CEST
Centro de Estudos Sociedade e Tecnologia
USP
Universidade de São Paulo
Bulletin - Volume 3, Number 4, May/2018

Use of artificial intelligence to promote social inclusion

Eduardo Bertassi

Recently, we have experienced several advances in the artificial intelligence and robotics area which, fifty years ago, had only been described in science fiction books: facial recognition made by telephones, autonomous cars and trucks replacing drivers, space rockets landing without the need for pilots, use of computerized systems to perform simple day-to-day tasks (such as booking a restaurant or scheduling a hairdresser), among others.

It is very easy to be marveled at technological achievements in the artificial intelligence area, but we must be aware that certain advances might hide the threat of creating highly unequal societies.

The misusage of certain technologies might contribute to the development of less inclusive societies, even if unintentionally.

Unpredicted threats to diversity and inclusion

The artificial intelligence usage will benefit society in a variety of forms, bringing greater comfort, simplification, cost savings, service delivery improvements, and will become essential for the creation of increasingly intelligent products that will fit buyers' individual needs.

For companies, the artificial intelligence usage might give them the ability to survive in a highly competitive market, in which new companies with disruptive technologies emerge each year from garages, as it has happened with companies like Amazon, Apple, Google and HP that were once born in ordinary home garages.

But is the artificial intelligence usage being applied correctly? That is a question that many university researchers and practitioners from worldwide organizations have been asking themselves, starting from the concern that the misusage of certain technologies might contribute to the development of less inclusive societies, even if unintentionally.

Examples of inappropriate use of technologies containing artificial intelligence may occur, for example, in the following situations: use of inadequate data sampling when training intelligent systems; use of biased algorithms; and occurrence of cases in which there are reinforcement loops.

A study carried out by an MIT Media Lab's researcher Joy Buolamwini showed that the machine learning algorithms of facial recognition systems from three large companies could lead to discrimination by gender or race. In one of such systems, developers claimed that the success rate when detecting a person's gender or race was higher than 97%, however Buolamwini found out that the data used for training the system contained a sampling in which there were 77% male faces, which 83% of them had white people's faces. That is an example of a case that might require a training database review so that the faces of people from different genders and ethnicities could be more representative. The researcher told in a TED Talks that one of the reasons that motivated her research was the fact that several face recognition algorithms she had tested were unable to recognize her face.

In another situation, the Brazilian non-governmental organization *Desabafo Social*, which fights for better black people representativeness in society, carried out a study that showed that in several photo-commercial sites, it was necessary to include the word "black", or similar words, in their keywords to find photos of black people in the search results within those

same sites. For example, when entering the word "babies" in the photos search engine, the white skin babies' photos search results was greater than black skin babies' photos search results. Until today, it is necessary to type "black babies" so that the photos search engine brings more "accurate" results. Naturally, the search results of these algorithms do not necessarily demonstrate an intrinsic bias, but it is possible to improve them so that companies help to promote more diversity and social equality.

In 2016, The Washington Post journal published a story about a "predictive policing" system being used by the Los Angeles police called PredPol. The system basic working principle was that an intelligent algorithm, fed with data from various parts of the city, could indicate to the police department potential locations where a crime was about to occur. The idea, which appears to be taken from Philip K. Dick's 1956 science fiction short tale Minority Report, aims to maximize the police results and effectiveness by concentrating the few officers from the contingent in the most critical city locations. The reasons for using the system are noble, but there are risks that the system might be "stuck" or "limited" on reinforcement loops, i.e., the concern that some communities could be unfairly labeled and that policing routines might end up distorted due to the way the system works, potentially indicating the same locations as risky areas.

From the cited examples it is possible to notice that the artificial intelligence usage to solve a given problem is not as trivial as it seems, because sometimes, in the good intention of trying to provide a solution, one might end up creating another unnoticed problem.

Researchers and developers' awareness

Fortunately, several universities and institutions, both public and private, are aware of the need of responsibly using artificial intelligence for good and to make the researchers and developers community aware of such objectives.

In November 2017, the Global Network of Internet and Society Centers (NoC), the Institute of Technology and Society (ITS-Rio), and the Berkman Klein Center for Internet & Society of Harvard University organized the event "Artificial Intelligence and Inclusion" in the Museum of Tomorrow, in Rio de Janeiro. Researchers and practitioners discussed the potential challenges and risks that the artificial intelligence usage could entail regarding, for instance, the future of labor, the emergence of new power structures, and the possibility of widening social inequality.

Most recently, the XPRIZE Foundation hosted the second edition of the "AI for Good" event in partnership with UNESCO, UNICEF, the World Food Program, the World Bank, and other global entities in May 2018. The objective was to identify practical applications using artificial intelligence that could improve the life quality and sustainability on our planet.

As it can be noticed, the artificial intelligence usage in systems requires a broader vision that goes beyond reaching specific benefits when trying to solve particular problems. The choice to join the various researchers and entities that seek to contribute to transforming our planet into a world that is more equitable, inclusive, and sustainable depends only on our will.

Eduardo Bertassi *is a master's degree student in computer engineering at Escola Politécnica da USP, and researcher at CEST-USP.*

Academic Coordinator: Edison Spina

This article is a result of the author's ascertainment and analysis, without compulsorily reflecting CEST's opinion.