OLIVIA WHITCROFT

“All providers and deployers have an obligation to ensure a sufficient level of AI literacy”

After last month’s introduction to AI and the law, Olivia examines the Getty Images judgment, the EU’s AI Act and the latest on Clearview AI’s case

Last month I considered legal and contractual liabilities in the AI supply chain, and the use of AI to create deliverables. This time I’ll be sharing my thoughts on the huge Getty Images judgment, as well as the EU AI Act and data protection updates.

Getty Images vs Stability AI

Getty Images has been in an intellectual property battle with Stability AI over its use of Getty’s images for its Stable Diffusion AI model. Action was taken on several grounds, including copyright infringement when training the model and in generated outputs, secondary copyright infringement in deploying its (allegedly infringing) model in the UK, and trade mark infringement relating to watermarks appearing in output images.

Key copyright claims were dropped due to lack of evidence of training in the UK, and when Stability AI made technical changes to limit generation of infringing content. The secondary copyright and trade mark claims did proceed, and the High Court gave its exciting judgment on 4 November 2025 ([2025] EWHC 2863 (Ch)).

Getty had some success in the trade mark claim, though the findings were “both historic and extremely limited in scope”. It didn’t succeed in the secondary infringement claim. A key issue was whether (under the Copyright, Designs and Patents Act 1988) the Stability AI model was an “article” which is an “infringing copy” of copyright works. The court agreed that an article may be an intangible object. However, while the parameters of the model have been exposed to copyright works, the model doesn’t store or reproduce those works. The parameters are “purely the product of the patterns and features which they have learnt over time during the training process”. Therefore, there was no infringing copy.

Olivia is principal of the law firm OBEP, which specialises in technology contracts, IP and data protection
@ObepOlivia

“Key copyright claims were dropped due to lack of evidence of training in the UK”

As well as having a big impact in an intellectual property context, I wonder whether it helps with the much-debated issue of whether an AI model in itself processes personal data. I’m coming back to this below.

EU AI Act

I’ve been a bit lax at getting on top of the EU AI Act, because, hey, it’s an EU law, and the UK blew its chances of joining in with it when it left the EU. But unsurprisingly there’s a lot of talk about it in the data protection and tech law circles that I frequent. It’s being seen as a potential baseline for UK organisations, for two important reasons: we don’t have an alternative baseline; and if we don’t meet its standards, products incorporating AI may be excluded from the EU market.

The Act came into force in August 2024, though not all its provisions apply yet. In 2025, Articles on prohibited AI and AI literacy came into effect, as well as obligations relating to General-Purpose AI (GPAI). As from August 2026, the rules for most high-risk systems will kick in, though the deadline for some extends to August 2027. There are transitional provisions for systems put on the market before August 2026, which could prompt a rush to launch before this date. I’m focusing here on the matters now in force.

The EU AI Act has the concept of “providers”, who (broadly) develop AI systems, and “deployers”, who use AI systems. All providers and deployers have an obligation to ensure a sufficient level of AI literacy of staff dealing with the operation and use of AI systems. In short, they need some AI experts on the team, rather than stopping the first person you see in the corridor and giving them a shiny new lanyard saying “Head of AI” (which is the way many data protection officers were appointed when the GDPR was new).

GPAI models are those that are trained on large amounts of data and can be used within other products down the supply chain (such as GPT-4 and other LLMs). Providers must maintain technical documentation, facilitate others to understand the model’s capabilities and limitations (with some exceptions for open-source models), and have policies to comply with intellectual property laws. If the model has high impact capabilities, it’s classified as having systematic risk, and the provider must evaluate and mitigate those risks, and ensure cybersecurity protection.

Prohibited AI practices include those that sound dodgy and we can all rally behind, such as AI systems deploying subliminal techniques or exploiting human vulnerabilities based on age or disability. Then there are some that sound familiar: AI systems that create facial recognition databases through untargeted scraping of images from the internet. (Isn’t that what Clearview AI does? See below.) And AI systems that infer emotions in the workplace – I attended a talk last year about use of emotion-detecting AI in recruitment. Some firms may therefore need to stop in their tracks, at least for EU activities.

Using personal data to train AI

In September 2025, LinkedIn sent me an email to let me know that, as from 3 November, it would be using my data (as a member) to “improve content-generating AI that enhances your experience and better connects our members to opportunities”. It wasn’t asking for my consent, but it went on: “You can opt out anytime in your settings if you’d prefer not to have your data used in this way.”

A year before that, the Information Commissioner’s Office (ICO) had raised concerns about LinkedIn’s approach to training GenAI models with personal data, and LinkedIn had suspended this training. Fast-forward to September 2025, when the ICO was apparently pleased with the

BELOW The High Court ruled that Stability AI hadn’t infringed Getty’s copyright

Wednesday 5 November 2025 The Guardian

National

AI firm wins high court ruling over copyright claim by photo agency

Robert Booth
UK technology editor

A London-based artificial intelligence firm has won a landmark high court

generate images with text prompts. Stability was found to have infringed Getty’s trademarks in some cases.

The judge, Mrs Justice Smith, said the question of where to strike the balance between the interests of

improvements to transparency (LinkedIn sent me an email), a simple way to opt out (in my settings) and the window to do this (I had until November).

LinkedIn needs to have a valid lawful basis under data protection law for using member data in this way (as I discussed in issue 361). It has gone for the option that the training is necessary for its legitimate interests, meaning it must demonstrate the necessity, and ensure the rights of individuals aren't compromised. Common safeguards include providing a clear and timely opt-out mechanism. From LinkedIn's perspective, this is likely to be preferable to the alternative of seeking consent, not least because us lazy customers must put in some effort to stop them doing it, and we might not bother.

Meanwhile, in the EU, the EU Commission published its "Digital Package" in November. This includes plans to reform the GDPR to explicitly allow the legitimate interests lawful basis to be used for development and operation of an AI system, provided safeguards are in place, such as data minimisation, transparency and an unconditional right to object.

Clearview AI

Another big AI-related decision considered the meaning of "monitoring of behaviour" of individuals, in the context of the extra-territorial reach of UK data protection law. In 2022, the ICO fined the US company Clearview AI £7.5 million for failure to comply with UK rules for activities involving the monitoring of the behaviour of UK individuals.

Clearview AI sells its facial recognition tool primarily to law-enforcement customers. It scrapes images of faces and associated meta data from the internet to build up a database against which customers can match their own images.

Clearview appealed the fine, and, in 2023, the First Tier Tribunal determined that the ICO shouldn't have taken enforcement action, as it didn't have jurisdiction. Although Clearview was conducting activities related to its customers' monitoring of UK individuals (which brought it within the territorial scope), Clearview's activities were outside the material scope of UK data protection law, due to such customers' law enforcement connection (though, as the Upper Tribunal commented, the reasons given were "very sparse indeed").

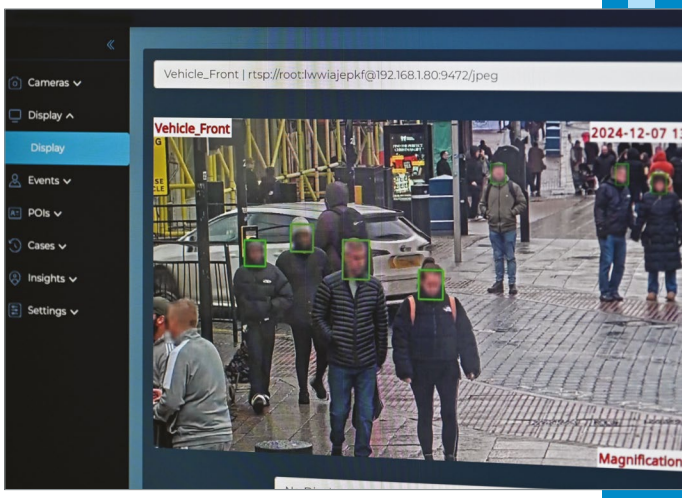
The ICO appealed on upwards and, in October 2025, the Upper Tribunal determined that the activities were within the material scope and, importantly, that Clearview *itself* was monitoring the behaviour of UK individuals (and not just its customers). It gave a broad interpretation to the meaning of "monitoring", and referred to Clearview's gathering, sorting and storing of "behaviourally rich" data.

I'm often asked what activities count as behavioural monitoring. Some online activities may be focused on UK data subjects, so it makes sense that the rules should apply. However, they don't always comfortably sit within the literal interpretation of "monitoring of behaviour". The broad interpretation from this decision provides some clarity that use of AI or other automated tools to gather personal data will be caught by the rules.

Anonymous or pseudonymous data

Another thought is whether removing identifiers from data used to train AI brings it outside the scope of data protection law. The rules don't apply to truly anonymous data, though companies should bear in mind risks of contextual identification within large data sets, or an individual re-identifying themselves (in which case, special rules kick in under Article 11 UK GDPR).

Pseudonymisation, in contrast, is where identifiers are removed from a set of personal data so that the information by itself can no longer be attributed to a specific individual, but the controller separately retains a key to enable it to re-identify individuals. This can be an effective security measure, but the data is still personal data. A recent judgment of the Court of Justice of the European Union (EDPS v SRB Case C-413/23 P) considered whether pseudonymous data necessarily remains personal data when shared with another party.



ABOVE Clearview AI's facial recognition tool provides insights for law enforcement

"It can be hard to explain how or why AI models make the outputs they do"

BELOW The CJEU has recently considered the thorny issue of pseudonymisation

The Court decided pseudonymisation *could* effectively prevent parties other than the controller (who holds the key) from identifying individuals, such that the data is not personal data in the recipient's hands.

This case didn't relate to an AI model, but it makes me wonder how it impacts the AI training context. If pseudonymised data is shared with the provider of a model, who then (for its own purposes) uses that data to train the model, perhaps it's not personal data in the AI company's hands. Though with vast amounts of information fed into a model, there may be risks of re-identification, and consider whether data should in any case be anonymised instead (and the special rules under Article 11 are also relevant here).

GenAI outputs

A couple of years ago (*see issue 353, p116*) I mentioned a complaint made to the Polish data protection authority when ChatGPT generated false information about the complainant. OpenAI subsequently blocked generation of this information, but it appears unresolved how it was generated to start with. The Law Commission paper that I discussed last month refers to the "opacity" of AI models, meaning that it can be hard to explain how or why they make the outputs they do.

Separate to the use of data in the training process is the question of whether there's personal data in the model itself, and whether it's possible to access or remove this if needed (such as where an individual seeks to exercise their rights). The Getty case may lead us to say "no", though the proposals to amend the GDPR say there should be safeguards "to protect against non-disclosure of residually retained data in the AI system or AI model". What does this mean? I don't know yet, and it will have to wait as I've no more space here.

olivia.whitcroft@obep.uk