorship, and not capable of copyright protection.

- a work may be registered if it qualifies as an "original work[] of authorship fixed in any tangible medium of expression." 17 U.S.C. § 102(a)
- "original" in this context consists of two components: independent creation and sufficient creativity. See Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991) Supreme Court
- the work must have been independently created by the author
- the work must possess sufficient creativity
- US Courts have interpreted the phrase "works of authorship" have uniformly limited it to the creations of human authors[7]
- where non-human authorship is claimed, appellate courts have found that copyright does not protect the creation[8]
-humans not creating work means no copyright protection is available[9]
- As Midjourney offers an artificial intelligence technology capable of generating images in response to text provided by a user and users also have the option to include (1) a URL of one or more images to influence the generated output, or (2) parameters directing Midjourney to generate an image in a particular aspect ratio or providing other functional directions, which are converted to an algorithm meaning that the process by which a Midjourney user obtains an image differs from a human artist, writer, or photographer.
- the Office concludes that the images generated by Midjourney contained within the Work are not original works of authorship protected by copyright- produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author and Midjourney originated the "traditional elements of authorship" in the images, in an uncontrolled and unpredictable way and Midjourney users lack sufficient control over generated images to be treated as the "master mind" behind them.
- even though Ms. Kashtanova's contention that she expended significant time and effort working with Midjourney. But that effort does not make her the "author" of Midjourney images under copyright law, Ms. Kashtanova is not the author for copyright purpose of the individual images generated by Midjourney
- tweaks to the image by Ms. Kashtanova are too minor to create copyright protection[10]

---

[4] https://www.copyright.gov/docs/zarya-of-the-dawn.pdf

[5] In the Monkey Selfie (Naruto) case, it was recognised that animals can execute cognitive tasks such as pressing a camera shutter but that they did not meet the standards of creativity required to own the copyright and that if copyright existed, it lay with the photographer who made the arrangements for the photograph to be taken.

[6] https://www.copyright.gov/title17/

[7] Burrow-Giles Lithographic Co. v. Sarony 111 U.S. 53, 57–59 (1884)

[8] Urantia Found. v. Kristen Maaherra, 114 F.3d 955, 957–59 (9th Cir. 1997). The Urantia court held that "some element of human creativity must have occurred in order for the Book to be copyrightable" because "it is not creations of divine beings that the copyright laws were intended to protect."

[9] U.S. Copyright Office, Compensium of US Copyright Office Practice § 313.2 (3d ed. 2021) ("COMPENDIUM (THIRD)")

[10] See doctrine in Contra Eden Toys, Inc. v. Florelee Undergarment Co., 697 F.2d 27, 34–35 (2d Cir. 1982) (

- copyright registration explicitly excludes "artwork generated by artificial intelligence."
- And in the current edition of the Compendium, the Office states that ''to qualify as a work of 'authorship' a work must be created by a human being'' and that it ''will not register works produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author[11].''

Curiously whilst the Monkey selfie case mirrored the law in the UK, the Zarya of the Dawn case didn't. Although on 16th march 2023, the US Copyright office issued a review[12] to look further at this issue.

In the UK, even despite s9(3) there are potential pitfalls under s9(40 where "a work is of "unknown authorship" if the identity of the author is unknown' or in the context of AI generation ..if the identity of the person by whom the arrangements necessary for the creation of the work cannot be identified and it is likely that challenges to AI-generated infringement claims may arise as a result of this lack of ability to identify the maker of the arrangements, although in may cases it will be the employer under the terms of employment. It is also possible that as things progress in AI, this test may become harder and harder to determine.

There is also problematic latitude in the UK law because whilst s9(3) "necessary arrangements" applies it is open to argument about whether the person making the arrangements is the developer company (so that, for example, Midjourney owns all of the copyright in their images) or whether the person making the arrangements is the person who is creating the keywords from which the AI generates the image. In this respect, some of the arguments raised in the Zarya case and in the subsequent guidelines of the US Copyright office may have applicability in the UK, particularly for commercial use of images.

Like the early days of the internet, the answer will arise from judicial interpretation.

3. The Problem of Originality – is it a human construct?

In addition, even in the UK, the question of "originality" is also open to interpretation in relation to AI-generated works as *"Creativity is a quintessentially human faculty"*[13]. The Berne convention mentions originality and original works without defining these and the law is not assisted by TRIPS and WIPO treaties which refer to an *idea v expression* dichotomy and by caselaw phrases such as "*sweat of the brow"* (US), "*skill, labour and judgment*" (UK) and "*exercise of skill and judgment*" (Canada) and "Schöpfungshöhe" (level of creativity) (Germany). [14]

As such, the current caselaw[15] appears to open a challenge to the question of an "original work" or "originality" and whether there are challenges based on a lack of originality for AI-generated work where the algorithm with the same instructions would necessarily produce (as matters currently stand) the same work, making it non-original or where there is a merely mechanical mathematic (albeit very complex) result being produced. In Australia, the Court ruled that if there is no human author, then

[11] U.S. Copyright Office, Compendium of U.S. Copyright Office Practices sec. 313.2 (3d ed. 2021) (''Compendium (Third)'').

[12] https://www.govinfo.gov/content/pkg/FR-2023-03-16/pdf/2023-05321.pdf

[13] Vanherpe: AI and Ip- a tale of 2 acronyms. See also De Bruyne and Vanleenhove "Robota AI and the Law in Belgium (2020),221.

[14] Pre-Feist caselaw as Feist was quashed after the US joined the Berne Convention in 1988, Ladbroke v. William Hill [1964] *All E.R.* 465 at 469, CCH Canadian Ltd. v. Law Society of Upper Can, Gerhard Schricker, 'Farewell to the "Level of Creativity" (Schöpfungshöhe) in German Copyright Law? ' (1995) *IIC - International Review of Intellectual Property and Competition Law* 41

[15] Compare to Case C-5/08 - Infopaq International A/S v. Danske Dagblades Forening, Judgement of the Court (Fourth Chamber), 16 July 2009 (ECLI:EU:C:2009:465); Case C-145/10 - Eva-Maria Painer v. Standard Verlags GmbH and Others, Judgment of the Court (Third Chamber), 1 December 2011 (ECLI:EU:C:2011:798); Joined Cases C-403/08 and C-429/08 - Football Association Premier League Ltd and Others v. QC Leisure and Others, and Karen Murphy v Media Protection Services Ltd., Judgment of the Court (Grand Chamber), 4 October 2011 (ECLI:EU:C:2011:631); Case C-604/10 - Football Dataco Ltd and Others v. Yahoo! UK Ltd and Others, Judgment of the Court (Third Chamber), 1 March 2012 (ECLI:EU:C:2012:115); Case C-310/17 - Levola Hengelo BV v. Smilde Foods BV, Judgment of the Court (Grand Chamber), 13 November 2018 (ECLI:EU:C:2018:899); Case C- 683/17 - Cofemel – Sociedade de Vestuário SA v. G-Star Raw CV, Judgment of the Court (Third Chamber), 12 September 2019 (ECLI:EU:C:2019:721); Case C-833/18 - SI, Brompton Bicycle Ltd. v. Chedech / Get2Get, Judgment of the Court (Fifth Chamber), 11 June 2020 (ECLI:EU:C:2020:461)

there can be no originality as defined under Australian Copyright law[16] and seems to reflect a difference between uniqueness and the creative element of originality, creativity being a purely human attribute and not arising where a complex mathematical construct produced a work.

Following its recent review, the IPO looked at removing AI-generated copyright or changing the law and it was felt unnecessary to make changes because we are too early in the AI development cycle to change the law and s9(3) was currently adequate; however as the GoLLMs develop, we are likely to find that the law is inadequate. In an infringement case, expect that a defendant may argue that if the algorithm produces the same result for a given input, then it lacks originality and is the mere product of a mechanical input and something therefore incapable of copyright protection.

What is certain, however, is that AI cannot **own** the copyright and it must be a legal person who holds the ownership, just as the Thaler case has confirmed that AI cannot be the inventor of a patent.

This is an area of potential challenge and defence to infringement to watch as the AI develops and as we understand more about the workings of GoLLMs, but the early indication is that the Courts will want detailed analysis about how the algorithm works as indeed the US Copyright Office required when assessing the Zarya case and caselaw references to the "creation of the mind"which must clearly be a human one[17]. In Europe we have long allowed rights to producers of sound recordings, films etc as this also reflects the product of the human mind.

With the diversity of laws where AI-generated images are protectable by copyright in some countries and not others, we also have the added complexity of working out where the image is created. If a user is sitting in the UK using Midjourney's US servers, and applying some of the logic in the Zarya case and the subsequent commentaries of the US Copyright Office, the question is where does the image arise, is it on the servers of Midjourney in the UK which is then transmitted to the UK or is it on the laptop of the user. This may create further arguments and defences of non-protection by copyright. Despite the use of the internet over the last 30 years, there is very little, if any, reliable caselaw on this point [18].

The policy issue faced by the UK IPO was whether to amend copyright to expressly protect machine creativity or stay with the existing state of UK copyright protection, and if changed whether copyright law should treat AI-generated works differently from other computer-generated works (i.e. is there a difference between AI generated art and other computer-generated works which are not artificially intelligent). The IPO ducked the question, deciding to stay with the status quo until GoLLMs were better understood.

It should also be noted that, even though there is protection under s9(3) of the CDPA in the UK, this protection is limited to 50 years and excludes moral rights.

A number of reports have also arisen since the IPO consultation about GoLLMs which have indicated that the AI's may be skiing off-piste. For example, an AI system trained only in English and programmed only to look at English texts was found, because it had been released on the internet as an additional data model, was exposed to the foreign languages and (although the programmers don't know how or why), the GoLLM managed to learn the other languages and therefore took a new dataset of works in other languages. What we don't know is how the law will look at this because whilst ownership is certain, and the liability for foreseeable infringement may sit with the legal person making arrangements for the machine, we are not entirely

[16] Achohs Pty Ltd. v Ucorp Pty Ltd. [2010] FCA 577

[17] Stanford Technology Review: 1, 3-9; gervais 2073-2085 Coding Creativity : Copyright and the Artificially Intelligent Author (Bridy).

[18] Nova Productions Ltd v Mazooma Games [2007] EWCA 219]. While Kitchin J's analysis in Nova is sound (as between the programmer and player in the video game) it does not have direct parallels to Generative AI images as the scope for 'the arrangements necessary for the creation of the work" are unclear and the capability, scale, autonomy and automation is now far beyond anything envisioned by the draftspersons of the CDPA 1988.

sure where infringement will arise for these off-piste activities and again this will be decided by caselaw, initially on a case by case basis.

On a purely practical note, contracts should specify how AI-generated work is to be incorporated into contracts as well as specifying where liability for infringement will arise.

Contractor shall not use or incorporate, and shall not allow any third party to use or incorporate, any artificial intelligence (AI), including but not limited to any generative AI or any AI tools (ChatGPT, Midjourney, etc.)(or their work products) in the creation or development of any work product without company's prior written approval.

Contractor shall be liable for all infringement claims arising out of any use, training or output of any artificial intelligence (AI), incorporated into the work product including but not limited to any generative AI or any AI tools (ChatGPT, Midjourney, etc.).

4. Orphan Works

An orphan work is a copyright-protected work for which rightsholders are positively indeterminate or uncontactable. Sometimes the names of the originators or rightsholders are known, yet it is impossible to contact them because additional details cannot be found.

A work can become orphaned through rightsholders being unaware of their holding, or by their demise (e.g. deceased persons or defunct companies) and establishing inheritance has proved impracticable. In other cases, comprehensively diligent research fails to determine any authors, creators or originators for a work.

Since 1987, as the internet expanded, the amount of orphan works has increased dramatically since some works are published anonymously, as assignments of rights are not required to be disclosed publicly, and (where registration is required) due to failed or missing registration and, thus, there are many unknown orphan works released to the public even when those rights are being actively exploited by authors or other rightsholders.

For the purposes covered by this talk, the orphan works are synonymous with other works.

5. Text and Data Mining – TDM

Finally, we come to the question of text and data mining (TDM).

A class-action lawsuit filed in the US District Court for the Northern District of California alleges that AI art generators Stability Diffusion, Midjourney, and DreamUp were trained on billions of copyrighted materials without credit, compensation, or consent of content owners, but in the UK the training of AI is legal as s29A CDPA exempts use for text and data analysis (TDA) for non-commercial purposes where there is a computational analysis of the works and this is exactly what an AI being trained on a LLM is doing.

> *(1)The making of a copy of a work by a person who has lawful access to the work does not infringe copyright in the work provided that–*
>
> *(a)the copy is made in order that a person who has lawful access to the work may carry out a computational analysis of anything recorded in the work for the sole purpose of research for a non-commercial purpose, and*
>
> *(b)the copy is accompanied by a sufficient acknowledgement (unless this would be impossible for reasons of practicality or otherwise).*



Whilst it is claimed by some that the development of AI is commercial and therefore not within the exemption under s29A, it is equally arguable that the copy is made for the purposes of training the AI which then results in an analysis and the AI stores an algorithmic analysis and not the copy of the work and therefore the use falls within s29A. This reflects the analysis of the US Copyright Office in the Zarya case. The result of this is that having trained the GoLLM system, and as it is only then deployed commercially, there is no infringement because at the time of the commercial release of the AI, the AI is no longer storing a copy of the work but only the computational analysis of the work so only s29A(1) applies.
This parallels what arises in classrooms and universities where for many years, students have studies their forebears and copied their particular styles, whether a Hockney collage or a Picasso or Guillermo Del Toro,with the student ultimately graduating with a style of their own. In the context of AI, AI image GoLLMs can produce works in any style, and every image is unique and given the nature of randomness built into Generative Ais, it would be extraordinary if the AI produced an exact copy of the underlying artist and as developers are fully aware of the risk, the AI system checks against its algorithmic analysis of original works to ensure that no exact copies of the original works are being produced. Leaving only the question of a substantially similar work to be open to litigators to argue about (although the substantially similar work relies upon proof of knowledge of the original work which is not a sustainable argument against the AI's analysed algorithm if the original work cannot be reproduced from this.

Whilst for artists, there is a huge concern and legitimate concern that simply using a prompt of "in the style of X" will result in a myriad of AI-generated images that means that the original artist's works will not be able to be found amongst the mass of copies[19], we are fortunate in the UK as the IPO have clearly taken the view that it is in the public interest that AI training is promoted as accurately as possible and that AI systems should be able to be fed text, art, music etc all of which are copyright

[19] MT Technology Review 16th Sept 2022 "AI: This artist is dominating AI-generated art. And he's not happy about it."

with the result that this is now a moot point in the UK as the UK IPO has however announced that to avoid this confusion, the provision of "for a non-commercial purpose" will be shortly removed creating a general exemption for TDA and this will also apply to databases.

The solution in the US and wider internet are possibly:
a) a collective licensing scheme; and
b) the adoption of a ***NoTrain*** meta-tag.

Whilst the UK has a collective licensing team which is effective, this does not currently cover internet data or content and any collective licensing scheme for training AI would need to cover this data and would require a novel and widespread registration scheme for internet content.

Whilst *NoTrain* tags are expected to be easily possible, with Artificial Intelligence systems ignoring the content, this has an economic impact for the user as their content will be excluded from the next generation of AI tools and AI-generated data as well as excluded from the AI training data. The *NoTrain* tag would easily be compromised by major AI players such as Google and Microsoft who will create an incentive not to use these by stating that their search engines are AI-based and therefore that use of the *NoTrain* tags will also mean that the user disappears from critical search engines.

And the next challenge is privacy and image rights (and autonomous weapons)

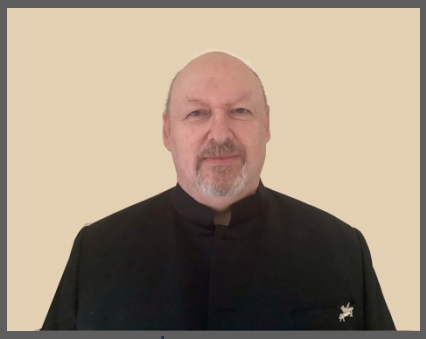
..but we're not out of the woods yet as there are concerns that it now only requires a few seconds of voice for AI to be able to produce an exact voiceprint and the same is true of real-time video fakes which means that the millions spend on voice and face identification for banks has overnight become both obsolete as a technology and also a major source of fraud and (coupled with smartphones and "undressing" software) this means that privacy and publicity rights will now become a significant issue.

ADL | ADL SOLICITORS

LONDON

TECHNOLOGY

AI & Intellectual Property: COPYRIGHT and the fruit of the inventive tree.



INSIGHT SERIES 2023

ADL SOLICITORS, 131 -133 CANNON STREET, LONDON, EC4N 5AX
WWW . ADLSOLICITORS . COM
Tel: 0200 888 0300

A short summary of the impact of Artificial Intelligence on copyright and intellectual property as delivered to the international conference on Intellectual Property and Artificial Intelligence at King's College London in April 2023.
bv Nick Lockett

