2019 Annual Conference
March 13 -15 2019
Orlando, FL

Artificial Intelligence - Real Risks and Coverage Implications of Things Previously Reserved for Sci Fi Movies

I. The History and Direction of Artificial Intelligence

A. Brief History

i. Old AI

The Encyclopedia Britannica states that "artificial intelligence (AI) [is] the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings [meaning those] that can adapt to changing circumstances." AI began in the 1940s with the development of digital computers, which could be programmed to perform highly complex tasks with great proficiency, such as playing chess and discovering proofs for mathematical theorems. However, for decades, computers were too slow to process the amount of data needed for "learning" and funding for AI research became scarce until the late 1990s when interest in AI resumed and companies like IBM developed "Big Blue," the chess machine that beat Gary Kasparaov in 1997. This led to a renewed interest in development of AI.

ii. Current AI

Since that time, serious advances in computer processing power and storage enabled companies to store and process vast quantities of data for the first time. While robotics has been used in manufacturing for a long time, in the past 15 years, companies like Amazon, Google and others have leveraged machine learning not only to process user data to understand consumer behavior, but to develop computer "vision," natural language processing, and other AI applications that have become an integral part of online user services and everyday home, office and remote devices.

AI now includes everything from Siri and Alexa to self-driving cars to IBM's Watson (which beat humans at *Jeopardy*) to search engine algorithms to drones and autonomous weapons. Presently, AI is limited to narrow AI, which is designed to perform only a narrow task (e.g., only driving a car, or only facial recognition, or only internet searches).

iii. Emerging AI

While narrow AI may outperform humans at the specific task for which it is created (problem solving, playing chess or *Jeopardy*), many researchers have a long-term goal of creating general AI, which would ostensibly outperform humans at nearly all cognitive tasks.

B. Why We Care

As with all technology, the insurance industry cares because it raises new and unforeseen risks in multiple sectors of the economy that are likely to escalate as AI continues to evolve, especially into broader general AI. Most researchers agree that super-intelligent AI is unlikely to exhibit human emotions (love, hate etc.), or that AI would become intentionally benevolent or malevolent, but experts generally agree that two risk scenarios are most likely:

(1) AI is programmed to do something beneficial, but it develops a destructive method for achieving that goal.

This risk arises when developers fail to properly and completely align the AI's goals with the human intent, which is extremely difficult to do. A "not just sci-fi anymore" example is asking an autonomous car to take you to the airport as fast as possible, but it goes so fast it results in a high-speed police chase including helicopters, and you arrive at the airport covered in vomit. The car will have literally done what you asked, but not what you wanted. Extrapolating that out, if a super-intelligent AI system is created to perform an ambitious geoengineering project, a potential risk is that it might wreak havoc on the surrounding environment and ecosystem but interpret human efforts to stop the adverse ecological effects as a threat to the project.

This kind of risk already exists with narrow AI. A real-life example, which was caught before it caused any actual harm, made headlines in 2018 when Amazon scrapped its secret human resources recruitment tool because it exhibited a clear hiring bias against women.

In 2014, the company developed an experimental hiring tool using AI to give job candidates scores ranging from one to five stars (like the shopper product rating tool on Amazon), with the intention that it would review massive numbers of resumes and quickly identify the most qualified candidates for each position. However, by 2015, the company realized the system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way.

Amazon's investigation determined that its computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most of the resumes came from men, which reflected historical male dominance in the tech industry. Amazon's system taught itself that male candidates were preferable, penalizing resumes that included the word "women's" (such as "women's chess club captain"), and downgraded graduates of all-women's colleges. Amazon edited the programs to make them neutral to these specific terms but found that it could not guarantee that the AI would not devise other discriminatory ways of sorting candidates based on inherent biases in the initial programming. Ultimately, Amazon abandoned the project and disbanded the development team.

According to Reuters, "55% of U.S. human resources managers said artificial intelligence, or AI, would be a regular part of their work within the next five years, according to a 2017 survey by talent software firm CareerBuilder," and that huge companies including Hilton Worldwide, and Goldman Sachs are actively seeking to use AI to automate the hiring process. (From a risk/insurance perspective, EPLI carriers take note.)

(2) AI is programmed to do something harmful/devastating.

This "not just for sci fi anymore" scenario is the concern that autonomous weapons (AI systems programmed to kill), in the hands of the wrong people, could be programmed to cause mass casualties, and that an AI arms race could inadvertently lead to an AI war that also results in mass casualties. To avoid being thwarted by an enemy, these weapons would have to be designed to be very difficult to "turn off," so humans could plausibly lose control of such a situation. This risk is also present in narrow AI. For example, Uber recently pulled its self-driving cars off the road in Arizona after an accident in which a pedestrian was killed, because the car's controls could not be overridden to avoid conflicts with pedestrians and bicyclists in local driving and Uber was having difficulty programming the AI to address local driving challenges which do not exist in highway driving. (CGL/products-completed operations and auto insurers take note.)

II. Potential Risks

A. Hypotheticals

An autonomous freight train moving at high speed derails when the train fails to break going into a curve, with a few cars crashing into cars on the adjacent highway, causing a fire that ignites a runaway fire that decimates the next town and releasing hazardous chemicals into the air and adjacent river.

B. Potential Claims

The potential claims include third party claims by the injured vehicle and building owners for bodily injury and property damage arising out of the derailment as well as the homeowners, business owners, property owners, municipalities and others affected by the fire. BI claims are also foreseeable as a result of the potential toxic chemical exposure both for immediate exposure (sore throat, cough, eye irritation etc.) as well as potential long-term exposure and medical monitoring depending on the substances released. Claims for recoupment of response costs are likely by the municipalities responding to the crash scene, fire and hazmat emergencies. CERCLA and other environmental cleanup claims are also foreseeable for air, soil and groundwater contamination. Because the incident involved a freight train, claims for the lost lading are foreseeable. Business interruption claims are also likely, as are subrogation claims.

i. Potentially liable parties

If the train were not autonomous, the old loss scenario liability evaluation would focus on the train operator, manufacturer, brake manufacturer, and track owner. Those entities would likely dispute liability, but the options are limited. With the addition of AI, however, the liability picture becomes significantly more complex. Issues now include whether the AI circuitry was faulty; whether the computer chips were defective; whether the programming was proper; whether there was a connectivity failure; if the system was hacked; whether the AI elected not to apply the brakes because a specific set of circumstances existed. This implicates the designer and developer of the AI, the programmer(s), installers (if different), the AI system component part manufacturers; software manufacturers; and whatever entity may have responsibility for cyber security to prevent hacking.

ii. Potential damages

The damages arising from such a scenario are the same regardless of the existence of AI.

III. Coverage Implications

Under the above hypothetical, there is no question that CGL and products-completed operations liability policies will be implicated as respects the train manufacturer, owner/operator, railway that owns the tracks. However, the presence of

AI raises a question as to potential exposure under professional liability errors & omissions policies issued to the designers, programmers and installers of the AI system. Cyber coverages may also come into play to the extent that there is a hacking issue. If AI blocks email that should have been allowed to get to a server, Tech E&O policies designed to cover losses from faulty software and other tech products and services may cover some costs.

However, the insurance implications of AI are far broader than traditional products liability and tort scenarios. As noted earlier in this paper, broad applications of AI can result in discriminatory hiring practices which would implicate EPLI policies. Use of AI in the medical and dental professions (e.g., to assist in reading MRIs and other imaging for other diagnostic purposes, or for programming medicine compounding or delivery systems) also implicates medical malpractice and other professional liability policies for physicians, surgeons, nurses, hospitals, nursing homes and other patient care facilities. Use of AI in home security systems implicates first party coverages (homeowners). AI in self-driving cars clearly impacts the auto insurance market, both personal and commercial. E&O policies may also apply if some products performed as intended but produced poor results because it learned from bad data.

Coverage is less clear if the AI failure did not cause any physical damage, or if a company's first party loss stems from its own use of AI. Changing the hypothetical, let's say a programming error caused a security flaw in the AI software and a hacker exploited the flaw, disabling the brakes. The crash disabled the train fleet while the network was restored, causing business interruption loss to the company itself. The company using the train to haul its goods suffered reputational damage because it breached its contractual obligations. The train company's first party policy may or may not cover the physical damage to the train and the business interruption if caused by a cyber-attack. The company's cyber policy may cover lost data, but not the physical damage or business interruption. The manufacturers' products policy or the software developer's E&O policy may respond but not if the damage was caused by the hacker rather than directly by the programming error. Bottom line: AI adds many layers of potential liability and many layers of potential coverage issues across multiple lines of coverage. Coverage issues under every policy will include policy definitions, insuring agreement terms, exclusions, limits/sublimit, anti-stacking provisions and other insurance clauses.

IV. AI in the Insurance Industry/Insurtech

Not only is AI an issue for underwriters and claims handlers when evaluating policyholders and policies, the insurance industry is exploring and adopting AI for its own use within the industry.

One example is behavioral policy pricing. Some insurance companies (think Progressive and other auto and healthcare insures) have adopted usage-based insurance models that involve policyholders using telematics/wearable sensors (e.g., safer drivers (*OctTelematics*), healthier lifestyles (*FitSense*)) to provide information about safe driving or healthy lifestyles practives. These can result in lower premiums for policyholders who record better behavior patterns (or, conversely, higher premiums or cancellation for poor practices).

Another example is bundling homeowners' policies with loss prevention hardware. In this scenario, "smart home" companies offer discounts on homeowners' premiums for the use of sensorized loss prevention technology. Note: this can come with other risks, such as hackers accessing chatbots created by the interplay between sensorized equipment in the home to track usage patterns and determine when people are home or away, or to infiltrate wireless servers and access personal information.

Another way in which insurance companies are leveraging AI include using the Internet of Things (IoT) data markets ("big data") to verify risk management info instead of costly, slower traditional audits and risk assessments. Companies are also looking at Ai to assist in expediting and managing costs associated with the claim verification and settlement processes and increase fraud detection. Efforts in this regard include automated claims intake/setup systems, Chatbots which act as virtual claims adjusters (e.g., *Cognicor*, *Conversica*), and automated claims management systems.

CONCLUSION

While the Terminator and The Matrix remain pure fiction, and experts agree that super-intelligent AI does not yet exist and is highly unlikely to replace humans anytime in the foreseeable future, there is no question that narrow AI is here to stay and is presenting new, novel and potentially significant risks. Underwriters and claims adjusters across all lines of business should consider these aspects and potential issues when assessing policyholders and claims.