

Review of: "Do Androids Dread an Electric Sting?"

Kamil Muzyka¹

¹ Polish Academy of Sciences

Potential competing interests: No potential competing interests to declare.

Do androids review

The work presented aims to make a case of considering animal welfare and rights as a basis for future AI law and philosophy. It presents the current state of animal welfare, AI technology which includes especially GenAI, the concept of consciousness and attempts to provide the reader with a visible connection between the basis for animal welfare and AI welfare. What is especially noteworthy is the step by step process of going through milestones for AIs and its implications in the legal and societal context. However, in the opinion of the reviewer, the text itself needs to add at least a paragraph of further elaboration on industrial level of exploitation of AIs (working as basis for commercial products and treated as means of production), and a paragraph on the thought experiment regarding conscious AIs - or on how invisibility (invisible women and social invisibility) might add to neglect, abuse and maltreatment of AIs.

The review has been passed positively, however, there are some remarks, that in the eyes of the reviewer would improve the manuscript and help the authors correct errors or make a better case for AI welfare.

Major points to be considered.

+The term "welfare goods" has a completely different meaning in animal studies. It would be wise to consider a different term for welfare rights or basic legal protection under AI welfare.

(<https://www.frontiersin.org/articles/10.3389/fanim.2022.930930/full>)

+It is implied that animal welfare and animal rights stem from the natural philosophy of law, thus they seek to provide non-humans with legal protection against cruelty based on our current understanding animal behavior, biology and seeing similarities between our suffering and theirs. However, these acts can be viewed mostly as acts of Positive law, where the legislator decides on the categories and arbitrary catalogues of protected species and clades. Thus, as recognized by the authors, these protections seem "jagged" and that is the legislative body's decision and governmental policy to include or or exclude certain groups (like fish) from protection granted by national welfare laws. The example of fish abuse might be an excellent point to hammer down and nail the comparison even more in regards to AI that is unable to express feelings and sensations by design.

+If Chat GPT gains consciousness and sense of self - can it feel harm? It might understand the context or subtext, but is consciousness without pain receptors susceptible to psychological pain? It is understandable, that it is only a thought

experiment, aiming to show reasons for protection of any entity against malicious behavior. The building blocks include the element of embodiment. Without the physical or quasiphsical substrate of nocipercetion, will the information be treated as pain, or just information, which is being categorized as “pain” by internal programming?

+The problem with animal welfare in relation to AI is the liablilty and responsibility for damage generated by AIs going rogue (like animals on rampage). Therefore, in the view of the authors, who will be the party liable for AIs acting in self defence or going “berserk” from maltreatment?

Suggestions

- in introduction: add a reference to at least one instance of criticizm of the Apollo program or flying machines pre-Kittyhawk (suggested - Arthur C Clarke, Profiles of the future, 1960)
- add a reference atthe end of the dogs and chase better than the catch paragraph with a study or two on dopamine rush
- add references or mention Kate Darling’s study on human empathy towards machines that imitate animals (pets) and dinosaurs.
- The part on roomba’s might benefit from referencing Cory Doctorow’s “I Row-boat”. A hard sc-ifi story, which nonetheless deals with scenarios of robots opting out of consciousness. Further suggested readings - Rudy Rucker “Software”, as well as his non-fiction Infinity and the mind.
- Page 7, paragraph 2, with regards to reference 68 requires editing (missing entity responsible for identifying the building blocks)