ARTIFICIAL INTELLIGENCE, EMPATHY AND INCLUSION: A DESIGN PROBLEM

Cid Boechat[1], M.Sc

(1) Fundação CECIERJ - Praça Cristiano Ottoni, S/N – 6º andar, sl. 650 - Centro – Rio de Janeiro – RJ, Brazil. 20221-250
cboechat@cecierj.edu.br

Design, Artificial intelligence, Human-computer interaction

1. Introduction

Our daily interactions with Artificial Intelligences (A.I.) are increasing. At first, many companies and researchers advocate the use of these tools without replacing people's decision-making power. But if these machines become unreachable in performance and economically cheaper than professionals, will this option remain? In this scenario, how will be the experience of interacting with decision-making systems? Will Artificial Intelligences be able to understand the human being in all its complexity of emotions, differences, diversity and contradictions?

Design is, also, a search for significant improvements of human life (BUCHANAN, 2015) focused on people's fundamental needs (BROWN, 2010). Thus, this paper will seek to show that Design, through its approach and precepts, must analyze A.I. issues in a macro way, joining current and future discussions about the impact of extensive decision-making Artificial Intelligences in human life.

2. Artificial Inteligencies in everyday life

A.I. strives to build entities capable of thinking like humans. It seeks to create programs that can emulate our reasoning and decisions (RUSSEL; NORVIG, 2003). So, specifically in machine learning and deep learning subfields, the system is programmed with algorithms - which give it the ability to learn - and is fed with data that will allow it to "teach" itself (RUSSEL; NORVIG, 2003).

These processes allow the detection of patterns that humans do not perceive or would never think of. Thus, these applications could be used in virtually any area where there are sufficient data and/or interactions avaiable (VINCENT, 2019).

Currently, A.I. and its subfields are present in people's lives at different applications: transportation, education, academic life, entertainment, work, and health, among others (NARULA, 2019).

3. Problems and issues of Artificial Intelligence

As A.I. tends to reflect the data used to train it, any data bias, non-conformity or error will reflect on its “learning”. If a developer is biased about something, this can infiltrate the system and affect the result of its responses (PTI, 2020).

Deeper algorithms that learn through experience often cannot be questioned (the so-called “A.I. black box”). As it “teaches” itself, there is no way to tell what it effectively knows or what parameters were used to come with its results (MUKHERJEE, 2017). How does it draw responses from millions of resources and possibilities? There is no way to know, and it does not know how to tell us (MUKHERJEE, 2017). The more powerful the system, the opaquer it can become.

This scenario leads to a big question: how will A.I. be used? Banks, military personnel and employers, among others, seek more complex machine learning approaches, with automated decision making that cannot be evaluated (KNIGHT, 2017).

There is another important problem: Artificial Intelligence does not have empathy. Recognizing an emotion is different from understanding it. Human beings tend to experience emotions differently. Currently, learning algorithms recognize some physiological or behavioral characteristics associated with specific emotional states (MORGAN, 2018). However, more sophisticated emotions are much more complicated. It is not as simple to identify them as it is to recognize voice tones, posture or facial expressions (DEDEZADE, 2019).

Unfortunately, we can already see examples of debatable use of A.I. In China, a facial recognition system monitors students' attention in real time. If it concludes that the student has other thoughts during classes, it sends a notification to teachers to take action (CONNOR, 2018). Amazon uses a fully automated productivity control system in its warehouses (LECHER, 2019). It tracks the individual production rates of each employee, measures the time spent off task and automatically generates warnings or even layoffs, without humans taking part in the process. According to reports, workers complain of being treated like robots, because they are monitored and supervised by these automated systems (LECHER, 2019). In China, leaked documents show that a predictive surveillance and policing program collects large amounts of personal information about citizens and then uses artificial intelligence to formulate lists of so-called suspicious people. Through large-scale data collection, machine learning is used to predict in advance where possible incidents may occur and to identify possible populations that are likely to be involved in anti-regime and anti-state actions (ALLEN-EBRAHIMIAN, 2019).

Some of the cases shown raise an especially dangerous use of A.I.: a decision making tool without human supervision. If Artificial Intelligence becomes much more accurate, efficient and cheaper than hiring humans, will we have important decisions that impact our lives taken without an human interlocutor? Will ethical and humanistic discussions have enough room in a billion dollar market that keeps growing? How can Design act in this scenario?

4. The use of Artificial Intelligence from the perspective of Design

Issues of empathy, human centrality, and inclusion make A.I. a problem especially tailored to a Design approach. A.I. was not created and developed out of people's demands and needs. It is a technology created *a priori* and, as it evolves, people seek to use it according to needs or opportunities. A design approach is an alternative to invention (PRESTERO, 2010). The difference is that an invention approach usually starts with technologies and then searches for users. A design approach starts with the user and then seeks for technologies.

Design try to solve problems using empathy to identify and contextualize needs and then converting them into specifications (PRESTERO, 2010). Empathy is a defining characteristic of designer-user relationships when Design is concerned with user experience (WRIGHT; McCARTHY 2008). Without a mediator in an Human-AI interaction, Artificial Intelligence becomes an interlocutor and, in a way, the designer of the experience - since it has the decision-making prerogative. Thus, it needs the same empathy and contextual understanding that we demand from designers. These A.I. Will have that ability by themselves? Is it possible, or advisable, to subtract the human interlocutor from this interaction?

Wright and McCarthy (2008) argue that adopting a dialogical perspective does not diminish the importance of intuition and the individual agency. Quoting Buber (1947/2006), Cipolla and Bartholo (2014) define the difference between empathy and inclusion, the latter involving encounters in which the other is not an "It" but rather a "He/She". Quoting Friedman (2002), the authors remember that if the "It" is not a "He/She", it is also not "another me".

Inclusion demands full presence of designers in the scenario in which they are participating. This means that each designer needs to perform two functions: as a facilitator who guides the design process and,

simultaneously, as someone included, who engages into relationships with others, seeking solutions to problems shared by those involved, including himself (CIPOLLA; BARTHOLO, 2014).

In other words, is a system that provides answers to make people's lives easier - but that do not understand humans or share their experiences - complete enough to autonomously take decisions that impact people? Can it be the "designer" of the "experiences" to which these humans will be subjected?

How to measure the extent of its responsibilities in case of any bias, if the system is not able to say how and on what basis it comes up with its decisions? How to correct a distortion? Will an A.I. be able to find the dialogical and inclusive term that a designer needs? How to assign a sense of responsibility to these tools?

5. Conclusion

This paper understands that the Human-AI interaction could be one of the most important aspects of the next decade. Considering all the positive and negative potential of A.I., Design should approach this issue in its full extent, in a macro, conceptual and systemic way. It will be necessary to collectively propose frameworks and guidelines for the operation and design of those tools. Especially due to empathy and human centrality issues, it is essential that Design, as a discipline, becomes even more active in multidisciplinary discussions that already take place in various forums.

Design can contribute to a better system design and data collection, taking users and their context into account. However, it is worrying that in the official Brazilian government document about the subject, Design, as a discipline, is not mentioned once.

It is undeniable that machine and deep learning processes will bring benefits that will improve our lives, and this article does not intend to demonize it. Even so, it is designers' role to constantly remember the importance of human issues - especially in decision-making and monitoring cases.

This is a design problem because when one turns a tool into a decision-making interlocutor, people will be interacting with systems that 1) often do not provide complete feedback; and 2) are not able to understand human and contextual issues. Two serious problems that go radically against the practices of interaction and interface design.

6. References

ALLEN-EBRAHIMIAN, B. Exposed: China's Operating Manuals for Mass Internment and Arrest by Algorithm. In: icij.org, 2019.

ANDERSON, J. A British start-up will put AI into 700 schools in Belgium. In: Quartz, 2019.

ASADA, M. Towards Artificial Empathy. International Journal of Social Robotics, v. 7, p. 19–33, 2014.

ASADA, M. Development of artificial empathy. Neuroscience Research, v. 90, p. 41–50, 2015.

BOECHAT, C.; MONT'ALVÃO, C. Refletindo Sobre o Futuro da Aprendizagem Através da Abordagem do Design. In: 13º Congresso Pesquisa e Desenvolvimento em Design, 2018, Joinville. Anais [...]. São Paulo: Blucher, 2019. p. 6151-6165.

BROWN, T. Design Thinking. Rio de Janeiro: Elsevier, 2010.

BUCHANAN, R. Worlds in the Making: Design, Management, and the Reform of Organizational

Culture. Xangai: She Ji: The Journal of Design, Economics and Innovation, 2015.

CANO, R. J. C. O robô racista, sexista e xenófobo da Microsoft acaba silenciado. In: El País, 2016.

CASTELLANOS, S. Siri Contributor Tackles Software That Detects Emotional States. In: The Wall Street Journal, 2018.

CONNOR, N. Chinese school uses facial recognition to monitor student attention in class. In: The Telegraph, 2018.

CIPOLLA, C.; BARTHOLO, R. Empathy or inclusion: A dialogical approach to socially responsible design. International Journal of Design, v. 8, n. 2, p. 87–100, 2014.

DAVIS, M. Global Artificial Intelligence Boom Predicted in Education, Particularly in China. In: EdWeek Market Brief Marketplace, 2019.

DEDEZADE, E. Jobs of the future: teaching empathy to artificial intelligence. In: Microsoft.com, 2019.

HAO, K. What is AI? We drew you a flowchart to work it out. In: Technology Review, 2017.

KNIGHT, W. The Dark Secret at the Heart of AI. In: Technology Review, 2017.

LECHER, C. How Amazon automatically tracks and fires warehouse workers for 'productivity'. In: The Verge, 2019.

MORGAN, L. AI Challenge: Achieving Artificial Empathy. In: Information Week, 2018.

MUKHERJEE, S. A.I. Versus M.D. What happens when diagnosis is automated? In: The New Yorker, 2017.

NARULA, G. Everyday Examples of Artificial Intelligence and Machine Learning. In: Emerj.com, 2019.

PARTNERSHIP ON AI. About us. In: Partnership on AI, 2019.

PRESTERO, T. Better by Design: How Empathy Can Lead to More Successful Technologies and Services for the Poor (Discussion of Design Case Narratives : Rickshaw Bank, Solar-Powered Tuki, FGN Pump). Innovations: Technology, Governance, Globalization, v. 5, n. 1, p. 79–93, 2010.

PTI. How to make sure Artificial Intelligence isn't biased as people? Satya Nadella's advise to techies. In: Financial Express, 2020.

RUSSEL, S.; NORVIG, P. Artificial intelligence: A modern approach. Nova Jersey: Prentice Hall, 2003.

SHOHAM, Y. et al. AI Index 2018 Annual Report. Stanford: Stanford University Press, 2018.

TWENGE, J.; MATEOS-GARCIA, J.; JOHN, J. Por que a falta de diversidade é um obstáculo à inteligência artificial. In: El País, 2019.

VINCENT, J. DeepMind's AI can detect over 50 eye diseases as accurately as a doctor. In: The Verge, 2018.

VINCENT, J. The state of AI in 2019. In: The Verge, 2019.

WRIGHT, P.; MCCARTHY, J. Empathy and experience in HCI. In: Conference on Human Factors in Computing Systems, 2008, Florence. Anais [...]. Florence: ACMPress, 2008. p. 637–646.

ZARLEY, D. Meet The Scientists Who Are Training AI To Diagnose Mental Illness. In: The Verge, 2019.