



**And The Oscar Goes to ChatGPT
Interview with Paul Sweeting**

**For podcast release
Monday, May 8, 2023**

KENNEALLY: Writers in Hollywood are the latest to declare concerns that technology based in powerful artificial intelligence tools may jeopardize their livelihoods. As of this recording, members of the Writers Guild of America are on picket lines in Los Angeles, largely over worries that studio bosses will use ChatGPT to write jokes and dramas.

Welcome to CCC's podcast series. I'm Christopher Kenneally for Velocity of Content.

When the Writers Guild voted to strike on May 2nd, the union demanded of producers that AI can't write or rewrite literary material, can't be used as source material, and that contract-covered material can't be used to train AI. Questions about the role of generative AI technology invariably focus on intellectual property law, both in the so-called training of large language models like ChatGPT and in the output of works, including text, images, and even videos.

Variety has just published a special report on gen AI and IP law, and the report's author, Paul Sweeting, joins me now. Welcome to Velocity of Content, Paul.

SWEETING: Thanks, Chris. Good to be here.

KENNEALLY: Are you surprised, Paul, that AI and authorship is a central issue in the WGA strike? ChatGPT only became a household word months ago. Did that anxiety erupt overnight, or has this been coming to a slow boil for a while?

SWEETING: Well, first of all, I'm not surprised at all that it would be an issue in the strike. It's sort of the issue in just about all of the creative industries at this point. It's hit everyone like a ton of bricks. But as you suggested, it did just become sort of a popular phenomenon with the release of the ChatGPT bot, which happened in late November of last year, and it very quickly took off. I mean, it had 30 million users within the first month, and it now has over 100 million users. Since then, other similar, related tools have become available commercially – Stable Diffusion and Midjourney, which are used primarily for generating images from textual prompts.

But all of these have been sort of percolating in the background for a while. ChatGPT is based on a model that was actually developed two years earlier, and the company behind it,



OpenAI, had made it available only on a selective basis. What's changed was late last year, for reasons they never really articulated that I've heard, they decided to commercialize this technology in a big way. So they sort of combined the model that they had built with a interactive chatbot so that people can put in prompts and have it respond and released that to the public. And it became overnight a phenomenon. But it's like they say in the music business – it takes 10 years to be an overnight phenomenon.

KENNEALLY: Well, let's look at this first from the writers' perspective. What is it about generative AI that has sent them to the picket lines?

SWEETING: Well, there are a number of concerns. The primary concern is that this technology can be used – and in fact, a lot of writers are themselves using it at an early stage of the process for ideation, the development of ideas, as are artists working in other media. But the concern for the Writers Guild is that studios will rely on this sort of technology to generate the basic script for a movie – the story, the characters, the basic dialogue – and then just hire writers on a basically day labor basis to come in and punch up what the machine has created. It would significantly devalue what writers do and would turn that sort of work – or their concern is that it would turn what they have been doing for a living for years into a form of day labor, basically.

KENNEALLY: What about the studios, Paul Sweeting? Do they see generative AI as a real opportunity for them?

SWEETING: Well, if you believe the writers, (laughter) it's an opportunity for them to save money on writers. But generative AI is already playing a role within the movie production process, and it's only going to play a more prominent role as the technology advances. It's possible today with the technology that they have to create an entirely synthetic performance by an actor. You can have an actor appear in a scene – or appear to appear in a scene – without that actor ever having been on the set using previous footage as a sort of starting point and then using AI to create an entirely realistic-looking performance by somebody who was never in front of the camera. And it's also, of course, being used in special effects and all sorts of areas.

It's also being used sort of before the production process has begun and has been even before these current generative AI models came on the market. Studios have been very quietly using AI to essentially test-read scripts and asking the AI to make predictions about its commercial possibilities or potential and even weighing in on casting and various other aspects of it. So AI has been there for a while in Hollywood, and it's only going to become a more prominent part of the production process and the pre-production process as the technology continues to advance, which it is at lightning speed.



KENNEALLY: Recent decisions by the US Copyright Office do make for uncertainty about the value of these AI-generated works, because the Copyright Office has ruled they cannot be copyrighted. And the basis for that decision is the black-box nature of tools like OpenAI's ChatGPT. Help us understand that issue.

SWEETING: There are a number of reasons why the Copyright Office has taken that position. The fundamental reason is the longstanding principle, which has been upheld by the Supreme Court many times, that human authorship is required for a work to be eligible for copyright. The whole purpose of copyright is to incentivize humans to create more works by providing them a means of monetizing that. You can't incentivize a machine. It's a non-sequitur. It makes no sense. So the office has long held, and the courts have long upheld, that human authorship is the sort of sine qua non for copyright eligibility.

Where things get really tricky these days is exactly how much human authorship or human involvement and of what type is required for a work to be eligible for copyright. Because as these generative AI tools become a bigger part of the process by which artists create and writers write, where do you draw the line? What exactly does the human have to do, at what stage, to invest the work with the requisite amount of human authorship, whatever that is? That's where the black-box problem comes in.

Because these tools – there's an inherent unpredictability in a generative AI model. These models – the way they're created is they suck up an immense amount of content. ChatGPT, for instance, was trained on something on the order of 45 terabytes of data, which is basically the entire textual content of the publicly available World Wide Web. But what it does when it ingests all of that material is it doesn't copy it. It essentially reads it the way you and I would read a text, and it extracts from the text various data points – a whole lot of data points – about where words appear in sentences, how often they appear, how often they appear in relation to other words, and how similar or dissimilar they are to other words.

It uses all that data to create an immensely complex statistical model of language. It's estimated, for instance, that GPT-3, which is the sort of engine of ChatGPT, built a model on 175 billion parameters. So it's effectively impossible for a human, including the people who designed these things, to know exactly how the system built the model and then what is the model doing? It's engaging in a highly complex, very involved mathematical process. But exactly what that computational process is is effectively unknowable. It's a black box.

So when you put in a prompt, there is a certain unpredictability as to what the output will be. And that raises a question about is there human agency in there, and is there a sufficient amount of human agency? If I can't predict with great accuracy or certainty



what is going to come out the other end from my prompt, can I say that I caused that output? Was my agency responsible for that output? That's where things have gotten very messy right now, because the Copyright Office has issued some guidance around that, but there's a fair amount of ambiguity in the guidance. So we don't have an answer right now is the problem. It comes down to how much human authorship is required, and can you isolate that human authorship from the process when the process involves the use at some point of a generative AI model?

KENNEALLY: Paul Sweeting, if these new AI works cannot be copyrighted, what's the potential impact on the value of movies and music? As you say, AI is quickly becoming established in the workflows of many types of media. So they are there. They are using the AI. What's it going to mean to the business?

SWEETING: Well, that's a very good question, Chris, and I wish I had a firm answer for you. (laughter) But it's hard to say. I mean, these things – the scale at which a generative AI model can churn out content is staggering. It takes only a few seconds to generate a song, for instance, or a poem. And it can do that almost infinitely.

It's already happening. The amount of this sort of ambiguously owned content – there's a real danger that that's going to sort of flood the market and crowd out works that were genuinely created by humans. If you're a graphic artist, and you spend two weeks trying to come up with something for an advertisement, and a generative AI can produce 4 million of those in a matter of minutes, what is the value of the work that the human artist has put into it? It's a real problem, and there isn't a good answer to it yet. This is very, very early days with this technology in widespread use, and everybody is just sort of feeling their way in the dark.

KENNEALLY: Paul Sweeting, author of the new *Variety* special report on gen AI and IP law and co-founder of the RightsTech Project and editor of the RightsTech blog, thank you for joining me today.

SWEETING: My pleasure, Chris.

KENNEALLY: That's all for now. Our producer is Jeremy Brieske of Burst Marketing. You can subscribe to this program wherever you go for podcasts, and please do follow us on Twitter and on Facebook. You can also find Velocity of Content on YouTube as part of the CCC channel. I'm Christopher Kenneally. Thanks for listening.

END OF FILE