Ethical Issues Concerning Brain-Reading A.I. in Criminal Justice and Forensic Psychiatry

Leda Tortora ("Sapienza" University of Rome), Stefano Ferracuti ("Sapienza" University of Rome), Enrico Tronci ("Sapienza" University of Rome), Gerben Meynen (Utrecht University)

Brain-reading technologies in which neuroimaging and A.I. are combined may have many applications, such as lie detection, neuromarketing, brain-computer interfaces, etc. Here, we are focusing on its use for crime prevention and, more specifically, risk assessment in forensic psychiatry. We will provide some examples of current possibilities and significant technical limitations of these techniques. Next, we will consider two important ethical issues regarding the use of neuroimaging combined with AI in criminal justice, and particularly in forensic psychiatry:

1.Neurodeterminism: modern neuroscientific techniques might increase the ability to make accurate predictions about various forms of future behaviour, and therefore they may challenge the notion of "free will" in our criminal justice system and the concept of responsibility.

2.A.I. Algorithms may entail ethnic and gender biases. We address the possibility that a hypothetical data-driven A.I. system used to help judges to determine the probability of recidivism of a convicted criminal may be biased, e.g., against gender and minorities.

Use of software-based control systems in safety-critical applications (e.g., as automotive, aerospace, biomedical devices) is highly regulated since software errors may damage (physically) humans. Along the same line, besides the alluring potential of these technologies, we argue that the use of A.I.-based systems in criminal justice and forensic psychiatry should be subjected to substantial regulation to protect citizens from system errors or misuse. For example, there are situations where changing just one single pixel may lead deep neural networks to misclassify images. Recently, Amazon facial "Rekognition" software mistook congressman for criminals. On such basis, we highlight the importance of accurate harms/benefits analyses not only when these technologies will be fully available, but also while they are being researched and developed.

References

- Meynen G. Brain-based Mind Reading in Forensic Psychiatry: Exploring Possibilities and Perils. Journal of Law and the Biosciences, 4(2), 311–329.
- Parmigiani G, Mandarelli G, Meynen G, Tarsitani L, Biondi M, Ferracuti S. Free will, neuroscience, and choice: towards a decisional capacity model for insanity defense evaluations. Rivista Di Psichiatria. 52: 9-15.
- Meynen G (2016). Legal Insanity. Explorations in Psychiatry, Law, & Ethics. Series: International Library of Ethics, Law, and the New Medicine, Vol. 71, Weisstub DN & Cooley DR (Eds). Springer.
- Scarpazza, C., Ferracuti, S., Meng, F., & Sartori, G. (2018). The charm of structural neuroimaging in insanity evaluations: guidelines to avoid misinterpretation of the findings. Translational Psychiatry.
- Kiehl, Kent A et al. "Age of gray matters: Neuroprediction of recidivism"NeuroImage. Clinical vol. 19 813-823. 3 Jun. 2018.
- Strano, M. (2004). A Neural Network Applied to Criminal Psychological Profiling: An Italian Initiative. International Journal of Offender Therapy and Comparative Criminology, 48(4), 495–503.
- Steele, Vaughn R et al. "Machine learning of structural magnetic resonance imaging predicts psychopathic traits in adolescent offenders" NeuroImage vol. 145, Pt B (2015): 265-273.
- Crawford, K. Artificial Intelligence's White Guy Problem. The New York Times, June 25, 2016.
- Morse, Stephen J., Neuroethics: Neurolaw, Oxford Handbooks Online (2017).
- Roskies, A. L. (2015). Mind reading, lie detection, and privacy. In J. Clausen & N. Levy (Eds.), Handbook of Neuroethics (pp. 679-695). Dordrecht: Springer.