

When self-driving cars crash, who's responsible? Courts and insurers need to know what's inside the 'black box'

May 25 2022, by Aaron J. Snoswell, Henry Fraser and Rhyle Simcock



Credit: AI-generated image ([disclaimer](#))

The first serious accident involving a self-driving car in Australia occurred in March this year. A pedestrian suffered life-threatening injuries when [hit by a Tesla Model 3](#) in "autopilot" mode.

In the U.S., the highway safety regulator is investigating a [series of accidents](#) where Teslas on autopilot crashed into first-responder vehicles with flashing lights during traffic stops.

The decision-making processes of "self-driving" cars are often [opaque and unpredictable](#) (even to their manufacturers), so it can be hard to determine who should be held accountable for incidents such as these. However, the growing field of "explainable AI" may help provide some answers.

Who is responsible when self-driving cars crash?

While self-driving cars are new, they are still machines made and sold by manufacturers. When they cause harm, we should ask whether the manufacturer (or [software developer](#)) has met their safety responsibilities.

Modern negligence law comes from the famous case of [Donoghue v Stevenson](#), where a woman discovered a decomposing snail in her bottle of ginger beer. The manufacturer was found negligent, not because he was expected to directly predict or control the behavior of snails, but because his bottling process was unsafe.

By this logic, manufacturers and developers of AI-based systems like self-driving cars may not be able to foresee and control everything the "autonomous" system does, but they can take measures to reduce risks. If their risk management, testing, audits and monitoring practices are not good enough, they should be held accountable.

How much risk management is enough?

The difficult question will be "How much care and how much risk

management is enough?" In complex software, it is [impossible to test for every bug](#) in advance. How will developers and manufacturers know when to stop?

Fortunately, courts, regulators and technical standards bodies have experience in setting standards of care and responsibility for risky but useful activities.

Standards could be very exacting, like the European Union's [draft AI regulation](#), which requires risks to be reduced "as far as possible" without regard to cost. Or they may be more like Australian negligence law, which permits less stringent management for less likely or less severe risks, or where [risk management](#) would reduce the overall benefit of the risky activity.

Legal cases will be complicated by AI opacity

Once we have a clear standard for risks, we need a way to enforce it. One approach could be to give a regulator powers to impose penalties (as the ACCC does in competition cases, for example).

Individuals harmed by AI systems must also be able to sue. In cases involving self-driving cars, lawsuits against manufacturers will be particularly important.

However, for such lawsuits to be effective, courts will need to understand in detail the processes and technical parameters of the AI systems.

Manufacturers often prefer not to reveal such details for commercial reasons. But courts already have procedures to balance commercial interests with an appropriate amount of disclosure to facilitate litigation.

A greater challenge may arise when AI systems themselves are opaque "[black boxes](#)." For example, Tesla's autopilot functionality relies on "[deep neural networks](#)," a popular type of AI system in which even the developers can never be entirely sure how or why it arrives at a given outcome.

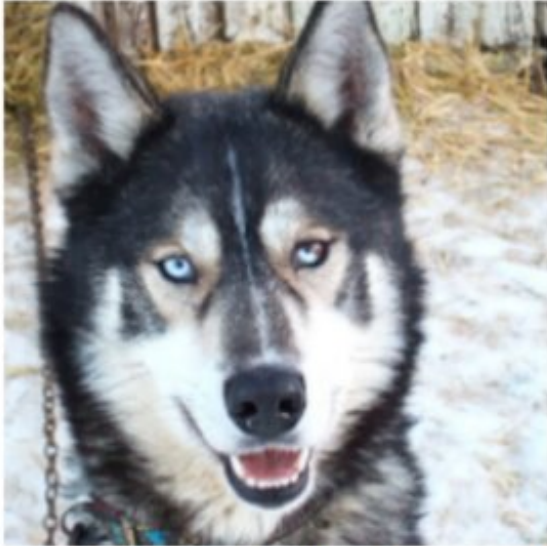
'Explainable AI' to the rescue?

Opening the black box of modern AI systems is the focus of a [new wave](#) of computer science and humanities [scholars](#): the so-called "explainable AI" movement.

The goal is to help developers and end users understand how AI systems make decisions, either by changing how the systems are built or by generating explanations after the fact.

In [a classic example](#), an AI system mistakenly classifies a picture of a husky as a wolf. An "explainable AI" method reveals the system focused on snow in the background of the image, rather than the animal in the foreground.

How this might be used in a lawsuit will depend on various factors, including the specific AI technology and the harm caused. A key concern will be how much access the injured party is given to the AI system.



(a) Husky classified as wolf



(b) Explanation

Explainable AI in action: an AI system incorrectly classifies the husky on the left as a ‘wolf’, and at right we see this is because the system was focusing on the snow in the background of the image. Credit: [Ribeiro, Singh & Guestrin](#)

The Trivago case

Our [new research](#) analyzing an important recent Australian court case provides an encouraging glimpse of what this could look like.

In April 2022, the Federal Court penalized global hotel booking company Trivago \$44.7 million for misleading customers about hotel room rates on its website and in TV advertising, after a case brought on by [competition watchdog the ACCC](#). A critical question was how Trivago's complex ranking algorithm chose the top ranked offer for hotel rooms.

Trivago fined \$45 million for misleading customers on hotel

pricing claims <https://t.co/xLpzBeIUNz>

— ABC News (@abcnews) [April 22, 2022](#)

The Federal Court set up rules for evidence discovery with safeguards to protect Trivago's intellectual property, and both the ACCC and Trivago called [expert witnesses](#) to provide evidence explaining how Trivago's AI system worked.

Even without full access to Trivago's system, the ACCC's expert witness was able to produce compelling evidence that the system's behavior was not consistent with Trivago's claim of giving customers the "best price."

This shows how technical experts and lawyers together can overcome AI opacity in court cases. However, the process requires close collaboration and deep technical expertise, and will likely be expensive.

Regulators can take steps now to streamline things in the future, such as requiring AI companies to adequately document their systems.

The road ahead

Vehicles with various degrees of automation are becoming more common, and fully autonomous taxis and buses are being tested both [in Australia](#) and [overseas](#).

Keeping our roads as safe as possible will require close collaboration between AI and legal experts, and regulators, manufacturers, insurers, and users will all have roles to play.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: When self-driving cars crash, who's responsible? Courts and insurers need to know what's inside the 'black box' (2022, May 25) retrieved 7 May 2024 from

<https://techxplore.com/news/2022-05-self-driving-cars-responsible-courts-black.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.