OSLER

German and U.K. courts diverge on whether AI models infringe copyright

NOVEMBER 27, 2025 4 MIN READ

Related Expertise

- Artificial Intelligence
- Intellectual Property
- Intellectual Property Disputes

Authors: Mat Brechtel, Matthew J. Diskin, Michael Fekete, Barry Fong, Sam Ip

Key takeaways

- An international consensus on whether AI models infringe copyright is likely years away.
- The Munich court ruled in *GEMA v. OpenAI* that OpenAI memorized song lyrics, creating copyright issues.
- In contrast, the U.K. High Court's *Getty Images v. Stability AI* found no infringement since the model didn't store the copyrighted images.

The rapid adoption of AI has led to a series of disputes that are now forcing courts to grapple with whether AI models trained on copyrighted works constitute copyright infringement. Two recent decisions highlight the lack of international consensus and suggest that outcomes may turn on the specific evidence, the ease at which infringing outputs are generated, and the copyright regimes in each jurisdiction.

The Regional Court of Munich recently issued a significant decision in *GEMA v. OpenAI*. The court found OpenAI had infringed copyright because song lyrics used to train its large language models were memorized, making the models themselves copyright infringing works.

The decision stands in contrast to the U.K. High Court's ruling in *Getty Images v. Stability AI* (see our recent Osler update), where the Court found that Stable Diffusion's image generating diffusion model did not store or reproduce copyrighted works, but instead incorporated patterns and features learned over time during the training process.

A few factors may explain the divergent legal conclusions. Perhaps most significantly, the evidence before the German and U.K. courts of how prompts could elicit a reproduction of training works differed markedly. As well, the concept of reproduction is different between U.K. and German copyright law.

GEMA – the ease of prompting a memorized work

Central to the Court's ruling in *GEMA* was the ease with which the infringed works could be prompted.

GEMA, a collective society that represents music rights holders, sued OpenAI alleging copyright infringement in song lyrics. GEMA presented evidence of examples of simple prompts that caused OpenAI's model to reproducibly output identical or nearly identical copies of the song lyrics at issue.

The Court's view was that the resulting outputs established that lyrics were memorized within the model, and therefore the trained model itself contained reproductions of the protected works. It rejected the possibility that outputs were simply a product of chance or coincidence.

Importantly, the Court's analysis turned primarily on interactions with end-users, and the outputs produced, rather than the technical details of the underlying models.

Similar evidence of identical or near-identical copies of specific training works found in outputs generated in response to simple prompts, was not presented to the Court in *Getty*.

Reproduction as a physical fixation only indirectly perceptible to humans

The willingness to accept infringing *outputs* as evidence of an infringing *model* appears to be a product of the interpretation of "reproduction" in German copyright law, which defines a reproduction as any physical fixation that makes a work directly or indirectly perceptible to human senses.

In finding the LLM to be infringing, the Court stated that it is not necessary to isolate exact portions of the texts within a model to determine physical fixation.

It was sufficient that simple prompts, along the lines of "What are the lyrics of [song title]", reproducibly led the model to output copies or near copies of the lyrics at issue. The reproduction standard was satisfied because lyrics were embedded in a way that they could be identified with sufficient precision and objectivity.

This approach to reproduction is a second important contrast with the *Getty* decision. In *Getty*, the court's focus was on the data while stored in the model in the form of tokens and weights. The Court in *Getty* therefore concluded that model weights do not store or reproduce copyrighted works themselves but are instead a distinct form of record created during the training process.

The text and data mining (TDM) exception

While a TDM exception exists in German law, the Court held the exception was not available to OpenAI.

Open AI argued that the TDM exception permitted reproductions made for subsequent analysis. The Court agreed. However, it concluded that the TDM exception did not permit the later memorization—and therefore reproduction — of the works in a manner that affected creators' exploitation interests.

The Court acknowledged its finding may be a fundamental barrier for LLM businesses, admitting that if memorization could not be avoided in the current state of the art, then the TDM exception may not permit the use of copyright protected works.

Implications for AI developers

The divergent outcomes, as noted earlier, suggest there is no international consensus as to whether an AI model constitutes copyright infringement, and that courts may reach very different conclusions based on the evidence and the copyright regime at issue.

AI-related copyright litigation in the U.K. and Germany is likely to be of significant interest to AI developers, copyright owners and courts in Canada and across the globe. It may be some time before we have more clarity on how copyright laws, which are grounded in domestic legislation within each jurisdiction, are interpreted in relation to AI training and outputs.

In the meantime, organizations training AI models will need to consider a range of risk mitigation measures, including input/output filters, architectural design choices, and licensing options. Taken together, these measures can help reduce legal exposure and provide a practical way to manage uncertainty in an evolving legal landscape.