Ethical Challenges in Artificial Intelligence and Neuroscience

Fereshteh Azedi [1, 2*]

1. Cellular and Molecular Research Center, Iran University of Medical Sciences, Tehran, Iran
2. Department of Neuroscience, Faculty of Advanced Technologies in Medicine, Iran University of Medical Sciences, Tehran, Iran

Original Article

Abstract

Background: Neuroscience merges with technology in a way which will have a huge impact on society. It will be more than just improving brain health or brain function. There will also be significant ethical challenges. It is conceivable that old understanding of human beings should be redefined. Artificial intelligence (AI) and robotics are novel technologies in the neural sciences. It is well established that AI has a significant impact on the development of humanity in the near future. They have raised fundamental doubts around how to get along with these new systems, what the systems themselves should serve, what risks they involve, and how to control them. This review paper includes questions for a general explanation of ethical issues and provides current viewpoints on these technologies.

Methods: The major challenges of ethics in AI and neuroscience were investigated in order to find solutions for the problems. The number of studies in this field reviewed and the major challenges of ethics in AI and neuroscience were explained.

Results: We summarized the information and the relationship between the obtained data and information in this field.

Conclusion: Solving challenges of ethics in AI and neuroscience is not an easy task. Indeed, collaboration among neuroscientists, businessmen, clinicians, engineers and ethicists is critical to leverage the maximum benefits of AI and neurotechnologies in all aspects.

∗ Corresponding author: Fereshteh Azedi, Tel: 86704720, Fax: 86704698, Email: fazeditehrani@gmail.com; azeditehrani.f@iums.ac.ir

Introduction

Artificial intelligence (AI) is a recent digital technology that will have a major impact on the advancement of numerous science issues mainly neural sciences (1-3). AI provides us with excellent opportunities for finding new brain discoveries in neurology and psychiatry and helps us in developing neurotechnologies (4-6). There are also a number of emerging ethical concerns that need to be addressed. It remains uncertain how AI-based approaches to the subject of the human brain will meet enough standards of scientific validity and affect prescriptive instruments in neuroethics and research ethics (7-9).

Indeed, the ethics of AI is a very young domain within applied ethics, with considerable dynamics; however, there are few well-established subjects and no authoritative overviews despite emergence of promising outlines and societal impact and policy recommendations (10). While it is the important responsibility of AI scientists to assure that the impact of AI and similar new technologies in the future should be more positive than negative, ethicists have important roles in the development of such technologies from the beginning (3, 8). It has raised fundamental doubt around how to get along with these new systems, what the systems themselves should serve, what risks they involve, and how to control them (11). In this study, the major challenges of ethics in AI and neuroscience were investigated in order to find solutions for the problems.

Methods

We summarized the current state of understanding on ethical challenges in artificial intelligence and neuroscience. The major challenges of ethics in AI and neuroscience were investigated in order to find solutions for the problems.

Results

In AI and neuroscience, the major challenges of ethics can be included as bias, responsibility and identity, brain enhancement, definition of human, privacy and morality. All of these challenges have to be considered as the main problems in the future of this field.

Bias

Data analysis is regularly used in predictive analytics in neuroscience and other fields, to anticipate imminent developments (12). Programmed algorithms resulting from AI are coded in a very logical method, with data fed into them and clear outcomes defined. The most important role of AI is to find certain patterns among various huge data sets, so it would appear that they are unbiased machines. But their regulatory operations are encoded into the programs, and the data fed to them originate from humans, who are not without any biases (13). Unluckily, AI reflects the biases of their human creators as well (14).

The complex human cognitive system is prone to have various kinds of cognitive biases, *e.g.*, the confirmation bias; humans tend to interpret and confirm new information based on what they already believe (15). This second form of bias is often known as hindrance effect in rational judgment, though at least some cognitive biases make an evolutionary advantage, e.g., economic use of resources for instinctive judgment. There is a question whether such cognitive bias could or should exist in AI systems (16). A third form of bias is present in the data where there is a systematic error, for instance statistical bias. In fact, any given dataset will only be unbiased for a single kind of theme, so the mere new function of a dataset has the main risk of being applied to a dissimilar type of subject area, and then turns out to be biased for that area. Machine learning about the origin of such data would then not only fail to identify the bias, but it codifies and computerizes the historical bias (17).

Responsibility and identity

Human actions are always controlled. But recent technologies that alter our brain activity have the potential to change this manner (18). For example, they may compel an individual to commit an out-of-character crime. Antidepressants or other medications which also affect brain activity may lead to such situations (19).

Besides, identity can be altered by these devices through any transformations in the patterns of our brain function. With respect to these questions, it is not the matter of questioning the importance of brain stimulation devices and neurotechnologies as medical therapies because it is well established that they are safe and effective treatments and the patients are thrilled to see their quality of life back on track. The main point is that as the new devices and technology progress, legal and ethical guidelines should be observed as well (20).

Brain enhancement

A chief goal of future research in neuroscience certainly will be enhancing brain activity, and the

military may be the best example of it. The United States' Defense Advanced Research Projects Agency, DARPA, is already trying to achieve main advancements in this field. In fact, they attempt to boost combat readiness, performance and recovery of military staff (21). However, it is unclear whether cognitive enhancement should be permitted or not. The fact is, some people already take steps to obtain it; drugs like Ritalin and Modafinil can enhance focus and increase attention (22). Moreover, Prozac changes feelings of depression and anxiety (23). The prospect of cognitive improvement raises questions of equality and equity; the challenges of accessibility and affordability to use such technology are a matter of debate. Moreover, it is not clear whether a high score in a test is the right criterion for using cognitive enhancement drugs (24).

The definition of human

Some modern robots are manufactured by AI technologies to have an accurate human appearance; however, it is not clear to what extent they should be treated like humans. In October 2017, Sophia, a social robot capable of producing more than 50 facial expressions, was named a citizen of Saudi Arabia, to the fear of many experts across the globe. A recent study also revealed that people were unconsciously treating robots in a very humane manner. A robot pleading "Please don't turn me off" caused nearly 30% of people to conform, even if the researchers had asked them to turn the robot off (25).

Privacy

Today's largest neurotechnology companies gather huge data of personal information. These companies can sell these data banks for commercial gain. Using a device like the wearable EEG headset records the brains activities and helps companies access valuable brain activity information. For example, if an individual imagines buying a new smartphone while wearing an EEG helmet, an online salesman might access his brain information. The vendor might apply AI software to call right away with their latest offer of smartphones. Therefore, there are more important privacy concerns regarding who would or should have access to people's brain activity (26).

Morality

Robotics researchers aim at building robots with general intelligence (Managed by AI), guided by a sense of morality. However, the type of morality for a robot, the rules for guiding robot decisions,

humans' deliberately dishonorable behavior towards robots are all unsolved challenges (27). Teaching general moral principles to robots and allowing them to infer appropriate decisions will also not work all the time; there will always be exceptions to the rule or ambiguity. The other alternative is helping the robot learn through experience, just as humans perform, under supervision of ethicists. However, it is challenging what code of ethics a robot has to learn. Humans often continue to disagree about what good moral decisions are, even in a single country or culture (28).

Discussion

Our understanding of the brain has developed very quickly and discoveries made by scientists in AI and neuroscience can certainly have a significant positive impact on quality of life, but unforeseen consequences can also take place. Accordingly, it is important to keep in mind the many possible uses and the motives for the development of AI technologies. To avoid pitfalls and harness potential for neuroscience, patients and health care in general, critical ethical challenges must be addressed. Here, we discussed the ethical principles that are primarily affected by neurotechnology and AI approaches to human neuroscience, and the normative safeguards that should be applied in this area.

Conclusion

Generally speaking, the ethical challenges in AI and neuroscience are not easy to address. Indeed, collaboration among neuroscientists, businessmen, clinicians, engineers and ethicist is critical to leverage the maximum benefits of AI and neuro-technologies in all facets.

Conflict of Interest

Authors declare no conflict of interest.

References

1. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet. 2020;395(10236): 1579-86.

2. Amari SI. [Brain and artificial intelligence]. Brain Nerve. 2019;71(12):1349-55. Japanese.

3. Lawrence DR. Advanced bioscience and AI: debugging the future of life. Emerg Top Life Sci. 2020;3 (6):747-51.

4. Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. Neuron. 2017;95(2):245-58.

5. Ienca M, Ignatiadis K. Artificial intelligence in clinical neuroscience: methodological and ethical challenges. AJOB Neurosci. 2020;11(2):77-87.

6. Tortora L, Meynen G, Bijlsma J, Tronci E, Ferracuti S. Neuroprediction and A.I. in forensic psychiatry and criminal justice: a neurolaw perspective. Front Psychol. 2020;11:220.

7. Liu TYA, Bressler NM. Controversies in artificial intelligence. Curr Opin Ophthalmol. 2020;31(5):324-8.

8. Safdar NM, Banja JD, Meltzer CC. Ethical considerations in artificial intelligence. Eur J Radiol. 2020;122:108768.

9. Nudeshima J. [Ethical issues in artificial intelligence and neuroscience]. Brain Nerve. 2019;71(7):715-22. Japanese.

10. Gruson D, Petrelluzzi J, Mehl J, Burgun A, Garcelon N. [Ethical, legal and operational issues of artificial intelligence]. Rev Prat. 2018;68(10):1145-8. French.

11. Gruson D. [The ethical risks associated with artificial intelligence must be identified and regulated]. Soins. 2019;64(838):48-50. French.

12. Weber C. Engineering bias in AI. IEEE Pulse. 2019;10(1):15-7.

13. Altman RB. Artificial intelligence (AI) systems for interpreting complex medical datasets. Clin Pharmacol Ther. 2017;101(5):585-6.

14. Friele M, Brokerhoff P, Frohlich W, Spiecker Genannt Dohmann I, Woopen C. Digital data for more efficient prevention: ethical and legal considerations regarding potentials and risks. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2020;63(6):741-8.

15. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Global Health. 2019;9(2):010318.

16. Wang F, Preininger A. AI in health: state of the art, challenges, and future directions. Yearb Med Inform. 2019;28(1):16-26.

17. Benke K, Benke G. Artificial Intelligence and Big Data in Public Health. Int J Environ Res Public Health. 2018;15(12):2796.

18. Ashrafian H. Artificial intelligence and robot responsibilities: innovating beyond rights. Sci Eng Ethics. 2015;21(2):317-26.

19. Harris J. Who owns my autonomous vehicle? ethics and responsibility in artificial and human intelligence. Camb Q Healthc Ethics. 2018;27(4):599-609.

20. Cole D. Artificial intelligence and personal identity. Synthese. 1991;88(3):399-417.

21. Fouse S, Cross S, Lapin Z. DARPA's impact on artificial intelligence. Ai Magazine. 2020;41(2):3-8.

22. Mohamed A. Neuroethical issues in pharmacological cognitive enhancement. Wiley Interdiscip Rev Cogn Sci. 2014;5(5):533-49.

23. Husain M, Mehta M. Cognitive enhancement by drugs in health and disease. Trends Cogn Sci. 2011;15(1):28-36.

24. Sahakian B, Morein-Zamir S. Neuroethical issues in cognitive enhancement. J Psychopharmacol. 2010;25(2):197-204.

25. Rocha E Sophia. Exploring the ways AI may change intellectual property protections protectio. DePaul J Art Technol Intell Property Law. 2018;28(2):126-46.

26. Farah MJ. Neuroethics: the practical and the philosophical. Trends Cogn Sci. 2005;9(1):34-40.

27. Jeste DV, Graham SA, Nguyen TT, Depp CA, Lee EE, Kim HC. Beyond artificial intelligence: exploring artificial wisdom. Int Psychogeriatr. 2020;32(8):993-1001.

28. Farah MJ. Neuroethics: the ethical, legal, and societal impact of neuroscience. Annu Rev Psychol. 2009;63:571-91.

چالش‌های اخلاقی در هوش مصنوعی و علوم اعصاب

فرشته عاضدی [1، 2]*

۱. مرکز تحقیقات سلولی و مولکولی، دانشگاه علوم پزشکی ایران، تهران، ایران

۲. گروه علوم اعصاب، دانشکده فن‌آوری‌های پیشرفته پزشکی، دانشگاه علوم پزشکی ایران، تهران، ایران

چکیده

مقدمه: مباحث مرتبط با حوزه علوم اعصاب به گونه‌ای با فناوری‌های نوین درآمیخته است که در حال حاضر تأثیر زیادی در جوامع بشری داشته و خواهد داشت. تحقیق در این حوزه نشان داده است که گستره مباحث مرتبط با علوم اعصاب فراتر از ارتقای سلامت مغز یا بهبود عملکرد مغز است و در حال حاضر با توجه به گسترش روز افزون کاربرد آن، شاهد چالش‌های اخلاقی قابل توجهی در این حوزه هستیم. حقیقت آن است که درک قدیمی از مفهوم انسان و بشر در حال تغییر است. همچنین هوش مصنوعی و رباتیک، فناوری‌های جدیدی در این حوزه هستند که تأثیر بسزایی در توسعه بشریت در آینده نزدیک خواهند داشت. در کنار توسعه این فناوری‌ها، ابهامات اساسی در مورد چگونگی مواجهه با این سیستم‌های جدید، خدمات قابل ارائه توسط آنها، مخاطرات پیشرو و چگونگی کنترل آنها به خصوص از منظر اخلاق پزشکی مطرح است. در این مقاله سعی شده است تا به مبحث چالش‌های پیشرو در مسائل اخلاقی حوزه و دیدگاه‌های فعلی در مورد این فناوری‌ها بپردازد.

روش‌ها: در این مقاله، چالش‌های مهم اخلاقی در هوش مصنوعی و علوم اعصاب با مرور چندین مطالعه به منظور یافتن راه حل برای مشکلات پیشرو پرداخته شده است.

نتایج: پس از کسب و جمع‌آوری داده‌ها از مقالات مطرح، اطلاعات بدست آمده خلاصه و چالش‌های موجود در هر حیطه از علوم اعصاب و هوش مصنوعی جمع‌بندی و نتایج آن ارائه گردید.

نتیجه‌گیری: قطعاً حل چالش‌های اخلاقی در حیطه نوین هوش مصنوعی و علوم اعصاب کار ساده‌ای نیست. در واقع، همکاری میان دانشمندان حوزه علوم اعصاب، فعالین در بخش خصوصی و شرکت‌ها، پزشکان، مهندسان و متخصصین در حوزه اخلاق پزشکی برای استفاده از حداکثر مزایای کاربرد هوش مصنوعی و فناوری‌های حوزه علوم اعصاب در تمامی جنبه‌ها حایز اهمیت است.

* نویسندهٔ مسئول: فرشته عاضدی، تلفن: ۸۶۷۰۴۷۲۰، نمابر: ۸۶۷۰۴۶۹۸، رایانامه: fazeditehrani@gmail.com, azeditehrani.f@iums.ac.ir