

How understanding animals can help us make the most of artificial intelligence

March 31 2017, by Heather Roff



Autonomous cars aren't smarter than dogs. Credit: X posid

Every day countless headlines emerge from myriad sources across the globe, both [warning of dire consequences](#) and [promising utopian futures](#) – all thanks to artificial intelligence. AI "is transforming the workplace,"

writes the [Wall Street Journal](#), while [Fortune](#) magazine tells us that we are facing an "AI revolution" that will "change our lives." But we don't really understand what interacting with AI will be like – or what it should be like.

It turns out, though, that we already have a concept we can use when we think about AI: It's how we think about animals. As a former animal trainer (albeit briefly) who now studies how people use AI, I know that animals and animal training can teach us quite a lot about how we ought to think about, approach and interact with artificial intelligence, both now and in the future.

[Using animal analogies](#) can help regular people understand many of the complex aspects of [artificial intelligence](#). It can also help us think about how best to teach these systems new skills and, perhaps most importantly, how we can properly conceive of their limitations, even as we celebrate AI's new possibilities.

Looking at constraints

As AI expert [Maggie Boden explains](#), "Artificial intelligence seeks to make computers do the sorts of things that minds can do." AI researchers are working on teaching computers to reason, perceive, plan, move and make associations. AI can see patterns in large data sets, predict the likelihood of an event occurring, plan a route, manage a person's meeting schedule and even play war-game scenarios.

Many of these capabilities are, in themselves, unsurprising: Of course a robot can roll around a space and not collide with anything. But somehow AI seems more magical when the computer starts to put these skills together to accomplish tasks.

Take, for instance, [autonomous cars](#). The origins of the driverless car are

in a 1980s-era Defense Advanced Research Project Agency project called the [Autonomous Land Vehicle](#). The project's goals were to encourage research into computer vision, perception, planning and robotic control. In 2004, the ALV effort became the first [Grand Challenge](#) for self-driving cars. Now, more than 30 years since the effort began, we are on the precipice of autonomous or self-driving cars in the civilian market. In the early years, few people thought such a feat was impossible: Computers couldn't drive!

Yet, as we have seen, [they can](#). Autonomous cars' capabilities are relatively easy for us to understand. But we struggle to comprehend their limitations. After the [2015 fatal Tesla crash](#), where the car's autopilot function failed to sense a tractor-trailer crossing into its lane, few still seem to grasp the gravity of how limited Tesla's autopilot really is. While the company and its software were [cleared of negligence](#) by the National Highway Traffic Safety Administration, it remains unclear whether customers really understand what the car can and cannot do.

What if Tesla owners were told not that they were [driving a "beta" version of an autopilot](#) but rather a semi-autonomous car with the [mental equivalence of a worm](#)? The so-called "intelligence" that provides "[full self-driving capability](#)" is really a giant computer that is pretty good at sensing objects and avoiding them, recognizing items in images and limited planning. That might change owners' perspectives about how much the car could really do without human input or oversight.

What is it?

Technologists often try to explain AI in terms of how it is built. Take, for instance, advancements made in deep learning. This is a technique that uses [multi-layered networks](#) to learn how to do a task. The networks need to process vast amounts of information. But because of the volume of the data they require, the complexity of the associations and

algorithms in the networks, it is often unclear to humans how they learn what they do. These systems may become very good at one particular task, but we do not really understand them.

Instead of thinking about AI as something superhuman or alien, it's easier to analogize them to animals, intelligent nonhumans we have experience training.

For example, if I were to use [reinforcement learning](#) to train a dog to sit, I would praise the dog and give him treats when he sits on command. Over time, he would learn to associate the command with the behavior with the treat.

Training an AI system can be very much the same. In reinforcement deep learning, human designers set up a system, envision what they want it to learn, give it information, watch its actions and give it feedback (such as praise) when they see what they want. In essence, we can treat the AI system like we treat animals we are training.

The analogy works at a deeper level too. I'm not expecting the sitting dog to understand complex concepts like "love" or "good." I'm expecting him to learn a behavior. Just as we can get dogs to sit, stay and roll over, we can get AI systems to move cars around public roads. But it's too much to expect the car to "[solve](#)" the [ethical problems that can arise in driving emergencies](#).

Helping researchers too

Thinking of AI as a trainable animal isn't just useful for explaining it to the general public. It is also helpful for the researchers and engineers building the technology. If an AI scholar is trying to teach a system a new skill, thinking of the process from the perspective of an animal trainer could help identify potential problems or complications.

For instance, if I try to train my dog to sit, and every time I say "sit" the buzzer to the oven goes off, then my dog will begin to associate sitting not only with my command, but also with the sound of the oven's buzzer. In essence, the buzzer becomes another signal telling the dog to sit, which is called an "accidental reinforcement." If we look for accidental reinforcements or signals in AI systems that are not working properly, then we'll know better not only what's going wrong, but also what specific retraining will be most effective.

This requires us to understand what messages we are giving during AI training, as well as what the AI might be observing in the surrounding environment. The oven buzzer is a simple example; in the real world it will be far more complicated.

Before we welcome our AI overlords and hand over our lives and jobs to robots, we ought to pause and think about the kind of intelligences we are creating. They will be very good at doing particular actions or tasks, but they cannot understand concepts, and do not know anything. So when you are thinking about [shelling out thousands](#) for a new Tesla car, remember its autopilot function is really just a very fast and sexy worm. Do you really want to give control over your life and your loved ones' lives to a worm? Probably not, so keep your hands on the wheel and don't fall asleep.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: How understanding animals can help us make the most of artificial intelligence (2017, March 31) retrieved 9 April 2024 from <https://phys.org/news/2017-03-animals-artificial-intelligence.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.