Moral Neurolearning by Machines: Artificial Values, Intelligences and Neural Networks

Chapter · January 2019

DOI: 10.1007/978-3-030-22562-9_13

1 author:

Patrici Calvo
Universitat Jaume I
75 PUBLICATIONS **277** CITATIONS

SEE PROFILE

Chapter 13
Moral Neurolearning by Machines: Artificial Values, Intelligences and Neural Networks

Patrici Calvo

13.1 Introduction

Artificial Intelligence (AI) is booming as never before. The initial development of this technology took place in the summer of 1956, when ten researchers interested in neural networks held a six-week workshop to try "(…) to discover how to make machines that use language, form abstractions and concepts, solve problems so far reserved to human beings, and improve themselves" (Bostrom 2014). However, during the 1980s, AI's target audience all but succumbed to a lack of interest, as it gradually dawned on both its potential developers and consumers that the advances predicted and the expectations generated for AI's practical applications were, in most cases, overly optimistic (Bostrom 2014). However, in the last decade there has been a major resurgence of AI thanks, above all, to the emergence of neuroscience, and studies in this field have provided a better understanding of neural learning processes. Furthermore, the consolidation and expansion of the Fourth Industrial Revolution, which has had an important impact on sectors as disparate as retail, cosmetics, communication, education and health care, has revealed how the potential of these sectors can be achieved by optimizing the use of scarce resources and improving production, communication, education, health care and decision-making processes.

Neuroscience is a branch of physiology that studies the cerebral dimension of human or animal behaviour by using invasive and non-invasive techniques to capture images of brain activity, such as functional Magnetic Resonance Imaging (fMRI), Repetitive Transcranial Magnetic Stimulation (rTMS), magnetoencephalography (MEG) and Positron Emission Tomography (PET). Since its emergence at the end of the 1990s, the main aim of this field of research has been to explore and analyse the cerebral processes underpinning human behaviour, especially those

P. Calvo (✉)
Universitat Jaume I, Castellón de la Plana, Spain
e-mail: calvop@uji.es

P. Calvo, J. Gracia-Calandín (eds.), *Moral Neuroeducation for a Democratic and Pluralistic Society*, https://doi.org/10.1007/978-3-030-22562-9_13

involved in decision-making processes across the diverse contexts of human activity. The improved knowledge of the brain and the huge quantity of data produced by neuroscientific studies have also led to considerable progress in AI. Neuroscience has, firstly, made it possible to improve the design of artificial neural networks and, secondly, it has provided a seminal source of Big Data on the emotions, feelings, values and other factors and elements involved in the learning and decision-making processes employed by intelligent machines.

With regard to the Fourth Industrial Revolution, this refers to the digital transformation process that was begun by the European car industry during the first decade of the twenty-first century, and which was subsequently applied to other areas of human activity (Calvo 2017, 2018, 2019a, b, c; Zamagni 2018). Two factors have underpinned this process of digital transformation, Key Enabling Technologies (KETs) and the convergence of the following three technological areas: the Internet of Things (IoT), Big Data (BD) and Artificial Intelligence (AI). The combination of these provides the potential for the *hyperconnectivity*, *datafication* and *algovernance* of everything: machines, people, animals and processes. The IoT makes it possible to recreate a cyber-physical space capable of digitally *hyperconnecting* things, processes, people and animals. Big Data converts, compiles and stores the huge amounts of data created by *digital hyperconnectivity*. Finally, AI processes and transforms Big Data into relevant information and then into knowledge that can be applied in making rational decisions, which in turn make it possible to optimize and improve all the processes involved in a specific area and organization. The incredible results provided by the digital transformation of sectors such as industry, retail and health care, and the specific application of AI as a part of this, has given rise to an unprecedented interest in AI's potential for improved development and application among academics, researchers, professionals and investors.

One of the most relevant concerns currently being raised in the field of AI regards algorithmic governance in politics and business. The expectations currently being generated about the benefits of the application of algorithmic governance are so high—the eradication of corruption, nepotism and selfishness in decision-making processes; optimization of resources; greater competitiveness; increased profits, etc.—that management and/or elected representatives are being replaced by mathematical decision-making models. For example, some companies have already included algorithms with the right to speak and vote on their boards of directors, and there are countries where people have been allowed to choose an algorithm as their political representative. Outstanding cases include Michihito Matsuda, the robot created with a female appearance, which came third in the elections for mayor of Tama New Town (Tokyo) in April 2018; and SAM (Semantic Analysis Machine), the algorithm being prepared for the New Zealand's 2020 general election (Calvo 2019a, c).

Nonetheless, processing Big Data about people's preferences, opinions and habits with regard to a specific activity in order to generate statistics is one thing; actually governing such areas is a wholly different matter. Among other things, this is because governance requires an understanding of autonomy, prudence, justice, the common good, recognition, tolerance and responsibility, and currently