



Feb 07, 2018

Categories:
[Publications](#)

You Can't Sue a Robot: Are Existing Tort Theories Ready for Artificial Intelligence? (Part 3 of 3)

This is the third article in a series on the legal issues surrounding artificial intelligence (AI), based on an fbtTECH webinar held in November 2017.

Part 1 – [AI is Surging: Are We Ready for the Fallout?](#)

Part 2 – [Artificial Intelligence and Data Privacy: Are We Sufficiently Protected?](#)

Part 3 – You Can't Sue a Robot: Are Existing Tort Theories Ready for Artificial Intelligence?

As artificial intelligence (AI) becomes more ubiquitous and sophisticated, the question of how to provide remedies to harm caused by AI is becoming more important. Google subsidiary Waymo, for example, recently launched the first large-scale test of fully-autonomous vehicles, with no safety driver, in Phoenix. And complex AI are popping up in appliances, communications technology, and professional services, among others.

AI has advanced to the point where it is behaving in surprising ways. A Chinese app developer, for example, pulled its instant messaging "chatbots"—designed to mimic human conversation—after the bots unexpectedly started criticizing communism. Facebook chatbots began developing a whole new language to communicate with each other—one their creators could not understand. And Google subsidiary Deepmind built an AI that taught itself in only a few days, with no human direction whatsoever, to be the best player ever of the board game Go, handily beating its human-trained AI predecessor. The troubling underlying truth is that AI has become so sophisticated that its creators don't always understand how it works or why the AI makes the decisions it does.

Who is Responsible for AI Harm?

As AI gains more control over more common objects and services, but also becomes more sophisticated and potentially unpredictable, it is inevitable that AI will cause harm. When that happens, is our legal system up to the task of providing a remedy?

We already have robust legal frameworks to deal with simple machines that make no decisions. If a factory robot injures a worker, for example, we don't consider blaming the robot any more than we would think to blame a couch or faulty oven. The employee's injuries are covered by workers' compensation schemes. We ask whether the employer had sufficient safety protocols in place. We look to products liability principles and consider whether there is a manufacturing or design defect with the robot, or whether the manufacturer failed to provide adequate safety warnings. These principles let us apportion blame and provide a remedy.

What about simple AI decision makers? Existing legal concepts may still suffice. If an AI is programmed to make narrow autonomous decisions along specific lines and injures someone then we could look to the AI's creators and to negligence principles. Was the injury reasonably foreseeable? Was the victim also negligent in some way? If the AI's creators reasonably should have anticipated this harm, then they could be responsible. Problems may still arise: What if the AI's creators are no longer in business? Do we also look to the owner of the AI to bear some responsibility, even though the owner had no input into the programming that caused the harm?

These problems become magnified as the AI gets more complex and potentially starts behaving in unpredictable ways. When AI make complex decisions and take actions completely unforeseen by their creators, where do we look to apportion blame? After all, you can't sue a robot. (And more importantly, even if you could, it doesn't have any assets.) There is no perfect analogy under the law to deal with a fully autonomous and somewhat unpredictable AI, but we can still look to existing legal frameworks for ideas.

Spreading the Risk

Some commentators have proposed expanding products liability theories to cover autonomous robots, on the theory that, if a robot

causes harm, this is implicit proof of some defect with the robot. In many ways this resembles a strict liability standard—a robot caused harm, and the creator must pay for it. All AI should have limits placed on their ability to cause harm, and the creator is in the best position to prevent any harm and to absorb any economic losses stemming from such harm. This approach may go too far, however, in removing any inquiry into human fault for the harm, and may stifle innovation with autonomous AI.

Another approach is to modify negligence principles. We would analyze what happened, whether human actors contributed, and what degree of fault to assign to the AI, for which the creator would then be liable. This would be true even if the AI's actions were not intended or reasonably foreseeable by the creator. Many industry experts, of course, predict that AI that take over human tasks (like driving) will be much safer than humans. If that prediction bears out, then this kind of framework could eventually morph into one that defaults to blaming human actors, or that carries a presumption that the human actors involved are at fault.

Modern worker's compensation schemes point to another possible solution. The purpose of workers' compensation laws is to avoid endless litigation over who is at fault when an employee suffers an injury at work. Workers' compensation insurance (or state-run workers' compensation schemes) spreads the risk across all employers, provides remedies for individual workers, and shields individual employers from catastrophic damages. Employers are still required to pay a portion of the damages awarded to employees, however, so as to incentivize employers to prevent workplace injuries. Perhaps a system modeled on worker's compensation could work to mitigate potential harms caused by AI and provide a remedy for victims.

Endless Possibilities and Challenges Ahead

But how would such a system work? Would AI creators be required to pay into such a system? Absolving the creators from any liability, after all, removes any incentives for creators to continue to refine safety measures. What if the creators go out of business and their AI continue operating out in the world? Should the costs be pushed

on to AI owners, so they bear some responsibility for the actions of an AI they send out into the world on their behalf? Should AI be required to be insured, either on the creator's end or the owner's end? Or should every autonomous AI/robot be taxed annually to fund an AI-compensation system? Perhaps placing all responsibility on the AI owners would be sufficient, as market forces would weed out AI creators who give short shrift to safety. If we are taxing autonomous AI to cover the costs of any harms, how do we determine if an AI is sophisticated enough to require coverage?

Another problem to sort out with all these approaches is, when we say the creator should be liable or should pay into the system, who do we mean? Do we mean only the software developer—the company that created the AI's "brains?" Or do we also mean the hardware manufacturer, or component suppliers? And as discussed above, what role in any of these approaches, if any, should the owner play?

A world saturated with smart phones and on the cusp of fully autonomous vehicles was difficult to imagine 15 years ago. If the pace of technological development has taught us anything, it's that the future of AI will bring undreamt of challenges. We cannot rely on legal frameworks developed in the 20th century to meet the unforeseen dilemmas of a dawning new world. Although existing legal theories may help point the way, the law will need to evolve to keep pace with developments in AI.

© 2026 FBT Gibbons LLP. May be considered attorney advertising in some jurisdictions.