Knobbe IP+
PODCAST

AI and Copyright: Navigating the Legal Maze with Liz Rothman

Mark Lezama
Partner
Knobbe Martens

Elizabeth Rothman
Attorney and Advisor on
Emerging Technology

ML Welcome to this episode of Knobbe IP+. I'm Mark Lezama, a partner at Knobbe Martens. Today I'll be speaking with Liz Rothman, who's an attorney in private practice as well as the Senior Legal Affairs and Digital Trust Advisor for X Reality Safety Intelligence. Liz is a leading voice on legal issues surrounding AI, and I invited her on to discuss the intersection of generative AI and copyright and to explore the applicability of fair use in that context. Let's jump right in. Liz, welcome to the podcast.

ER Thank you so much for having me, Mark. It's great to be here, and I'm excited to talk about this topic.

ML Yeah, it's good to be talking to you again. You and I were co-panelists back in 2023 for an AIPLA conference, and I remember you gave this terrific presentation on emerging issues, emerging IP issues with AI. And there was the slide where you showed how, in the span of just a few months, from the end of 2022 to the beginning of 2023, the number of generative AI companies had just exploded. And so here we are two years later, and I guess I'm wondering, how do you see the generative AI landscape today? Because, on the one hand, I'm guessing there are a lot more generative AI companies today. But, on the other hand, it feels like in the—at least in the media coverage, the podcasts I listen to—the space has kind of coalesced around a handful of clear leaders.

ER Yeah, thank you for that question. That was a really fun conference, especially at that time, given the—how new to a lot of people generative AI was, and how many new products were coming out, and how much the number of AI companies were just sort of exploding. And there was a really big era of growth during that time. And I think most of us know that that has continued for the last couple of years, but that things are slightly different now. During that time period a few years ago, sort of hundreds of generative AI companies sprung up, it seemed like, over a matter of months. And it was a bit of a gold rush of who could claim ground for certain purposes, or claim to be doing things for the first time, or, you know, generate more high fidelity images or create text that was more human-like, or things like that.

And then, if we're fast forwarding to today, two years later, the whole conversation around AI has changed. Back then, I think people were not—it was just kind of getting onto some people's radar. Students were starting to recognize that this was a thing that perhaps they could utilize in the classroom. But it wasn't, really, you know, omnipresent in everything. And now it is the center of a lot of conversation. It's the center of our cultural discourse in a lot of ways. And the landscape today does look a bit different. On the surface, there are even more companies. There are thousands and thousands and thousands of AI companies now, but underneath it, there has definitely been some consolidation around a few dominant players. And those players largely right now are Open AI with ChatGPT, Anthropic with Claude, Google with Gemini, Meta with Llama, and there are a few others, especially in the image generating space. But really those are the foundation models that are really anchoring the development underneath.

So, a lot of what's happening today is happening downstream of those models. So we're seeing a shift from this model training to model integration and then really embedding AI into everything. So whether it's legal workflows or enterprise tools or creative software, it seems, you know, even places where I don't know that AI is particularly helpful, we're integrating it, and we're trying to figure out where it can be useful. And in a lot of cases, where money can be made in the situation, I think, is the reality of what's happening right now with AI.

And the truth is that the barrier to entry for training large models has actually gotten higher and not lower over time, because of the enormous computational costs involved. And so to catch up with any of these companies is going to be very difficult and take a lot of resources.

ML Sure. Yeah, I imagine the investment in infrastructure to kind of scale up is very high at this point to like you say, to compete with OpenAI or Anthropic.

ER Yeah, yeah, absolutely. And I think we're almost, we're getting to the point where we've narrowed down to these few players with a lot of technical expertise and the deep pockets to train the models. But that leads us to somewhere where we can no longer just sort of be fascinated with what the models can do, but we can start to really examine who the models are serving, what purposes they serve, at what cost, and who can, you know, benefit in society from these, and how we can regulate and look at these issues a little deeper.

ML Alright. Well, I think that that that idea that we're seeing a handful of leaders in this space lets me turn to my next question, which is—I want to turn to some of the copyright issues that are being litigated, and there have been quite a few lawsuits already, and some of the players you were just mentioning are naturally the defendants in these lawsuits. So you've got The New York Times and several prominent authors suing OpenAI and Microsoft for use of written works in ChatGPT. There are music publishing companies suing Anthropic for use of their copyrighted song lyrics with Claude. Getty Images and visual artists have sued Stability AI for use of their photos and art with Stable Diffusion and DreamStudio. Record labels are suing Udio and Suno for use of copyrighted musical works and sound recordings. And that's

just a few of the lawsuits out there, but those are some of the big ones. And of course every case is different and has its own unique facts, but I guess I've seen at least some common themes, and I'd like to get your take on it.

First, in each case it seems like there's a claim that copyrights were infringed at the training stage, when the generative AI company made unlicensed copies of the works to train their models. Second, a lot of the cases, but not all of them, also include a claim that copyrights have been infringed a second time at what you might call the "output stage," when the generative AI tool presents an output that arguably or allegedly is similar to the copyrighted work or a portion of that work. And then third, and this is kind of a smaller category, I would say some of these cases also include some allegations of DMCA violations, often around omitting copyright attribution or copyright management information. So let me just pause there. First of all, does that fairly capture most of the issues being litigated?

ER Yeah, I think your summary captures the main categories very well. Right, so we have the infringement at the training stage, the output stage infringement. And then these violations around attribution as well.

There are a few other areas that I think are quite interesting. As we examine the technologies more and get a clearer grasp on exactly what the training entails and how companies are using other creative works to train these systems, and also how those architectures are evolving, I think that there are some interesting issues that come up. One is the role of something like RAG, which is Retrieval Augmented Generation, where you're actually pulling from external sources during the generation process. So you're not just relying on what's been trained into the model initially, but you're kind of grabbing new sources in real time, and sometimes either from the web or other repositories. And this means that you could have potentially fresh copyright exposure every time that you're running a query, depending on what's retrieved, how it's displayed, and whether attribution is preserved in a variety of ways.

The second one that I think we're going to start to see more of the conversation revolve around is things like reinforcement learning with human feedback, where the—it's a technique used to make models more aligned with human preferences. So we have trained these models on all of the things that we—we've trained them to mimic us, to be helpful assistants to us, to mimic our culture, to do all these things. But sometimes the models aren't performing exactly the way that a human would, or the responses aren't exactly what a human would respond. So this is a technique where human beings are actually scoring the output of a model to show if it is what a human wants to see, if it's something that is along the lines of what we want for the output of the model. So if human labelers are reinforcing AI responses that closely mimic copyrighted works, or books, or lyrics or scripts, or whatever it is, that could be another way, that the model is further relying on copyrighted expression, perhaps. And so that's something that I think we might start to see in the future here. And it's not "training" in the classical sense of these—training these large foundation models with scraped repositories of books, and all of this. But it is another, maybe more curated, way of training that could come up.

And then, the last area that I think—it's a collection of areas I think that we're starting to see more cases around—are publicity rights, voice likeness, things like that, where the models are producing content that can really compete in the market with the original content.

ML Gotcha. Yeah, I think that those are all definitely very interesting areas. And let me go back to the first one for a moment. So just to make it concrete for our listeners. So this this could be, for example, I'm typing in to ChatGPT, “Tell me what's going on with the tariff situation today.” And so ChatGPT might be relying kind of on the base model, to generate a portion of the output, but part of it might be looking at very recent—basically searching the Internet—for articles and maybe summarizing that in the output. That's part of what you're talking about?

ER Yes. Yeah, absolutely, anything that is retrieved kind of in real time in that search. And whether it's in repositories that are set aside for this purpose that whatever company has curated and left to the side, or if it's actually searching the web and pulling data in real time.

ML Gotcha. And do you know whether we've—that sort of stage of the output has come up as the basis for an allegation of infringement in any of these lawsuits?

ER So there was a case I'm now blanking on the name of. It was against Perplexity AI. It might have been The New York Times. But I think that that was a case based on retrieving large amounts of published articles and materials from, I'm going to say, The New York Times for purposes right now, but it is from a newspaper, and that was based at least loosely on RAG systems.

ML Gotcha. Alright, let’s turn back to the first two kinds of claims that I mentioned: the alleged infringement at the training stage and the alleged infringement at the output stage. So assuming that copyright holders never authorized any use in these cases, how do you think about these claims? What's your take at a high level?

ER Well, I think both of them present very interesting challenges, right? And they're very different. So training stage claims really hinge on this question of whether using copyrighted material as input for training is transformative, in a lot of the cases, under the fair use doctrine. And courts are still—there's so many cases going on—but still undecided on that issue for this purpose. And there are a lot of arguments in favor of AI companies, a lot in favor of the rights holders. But I think we're going to have to see, and it'll be a more nuanced discussion than whether it's there's just a broad exemption for it. I don't think that that's going to happen. I think that we’ll have a really nuanced conversation around what kind of data is used and for what purpose for those cases. And then the output stage, especially where outputs really closely mimic something that is a protected work, it can look more like classic infringement. So particularly if there's no meaningful transformation, or the output is commercially exploited, and it looks very, very similar to something that could have been input into the system, then I think those cases are a little, can be a little bit more clear-cut, but obviously very fact-intensive.

ML Okay. So yeah, I definitely agree that on the output side, the facts can vary widely. Thinking about the training question though, at least at a certain level, the question seems to be very similar. At least as far as the allegations go, right? Obviously the facts might be different than what's alleged. But at least on a certain level, it seems like in all these cases, or at least almost all these cases, there's an allegation that copyrighted works were copied wholesale without authorization for use at the training stage. And so, I'm curious whether you think—whatever the outcome ultimately is, whether it's infringement or non-infringement, I'm not asking you to take a position on that—but do you think as a sort of theoretical matter I guess, should the answer be likely the same across all these cases? Or do you think there are real differences in each case that could materially affect the analysis?

ER Well, yeah, I think that you're correct, that the training claims are a lot more uniform. And like I said, there's some, maybe newer systems and models that could come in and training mechanisms that may change that. And there might be more nuanced analysis going on with those. But yeah, essentially, every generative AI company has done large scale copying to create their training databases or "training with the data," if you don't want to call it copying. They have input this data into their systems. But the claims face a hurdle where the copying is an intermediate step, and it doesn't result in the public distribution of those copyrighted works in their original form. So the argument is that the training process has transformed the works into these mathematical representations.

But while the basic training questions are similar, I think there's important distinctions that could, as you suggest, change the analysis. And the nature of the copyrighted works would matter in that case. Using news articles for training might be analyzed differently than creative fiction or visual arts. The training methodology also matters, so copying indiscriminately with web scraping, or it wasn't curated and selected and used for a certain purpose. The purpose and character of this could vary a bit, too, and a model trained to compete with original works might face different analysis than one that's trained for transformative purposes.

That just reminds me of, there – this was many years ago – but in like 2015 or '16, there was The Next Rembrandt project. Not something that created a copyright issue at the time, but was actually just trained on Rembrandt's work to try to output the next Rembrandt. And so, if you're really training on curated material, that could make quite a difference as well across the cases.

ML I'm Mark Lezama, partner at Knobbe Martens. And this is Knobbe IP+. We're talking with Liz Rothman about the intersection of generative AI and copyright.

And so just a moment ago, we were starting to get into fair use. And so I want to hold on that a little bit. So just for our listeners, in case anyone's not familiar, fair use is a pretty typical defense in copyright litigation. The gist of it is, even if I otherwise infringed your copyright, I shouldn't be liable, because my use of your work is fair game in some sense. And maybe a quintessential example might be parody where I might have blatantly copyrighted—might have blatantly copied your copyrighted work, but the

argument is that my new work is parodying yours, and that's something that society wants to promote. And so, even though there was clear copying, we're not going to create liability for me, at least under copyright infringement. And deciding fair use involves looking at a host of factors that often play out in the facts specific to a particular case. But even though there is that sort of variability from case to case, Liz, how are you seeing fair use arise in the generative AI context?

ER Thank you for that question. Yeah, fair use is really central to these cases, and it's playing out in some predictable and some surprising ways. Defendants are really—have been, at least until this point—primarily hanging their hats on the transformative use prong of fair use. That they are just—there was a real attitude of "move fast and break things" I think, when these models were being trained. And they are really pushing the point that this was fair use to train on all of these materials, and that's the end of the sentence. And I don't know, now we're getting into more nuanced, as these cases go forward, we have to get into more nuanced discussion of what that means.

But they're basically arguing that the AI models transform copyrighted works into something fundamentally different, and that there are mathematical weights and parameters that enable a new creation, original content. And it's not reproducing works, but learning from them the way that a human artist would, or that anybody else studying it would, and then developing their own style. On the other side though, what we're seeing is that rights holders are saying, "Wait, this isn't as transformative as you're saying because some—a lot of the outputs are mimicking human output and really even imitating some of the original works."

So there's many different ways that these are being approached in court cases. And I think, there's a big fight over market harm here and how much that's coming into play. Creators arguing that the tools can replace their works in the market, whether it's art or music, or summaries and things like that that we've talked about. And then the AI companies on the other side saying that these are distinct. So it's a kind of back and forth. And we're going to continue to have these arguments, I think, as these court cases progress a little bit more. But it's important to note what really sets these cases aside from traditional ones is the scale and the automation of these, of the systems. So traditional fair use analysis, which was really developed for human-scale copying and commentary, and it has a hard time adapting to this kind of scale. And courts are looking at whether principles developed for cases involving dozens or hundreds of works can apply, when we're talking about training millions and billions—training on millions and billions of copyrighted works.

ML So not that long ago, there was a decision by a Third Circuit judge, sitting as a trial judge, in *Thomson Reuters v. ROSS Intelligence*. And this was a case that included a decision on fair use that I think may have been surprising to some of the AI companies. So tell us a little bit about this case.

ER So interestingly, this case was really instructive, even though it kind of predates the current wave of generative AI litigation. This case created – ROSS Intelligence, in this case, they created a legal research tool to analyze and summarize legal documents,

and Thomson Reuters alleged that ROSS had improperly copied or used content from their head notes in *Westlaw*. And the facts were that ROSS, through a third party, had scraped large amounts of legal content to train their system, and the tool would then provide summaries and analysis that Reuters said were substantially similar to their copyrighted headnotes and summaries. And the court emphasized that ROSS's use wasn't transformative enough. They were essentially using Thomson Reuters content to create competing legal research products, and they found that the commercial use had harmed the market for the original works, and ROSS really hadn't sufficiently transformed the copyrighted material.

So it's very interesting, because they also made a point to say that this was not applicable necessarily to generative AI cases. And when a model is trained on copyrighted material in any other sense, so the generative AI cases, there's a lot of the same issues that do arise for those cases. So we'll see how much of this will actually be applicable in the future to when more decisions come down on the current cases involving generative AI. But when models are trained on other copyrighted material in the same way, so news articles, and then they're generating news-style content, or if they're doing creative writing after being trained on creative works, you have the same basic dynamic where you're using copyrighted content to train a competing product in the market.

But the main distinction that I would see from this case than involving generative AI systems currently is that the generative AI systems now may have a stronger argument of transformation because of their argument that they're creating genuinely new content, and that they're not just reorganizing and summarizing what was already there. So I think that the core questions, though, are kind of the same, of how much was taken, how it's used, how does it compete? They're all kind of the same, but it'll remain to be seen how much this is applied in generative AI decisions, even though they said that it shouldn't apply, or that it doesn't necessarily apply.

ML Yeah, so I want to hold on that for a second because I mean, you anticipated some of my questions there. It was very interesting. The judge initially ruled that fair use would go to a jury. And then he changed his mind, and ruled that, on summary judgment, that as a matter of law, ROSS had not met its burden to show fair use, and this wasn't going to go to the jury. He could just decide it as a matter of law. I guess, on the training side of things, it seems to me like the questions are quite similar to generative AI, right? Whether the use is transformative—as you were pointing out, in the generative AI context, there might be more transformative use. But I take that to be more relevant at the output stage. And so to the extent you're arguing that there's copyright infringement just at the training stage by making these copies for purposes of training these models, it seems to me that the Thomson Reuters decision's reasoning could apply quite closely to generative AI case. What's your take on that?

ER Yeah, I think absolutely. But I think that the question will be a little more nuanced. So in this case, they decided that there was enough evidence there. The judge decided,

"There's enough evidence here. I don't even need this to go to a jury. This is clear – clear copying of this information. And it was close enough."

But in the generative AI case, I think it can be a lot more complicated because of the amount of data that is trained in the systems. And so you don't have this clear, you know, pathway. Like there, there was a very clear pathway as well from they actually used that data, and they were creating a competing product on the market. But I think it's just – it'll be a much more nuanced conversation when talking about specific models, their use cases, what has been input, what's output from the models, and how that affects the human creators.

ML Yeah. So picking up on the market factor of fair use. Certainly we've seen arguments that the generative AI models' use doesn't necessarily affect the market for the original work, and that favors fair use. But it seems to me that copyright owners can easily argue that any unauthorized copying for purposes of training these models or in the output, interferes with their ability to market their work for training or quotation purposes, kind of like the RAG concept you were talking about earlier. And we've seen OpenAI entering into deals to license content from various publishers like Washington Post, Hearst, Conde Nast. I think The New York Times recently entered into a deal with Amazon. It seems like those moves would strengthen that argument, that the market factor weighs against fair use. At least there's an argument there. What do you make of that?

ER Well, see, I think it's interesting, because some of these licensing deals are a real double-edged sword, right, for the claim of fair use. And they may actually, ultimately really strengthen the copyright holders' perspective on what the market is in this case. So some of the deals—it seems that AI companies are recognizing the value of the content that they are licensing, and in that case they're willing to pay for it. And they have in many cases. And we've seen a lot of deals, you know, coming up between news organizations and AI companies and licensing among artists and large databases being collected to use licensed content to train new models. Interestingly, some of those licensed models, where they said that everything was all licensed content, there was found to be some images from Stable Diffusion and other things. So it's a little bit, you know, murky, even at that point. But it is—these deals are important, I think, and it undermines the argument that there's not commercial value, and that it should be fair use to utilize these materials.

From a policy perspective, I think it's interesting to look at it, because the deals also suggest that maybe the industry is moving towards permission-based systems rather than just relying on fair-use exceptions. So there could be a lot of different ways that this kind of plays out as the litigation plays out. And there's a world where companies are licensing all of this data, and where maybe we preserve the fair use exceptions for things that are non-commercial or research or education, but there's a different standard that's being held for AI companies.

ML Alright, and so before we go, I just want to ask you one last question. So to a certain extent the genie is already out of the bottle. Use of copyrighted works to train AI models

has already occurred on a large scale, like we've been talking about. Do you think the lawsuits under existing copyright law can provide effective remedies? And if not, what might need to change?

ER Yeah. So I think the existing copyright law in general faces real limitations in addressing the scale and nature of AI training. Traditional copyright frameworks were built for, again, human-scale copying, distribution, artistic endeavors, not for the ingestion of billions of works to create these models. So the remedies under current copyright law are pretty blunt often, right? It's an injunctive relief to shut down an AI model entirely, possibly, which in our current environment, would seem pretty drastic. A couple of years ago, if you said that, people would be like, “Oh, shut down one AI model. That's not a big deal.” But if you try to shut down OpenAI right now, you are crippling millions or thousands and thousands of startups, and there'd be big implications of it. And perhaps—this is something that would have to be borne out with market research—but perhaps that would be disproportionate to some of the harm in cases. But we do know that these models were trained in a manner where they did not—they were scraping everything on the web to train them.

So the other thing that the law can provide are monetary damages. And to calculate that when you're dealing with the statistical use of millions of works, apportioning damages for any single AI output might be really not impactful for the person that is receiving that remedy. So I do think that we really need to think about this kind of, perhaps a nuanced approach going forward where we have a couple different rails of looking at these for big AI companies and for licensing regimes that will need to be established. And then, perhaps, some of these exceptions where there are researchers and truly transformative uses that could fall under current fair use exceptions.

ML Well, that wraps up today's episode. A big thanks to our guest, Liz Rothman, for joining us today. Be sure to visit Knobbe.com to listen to or view written transcripts of this and other episodes of Knobbe IP+. Until next time.