

The ethics of machines

March 19, 2018

The ability to create intelligent machines gives rise to many ethical issues for human beings, some very profound, others more "practical". Here's what we discussed at the "Co.Scienza" festival on March 8 in Trento

We live in an era where we are able to build **machines** that imitate and increase our cognitive abilities in an ever more precise way. An era in which discussions that have been going on for thousands of years come to meet reality and force us to deal with them in a systematic way.

What is **consciousness**? What is **individuality**? What makes us human and differentiates us from other creatures existing in the Universe? Questions that challenge us from many points of view. Alongside these questions, there are several practical issues that arise from the possibility of creating machines that can talk, understand, interact with us and take action.

For the first time we are faced with the **responsibility** of those who create something that is not totally under their control, not because they have their own will, but because they are programmed to carry out activities of various kinds, without our supervision and without precise instructions.

All this goes beyond automation as we have imagined and perceived it to date. It is no longer a matter of feeding the driving force of a machine with electricity, or of programming in detail an "If This Than That" (logic applied to the programming language); it is a matter of programming a machine that will be able to act/interact within an environment without precise instructions. This entails a whole series of **consequences** on which we ought to reflect in this phase of rapid advancement of artificial intelligence, so as to be able to manage critical issues and be able to seize the opportunities at best.

In particular, the following areas are problematic from an ethical point of view when we address this topic:

1. Immoral use

The first issue arises, as with any technology, with the use we make of it. It is not a matter of fearing that machines will take the upper hand and begin to act deliberately against humans: competitive behavior is typical of our species but not of machines, devoid of willpower. Instead, it is about defining the way in which we human beings decide to make use of them. It is in the moment

in which we act that the moral value of everything is determined.

In the case of artificial intelligence, we are dealing with a very powerful tool and the consequences of misuse can be very serious. Let us take the ability to make a [fake video](#) where we make VIPs talk in a convincing way with their own voice, but with words not necessarily spoken by them. On the one hand, we could shoot movies with actors that are no longer alive, on the other hand we could make hoaxes of global reach. And even in the case of using this tool to make a film, who would authorize me to use the image of the actor? Who could assure me that they would like to be part of that film?

The power that arises from a well-constructed artificial intelligence is such that Elon Musk (the South African entrepreneur founder of [SpaceX](#)) has compared this technology to nuclear weapons, arguing that it should be as open source as possible and available to all, so as not to concentrate too much power in the hands of a few. This is why he is among the creators of [OpenAI](#), a non-profit organization that aims to make available to everyone training, huge amounts of test data and computational skills needed to create and train an artificial intelligence algorithm.

2. Artificial intelligence as a source of social inequality

Precisely because it is a powerful tool, this technology is able to create profound social inequality. In fact, initiatives like OpenAI are not enough to avoid a concentration in the hands of a few, be they technology giants such as Amazon, Netflix, Google, IBM, Baidu, Apple, or countries like China or the United Arab Emirates, which are strongly investing in this field.

Margaret Chan, then Director General of the World Health Organization, pointed out this in 2017, stating: “Enthusiasm for smart machines reflect the perspective of well-resourced companies and wealthy countries. We need a wider perspective.” (“Enthusiasm for smart machines reflects the prospects of companies in countries with more resources. We need a broader perspective”).

Let us think about the inequalities created since its inception by the technology of writing: access to this tool or the lack of it has been (and still is) a cause for widespread discrimination. It is easy to understand the point raised by Ms Chan from this perspective.

3. Tendency to “humanize” machines

Since the first chatbot, Eliza, created at MIT by Professor Joseph Weizenbaum in 1965, we have had issues restraining our empathy and our emotions towards something able to talk with us. After all, for millennia, we have been able to do it only with other human beings and it becomes easy to get confused.

Today we are (and we will be more and more) surrounded, by machines that interact with us through voice commands, the web is full of hilarious videos of children who are growing up with Amazon Alexa, Siri, Cortana, and learning the language necessary to be understood with “Hey Alexa, Play music! Hey Siri, read me a story! Hey Cortana, buy me candies! “.

In a situation like this, it may become necessary to establish some rules: if I visit an online customer service, do I have the right to know if I am talking to a human or to a software program? Chances are, the answer is yes. And if in the future I found myself in front of a humanoid robot able to deceive me even in appearance, should I claim to be warned in some way of its identity?

4. Teaching our prejudices to software programs

Technology enhances our capabilities. Technology increases our good qualities but also our faults. If we provide data software that is tainted by implicit prejudices, the result can only be the increase in the prejudices themselves. Here are a few examples.

- Microsoft's [chatbot released on Twitter](#) in 2016 that within 24 hours started writing Nazi and racist phrases: somehow it had deduced that that was the way to become popular on the famous social network.
- If we [google images for CEO](#), how many women and how many black people are shown?
- The [test](#) conducted in the United States for the use of artificial intelligence to estimate the probability of reoffending shows that the algorithm believes that in the case of black people the probability is 77%, while for others 45%.

What is certain is that, if these tools are great for the analysis of large amounts of data, at the moment, they must be supervised by a human able to notice and eliminate the biases they carry.

Responsibility

If a household robot causes damages in the house, if it cooks a deadly recipe because it uses a wrong ingredient or tramples on the neighbor's dog when he takes out the garbage, whose responsibility is it? The European Parliament has already undertaken the creation of a new legal personality precisely to manage the rights and responsibilities of software and machines: it is called [Electronic Personhood](#). The regulatory framework has not been completed yet, but certainly it is the first step towards greater clarity. In the meantime, some say we should apply Roman law: in ancient times if a slave killed a man, the slave owner was held accountable, so, if a robot (word that in its origin means work, slave) causes damage to someone or something, the owner should be held accountable. But if the owner knows nothing about programming and is not able to understand the logic of the software program?

In the UK, for example, the consumer has the right to request information on the software of the products they buy. If you were to buy a self-driving car tomorrow, would you not want to know how it was programmed?

A useful tool could be to create quality certifications on software and central banks dedicated to code storage, to make it available and accessible in a transparent way.

Furthermore, if it is a hardware technical problem? Which and how many are the subjects involved in the realization of the finished product? Who is to be held accountable?

Can common sense be programmed?

When we face the problem of self-driving cars and how should they "behave" in our streets, we often tend to ask ourselves the moral question: what is the value of a life? How do I choose, in an emergency, who to save and, consequently, how to maneuver the car? The MIT online experiment in Boston called "[The Moral Machine](#)" focuses precisely on this: a series of scenarios is shown and users are asked to indicate how the car should behave in that context. The experiment, however, does not give an answer to the problem: in general we all agree on the logic to follow in these cases, but if we personalize the context and we or our loved ones are the protagonists of the scenario, all rules change.

What The Moral Machine points out is that we do not have to put the autonomous cars in the condition of having to manage this type of scenario. In Germany, for example, some guidelines

have already been drawn up for the development of rules for these vehicles and the basic concept is that autonomous cars should not be put in a position to face such scenarios. If these vehicles had dedicated lanes, were in constant contact with each other through a communication network between robots (a “Robochain” as they call it at MIT), then these dilemmas would not be posed.

The basic issue arises from the fact that much of our (moral) actions are based on common sense, on our ability to use rules and laws as guidelines and to apply them in the best way based on the context. Thus, if we know that solid lines should not be passed when we drive, we also know that there are cases in which exceeding 10 cm not only does not create any problem, but can also be a way to avoid an accident or not get stuck in the traffic because there is a car parked in a double row. Software is rational logic and, to date, we are not yet able to program it to understand these differences. Some researchers try to teach common sense to algorithms in the same way we teach them to children: by example and imitation. Will they succeed? It is not easy to predict, but we can say that at present we should not entrust technology alone with activities that require common sense and decision-making skills.

PERMALINK

<https://magazine.fbk.eu/en/news/the-ethics-of-machines/>

TAGS

- #etica
- #future
- #Intelligenza artificiale
- #robot

RELATED VIDEOS

- https://www.youtube.com/watch?v=4izR1Y_xKBI

RELATED MEDIA

- "The pleasure of AI research" Conference : <https://www.fbk.eu/en/event/pleasure-research-ai/>

AUTHORS

- Cristina Pozzi