

Jasper Schellekens

When Technology Decides: A Normative Approach to Assigning Responsibility in Machines that Learn

People make choices and they are accountable for these choices, but when an algorithm makes the choice it is unclear where the onus of responsibility should lie. Research as well as legal cases related to autonomous decisions has focused on autonomous vehicles, whether this be self-driving cars or drones as their decisions have direct and obvious impact on the physical world (Awad et al. 2018, Sparrow 1994). However, algorithms' decisions already impact millions of people each day in more subtle ways (Kasperkevic 2015, Iqbal 2018, Spielkamp 2017) and particularly in machines that learn there can be a lack of transparency and some unpredictability (Augustsson et al. 2002, Chu et al. 2017).

Using ethics of technology related to accountability (Nissenbaum 1994) to inform a normative legal approach, this paper examines where the moral responsibility lies for harm caused by AI that learns (i.e. machine learning) with a view on understanding how legal systems can address these challenges. Building on work by Vladek (Vladeck 1994) and Pagallo (Pagallo 2013) the examination will serve to inform the legal analysis and help ensure the development of responsible intelligent systems.