Artificial Intelligence

Interview: Dykema's Michael J. Word Provides Clarity On Intellectual Property, AI Voices

By
Michael J. Word

Dykema
Chicago, IL

LexisNexis®

Commentary

Interview: Dykema's Michael J. Word Provides Clarity On Intellectual Property, AI Voices

[Editor's Note: Michael J. Word, a Dykema attorney in Chicago, specializes in intellectual property, technology and related litigation. His profile can be found at https://www.dykema.com/people/michael-j-word.html.]

Mealey's Litigation Report: Artificial Intelligence spoke with Michael J. Word after OpenAI Inc. recently found itself in trouble when people noticed that the voice it added to its ChatGPT artificial intelligence sounded like actress Scarlett Johansson. The trouble deepened when Johansson noted that she rebuffed OpenAI's attempts at licensing her voice for that very use. In a statement, Johansson said: "I was shocked, angered and in disbelief that Mr. [Sam] Altman would pursue a voice that sounded so eerily similar to mine that my closest friends and news outlets could not tell the difference. Mr. Altman even insinuated that the similarity was intentional, tweeting a single word 'her' — a reference to the film in which I voiced a chat system, Samantha, who forms an intimate relationship with a human." Johansson hired counsel, and OpenAI eventually removed the "Sky" voice from ChatGPT. In her statement, Johansson noted, "In a time when we are all grappling with deepfakes and the protection of own likeness, our own work, our own identities, I believe these are questions that deserve absolute clarity."

Mealey's: I'd like to start off with an overview of the key legal issues that surround the use of AI to recreate someone's voice.

Word: Yeah. For a lot of the AI issues, there's a whole body of case law that is analogous. So you know a lot of times you hear these concerns that the law is behind.

People talk about needing new legislation to address AI issues, and a lot of times that's jumping the gun a little bit just because these issues aren't necessarily new, right? There are people laying claim to their voices and people trying to pass off others. It's famous voices for promotional opportunities. It is an old issue there.

There's a very famous case that goes back to Bette Midler, the Midler v. Ford [849 F.2d 460] case back in the 1980s, late 1980s. Basically, they tried to get Bette Midler to perform one of her songs for some promotional materials for Ford, and she turned them down. So they hired a voice actor with a very similar voice to sing. And Bette Midler said you can't do that. You're basically trying to capitalize on my famous voice to sell things. And Ford got the hammer brought down on them. So there's a lot of similarities to what we're seeing here.

There is a body of case law that's out there that's been developed for protecting rights to your likeness, your face. There's a patchwork of state laws that provide those types of protections, and by and large, most of them protect voices of celebrities. And individuals as well. And so there are remedies to be had out there for people who are claiming that, hey, you've basically taken my voice and used it.

We start to get to the thornier issues here where we move away from a celebrity to individuals. People go in and voice actors who think they're voicing some individual lines and really their voice itself is being sampled and then being used to create a whole AI assistant or capabilities to say things beyond what they may have just recorded.

Mealey's: So one of the one of the questions I have is in a lot of these suits involving AI, it involves something tangible; it's artists, it's authors bringing the suit, people who can point to works the AI is allegedly reproducing. Discovery can show that they used that work to train the AI. How does that work with voice? While the voice may sound similar, I'd imagine it's very easy for a company accused of something to just argue, "Hey, we didn't use Scarlett Johansson's voice. We used this other actress that we found or it's a conglomeration of two or three voices." Even if the intent was eventually to come close to Scarlett Johansson's voice, for instance.

Word: It's interesting, right? Because the analogy starts to break down a little bit between kind of visual arts and voice arts. Because AI is a tool, right? So if I create an AI model for generating paintings, and I go out and I train it on all the art that's out there and I say, "OK make a painting in the style of this person." The analogy is: AI is just a tool, and if I have a really good artist out there and my artist goes and studies all the paintings out there and he's really good at making reproductions and he just paints something that happens to look like the style of an artist. Is that wrong?

That's kind of the argument with respect to visual arts, and there's good arguments on both sides of it. It comes down a little bit to intent, and it also comes down to what's being done with the final work of art. Is the person trying to pass off the AI-generated visual art as being this individual's? That's probably not OK. But if it doesn't have a commercial use or you know there some other type of reason for doing it, maybe it is OK.

That's the visual space.

When you get to the "voice space" we can call it if I'm sampling your voice, I'm not sampling it to . . . create a new work of art necessarily, right? I'm using it to be my voice assistant. And it's basically a reproduction. It's this person's voice saying things or something that sounds similar.

So that's kind of where the analogy breaks down, right? You can't say: "Well, I was just doing this in the style of somebody else." And their style is out there for the world to see, and that's OK. Now we've got

something where it's an AI assistant, and you know you're passing it off basically. Or trying to free ride in a more significant sense than in the visual arts. And that's where it becomes a real problem because there's less of a kind of line.

And is it OK?

Is it not OK?

If it really does sound like the person . . . well, there are degrees, right? Did I actually use this person's voice as a basis for [the AI voice], or am I using a sound alike? And then in that case . . . the intent gets a little bit messy. But if you're actually using the person's voice, it's pretty clear cut, right? You shouldn't be able to be doing that absent permission to do that. If you're using a sound alike, then you start to get into more subjective issues. Does it sound like the person, or doesn't it sound like the person? And then it gets to questions of what are the scope of the person's rights.

Like Scarlett Johansson. Certainly Miss Johansson has a right to prevent other people from, under a variety of laws, [taking] her voice and then [using] it and pass[ing] it off directly. There are a variety of laws, state laws that would cover that. Well, that that's kind of in the the the eye of the beholder, so to speak, right?

Does it really sound like this person? And how close does it have to be to potentially infringe on our rights? You know, we're moving away from the Bette Midler case. Because now you say, "I'm sorry that Miss Johansson thinks it sounds exactly like her, but that wasn't our intent. Or, point to 50 people over here that say it doesn't sound like her versus her and her friends who say that it does.

Mealey's: Right. So let's say I run an AI company, and I wanted to have a voice very similar to Scarlett Johansson's and I have this collection of audio collected from wherever — the internet and various sources — and I just asked the AI to produce a voice that sounds similar to Scarlett Johansson's. So the resulting voice sounds like hers because I asked the AI to mimic that voice, but I didn't use her voice per se.

Word: Yeah, that's a really thorny issue. I think the problem is how do you discover that, right?

Mealey's: How does discovery even work in this kind of situation?

Word: Again, you look at the variety of laws out there; they provide protections. You can't sample. You can't say, "OK, I'm just using your voice and now I'm creating something just like it." That's essentially your voice, right? It is your cadence. You get to questions of degree on something like this.

Part of it's going to depend on what the end result is. And you're going to get into that thorny issue of people fighting. Well, hey, that's my voice. And it's going to be a tough call.

You'll get 50 people [who] will say, "No, that person says their A's differently" or "They've got a drawl" or it's a slightly different accent, higher pitch, lower pitch, more feminine, more masculine, etc.

Mealey's: Right.

Word: So could you train on your AI model on some generic voices out there? A British accent? Like the British accent of Jarvis from Marvel [movies]. [Actor] Paul Bettany voiced that. Can he lay claim to all British male AI assistants for that? Probably not. You know, unless it's really spot on, it's really his voice. It'll be a tough call, but then that gets us into discovery, right? The intent, the commercial use some of these other factors that tend to come into play if I'm using it to spoof . . . somebody's voice.

OK, maybe that's fair use. I'm not collecting any money from it, but if I'm building it into my AI model, and I'm going to go out and commercialize it, and I trained it on somebody's specific voice and said make it almost identical to this, it's really going to be fact intensive and discovery intensive. And that is why the OpenAI issue with Miss Johansson is very interesting because you know the facts are coming out. We may never be privy to all the facts, but we know there was an outreach by OpenAI, and according to Miss Johannsen she turned it down both times.

OK, that's an interesting fact. But OpenAI is saying it already trained these models or hired these separate voice actors long before and didn't end up using her voice. There is probably some more facts behind that, but you know, how do you discover that? You could

imagine a set of facts where there's nothing there. It is just an individual voice and everybody kind of liked the sound of her voice.

Well then, how do you go about proving it other than the subjective point of view of well, it sounds like it versus it doesn't sound like Ms. Johansson?

Mealey's: I want to come at it from a little bit of a different angle. Scarlett Johansson already employs an army of lawyers and is kind of ensconced in the AI issue. It may be easy for her to fight a fact-intensive case. If she wants to file suit, she has the resources. What about at the individual level, someone who is just doing a podcast, who doesn't have those kind of resources? What kind of protections do they have, especially in a case that, as you said, is going to be fact intensive and is going to require some work to even get to discovery?

Word: It's interesting because the scope of protection is going to vary based on how recognizable the voice is, right? If I'm just a podcaster, most likely my scope of protection is narrower, but at the same time, if I'm a podcaster, even a relatively well known one, and I suspect that somebody else is using my voice, my suspicion may not help you very much.

There's going to be some suspicion.   So if I'm a podcaster and I have a reason to believe somebody's using my voice for promotional use, AI use or duplicate podcast, I've got 1,000 users or maybe 100,000 users. If this claim is being brought to my attention, there's got to be a reason for it, right? There are lots of people who sound like each other.

A lot of times where there's a suspicion, there's got to be some sort of reason behind it. It's going to be what? What is the reasoning that drew the attention to it? Was it the fact that you had prior contact with this other individual? OK, well now you've got some sort of fact that's going to make it more likely that something wrong is happening here. Are they a podcast operating in the same space as you?

You're right, it's going to be very hard to get your foot in the door and have enough reason to believe that there's a complaint out there. But you know we talked about discovery.

We are lucky that we have the discovery rules that we do in the U.S. courts and in the state courts because you can get to the bottom of that. It's a very, very powerful tool for these types of cases because if you were doing something wrong, you know that there's a likelihood you're going to get found out, right?

There are lawyers that work on a contingency fee basis that would be willing to hear your case. It all goes back to what's your basis for believing this to be the case. Even attorneys on a contingency fee basis aren't just going to take everybody's case. There's got to be a good reason for it, and if you have a good enough reason for it, especially to believe that it's happening, then someone out there will take your case.

Mealey's: But there are all of these devices out there. The app we are using right now for this interview is probably storing our conversation a server somewhere where our voice can be analyzed. And so much of AI is a black box.

Word: Yep.

Mealey's: It would be very easy to try to avoid discovery by being like, "Well, we put these voices in and this is what the AI spit out and you know, it's hard to say why it did what it did."

Word: You know in that case there's the user agreements, right? You know what you're signing up for. We may use your voice.

Mealey's: Yeah, I did want to get into that.

Word: But at the same time there may be user agreements to say, "We don't use your voice; we don't use it to train things," and then in that case, whistleblowers obviously become such an important part of that. That area of law has just in recent years really increased in terms of the protections that are afforded not to just government whistleblowers but private company whistleblowers. The incentives that can be paid to these individuals under variety of state laws are so great and the protection so great now that it really does promote that kind of protection for the average individual. So that's a useful tool to prevent the types of wrongdoing or misuse of voice data, along with other biometric data by private companies. So, it's

not as if these companies could necessarily just get away with it or there's no way to find them out.

The companies have to be worried about that. They have internal employees who are incentivized to call out such wrongdoing.

Mealey's: There are some states that forbid contracts of adhesion, that prohibit on some level one-sided contracts where the other side has no chance to negotiate the terms. Does that apply and protect anyone in a situation where a tech company updates its terms of service and inside paragraph nine, 7,000 words down, they've added a clause allowing them to record a user's voice?

Word: The . . . fights over this are kind of well known, right? You know, there's a lot of factors at play. I don't have to use Alexa; there's a number of different voice assistants out there, and it really is kind of voluntary. So even though Amazon says, "Here it's take it or leave it," with their agreement, I'm still availing myself of it; I still click through and say, yeah, I agree to all that stuff.

You know you can fall back and say, well, I didn't mean that I didn't allow you to use it, you had all the power. But you knew . . . what the agreement is.

We notify you, and you can stop doing it at any time. That whole rubric for the types of agreements that we're talking about and user license agreements makes it difficult to make that argument.

And the exceptions, even in the state laws, aren't as broad as we like to think. Now at the same time, you know if it's a significant injury or significant issue, then you know those types of adhesion contract protections or prohibitions against that are stronger. If I'm an attorney and I'm representing a client, . . . I'd much rather have a written contract and prove that the party acknowledged it then a client saying, "Well, yeah, I agree to that contract, and I kept using their services. But I didn't really mean to agree to it."

Mealey's: The path forward for those protections for voice and IP — is that in your mind more of a tort avenue or is it more a regulatory avenue? And relatedly, I know there are a bunch of states that have enacted various biometric protections. Is it

something that can be solved state by state, or is there going to have to be a federal solution?

Word: That's a really good question. Do we need something at the federal level? There is this patchwork of laws in the U.S., and you these demands all the time that the Congress needs to step in, that we need to have this, like a singular law or whatever to address all these instances. I think if you look at it, my sense is that individuals probably benefit more in a lot of ways from kind of the patchwork of laws as they are right now because a lot of the companies that . . . might be using this information . . . or might be wanting to use voice data aren't necessarily local; they operate across the U.S., and it's kind of the most strict regulations that are out there, they're going to be the ones that they're going to tailor their operations toward because they want to be able to operate in all the different states, and it becomes quite a problem for these companies to make sure that they are complying with the very stringent laws in California.

On this issue, the stringent laws in Illinois regarding the collection and use of biometric data have spawned millions of lawsuits. So, you know, even Tennessee passing that [law], the companies need to operate within the confines of that if they're doing business in Tennessee. And if they're not, they're subjecting themselves to liability. I suspect that a lot of the companies . . . or a lot of the voices that are pushing for kind of a centralized federal approach to this that would impose a regularity across all the different areas are probably the companies that would want to use the data. They want to have the most types of freedoms to do it because then they only have one standard to meet and can actually eliminate some of these more stringent state laws if possible, and under the guise of trying to get everything to be more uniform.

I don't really see the patchwork of individual state laws necessarily as being an issue. It is kind of like a rising tide lifts all boats situation where something that's pretty stringent — California or in Tennessee — is going to benefit people or have benefits for people in other states.

Mealey's: Can you speak a little bit about what damages would be available? I imagine it might be scattershot across the landscape, but what dam-

ages would be available to someone like a podcast for example?**

Word: It will kind of depend on a state-by-state basis, but usually if I'm in that position, my damages model — absent some kind of set statutory damages like you might have for copyright right where it's this amount per violation each time — is going to look at what was the economic damage to me. If somebody is podcasting and a podcaster's voice is getting ripped off, well, how much did they make based on my voice. And you would want to argue that [their popularity] essentially comes down to voice and the use of the voice. Otherwise, why would they do it?

So you know, you'd be looking at the revenue that would be attributable to your competitor coming over to you — disgorgement type scenarios.

That's a viable economic model for damages in these types of situations.

For example, a voice actor who goes in and thinks that he's just recording his voice for a single commercial, and it turns out that they use his voice for another commercial, another commercial, another commercial or in perpetuity, and for a variety of different scenarios. Well, how much would you have had to pay me for that? And then you know there can be punitive damages on top of that. How willful was this violation and misuse of my personal voice such that we need to probe punitive damages to prohibit others from doing this in the future? So you know, you could imagine scenarios where the damages can get quite high. Especially if you start to take into account some punitive factors.

Mealey's: I would imagine if you could prove like, hey, this company, this person has done this in 10 instances rather than just a single instance.

Word: That I think that would be one factor that that would come into play . . . but you could also envision scenarios where maybe even just a limited use of it . . . how egregious . . .were some of the misrepresentations maybe that were made that induced a person to provide their voice and allow you to record. Umm.

Mealey's: So what kind of experts would you envision calling in a case like this? Let's say the AI

company argues, "No, no, no, we didn't use this voice. This is a synthesis of 100 voices." We talked before about how 50 people may think it sounds like another person and 50 people may not think the two voices sound the same. Is there a type of expert who could testify to the voice, the way it modulates, the way a speaker says certain words?

Word: There's probably a variety of different experts that you could call upon or data that could be generated by experts to either prove or disprove it. One would be survey data, right? We asked 1,000 people, like you see in trademark cases, likelihood of confusion. You could easily see similar factors at play here. So surveys [that] go out and ask people; those types of experts are well known and are already available on out there. And then you could see a couple of different types of voices.

One would be more of a hard sciences-based where they're taking recordings of whatever is being accused of infringing and running some type of analytical software to say, here's the percent match in terms of these different statements and these reflections, etc., cadence all these different points.

And here's the hard data that you could see being fairly persuasive to a jury, or at the same time we're running that type of analysis and saying there was no match, or you take the average Joe off the street and they would match 80% of the time.

Or you could get somebody who holds himself out to be a speech expert or other type of expert to say, "I've analyzed it with respect to cadence, overall tenor pitch etc. and they're they're almost an exact match." Is that persuasive to injury? Well maybe it is, maybe it isn't,

but there's all different manners of experts out there, and usually the best one is the one that could be helpful to your case and is reaching the right conclusion.

Mealey's: We talked about a little about the protections individuals enjoy. How long do those protections reach back? Can the people who voiced Jarvis or the gentleman who voiced KITT from "Knight Rider" sit back and say, "There are voices out there that sound like mine."

Word: It is going to depend on when it came out. So if there is a something out there that has been around for 20 years that holds itself out to be KITT, you probably missed your chance to go after it. But if somebody comes out with that now? Well, that's a new, and they're holding themselves out to sound like KITT from "Knight Rider." Well, you probably have a couple of claims. Hasbro or whoever was behind "Knight Rider" will probably be out there, but now as a voice actor, do you get to say, "Hey, that was my voice, I'm associated with KTT, and you're trying to capitalize on my voice or you've used my voice even or you got a sound alike and this is new"?

So it I mean it's a new claim, right? If you voiced Hal from "Space Odyssey 2001" long ago, well, that memory is still fresh in people's minds, and if somebody tries to capitalize on it now, that's not going to preclude you from bringing complaint based on a statute of limitations issue. Maybe there's an argument that your voice lost its distinctiveness or it's now out into the public and maybe we'll see some new defenses come up based on novel issues like the age of the voice actor and when they first made.

This interview has been lightly edited for clarity. ∎