LAW360

Portfolio Media. Inc. | 230 Park Avenue, 7th Floor | New York, NY 10169 | www.law360.com
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

Employers Must Battle AI Bias, Fisher Phillips' AI Chief Says

By **Hannah Albarazi**

David Walton

Law360 (February 27, 2024, 8:49 PM EST) -- Employers embracing artificial intelligence and machine learning tools to automate time-consuming tasks, such as screening resumes and conducting interviews, must ensure those tools don't engage in "algorithm drift" that results in improper bias, which could put a company on the hook legally and ruin its brand reputation, says David Walton, chair of Fisher Phillips' artificial intelligence team.

In an interview with Law360, Walton discusses the legal challenges facing employers as their workforce adopts generative AI tools, the threats posed by malicious actors and the possible legal, reputational and societal risks that accompany an employer's use of automated employment decision tools.

There's regulatory risk, too. Walton said the Equal Employment Opportunity Commission has made clear that an employer, "even if they didn't develop the tool, if they use it, and it results in disparate impact on members of a protected class, the employer is on the hook for it."

This interview has been edited for length and clarity.

What kind of challenges are you seeing clients bring up regarding AI policies?

There are things you've got to worry about, like algorithm drift. If you create an AI product — especially based on a large language model, you have to monitor it and constantly test your model, so it's not coming up with improper bias or disparate impact. As you feed the data, the model changes. It might change in a direction that you don't want it to go.

Right now we have one law on the books: New York City's Local Law 144, which requires you to do bias audits if you use an automated employment decision-making tool. But I think there's going to be a lot more in the future.

What could also follow then is that if employers are using resume-screeners based on AI, interview tools based on AI, chatbots for interviews, I think you're going to see cases where it's going to be like a battle of the experts on whether this algorithm has a disparate impact on a protected category.

So, employers are navigating whether some of these AI tools are bumping up against employees' or applicants' civil rights?

Absolutely right. And the Equal Employment Opportunity Commission issued guidance right before the effective date of Local Law 144 in New York City basically saying that when it comes to the use of an automated employment decision tool, even if you didn't develop the tool, if you use it, and it has a biased impact, you're on the hook for it.

It's not a defense to point to the developer and say, "Well, I'm not an expert on this, I just bought their stuff." The EEOC said, "That's not a valid defense, you're responsible for the tools that you use, including these automated employment decision tools within your workplace."

As you advise clients on what their policies should be regarding AI hiring tools and employee use of generative AI, are you seeing blanket bans on using these technologies? Or are you seeing more nuanced approaches?

I'm seeing both. I think most employers are using a nuanced approach because they recognize that people are going to use genAI. You tell them [employees] not to, but because it makes their life easier, they're going to use it anyway.

I think a lot of employers are saying, "OK, we aren't going to stop this, so we have to embrace it." And the way to embrace it is to put some guardrails around it: Don't put certain types of information into a genAI tool, don't represent that something you got purely from the genAI tool is your work product.

You've got to keep in mind that if you use genAI products, like ChatGPT, and it's not an enterprise version, you're putting the clients' information into the public forum. If you use trade secret information, and you put it into ChatGPT, that's putting it out into the public, and it's going to lose its trade secret status.

What you're starting to see is this concept of "private A" — where ChatGPT will take its large language model, its foundational piece, and basically give you an enterprise version of it — so that if you want to use ChatGPT for work, nobody outside your workplace is going to have access to it.

The other thing that you've got to worry about, too, with the use of genAI in particular, is that a lot of your malicious actors now know how to dupe the system.

If they get access to, say, a bank or a stock trading firm's AI, and they get into that model, they can try to dupe it to make certain things happen, so they get false results. That's what you're going to see more of in the future. You're going to see a breach to get into an AI system to help either screw up the results of the AI system or to help shape the results by feeding in certain types of data.

One version of AI is a generative adversarial network, or GANS. It develops really, really good deepfakes. It takes, I think, 3 to 5 seconds of your voice for an AI tool to completely fake it perfectly if you want to replicate a CEO or a CFO's voice, to ask people to send money and do stuff like that.

We're about to enter an age where I don't think that we're mentally ready for it.

Everything that you see on the Internet right now, a picture or a video, we cannot assume it's real anymore because deepfakes have gotten that good.

You know, I have two daughters, 16 and 19. And I told my kids about the deepfakes. And I said, 'Listen, if there's anybody that ever does a deepfake against you and tries to extort you, tell us.' Because you're starting to see that now. You're starting to see it in high schools.

What's at stake in getting AI policies right?

If you have something bad happen with AI, it could ruin your company's brand, it could ruin your company's reputation. There are good plaintiffs lawyers out there who know that this is the next big thing in terms of class actions, for not only employment class actions, but consumer class actions.

But a lot of these risks are kind of nebulous because we're still at the beginning.

Because we don't have anything like the European Union AI Act, a lot of your governance issues for American companies are: What are you doing to identify risk with your use and development of AI? What are you doing to address the risk? What are you doing to stay on top of it, to watch out for things like algorithm drift and stuff like that?

I think a lot of AI governance is setting up a framework right now, a risk management framework, to guard against things that might happen in the future that we aren't even thinking about yet.

What issues with AI do you think employees may face in the future?

I think employees are going to be worried in the future about the use of their data because there is a tremendous opportunity to use employee data to help predict and manage human behavior.

Using data in AI and predictive analytics to try to anticipate which prospective employees are going to fit in the job after they start working for you, to predict if they're thinking about leaving, if they're thinking about going to a competitor, if they're going to be on a certain career path.

I think right now it's like a free-for-all in the United States because if you work for a private employer, there are very few restrictions on how they can use your employment-related data to help manage a workforce.

I think there's going to be a perception out there that AI tools are used to illegally discriminate, and so there's going to be more lawsuits based on that. You're going to have more lawsuits where there's discovery into the use of algorithms by the employer.

That technology is there. It's being used, but under the radar now. Because it's effective, it works, so that's why it's never going to go away. It's very effective, but it raises a lot of issues.

--Editing by Rich Mills.