19056   Our Future Robots Will Need Super-Smart Safety Checks

By Jamie Condliffe, *The New York Times*, March 22, 2019

Machines are more than ever controlled by software, not humans. Occasionally it goes fatally wrong.

On March 10, 157 people died when an Ethiopian Airlines Boeing 737 Max 8 jet crashed. Five months earlier, another crash of the same model of airplane killed 189 people. There are indications that software intended to prevent the jets from stalling may have played a role in both accidents. The Boeing software was designed to perform a simple task: nudge the airplane's nose down based on sensor readings if a stall was anticipated.

Reporting by *The New York Times* suggests that the software didn't receive a detailed review by the Federal Aviation Administration before it entered use: under new rules, the agency delegated much of the responsibility to Boeing. If the software was at fault, and the problem did slip through the regulatory net, it raises questions about how safety-critical technology is vetted.

Those questions will become more important over the next few years.

A year ago, an Arizona woman was struck and killed by one of Uber's autonomous cars. The vehicle's autonomy systems failed to brake, as did the safety driver behind the wheel.

Companies like Uber and Waymo, along with most of the auto industry, expect autonomous cars to proliferate over the next decade. Such advances aren't limited to cars — they will further automate everything from air travel to food delivery. They are built on technologies, like artificial intelligence, that will make split-second decisions for humans.

As harder tasks are handed to software, the stakes rise.

"As you create more advanced A.I. systems, the harm that can result from them failing can be really large," said a researcher at Oxford University's Center for the Governance of Artificial Intelligence.

But increasing the complexity of systems makes checking them more difficult. Hardware, from chips to special sensors, can be difficult to test. And it can be difficult for humans to understand how some A.I. algorithms make decisions.

Regulators need to be more aware of tail-end risks — the highly unlikely but catastrophic events that could occur if something malfunctioned. That might mean introducing more conservative rules that relax as technology matures, ideally developed in tandem with technologists who understand deeply how the systems work.

"Verifying the performance and safety of software is a really, really hard technical challenge," the researcher at Oxford University's Center for the Governance of Artificial Intelligence said. Nonetheless, it's a challenge that has to be addressed.

(405 words)