Three relevant bodies of the literature to understand the ethics, design, and governance of emotional AI in Confucian contexts

Manh-Tung Ho

Ritsumeikan Asia Pacific University

Beppu, Oita, Japan

September 7, 2020

In this short essay, I want to hypothesize on how researchers can systematically analyze the ethics and governance of emotional AI in Confucian contexts. Emotional AI refers to machines/algorithms that are increasingly getting better at reading, scanning, sensing, and classifying human emotions (Bakir & McStay, 2018; McStay, 2018). Social scientists have started to theorize about how this new technology might affect societies. One of the dominant discourses is the “surveillance capitalism” framework by Shoshana Zuboff of Harvard Business School (Zuboff, 2015, 2018; Cinamon, 2017). However, these discourses are limited when it comes to forming an educated understanding of emotional AI in East Asian countries with a strong presence of Confucianism such as China, Japan, South Korea, Singapore, Vietnam.

Kai Fu-Lee attempts to fill this gap in his book *AI Superpowers, China, Silicon Valley, and the New World Order*, in which he argues that China with an advantage on the volume of data and the intense competition among copycats will be key factors will lead the world in AI (Lee, 2018). Understandably, economic and legal conditions are important aspects of understanding AI in countries like China. But overlooking the pervasive influence of cultural factors such as Confucianism will not help researchers achieve a more holistic, comprehensive understanding.

For example, there exist complex behavioral patterns of the mixing of cultural values that have distinct, counter-intuitive manifestations in Confucian cultures. These patterns have been explored in-depth in the Bayesian analysis of “cultural additivity,”

indicating how Confucian, Buddhist, and Taoist values influence folktale outcomes for lies and violence (Vuong et al., 2018; Vuong et al., 2020). Another example is there are cultural forces behind the high tolerance of Chinese people with “copycats” or “shanzai” in Chinese. There is also a body of literature for this *shanzai* culture (Hua, 2011; Jiang & Shan, 2016).

Regarding the ethics, design, and governance of emotional AI in Confucian contexts, I suggest there are three bodies of the literature that provide the most nuances and allow effective cross-cultural understanding between the East and the West to take place. They are the literature on i) Confucianism and emotions; ii) Confucianism and privacy; iii) Confucianism and democracy.

Why *emotions*, *privacy*, and *democracy*? Because each of these words signifies subtly different things in Confucian culture compared with the West. In other words, each of them has a counter-meaning in Western culture.

Moreover, the above three bodies of the literature are each related to the ethics, design, and governance of emotional AI in different levels of discussion.

- **Data:** The literature on Confucianism and emotion will inform how emotional data should be managed and structured for functioning EAI systems.
- **Legal framework:** The literature on privacy and Confucianism will inform how people in Confucian culture see the problem of data collection and any broader governing issues.
- **Critical aspects of the moral sphere:** The literature on democracy and Confucianism will inform how people in Confucian culture form their visions of society. But for this to work, the current debates cannot escape a clear view of the moral vision, which would likely be widely adopted by society (Basourakos, 1998).

A comprehensive analysis of how the discourses on emotional AI is formed and communicated within Confucian contexts. The rise of East Asian economies in terms of economic power and population means there will be more and more people in these

cultures consume the EAI products. As with the increasing ease of computing and coding platforms, there will be more people in these cultures to participate in designing technologies. These trends appear to have, thus far, attracted inadequate attention among researchers, leaving a considerable gap in the future body of knowledge.

References

- Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, 6(2), 154-175.
- Basourakos, J. (1998). Exploring the Moral Sphere through Dramatic Art: The Role of Contemporary Canadian Plays in Moral Pedagogy. *Revue Canadienne de l'éducation*, 23(3), 265-280.
- Cinnamon, J. (2017). Social injustice in surveillance capitalism. *Surveillance & Society*, 15(5), 609-625.
- Hua, Y. (2011). *China in ten words*. Anchor.
- Jiang, L., & Shan, J. (2016). Counterfeits or Shanzhai? The role of face and brand consciousness in luxury copycat consumption. *Psychological Reports*, 119(1), 181-199.
- Lee, K. F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Houghton Mifflin Harcourt.
- McStay, A. (2018). *Emotional AI: The rise of empathic media*. Sage.
- Vuong, Q. H., Bui, Q. K., La, V. P., et al. (2018). Cultural additivity: behavioural insights from the interaction of Confucianism, Buddhism, and Taoism in folktales. *Palgrave Communications*, 4(1), 143.

- Vuong, Q. H., Ho, M. T., Nguyen, H. K. T., et al. (2020). On how religions could accidentally incite lies and violence: Folktales as a cultural transmitter. *Palgrave Communications*, 6(1), 82.
- Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75-89.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.