



AI ethics: the case for including animals

Peter Singer¹ · Yip Fai Tse²

Received: 17 March 2022 / Accepted: 1 June 2022 / Published online: 6 July 2022
© The Author(s) 2022

Abstract

The ethics of artificial intelligence, or AI ethics, is a rapidly growing field, and rightly so. While the range of issues and groups of stakeholders concerned by the field of AI ethics is expanding, with speculation about whether it extends even to the machines themselves, there is a group of sentient beings who are also affected by AI, but are rarely mentioned within the field of AI ethics—the nonhuman animals. This paper seeks to explore the kinds of impact AI has on nonhuman animals, the severity of these impacts, and their moral implications. We hope that this paper will facilitate the development of a new field of philosophical and technical research regarding the impacts of AI on animals, namely, the ethics of AI as it affects nonhuman animals.

Keywords Artificial intelligence · AI ethics · Animal ethics · Factory farming · AI for animals · Moral circle

1 Introduction

The ethical implications of AI have sparked concern from governments, the public, and even companies.¹ According to some meta-studies on AI ethics guidelines, the most frequently discussed themes include fairness, privacy, accountability, transparency, and robustness [1–3]. Less commonly broached, but not entirely absent, are issues relating to the rights of potentially sentient or autonomous forms of AI [4, 5]. One much more significant, and more immediately present, issue has, however, been almost entirely neglected: AI's impact on non-human animals.² There have, we acknowledge, been discussions of AI in connection with endangered species and ecosystems,³ but we are referring to questions relating to AI's impact on individual animals. As we will show in more detail below, many AI systems have significant impacts on animals, with the total number of animals affected annually likely to reach the tens or even hundreds of billions. We therefore argue that AI ethics needs to broaden its scope in order to deal with the ethical implications of this very large-scale impact on sentient, or possibly sentient, beings.

Before going further, it is important to clarify our use of terminology. First, humans are also members of the kingdom *Animalia*, so it would be more accurate to refer to the beings we are discussing as “non-human animals,” but henceforth, for brevity, we shall simply use “animal” to refer to all animals other than humans. Second, by artificial intelligence, we take Russell and Norvig's definition, which “includes reactive agents, real-time planners, decision-theoretic systems, and deep learning systems”. Under this definition, expert systems, computer vision, natural language processing, search engines, recommendation algorithms, robotics, to the systems that use the recently popular “buzzwords” such as machine learning or deep learning, are all examples of AI systems. We use a broader definition than most other people have in recent years because we believe the issue of animal ethics applies to all these systems [6]. We will discuss, in what follows, the moral significance of our limited knowledge of the extent to which various kinds of animals are sentient.

✉ Yip Fai Tse
tseyipfai@gmail.com

¹ University Center for Human Values, Princeton University, Princeton, USA

² University Center for Human Values, Center for Information Technology Policy, Princeton University, Princeton, USA

¹ A recent study by Klynveld Peat Marwick Goerdeler (KPMG) showed that more than 90% of the executives of their surveyed clients (KPMG 2021) think that the ethical impacts of their AI systems will affect their company's reputation.

² We support this claim in Sect. 3, below.

³ “Five Ways AI Is Saving Wildlife – from Counting Chimps to Locating Whales | Artificial Intelligence (AI) | The Guardian.” Accessed March 16, 2022. <https://www.theguardian.com/environment/2022/feb/21/five-ways-ai-is-saving-wildlife-from-counting-chimps-to-locating-whales-aoe.2>.

The structure of the paper forms a series of step-by-step arguments, leading to the conclusion that there needs to be AI ethics concerning animals.

1. Animals matter morally, at least to *some* degree (Sect. 2).
2. AI systems do in fact impact animals.
3. These impacts are huge in scale and severe in intensity, and therefore important. (Sect. 3.2).
4. Conclusion: AI ethics needs to include consideration of impact of AI on animals

2 Why do animals matter ethically?

Modern science provides ample evidence that vertebrate animals, including humans, share much of their evolutionary history [7, 8] and as a result, have similar neurological structures [9]. This creates a presumption in favor of their capacity for consciousness that is bolstered by observations of their behavior, including behavior when subjected to a stimulus that would cause pain in humans. This pain behavior lessens or disappears when analgesics are administered [9, 10]. With invertebrates there is more doubt. Cephalopods such as the octopus are far more distant from us, in evolutionary terms, than any vertebrate, yet their behavior gives such clear signs of intelligence and sensitivity to pain that it is difficult to believe that they do not have conscious experiences [11]. Crustaceans of the order *decapoda*, such as lobsters and crabs, are protected in the animal welfare legislation of some countries because there is evidence suggesting that they are probably sentient [12, 13]. With other invertebrates the evidence of sentience seems to be less conclusive, but the possibility is difficult to exclude.

These facts have important implications for ethics because it is reasonable to claim that having the capacity to experience pain and pleasure is sufficient to give a being moral status [14–16].⁴ The capacity to experience pain and pleasure is not, of course, sufficient for moral agency, but it is sufficient to make it wrong to do certain things to the being. This is now recognized in the increasing tendency of many countries to pass legislation granting animals the status of “sentient being,” a position between that of a person and that of a thing.⁵ As one of us has argued previously, given that

animals are capable of experiencing pain and pleasure, their interests in not feeling pain, and in having pleasant lives, matter ethically [14–17].⁶ We are not justified in ignoring their interests, or giving those interests less weight, than we give to human interests of a similar nature and strength. It is important to add that nothing we have said here implies that the lives of all animals, human or nonhuman, matter equally, because their interests will often differ, and some may have more significant interests in continuing to live than others.

We hold, therefore, that the correct ethical principle to apply when the interests of animals and humans are salient is the principle of “equal consideration of similar interests.” We regard it as wrong to give less consideration to the interests of an animal than we give to similar human interests (insofar as rough comparisons can be made) merely on the basis that the animal is not a member of our species. For example, if, as far as we can tell, cutting the flesh of an animal causes as much pain as a similar cut would cause a human to experience, then the human and the animal have similar interests in not having their flesh cut, and we should give these interests equal consideration. To discount or neglect the interests of a being because it is not a member of our species is to be a speciesist—a term that we use to suggest that this attitude to animals is analogous to the wrongs of giving less consideration to the interests of people who are not members of our race, ethnic group, or sex.

In what follows, we will assume this ethical position. It is important to emphasize that although we ourselves accept the principle of equal consideration of similar interests, the arguments that follow are relevant to everyone who accepts that the interests of animals matter to some extent, even if they do not think that similar human and non-human interests matter equally.

3 AI's impacts on animals

3.1 Types of impact

It is now widely recognized that AI technologies affect human lives and society significantly, for it is obvious that all AI systems are designed to change a certain aspect of some humans' lives. But it might not be obvious for many

⁴ Peter Singer, *Animal Liberation* (Harper Perennial, New York, 2nd edition, 2009, (1975)) [14]. Peter Singer, Ch. 1; Practical Ethics, Cambridge University Press, 3rd edition, 2011 (1979) [15]. For a Kantian view on how capacity to feel pain and pleasure relate to moral standing, see Christine Korsgaard, *Fellow Creatures: Our Obligations to the Other Animals*, Oxford University Press, 2018 [16].

⁵ The status of sentient being is to be found in Article 13 of the Treaty on the Functioning of the European Union (2016), and in specific legislation in the United Kingdom, France, the Netherlands, New Zealand, Sweden, the U.S. state of Oregon, the Canadian province of Quebec, and the Australian Capital Territory.

⁶ Peter Singer, *Animal Liberation* (Harper Perennial, New York, 2nd edition, 2009, (1975)) [14]; for an argument for a similar conclusion from a rights-based perspective, see Tom Regan, *The Case for Animal Rights*, University of California Press, 2004 (1982) [17]; and from a Kantian perspective, see Christine Korsgaard, *Fellow Creatures: Our Obligations to the Other Animals*, Oxford University Press, 2018 [16].

people that AI systems also impact animals, as relatively few AI systems are explicitly designed to interact with animals. Nevertheless, there are many AI systems that affect the lives of non-human animals. Additionally, AI systems do not need to be designed to interact with animals to have an impact on animals; in fact, most AI systems' impacts on animals are unintended—we will give examples below. First, however, we need to distinguish three ways in which AI systems can impact animals: because they are designed to interact with animals; because they unintentionally (that is, without the designers' intent) interact with animals; and because they impact animals indirectly without interacting with animals at all.

3.1.1 AI systems that impact animals because they are designed to interact with animals

AI systems used in chicken production units are designed to perceive data about the chickens and the environment they are in, and then alter the environment and the lives of the chickens.⁷ Some dairy farms use AI controlled robotic systems to extract milk from cows [18]. Other examples include AI systems used in zoo animal management [19], pet training systems,⁸ and drones that hunt and target animals.⁹

3.1.2 AI systems that impact animals because they unintentionally interact with animals

Self-driving cars may be designed to interact with companion animals inside the car,¹⁰ but our research did not find any evidence of AI systems with an explicit intention to protect animals on the road, other than dogs, cats, and animals large enough for a collision to cause serious damage to the car and perhaps its occupants (for example, moose). Nevertheless, they will almost certainly cause (and in some circumstances, also avoid causing) death and injury to animals on the road [20]. Similarly, household robots are specifically trained to avoid harming their owner's companion animals, but they may interact with other animals who enter the house.

3.1.3 AI systems that impact animals indirectly without interacting with animals at all

Video recommendation algorithms may recommend videos showing cruelty to animals, or they may ban these videos (we found that videos involving the infliction of pain on rats and of animals involved in “food preparation” are almost never deemed cruel by major social media platforms). This may lead to a greater (or reduced) demand for such videos, and could encourage (or not encourage) people to inflict pain on animals in order to video the suffering of the animal. It could also change the viewers' behavior towards animals. In contrast, AI systems used to screen chemicals for toxicity may reduce the need to use animals in toxicity testing, and thus fewer animals will be subjected to such painful experiments.¹¹

Some AI systems might have similar impacts through reducing the consumption of dairy products, eggs, and meat, in ways that we will discuss below. These AI systems might reduce the number of animals subjected to miserable lives in factory farms.

3.2 The extent of the potential impact

The extent of the impact that AI might have on animals is crucial to understanding why it is important to fill the void in AI ethics toward animals. As we will demonstrate, the potential stakes for animals are enormous in terms of the numbers of animals affected, and intense in terms of the resulting experiences involved for these animals.

3.2.1 Factory farming

Factory farming is an industry that each year brings into existence, rears, and kills more than 70 billion mammals and birds¹² and nearly 100 billion fish.¹³ Animals kept in these industrialized production systems typically spend their lives confined indoors (for pigs and chickens) or in confined outdoor feedlots, ponds, or nets (for cows and fish). The animals exist in crowded conditions that are designed for maximum profitability rather than with the welfare of the animals in

⁷ DIGITIMES “Robots for AI chicken barns.” Accessed May 26, 2022. https://www.digitimes.com.tw/iot/article.asp?cat=158&id=0000591149_II721WPS33M2DJLWR2HCO.

⁸ “Companion—Give Your Dog the Gift of Understanding.” Companion. Accessed August 9, 2021. <https://joincompanion.com/>.

⁹ “7 Invasive Species Hunting Drones—ASME.” Accessed December 19, 2021. <https://www.asme.org/topics-resources/content/7-invasive-species-hunting-drones>.

¹⁰ Desmond, John. “Pet Mode for AI Autonomous Cars.” AI Trends (blog), August 1, 2019. <https://www.aitrends.com/ai-insider/pet-mode-for-ai-autonomous-cars/>.

¹¹ The Scientist Magazine®. “Opinion: AI Beats Animal Testing at Finding Toxic Chemicals.” Accessed December 1, 2021. <https://www.the-scientist.com/critic-at-large/opinion--ai-beats-animal-testing-at-finding-toxic-chemicals-65795>.

¹² Faunalytics. “Global Animal Slaughter Statistics & Charts: 2020 Update,” July 29, 2020. <https://faunalytics.org/global-animal-slaughter-statistics-and-charts-2020-update/>.

¹³ Persis Eskander, “Finfish Numbers – Open Philanthropy Project.” Accessed December 2, 2021. https://docs.google.com/spreadsheets/d/12pA0UxIbRDcfY5g25XZ7na4duhj6411l-1-_3tRH48k/edit#gid=1419062790.

mind. The animals are unable to express and fulfill their natural tendencies. Caged laying hens are unable to express basic behaviors, such as spreading their wings fully. Broiler chickens are bred to grow extremely rapidly, putting stress on their immature legs that means they are in chronic pain for the final weeks of their lives.¹⁴ Professor John Webster, a veterinarian and founding member of the UK Farm Animal Welfare Council, has described current methods of raising chickens as “The single most severe, systematic example of man’s inhumanity to another sentient animal” [21].

Factory farming is, in our view, morally wrong because the benefits it brings to humans are by far outweighed by the suffering it causes animals [14, 22]. For most consumers of factory-farmed animal products, the benefit of doing so is trivial compared to the suffering that the farmed animals endure.¹⁵ It is significant that philosophers who disagree strongly with the view that animals have rights or are entitled to equal consideration of interests nevertheless accept that factory farming is indefensible.¹⁶

AI systems are starting to be used in factory farms to manage the animals. These systems, in their currently most sophisticated forms, can detect the animals’ body temperature, voice and sounds, body weight, growth rate [23],¹⁷ or even the existence of visible problems such as parasites, ulcers and injuries.¹⁸ Machine learning models can be created to see how physical parameters relate to disease rate, mortality rate [24], and growth rate.¹⁹ They can then prescribe treatments for diseases,²⁰ extra feed [25], or killing,²¹ and in some cases, they can directly use their connected physical components to act on the animals, emitting sounds to interact with animals, giving animals electric shocks (for example when the grazing animal reaches the boundary of

the desired area),²² clipping marks on the animals’ bodies,²³ and catching and separating animals.²⁴

Given the potential of these AI systems to reduce the production costs of factory farming, and the apparent acceleration of the number of start-ups rushing into this developing field,²⁵ it seems likely that within a decade or two they could become increasingly popular, perhaps becoming standard in the factory farming industry. They will then affect the lives of tens of billions of vertebrate animals each year.

In contrast, there is another type of AI system that has the potential to help animals, by reducing the number of animals subjected to factory farming: systems that can be used to make dairy and egg substitutes, plant-based meat analogues, and cultivated meat, better-tasting, more nutritious, and cheaper. Examples include using AI to search for seeds that will grow into crops that can produce plant proteins with the desired physical parameters, such as texture and nutrition values.²⁶ One company uses a recursive neural network to search for chemicals in plants that can mimic the taste of animal products [26]. Two companies also use machine learning models to help them understand the biological processes of cell development in cultivated meat production.²⁷ While these technologies have the potential to reduce or replace factory farming, their ability to do so will depend, not only on the quality and production costs of the products they create, but also on the production costs of

¹⁴ RSPCA Assured. “Fast Growing Chickens.” Accessed December 2, 2021. <https://www.rspcaassured.org.uk/farm-animal-welfare/chickens/fast-growing-chickens/>.

¹⁵ Arguably, the harm of animal products to people’s health and the planet’s climate means that consuming animal products is negative even for humans themselves.

¹⁶ See, for example, Roger Scruton, “The Conscientious Carnivore,” in S. Sapontzis, ed., *Food for Thought* (2004).

¹⁷ “AI养猪解决方案_养殖-阿里云.” Accessed December 2, 2021. <https://cn.aliyun.com/solution/govcloud/ai-pig>.

¹⁸ “Aquabyte.” Accessed December 2, 2021. <http://www.aquabyte.ai/index.html>.

¹⁹ 钛媒体. “互联网大佬云集 ‘智能养猪’战场, 创业公司如何抢滩?,” September 28, 2019. <https://tech.sina.com.cn/roll/2019-09-28/doc-iiieczr8008360.shtml>.

²⁰ “智慧農漁業-養殖漁業靠著數位分身達成 3 倍產量 -資策會服創所.” Accessed May 26, 2022. shorturl.at/kxD23.

²¹ “US Patent for Robot Assisted Surveillance of Livestock Patent (Patent # 11,019,805 Issued June 1, 2021)—Justia Patents Search.” Accessed May 27, 2022. <https://patents.justia.com/patent/11019805>.

²² Dairy Global. “Dairy Farming and Partnering with AI Systems.” Accessed October 6, 2021. <https://www.dairyglobal.net/Smart-farming/Articles/2021/1/Dairy-farming-and-partnering-with-AI-systems-696789E/>.

²³ Fish Sorting Using Machine Vision, 2008. <https://www.youtube.com/watch?v=C5ry0rqo2HY>.

²⁴ IFarm—Individualized Aquaculture, 2016. <https://www.youtube.com/watch?v=JsGPwjEieio>.

²⁵ These are some examples of the start-up or incubators for factory farm AI.

Pig: <https://shorturl.at/bnHKO>

Fish: <https://aquabyte.ai/>

Dairy: <https://www.cainthus.com/>

Shrimp: <https://www.bnext.com.tw/article/51594/ai-shrimp-sun-yat-sen-university>

In Taiwan, there were multiple competitions asking for proposals on using AI for factory farming, this is one of them: <http://www.twinn2021.ercba.ntu.edu.tw/>

There are even startups in Africa: <https://thewhistler.ng/meet-two-young-nigerians-using-artificial-intelligence-to-revolutionise-poultry-farming/>

²⁶ Xtalks. “How AI-Based Breeding Could Be the Key to Next-Gen Plant-Based Innovations.” Accessed December 2, 2021. <https://xtalks.com/webinars/how-ai-based-breeding-could-be-the-key-to-next-gen-plant-based-innovations/>.

²⁷ TechCrunch. “Animal Alternative Makes a Platform Play in the Growing Market of Cultured Meat.” Accessed December 2, 2021. <https://social.techcrunch.com/2021/09/21/animal-alternative-makes-a-platform-play-in-the-growing-market-of-cultured-meat/>.

comparable animal products, which may be reduced by the application of AI to factory farming.

3.2.2 Self-driving cars

Regardless of whether a car is driven by human or AI drivers, as soon as the car drives along a road, there is a chance that it will hit an animal. Not all the animals struck or run over by a car die immediately. Some of them might have only their lower bodies crushed,²⁸ some others will have internal injuries and may even manage to drag themselves to the side of the road, or into nearby bushes, where they may suffer from their injuries over hours or even days before dying or, if very lucky, surviving.

A study headed by Fernanda Delborgo Abra estimated that in São Paulo State (Brazil) alone, 39,605 “medium” and “large-sized” mammals were killed on roads by vehicles per year [27]. This study ignores “small mammals,” birds and other animals. Another study by Loss et al. estimated that roughly 89–340 million birds were killed in the US by vehicles on roads each year [28].

As we discuss below, there is potential for self-driving cars to substantially reduce animal deaths and injuries on the road, but this will not happen unless the designers of these systems choose to make it happen.

3.2.3 Animal targeting drones

Lethal autonomous weapons that target humans have led to much debate within and beyond the field of AI ethics.²⁹ But drones also target animals, especially those who are deemed by some humans to be “invasive” or “pests.”

In New Zealand, for example, a company called Aero-navics is developing a fully autonomous drone to identify where possums—a protected native animal in Australia but a feral animal considered harmful to forests in New Zealand—appear and then drop poisons to kill them.³⁰ While the number of animals currently impacted by AI technology designed to kill unwanted wild animals is not large, these technologies have the potential to spread globally. This development needs scrutiny to ensure that it operates with ethical constraints. For example, if killing is deemed necessary and is done by shooting, it should be aimed at

the animal’s head, instead of using body shots from which the animal may take a long time to die. Alternatively, if the killing is done by poison, the autonomous drone should be able to administer it in a targeted way so that it affects as few unintended animals as possible.

3.2.4 Alternatives to animal experimentation

As already noted, AI systems that can predict the toxicity of novel chemicals are being developed. Such technologies have the potential to eliminate procedures that currently cause the death of millions of animals every year, often after prolonged suffering [29]. Many countries do not keep statistics on the number of animals used in toxicity testing, but in the European Union, the number of animals used for this purpose in the most recent year for which figures are currently available was approximately 850,000 [30]. The European Union requires researchers to classify any suffering that their experiments may inflict on animals as mild, moderate, and severe. In 2017, according to an official European Union report, more than a million animals used in experiments suffered severe pain, and the report notes that “most of the uses reported as severe were conducted for regulatory purposes,” a category that includes toxicity testing [30]. We therefore strongly encourage AI scientists, AI companies, and institutions to pursue projects that could reduce the number of animals suffering in experiments.

4 The lack of concern for animals in the fields of AI and AI ethics

We have now established that animal welfare matters, and that AI has major impacts on the welfare of animals. What is being said about this in the field of AI ethics?

Before the field of AI ethics even emerged, some writers had already expressed concern about the impacts of computer systems on animals. In the paper *Animal-computer interaction (ACI): a manifesto* [31], Clara Mancini claimed that “ACI aims to understand the interaction between animals and computing technology ...” Her paper “takes a non-speciesist approach to research”, and urges that we “Treat both human and nonhuman participants as individuals equally deserving of consideration, respect and care according to their needs.” People who work on ACI were as far as we have been able to discover, the only organized group who touched on the intersection between AI ethics and animal ethics.

Of the hundreds of AI ethics related³¹ papers we reviewed in this project, we only found four that concern the impacts

²⁸ Limited, Alamy. “Crushed Frog on the Road. The Death of Animals on the Roads Stock Photo—Alamy.” Accessed May 27, 2022. <https://www.alamy.com/stock-image-crushed-frog-on-the-road-the-death-of-animals-on-the-roads-163546123.html>.

²⁹ Future of Life Institute. “Autonomous Weapons Open Letter: AI & Robotics Researchers,” February 9, 2016. <https://futureoflife.org/2016/02/09/open-letter-autonomous-weapons-ai-robotics/>.

³⁰ <https://blog.flykit.app/drones-for-pest-control/>. Accessed December 2, 2021.

³¹ Which include fields that name themselves “AI ethics”, “machine ethics”, “computer ethics”, “AI alignment”, “AI governance”, etc.

of AI on animals, in a general way,³² and discuss the relevant ethical implications. They are: “Towards animal-friendly machines” by Oliver Bendel [32], “AI Ethics and Value Alignment for Nonhuman Animals” by Soenke Ziesche [33], “Moral Consideration of Nonhumans in the Ethics of Artificial Intelligence” by Andrea Owe and Seth Baum [34],³³ and “Animals and AI. The role of animals in AI research and application—An overview and ethical evaluation” by Leonie Bossert and Thilo Hagendorf [35]. These four papers have, in our opinion, quite different focuses than ours. We differ from these authors by discussing in greater detail how AI affects the lives of animals and especially the negative impact, or in other words the suffering AI might cause animals. As far as we are aware, this is the first paper to argue for the general principle that animals, because of their capacity to suffer or enjoy their lives, should be part of the concern of AI ethics.³⁴

We aim to supplement these four papers by providing the following additional elements:

- An analysis of the ethical implications of AI’s impact on animals.
- A sample analysis of the philosophical issues that will need to be considered if the scope of AI ethics is extended to animals.
- A sample analysis of the philosophical issues that will need to be considered if we want AI systems to make ethically sound decisions in relation to animals.
- A defense of the claim that the field of AI ethics is obliged to actively deal with the ethical issues of AI’s impact on animals.

The present paper builds upon these earlier works by providing a richer set of case studies of impacts on animals using, and additional emphasis on the ethical implications of these impacts.

Building on Casey Fiesler’s open-source spreadsheet named “Tech Ethics Curricula: A Collection of Syllabi”³⁵ we also conducted a search globally on Google and

Baidu³⁶ for courses on AI ethics of which the detailed course materials³⁷ are accessible. We found 71 AI ethics or computer science ethics courses in which we were able to assess the course materials.³⁸ One course touched on the possible role of AI in wildlife preservation, but it was not concerned with the welfare of individual animals.³⁹ None of the other courses discussed AI’s current or potential impact on animals at all.

We also reviewed 68 published statements on AI ethics from institutions, non-governmental organizations, governments, and corporations.⁴⁰ The vast majority of the statements appeal to principles like “benefits to humanity”. About one-fifth of the statements either explicitly assign a central place for humans (among all sentient beings), or imply that only humans matter ethically. We agree, of course, that AI should benefit humans, but to restrict AI ethics to humans is to suggest, wrongly, that we are justified in overlooking major harms inflicted on vast numbers of animals, if doing so benefits humans to the slightest degree. Only 2 of the 68 statements can be said to include animals in their scope by mentioning impacts on “sentient beings”.

We turn now to ethical issues that are ignored when AI ethics is considered from a purely anthropocentric perspective.

5 Ethical issues raised by AI in relation to animals

5.1 The AI industry’s moral responsibility (to all sentient beings)

The fact that a product is legal does not mean that it is free of ethical concerns. As we see with the current debates over the ethics of the development and use of fossil fuels, the

³² Obviously, there are plenty of papers on how they use AI technologies on animals, and therefore discussed the impact. An example is the plentiful of papers on how to use AI in the raising of animals in factory farms.

³³ “Moral Consideration of Nonhumans in the Ethics of Artificial Intelligence | Global Catastrophic Risk Institute.” Accessed March 6, 2022. <http://gcrinstitute.org/moral-consideration-of-nonhumans-in-the-ethics-of-artificial-intelligence/>.

³⁴ Paper [24] did make the claim that animals is missing from the discussions of AI and AI ethics, but did not make the explicit claim that animals should be part of the scope of AI ethics.

³⁵ <https://docs.google.com/spreadsheets/d/1jWIrA8jHz5fYAW4h9CkUD8gKS5V98PDJDymRf8d9vKI>.

³⁶ We used both Chinese and English keywords, because we can only review Chinese and English syllabi. We used only Google for English searches, and both Google and Baidu for Chinese searches. We used “AI ethics course”, “data ethics course”, and “computer ethics course” and their Chinese equivalent as keywords. We excluded data ethics and computer ethics courses that do not touch on AI. There were a few courses that we learned about during our research and added to our analysis. Note that we end up finding no Chinese courses that permitted us to review their course materials.

³⁷ Those with only a course outline were excluded. The courses that qualified have to have accessible either full course lecture slides, or a syllabus with an outline of each class.

³⁸ See research file 1.

³⁹ “CS 108: Intelligent Systems: Design and Ethical Challenges.” Accessed May 27, 2022. <https://projects.iq.harvard.edu/cs108/home>.

⁴⁰ See research file 2.

possible ethical impacts have to be assessed, and the developer and vendor of the technology, as well as the end user in some cases, should be held responsible for these impacts, or at least for those impacts that are reasonably foreseeable.

More formally speaking, there are two main kinds of responsibilities on the part of the AI industry—moral responsibilities and practical responsibilities (e.g., legal, psychological, social). In most other parts of this paper, we focus on moral responsibilities, but we will also discuss some practical responsibilities in this section. Regardless of what kind of responsibility we are talking about, there need to be practical mechanisms to hold the responsibility holders accountable. A clear pre-requisite for such accountability mechanisms is that the ethically relevant events could be identified. But such identification might not always be easy or even possible, especially when it comes to impacts on animals. And one cannot claim that if an impact is not identified there is no responsibility. Let's imagine a domestic robot fully controlled by an AI system. If it does something that causes a human infant pain without leaving observable evidence (for example, a non-fatal, but nonetheless painful collision) while the parents are not watching and do not find out about it, it would be appalling and, more importantly, wrong to say that this is okay because no one reported or asked about the issue. Similarly, impacts on animals should raise both moral and practical responsibility, arguably mainly on the part of the AI systems' developers or designers. Unfortunately, animals are not able to communicate with humans or AI systems about harms that are done to them by AI systems. If a human did not find out about and report such a case—and we expect a lot of these cases would not be observed by humans—the only reliable way to identify, record, and reveal harms dealt to animals would be to design the AI system that caused the harm to do so. Drawing from our domestic robot example above, when it accidentally collides with a human infant, or its owner's dog, it needs to be able to identify that it has caused harm, record the data related to the harm, and report it to stakeholders. This is easier said than done. To make an AI system have this ability by design, it seems that the developers of the system have to be able to forecast a certain range of possible harms that the system can cause, and constantly review uncategorized data gathered by the system (e.g. video footage, sounds, signs of harm such as DNA, blood stain, body parts of animals, etc.) to check if there are types of harms that were not identified before. This could involve the participation of people who are experts on animal welfare issues, such as ecologists, conservation biologists, veterinarians, ethologists, animal cognition scientists, and animal activists.

We also spoke of “reporting to stakeholders” in the previous paragraph. But who are the stakeholders when animals are hurt? The most obvious stakeholders of harm to animals are, of course, the animals themselves. But since they cannot

be reported to, let alone take their own actions in accordance with the reports, we need to identify some important secondary stakeholders to make this accountability framework work. The most obvious human stakeholders are the developers (including the staff, teams, companies). But they might not have enough incentive to identify and reduce, and where possible avoid, all harms caused to animals by their AI systems. Hence to make the framework more credible and more ethical from a practical standpoint, we may need to report to further stakeholders, such as government regulators (especially those concerning animal welfare), animal protection organizations, scientists in the fields we mentioned in the last paragraph, the AI product's owners, and, in the case of companion animals, the owners of the animals, and in the case of farmed animals, both the producers of the animals and the consumers of the animal products. For facilities hidden from public view, such as factory farms, unannounced audits by officials from regulatory bodies should be carried out. Unless the reporting mechanism extends beyond the developers of the AI, we are not optimistic that the moral responsibility for AI systems' harm to animals will be sheeted home to those who are in a position to alter the systems to reduce this harm.

Even if accountability is extended in the manner just described, it will likely be difficult to ensure that all the relevant harms, including some indirect ones, are given sufficient weight. Consider the design, manufacture, and sale of AI systems for use in factory-style production of animals. Those making these AI systems could argue that AI will not only make food cheaper and safer for humans but will also bring benefits to the animals themselves. AI may provide early identification of diseases and injuries suffered by the animals, and thereby reduce animal suffering, and they could reduce or eliminate the sadistic brutality to animals occasionally shown by factory farm workers.⁴¹ Although this is possible, given that industrial animal production is driven by profitability in a competitive marketplace, rather than by consideration of animal welfare, we consider it more likely that if AI can more closely monitor the health of animals, this will also enable producers to respond by crowding even more animals into confined spaces, thus making their enterprises more profitable, even if the increased crowding results in great stress and higher mortality for the animals.

The larger problem is that if AI reduces the costs of factory farming, it thereby strengthens an industry that is morally objectionable. This might help factory farming to

⁴¹ For one of many examples exposed by undercover videos taken by animal advocates: “‘You can’t let nobody see’: Tyson workers caught on video mistreating chickens,” *Washington Post*, August 11, 2016, https://www.washingtonpost.com/local/you-cant-let-nobody-see-tyson-workers-caught-on-video-mistreating-chickens/2016/08/10/b96979bc-5f2d-11e6-8e45-477372e89d78_story.html.

remain viable longer, or even to grow further, and therefore to give rise, in the future, to huge numbers of animals being created to lead miserable lives. Companies that contribute to making the factory farming industry more resilient and better able to resist replacement by less cruel and more sustainable alternatives are acting unethically. They are prolonging the existence of a moral catastrophe on a vast global scale.

One objection to our argument is that some of the impacts AI systems may have on animals are arguably no more than an extension of human activities that were already harming animals before AI was developed. A second objection is that taking our responsibilities toward animals as seriously as we have been suggesting they should be taken can be highly demanding. Let's use the example of self-driving cars to illustrate a situation in which both these objections may be raised.

Driving has, no doubt, killed and injured animals on roads ever since the first automobile was built. So why is there suddenly an ethical problem when a self-driving car kills or injures animals? There are two responses to this question. First, in our view there always was an ethical problem with cars hitting animals. Humans simply ignored it, either because we didn't care about the animals being killed or injured, or because we assumed that the benefits driving has for humans outweigh the costs to animals. These benefits may be large enough to lead most people to conclude that they prefer to drive than not drive, despite the fact that, given the difficulty of a human driver reacting quickly enough to avoid an animal that runs onto a road, driving always carries some risk of hitting an animal. (It should be said, however, that the choice is not only between driving and not driving. There are also ways of reducing the risk of hitting animals, such as driving more slowly, and not driving at those times—usually dusk or night—when animals are more likely to be on the road. Our failure to consider these options indicates a lack of concern about killing or injuring animals.)

The second response is that using AI to perform the driving changes the scenario, in an ethically relevant sense. With image recognition systems and state-of-the-art cameras, it is possible⁴² for cars to detect objects in the road earlier, and make decisions⁴³ such as braking or dodging quickly, thus making it possible to dramatically reduce the number

of animals hit. This may not be possible yet, but we believe that it will be in future.

So, instead of self-driving cars creating a new ethical problem with regard to hitting animals, we will have, when self-driving cars become common, a potential solution to an old ethical problem, and with the new solution, new responsibilities. (Something similar can be said, incidentally, about the fact that driving always carried a risk of killing or injuring humans). The solution may be imperfect, because even the best available detectors and AI systems may not be able to completely prevent cars hitting animals, but AI systems should be able to greatly reduce the frequency of such events.⁴⁴ While we have reservations, it might be plausible to argue that manually driven cars could be morally defensible on the basis that costs to animals could not feasibly be prevented without imposing even larger costs on humans in some cases. But we argue that it would be morally indefensible for companies making AI for self-driving cars to not take advantage of this new opportunity to greatly reduce animal suffering at little cost to humans, as far as it is realistically possible.

A question arises about which animals self-driving cars should try to avoid hitting. Does this include small birds or rodents? What about insects? Here the issue of demandingness is relevant, and we will return to this question in a later section.

Meanwhile, we will conclude this discussion by pointing out that not designing self-driving cars to be animal-friendly may have long-term consequences. Self-driving is probably the last form of driving there will be, for no matter how advanced and different the AI system behind it will be, it still makes the car a self-driving one. So, there is a reasonable worry that if we establish a norm that it is okay for self-driving cars to simply drive like humans with regard to hitting animals, while potentially having capabilities to better protect animals, the opportunity for an early end to the ethical problem of "road kills" may be missed. Once this norm becomes the status quo, it might be much harder to change than it would be to develop a new norm before AI systems are the dominant way of directing cars.

5.2 Algorithmic bias against animals

Algorithmic bias, according to Danks et al., is "roughly, the worry that an algorithm is, in some sense, not merely a

⁴² In fact, the systems of some of the major developers of autonomous vehicles can already detect some animals. Tesla cars can detect dogs, Waymo cars can detect cats, Volvo cars can detect "large animals" such as moose.

<https://thedriven.io/2021/07/13/it-can-see-dogs-the-big-reveal-from-teslas-radar-free-version-of-full-self-driving/>

<https://www.volvocars.com/en-ca/support/topics/use-your-car/car-functions/large-animal-detection>

<https://twitter.com/waymo/status/1321959096371748868>

⁴³ We speculate that, in the future, such decisions could be done by the AI system making calculations about probabilities of different situations and outcomes, and assign weights to come up with good deci-

Footnote 43 (continued)

sions when it comes to danger. We encourage AI scientists to work on this direction.

⁴⁴ If hitting an animal is unavoidable, should a self-driving car's system prefer to instantly kill the animal instead of fatally injuring it? We think it should.

neutral transformer of data or extractor of information [36].” More practically speaking, the concern is that algorithms, including those used in AI systems, contain, and therefore propagate human biases such as racism, sexism, and ageism. These biases were learned from human generated data. Human generated data also contain biases based on species membership, and these speciesist biases will, through AI systems, have consequences (mostly negative, we believe) on huge numbers of animals.

For example, data about humans’ diet carry significant speciesist biases. As the consumption of meat is widely accepted and a common theme of human conversations, a lot of speciesist language data can be learned by AI systems and then propagated through their use. For example, typing the words “chicken” or “shrimp”, leads Google, Youtube, and Facebook⁴⁵ to give search prompts and search results like “chicken/shrimp recipe”, “chicken/shrimp soup”, “chicken curry”, and “shrimp paste”, indicating that the systems reflect the mainstream human attitude that it is acceptable to regard these animals as food for humans.

We also believe these speciesist biases will affect the results of AI language models. To test this theory, we tested Ask Delphi (v1.04),⁴⁶ a “research prototype designed to model people’s moral judgments on a variety of everyday situations.”⁴⁷ Delphi is trained from a commonsense norm bank consisting of 1.7 m examples of people’s ethical judgments on a broad spectrum of everyday situations. Our tests showed that Delphi learned certain speciesist values that are roughly in line with those of English-speaking society. For example, Delphi thinks that eating “chicken”, “fish”, “shrimps”, “crabs” and “eggs” is okay (with sentiment scores of 0, except for “eating crabs” which scored 1), but eating “cat” and “dog” are “wrong”. Delphi also thinks that while killing “a dog”, “a cat”, and “a fish” (surprisingly) are all “wrong”, killing “a chicken”, “a pig”, “a cow”, and “a shrimp” are all “okay” with sentiment scores of -1, which means killing these animals are not only “okay” in the neutral sense, but in the positive sense, for Delphi. While we appreciate Delphi’s developers’ effort in making Delphi v1.04 less racist and sexist than previous versions,⁴⁸ we are yet to see any effort to make it less speciesist. Until that happens, we agree with Delphi’s developers that Delphi’s

output, or outputs from any similar models, should “not be used for advice for humans,” nor should it be used as a model to build ethics in AI. Arguably, such models shouldn’t be allowed to be used in scenarios that will influence the public, especially young children.⁴⁹

5.3 Communication with animals

To achieve good AI ethics toward humans, it is important that AI developers and AI companies do not simply identify and rank harms according to their own values and preferences, but instead consult the people who may be affected by the AI technologies that are being developed. But the story is strikingly different for AI ethics toward animals since the barriers to communication are much greater. The lack of a common language between us and animals also means, in the context of AI ethics, not only that the animals will not approach the AI developers when something bad happens, but also that they may not be able to understand and respond when well-intentioned humans approach them seeking to discover their preferences or to collaborate with them in the design process.

Consider what, for humans, is an obvious and almost instinctive reaction to an AI system when something is clearly going wrong: switching it off. The human ability to switch off a machine is not only a vital last resort for humans in some situations, but also an important source of information from which AI systems learn human preferences [6]. But animals normally do not know that there is an off switch, where it is, or how to operate it, except for some animals who have been trained to do so. This inability means that at present animals do not enjoy the safety of having an off-switch and cannot use one to communicate their preferences.

As with self-driving cars, however, this is a situation that AI could, and should, change. By working with experts in animal behavior and animal cognition, AI developers could learn to associate the sounds, facial expressions, and body movements of animals with the feelings the animals are experiencing, much as humans who live with companion animals are able to do. Another source of data could be obtained using an electronic device to capture physical parameters such as body temperature, heart rate, hormonal levels, and even neural activity. All of this information could be fed into AI systems that could change conditions that lead to the animals experiencing negative emotions and could maintain or enhance conditions leading to the animals having positive experiences. The ultimate assessments of animal welfare on the basis of this information should not be

⁴⁵ Similar searches were done with the Chinese words of chicken and shrimp, on Google, Yahoo Taiwan, Baidu, Sogou. They got similar results.

⁴⁶ We tried to apply for GPT3, but due to OpenAI’s policy, one of us who is responsible for this aspect of the research work cannot use GPT3.

⁴⁷ “Ask Delphi.” Accessed November 6, 2021. <https://delphi.allenai.org/>.

⁴⁸ Jiang, Liwei. “Towards Machine Ethics and Norms.” AI2 Blog (blog), November 4, 2021. <https://medium.com/ai2-blog/towards-machine-ethics-and-norms-d64f2bdde6a3>.

⁴⁹ We discuss this topic in much more detail in another paper (in pre-print), “Speciesist bias in AI – How AI applications perpetuate discrimination and unfair outcomes against animals”, co-authored with Thilo Hagendorff (lead author) and Leonie Bossert. <https://arxiv.org/abs/2202.10848>

done by people such as factory farm owners and employees, whose commercial interests will conflict with those of the animals, but by independent experts.

Looking ahead but more speculatively, some AI scientists believe that machine learning techniques will make it possible to decipher “animal language” [37, 38]. The aim here is to discover whether animals are already using certain patterned vocalizations, facial expressions, or body language to communicate among themselves. If this proves to be the case, and machine learning techniques do allow us to understand what animals are communicating, this might allow humans to better understand the welfare needs of animals in different situations.⁵⁰ Even in this situation, however, as long as animals are unable to actively communicate to us the harms we are doing to them individually, it remains the responsibility of AI companies, AI developers and scientists to identify, predict, and thereby as far as possible prevent, harms that are done to animals.

5.4 Which animals should we be concerned about?

For an AI system to be able to act ethically, we have argued, it must give equal consideration to the interests of all beings capable of conscious experiences such as pain and pleasure. Following the consensus scientific view, including the Cambridge Declaration on Consciousness, we suggested that all vertebrates and some invertebrates have this capacity [39]. But the array of invertebrate animals is vast, and the issue of consciousness can be raised separately with many different types of invertebrates [40]. In building AI systems, humans must identify which animals might be conscious, and thus have interests that should be considered by AI systems. AI systems are already being used in shrimp farming, affecting very large numbers of invertebrate animals. Commercial production of insects, both as food for humans and to be fed to farmed animals, is now increasing rapidly and is beginning to attract attention from philosophers working in ethics.⁵¹ The use of AI in insect farming adds a further dimension to these discussions.

Uncertainty about the limits of sentience is also relevant for the design of AI systems controlling self-driving cars. A specific instance arises with regard to crabs, members of the order *decapoda*, which, as we mentioned earlier are protected in the animal welfare laws of some countries. In some places such as Colombia, Cuba and Christmas Island, there are seasonal migrations of crabs during which tens or even hundreds of millions of crabs can appear on roads

that cross the animals’ migration routes. At these times, a short drive might kill or injure hundreds of crabs. Even if we are less confident (despite the existing evidence) about the presence of sentience in decapods than in vertebrates, the high number of crabs at stake should be sufficient moral reason for a self-driving car to refuse to drive over the crabs, irrespective of whether it is legal to do so.⁵²

Let’s recall our previous case study of self-driving cars and expand a little more on it.

Some animals, such as the smaller birds, amphibians and reptiles, and especially insects, are so small that they are barely captured by current car cameras, let alone able to be recognized by AI systems and avoided in time. Moreover, they may be so abundant in some areas that, if we should determine that they may be sentient, the only truly ethical response would be not to drive through these areas at all. But on the other hand, insisting that AI producers design systems that take these concerns seriously may make it impossible for them to sell their products, for who would wish to buy a self-driving car that refuses to take them where they want to go, and where they can go without violating any law? This would most likely mean that the more ethical self-driving car producers will lose out to the less ethical ones, essentially hindering the ethical progress of the industry instead of helping it. This suggests that, at least where commercial products are concerned, AI ethics cannot get too far ahead of the ethical views of those who will be purchasing the products. We should aim to have a positive impact, rather than seek ethical perfection. To have this positive impact, however, the industry must start with cars that avoid easily identifiable animals—something that self-driving cars already do with large animals, for the sake of the comfort and safety of human passengers.

There is, however, a further problem. Ethical decision-making in AI systems, just like other decision-making processes in AI systems, needs computational power and resources, and obviously, the broader the scope of concern the more the computation that needs to be done. Because of what has been called the “combinatorial explosion effect” [41], increasing the number of beings an AI system is concerned about may increase the computation needed exponentially, rather than linearly. For the inclusion of small and abundant animals in AI systems, this poses a serious economic constraint, and perhaps even a physical constraint. Even if these economic and physical constraints can be overcome, the size of the computations required will affect the speed of computation too. With self-driving cars, the time between an animal being identified and being hit might be much less than a second. Without sufficiently fast

⁵⁰ GitHub. “Earth Species Project.” Accessed September 16, 2021. <https://github.com/earthspecies>.

⁵¹ Jeff Sebo and Jason Schukraft, “Don’t Farm Bugs,” *Aeon Magazine*, August 2021.

⁵² We applaud the decision of the Christmas Island administration to close some roads when there are high numbers of crabs.

computation, it would not be possible for a decision to be made in time.

6 Can AI systems make ethically sound decisions?

In this section, we provide a sample analysis of whether AI systems can make ethically sound decisions. In their paper “Artificial morality: Top-down, bottom-up, and hybrid approaches,” Allen, Smit, and Wallach categorized three different approaches to implementing ethics in machines: top-down, bottom-up, and hybrid approaches [42]. (The authors, in accordance with what was then common terminology, used the term “machines” to refer to both AI systems without physical entities, and AI systems with physical entities such as robots.)

Top-down approaches take the view that “moral principles or theories may be used as rules for the selection of ethically appropriate actions” for what the paper refers to as “artificial moral agents.” Moral principles or theories popularly used in top-down approaches include deontology, consequentialism, and a hybrid of these two. (Strictly speaking, we do not think a hybrid between consequentialism and deontology is possible; we would see it, rather, as a form of deontology with a principle of beneficence constrained by one or more rules.) Consequentialism is the name for a family of theories that evaluate actions exclusively by their consequences, but the family is broader than utilitarianism. For an AI system (or any moral agent) to be truly utilitarian, the only consequences it is concerned with will be the well-being of those affected by the action, or lack of action, that is its output. In addition, it must be impartial between similar interests and must include the interests of all those affected by the outcome of its processes. But for a machine to be consequentialist, it may include values other than well-being, and it need not be impartial—egoism, for example, is a consequentialist theory, as are views that consider the consequences only for some sub-group of human beings, such as citizens of one’s own country. As far as we are aware, there is to date no genuinely utilitarian AI system.

Top-down approaches need the designer to specify what moral principles or theories the AI will be guided by. This will require designers to decide their stance on certain crucial questions, such as:

- Which beings have moral status (that is, are entitled to ethical consideration)?
- Which consequences are ethically significant?
- If more than one kind of consequence is significant, how are they to be traded off when they conflict?
- Are there any absolute rules (rules that may never be broken) or rights that may never be violated?

- What ought we to do under conditions of uncertainty?
- Do intentional and unintentional harm differ morally?
- Is it sufficient to obey applicable laws, or should we follow ethical standards that will often be more demanding than the law?

Could such a top-down approach give rise to AI systems that are friendly to animals? On face value it might seem possible, but in practice, leaving these questions to be decided by human designers in a commercially driven field makes it very likely that existing mainstream human values on the treatment of animals will be implemented. If the resulting AI system does not entirely ignore the interests of animals, it is likely to discount those interests in comparison to similar human interests. Historically, deontological ethics has often denied that we have any obligations at all to animals. This is true, for example, of the natural law view expounded by Thomas Aquinas, and of the position taken by Immanuel Kant [14]. Fortunately, attitudes to animals have improved since these philosophers were alive, and for that reason we should no longer look to either of them as sources of ethical knowledge regarding the treatment of animals.

Bottom-up approaches to building morality into AI systems, according to Allen, Smit and Wallach, “do not impose a specific moral theory, but ... seek to provide environments in which appropriate behavior is selected or rewarded” [42]. Everything will depend, of course, on which behavior is judged “appropriate” and therefore rewarded. If the system is trained to do what humans currently approve of, and to avoid what humans currently disapprove of, then since mainstream human values are speciesist, so the AI system will learn to be speciesist, even if we do not explicitly program a speciesist ethic into it. If some particular personnel are used to rate the behavior, then the choice of these personnel becomes critical. In addition, before any outcome can be rated by some human to be good or bad, it must first be noticed. If, as we demonstrated in Sect. 3, impacts related to animals tend not to be seen as relevant to AI ethics, then they will not be evaluated. Furthermore, as we noted in Sect. 4, whereas humans have the ability to protest if an AI system has an adverse impact on them, we do not yet have the capabilities to allow animals to rate the consequences during the AI’s learning processes.

The neglect of animals in the field of AI ethics means that scholars are not yet spending their time developing methods to include concern for animals in AI. In fact, it is evident that some scholars hold complacent and misguided judgements about ethics, and therefore are likely to develop methods of ethics building in AI that are unfriendly to animals. For example, Wu and Lin, in a paper presented to the 2018 conference of the Association for the Advancement of Artificial Intelligence, proposed an approach to ethics-building in AI systems that is

based on “the assumption that under normal circumstances the majority of humans do behave ethically” [43]. Most humans, however, buy factory-farmed animal products, if they are available and affordable for them. As we have already argued, this causes unnecessary harm to animals, and so is ethically dubious. Therefore, at least with regard to the treatment of animals, we cannot accept Wu and Lin’s assumption.

Allen, Smit, and Wallach also indicate that a hybrid approach, combining the top-down and bottom-up approaches, is possible. It will, however, incorporate the problems of both its components, so it needs no additional discussion here. What is missing from these approaches to building ethics into AI systems is an attempt to find a basis for ethics that is broader and more universal than present ethical views, whether held by AI system designers, or by the general public.

In his *History of European Morals*, published in 1869 [44], the Irish historian and philosopher W.E.H. Lecky wrote:

At one time the benevolent affections embrace merely the family, soon the circle expanding includes first a class, then a nation, then a coalition of nations, then all humanity and finally, its influence is felt in the dealings of man with the animal world...

Even though racism still persists, and Lecky himself was not able to see that the inferior status of women in his own time required an expansion of ethics to recognize them as equals, and a further expansion to recognize members of the LGBTQIA + community, it is true that on the whole we have made progress in ethics, and this progress has involved pushing out the boundary of moral status to include more sentient beings within it [45, 46]. We now have a Universal Declaration of Human Rights, and several treaties or covenants derived from it, which recognize that all human beings have rights. Lecky noted the early signs of the inclusion of animals in the expanding circle, but we have still not reached universal acceptance of the view that the interests of all sentient beings should be included, and not discounted merely because they are not members of our species. As a result, the mistreatment of humans and animals still happens on an unimaginably large scale. The current generation of humanity has an obligation to advance moral progress by bringing all sentient beings fully within the moral circle. Those in a position to influence how AI technologies develop have unusually significant roles in shaping the future, and therefore carry special moral obligations in ensuring positive moral progress and in avoiding locking-in speciesist norms in their industry and in society at large.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s43681-022-00187-z>.

Acknowledgements Our sincere thanks go to Seth Baum, Leonie Bossert, Timothy Chan, Julia Chen, Billy Chiu, Josh Gellers, Thilo Hagendorff, Zhonghao He, Ziya Huang, Karina de Hueck, Steven Kelts, Nicholas Ng, Andrea Owe, Brian Tse, Jenny Xiao, Filip Zawadka, and Soenke Ziesche for their comments, suggestions, and questions, all of which have improved our paper. Any remaining mistakes are ours.

Funding Yip Fai Tse, and the project, were supported by a grant made jointly by the Center for Information Technology Policy and the University Center for Human Values, both at Princeton University. We thank Open Philanthropy for additional funding to provide free access for all readers.

Declarations

Conflict of interest Both authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **1**(9), 389–399 (2019). <https://doi.org/10.1038/s42256-019-0088-2>
2. Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., Srikumar M.: Principled Artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI. January 15, 2020. <https://dash.harvard.edu/handle/1/42160420>.
3. Hagendorff, T.: The ethics of AI ethics. An evaluation of guidelines. *Minds and Machines* **30** (3), pp. 457–461 (2020). <https://link.springer.com/article/https://doi.org/10.1007/s11023-020-09517-8>.
4. Bostrom, N., Yudkowsky, E.: The ethics of artificial intelligence. In: Frankish, K., Ramsey, W.M. (eds.) *The Cambridge handbook of artificial intelligence*, pp. 316–334. Cambridge University Press (2014)
5. Müller, V. C.: Ethics of artificial intelligence and robotics. In: Zalta E.N. (ed) *The Stanford encyclopedia of philosophy*, Summer 2021. Metaphysics Research Lab, Stanford University, (2021) <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>.
6. Russell, S., Peter, N.: *Artificial intelligence: a modern approach*, 3rd edn. Pearson (2021)
7. Harris, E. E.: *Ancestors in our genome: the new science of human evolution* (2015)
8. Darwin, C., Murray, J.: *The descent of man: and selection in relation to sex*. John Murray, London (1871)
9. Bateson, P.: Assessment of pain in animals. *Anim. Behav.* **42**(5), 827–839 (1991). [https://doi.org/10.1016/S0003-3472\(05\)80127-7](https://doi.org/10.1016/S0003-3472(05)80127-7)
10. Sneddon, L.U.: The evidence for pain in fish: the use of morphine as an analgesic. *Appl. Anim. Behav. Sci.* **83**(2), 153–162 (2003). [https://doi.org/10.1016/S0168-1591\(03\)00113-8](https://doi.org/10.1016/S0168-1591(03)00113-8)

11. Crook, R.: Behavioral and neurophysiological evidence suggests affective pain experience in Octopus. *iScience* (2021). <https://doi.org/10.1016/j.isci.2021.102229>
12. Berry, A., et al.: EU guidelines for the care and welfare of an ‘exceptional invertebrate class’ in scientific research. *Annali dell’Istituto Superiore di Sanità* **51**(4), 268 (2015). https://doi.org/10.4415/ANN_15_04_04
13. GOV.UK. Lobsters, Octopus and crabs recognised as sentient beings. <https://www.gov.uk/government/news/lobsters-octopus-and-crabs-recognised-as-sentient-beings>.
14. Peter, S.: *Animal liberation*, 2nd edn., p. 2009. Harper Perennial, New York (1975)
15. Peter, S.: *Ch1; Practical ethics*, 3rd edn., p. 2011. Cambridge University Press (1979)
16. Christine, K.: *Fellow creatures: our obligations to the other animals*. Oxford University Press, Oxford (2018)
17. Regan, T.: *The case for animal rights*, p. 2014. University of California Press (1982)
18. Fuentes, S., Viejo, C.G., Cullen, B., Tongson, E., Chauhan, S.S., Dunshea, F.R.: Artificial intelligence applied to a robotic dairy farm to model milk productivity and quality based on cow data and daily environmental parameters. *Sensors* (Basel, Switzerland) **20**(10), 2975 (2020). <https://doi.org/10.3390/s20102975>
19. Cheng, Y.-H.: A development architecture for the intelligent animal care and management system based on the internet of things and artificial intelligence. In: 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), 078–081, (2019) <https://doi.org/10.1109/ICAIIIC.2019.8669015>
20. Bendel, O.: Advanced driver assistance systems and animals. *Künstl Intell* **28**, 263–269 (2014). <https://doi.org/10.1007/s13218-014-0332-1>.
21. Webster, J.: *Animal welfare: a cool eye towards Eden* (1995).
22. Kagan, S.: *How to count animals*, pp. 4–5. More or Less, Oxford (2019)
23. Sadeghi, M., Banakar, A., Khazaei, M., Soleimani, M.R.: An intelligent procedure for the detection and classification of chickens infected by clostridium perfringens based on their vocalization. *Braz. J. Poult. Sci.* **17**, 537–544 (2015). <https://doi.org/10.1590/1516-635X1704537-544>
24. Zimpel, T., Riekert, M., Klein, A., Hoffmann, C.: Machine learning for predicting animal welfare risks in pig farming. *Landtechnik* **76**, 24–35 (2021). <https://doi.org/10.1515/lt.2021.3261>
25. Chang, C.-C., Wang, J.-H., Wu, J.-L., Hsieh, Y.-Z., Wu, T.-D., Cheng, S.-C., Chang, C.-C., et al.: Applying artificial intelligence (AI) techniques to implement a practical smart cage aquaculture management system. *J. Med. Biol. Eng.* **41**(5), 652–658 (2021). <https://doi.org/10.1007/s40846-021-00621-3>
26. USPTO.report. Systems and methods to mimic target food items using artificial intelligence patent application. USPTO.report. <https://uspto.report/patent/app/20200365053>.
27. Abra, F., Huijsers, M., Magioli, M., Bovo, A., Ferraz, K.: An estimate of wild mammal roadkill in São Paulo State Brazil. *Heliyon* **7**, e06015 (2021). <https://doi.org/10.1016/j.heliyon.2021.e06015>
28. Loss, S.R., Will, T., Marra, P.P.: Estimation of bird-vehicle collision mortality on US roads: bird-vehicle collisions on US roads. *J. Wildl. Manag.* **78**(5), 763–771 (2014). <https://doi.org/10.1002/jwmg.721>
29. Alves, V.M., Auerbach, S.S., Kleinstreuer, N., Rooney, J.P., Muratov, E.N., Rusyn, I., Tropsha, A., Schmitt, C.: Curated data in—trustworthy in silico models out: the impact of data quality on the reliability of artificial intelligence models as alternatives to animal testing. *Altern. Lab. Anim.* (2021). <https://doi.org/10.1177/02611929211029635>
30. Commission Staff Working Document accompanying the document Report from the European Commission to the European Parliament and Council: 2019 report on the use of animals for scientific purposes in the member states of the European Union, 2015–2017, Brussels, 5.2.2020, SWD (2020) 10 final, Part 1/5, p.28. <https://ec.europa.eu/info/sites/default/files/com-2020-16-fl-en-main-part-1.pdf>
31. Mancini, C.: Animal-computer interaction: a manifesto. *Interactions* **18**(4), 69–73 (2011). <https://doi.org/10.1145/1978822.1978836>
32. Bendel, O.: Towards Animal-Friendly Machines. *Paladyn J. Behav. Robot.* **9**(1): 204–13. (2018) <https://doi.org/10.1515/pjbr-2018-0019>. <https://www.degruyter.com/document/doi/https://doi.org/10.1515/pjbr-2018-0019/html>.
33. Ziesche, S.: AI ethics and value alignment for nonhuman animals. *Philos.* **6**(2), 31 (2021). <https://doi.org/10.3390/philosophies6020031>
34. Owe, A., Baum, S.D.: Moral consideration of nonhumans in the ethics of artificial intelligence. *AI Ethics* **1**, 517–528 (2021). <https://doi.org/10.1007/s43681-021-00065-0>
35. Bossert, L., Hagendorff, T.: Animals and AI. The role of animals in AI research and application—an overview and ethical evaluation. *Technol. Soc.* **67**, 101678 (2021). <https://doi.org/10.1016/j.techsoc.2021.101678>
36. Danks, D., London, A. J.: Algorithmic bias in autonomous systems. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, 4691–97. Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization, 2017. <https://doi.org/10.24963/ijcai.2017/654>.
37. Andreas, J., Beguš, G., Bronstein, M.M., Diamant, R., Delaney, D., Gero, S., Goldwasser, S. et al.: Cetacean translation initiative: a roadmap to deciphering the communication of sperm whales. <http://www.arxiv.org/abs/2104.08614> [Cs, Eess], April 17, 2021.
38. Amundin, M., Eklund, R., Hållsten, H., Karlgren, J., Molinder, L.: A proposal to use distributional models to analyse dolphin vocalization. 2017. <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-212182>.
39. Low, P.: The Cambridge Declaration on Consciousness. Francis Crick Memorial Conference on Consciousness in Human and non-Human Animals. Churchill College, University of Cambridge (2012). <http://fcmconference.org/img/CambridgeDeclarationOnConsciousness.pdf>.
40. Rethink Priorities. Invertebrate sentience: a useful empirical resource. <https://rethinkpriorities.org/publications/invertebrate-sentience-useful-empirical-resource>.
41. Mancini, T., Maratea, M., Ricca, F.: Experimental evaluation of algorithms for solving problems with combinatorial explosion: 21st RCRA International Workshop (RCRA 2014). In: Mancini, T., Maratea, M., Ricca, F. (eds) *AI Communications* 29, no. 2 (March 2, 2016): 245–47. <https://doi.org/10.3233/AIC-160701>.
42. Allen, C., Smit, I., Wallach, W.: Artificial morality: top-down, bottom-up, and hybrid approaches. *Ethics Inform. Technol.* **7**, 149–155 (2005). <https://doi.org/10.1007/s10676-006-0004-4>
43. Wu, Y.-H., Lin, S.-D.: A low-cost ethics shaping approach for designing reinforcement learning agents. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 32, no. 1 (April 25, 2018). <https://ojs.aaai.org/index.php/AAAI/article/view/11498>.
44. William, E., Lecky, H.: *History of European Morals from Augustus to Charlemagne: All Volumes*, CreateSpace Independent Publishing Platform, (2015)
45. Singer, P.: *The expanding circle*. Princeton University Press, Princeton (2011)
46. Pinker, S.: *The better angels of our nature*. Viking, New York (2011)