

How Ethical Design of Artificial Intelligences Systems is Possible in a Transcultural Perspective

Liang Wang¹; Kehao Zhang²

1. School of Marxism, Xi'an Jiaotong University, Xi'an 710049, China; E-Mails: wangg85@163.com. 2. School of Marxism, Xi'an Jiaotong University, Xi'an 710049, China; E-Mails: zhangkh@mail.xjtu.edu.cn.

Abstract: In the face of today's culturally diverse social reality, the ethical design of Artificial Intelligences Systems (AIS) no longer follows only one cultural tradition, but must take into account multiple cultural traditions. Therefore, it is necessary to study the possibility of ethical design of AIS from a transcultural perspective. This paper answers the question of the possibility of ethical design of AIS in transcultural perspective from four aspects: ethical value base, ethical consensus, designer's ethical responsibility, and internal ethical decision-making mechanism of AIS from both theoretical and practical levels.

Keywords: Transcultural; Artificial Intelligences Systems; Ethical Design

1. The Broadest Basis of Ethical Value: "Common Value"

Normativity can only be a function of ethics, not its foundation; the moral foundation of ethics is precisely value, which reveals what makes virtue possible. The value ethical path reflects more the problem of the foundation of ethics, so it is obvious that the value ethical path is more profound than the normative ethical path to explore the ethical problems of AI. Moreover, value is not only the moral foundation of ethics, but also the core issue of culture. Culture and values are intrinsically linked, and any culture carries certain values and is the embodiment of certain values. Since value is both the basic issue of ethics and the core issue of culture, it is deservedly the natural combination of culture and ethical issues. The study of transcultural AI ethics necessarily requires the penetration of value ethics theory, and its core is to explore whether ethical norms of different value orientations can be embedded in AIS.

Transcultural AI ethics has a global ethical nature, and it needs a globally shared value concept as a theoretical support. The community of human destiny is based on the basic value of "there is only one earth for human beings and one world for all countries", and believes that today's globalization, informationization, and the complex situation faced by all countries have linked the destiny of all countries into one, and no country can do it alone. It can be seen that this value consensus based on today's reality is not a waterless source or a wood without a foundation, but a consensus at the spiritual level among the stakeholders of the community of human destiny, which is an important embodiment of the concept of "common values". Therefore, it is more realistic and operable to take the "common value" concept of human destiny community as the value concept support of transcultural AI ethics.

2. The Biggest Ethical Consensus: "the Good Life"

"The good life" is a kind of value pursuit, one of the common values pursued by human beings. Since ancient times, what is the good life and how to achieve it have been ethical issues

that people all over the world have been pursuing and discussing. Therefore, it can be said that "the good life" is the greatest ethical consensus in different cultures. Coeckelbergh argues that the ethics of artificial intelligence should take into account cultural differences and focus on how artificial intelligence can help human beings achieve prosperity and well-being.^[1] Duan Weiwen believes that the development of artificial intelligence should be consistent with the value goal of human well-being.^[2] Capurro believes that robotics ethics is the use of robots to pursue "the good life" according to different customs and cultures.^[3] According to Triolo, "the obvious obstacle to the widespread use of AI is the extremely difficult balance of privacy issues," but there is also a broad consensus that "AI should benefit all of humanity" and "the principle of the public good. and other broad consensus."^[4]

The pursuit of "the good life" is diverse, for example, "material abundance", "spiritual abundance", "convenience of life", "eco-friendly", etc. For example, "material abundance", "spiritual abundance", "convenience of life", "eco-friendly", etc. When the application of AI technology is consistent with the pursuit of "the good life", "the good life" becomes a value leader, and people in most countries or regions are interested in the application of AI technology to create abundant material wealth and environmental protection. Most people in most countries or regions welcome the application of AI technology to create rich material and spiritual wealth to improve the quality of life and achieve the value goal of "the good life". Therefore, the ethical consensus of "the good life" has a good cohesive effect on transcultural values and is one of the prerequisites for the implementation of the ethical design of AIS from a transcultural perspective.

3. Designers' Sense of Ethical Responsibility and Transcultural Understanding

Edmond Awad et al. simulated a driverless car ethical decision-making scenario and surveyed the ethical decision-making preferences of millions of people in 233 countries and regions around the world online, concluding that cultural traditions, economies, and the state of the legal system in different regions are closely related to human ethical decision-making preferences.^[5] Obviously, the ethical design of AIS cannot be done without the role of designers and other engineers, and the ethical decision preferences of designers directly affect the ethical preferences of the designed AIS. Therefore, several scholars have explored the transcultural issues of AIS from the designers' perspective, respectively. Anniina Huttunen et al. argue that to address transcultural AIS ethical issues, in addition to developing universally accepted ethical provisions, relevant engineers should be required to learn a culturally competent "neo-Archimedean oath".^[6] Takeshi Kimura believes that robotics engineers should be familiar with different social, cultural and ethical connotations, and should have a good sense of ethical responsibility.^[7] Coeckelbergh believes that artificial intelligence ethics should shift from the "robot-centric" paradigm to the "human-centric" paradigm, combining cultural differences to examine the issue of virtue ethics, that is, "ethics of human good", and its connotation is to encourage people to use existing human-computer interaction.^[11] "Ethical experience" and "moral imagination" build a human-machine co-living life that enhances the prosperity and well-being of mankind.^[11] Peter-Paul Verbeek also puts forward responsibility requirements for designers from the perspective of "moral imagination". He believes that when designers try to imagine the intermediary role that the technology they have designed may play in user behavior, they can feed back expectations into the design process.^[8]

Therefore, it can be seen from these discussions that the ethical design of AIS from a transcultural perspective is very demanding for designers, and how to enhance and ensure that

designers have a strong sense of ethical responsibility and a deeper transcultural understanding is the key to the problem.

4. Cultural Inclusive Moral Decision-Making Mechanism of AIS

The design thinking and values of the AIS designers are ultimately reflected in the internal ethical decision-making mechanism of the AIS. Therefore, the success of designing a culturally inclusive internal ethical decision-making mechanism is the key to the successful implementation of the ethical design of AIS from a transcultural perspective. Wallach and Allen argue that in the face of transcultural diversity and complexity, the design of AI moral decision-making systems should combine both "bottom-up" moral learning and "top-down" moral evaluation.^[9] Although Wallach and Allen do not explicitly state that the design of their ethical decision-making system is specific to transcultural AIS, their design pathway nevertheless provides good ideas for the ethical design of transcultural AIS. However, further adaptation of their pathway is needed. A clear idea would be to analyse in greater depth how AIS can be made to accommodate cultural diversity in the process of 'bottom-up' ethical learning and 'top-down' ethical assessment. Some of the existing research by other scholars provides a good reference for the author. Xu Yingjin profoundly analyzed the relationship between the computational model of artificial neuron network in artificial intelligence and the metaphorical projection of the Confucian tradition of "virtue fumigation" and "Confucian virtue model library".^[10] Dehghani et al. advocate the creation of cultural databases to inform ethical decision making in artificial intelligence.^[11] Moon et al. argue that an open source model should be used to co-create an ethical knowledge base for artificial intelligence in the face of cultural differences.^[12] Therefore, one of the ideas is that the "bottom-up" intercultural ethical design path of AIS is feasible if the database of intercultural ethics is constructed to provide abundant intercultural ethical cases for the "bottom-up" moral learning of AIS. In addition, if we consider the "top-down" ethical assessment path, the value embedding study of AIS from a transcultural perspective provides us with good insights and theoretical pavement. Li Lun and Sun Baoxue believe that AIS should have built-in value correction mechanisms.^[13] Aimee van Wynsberghe believes that the ethical embedding of AIS should include the value-sensitive mechanism.^[14] Nagenborg argues that values are always consciously embedded in "ethical subroutines" in the design of artificial moral agents, and that one should consciously reflect on these (universal, Western or non-Western) values.^[15] Spiekermann regards the setting and understanding of value goals as the first step of AIS ethical design, and believes that this step is based on value ethics.^[16]

Therefore, another idea is to provide reasonable value reference and value guidance for "top-down" moral evaluation by embedding value consensus mechanism, then the "top-down" intercultural ethical design path of AIS is feasible. It can be seen that the combination of the "bottom-up" moral learning path based on transcultural ethics database and the "top-down" moral evaluation path based on value consensus greatly improves the possibility of the ethical design of AIS from the transcultural perspective on the practical operation level.

5. Conclusion

This paper mainly answers the possibility of ethical design of AIS from the perspective of theory and reality. Specifically, it can be divided into four aspects: ethical value basis, ethical consensus, ethical responsibility of designers, and internal moral decision-making mechanism of

AIS. In terms of the first aspect, "common value" is the most extensive ethical value basis in transcultural ethical issues. In terms of the second aspect, it is the biggest ethical consensus of artificial intelligence ethics from a transcultural perspective to make artificial intelligence technology create more benefits for human beings and create "the good life". In terms of the third aspect, "moral experience", "moral imagination" and the establishment of global ethical norms are all helpful to enhance and guarantee designers' awareness of ethical responsibility and transcultural understanding. In terms of the fourth aspect, the strong support of the transcultural ethics database for the "bottom-up" moral learning and the value consensus for the "top-down" moral evaluation provide a very useful reference for the ethical design of AIS from the transcultural perspective.

Acknowledgments

I would like to express my sincere thanks to all those who have helped me in the course of my writing this paper. This paper was funded by the National Social Science Foundation of China under the Youth Program: Research on the Ethical Embedding Mechanism of Artificial Intelligence in Transcultural Perspective (19CZX018).

References

- [1] Coeckelbergh M. Personal robots, appearance, and human good: a methodological reflection on roboethics[J]. *International Journal of Social Robotics*, 2009, 1(3): 217-221.
- [2] Duan Wei-wen. Value Reflection and Ethical Adjustment in the Era of Artificial Intelligence [J]. *Journal of Renmin University of China*, 2017, 31(6): 98-108. (in Chinese)
- [3] Capurro R. Intercultural roboethics for a robot age [EB/OL]. http://www.capurro.de/intercultural_roboethics_2017.pdf, 2017/2019-06-08.
- [4] Triolo P. The Challenge of Establishing Global Ethical Norms for Artificial Intelligence [EB/OL]. https://www.sohu.com/a/272210929_468736, 2018-10-30/2019-03-08. (in Chinese)
- [5] Awad E, Dsouza S, Kim R, et al.. The moral machine experiment[J]. *Nature*, 2018, 563(7729): 59.
- [6] Huttunen A, Brace W, Kantola V, et al.. Cross-cultural Application of Ethical Principles in the Design Process of Autonomous Machines [EB/OL]. http://www.iis.org/CDs2010/CD2010IMC/ICEME_2010/PapersPdf/FB476NT.pdf, 2010/2019-08-05.
- [7] Kimura T. Roboethical arguments and applied ethics: Being a good citizen [C]// Yoshiyuki Sankai, Kenji Suzuki, and Yasuhisa Hasegawa. *Cybernetics*. Springer, Tokyo, 2014: 289-298.
- [8] Verbeek P P. *Moralizing technology: Understanding and designing the morality of things* [M]. Chicago: University of Chicago Press, 2011: 99.
- [9] Wallach W, Allen C. *Moral Machines: Teaching Robots Right from Wrong* [M]. New York: Oxford University Press, Inc., 2009.
- [10] Xu Yingjin. Confucian Virtue Ethics, Neural Computation and Cognitive Metaphors [J]. *Wuhan University Journal (Social Science)*, 2017, 70(6): 49-59. (in Chinese)
- [11] Dehghani M , Tomai E , Forbus K , et al. An Integrated Reasoning Approach to Moral Decision-Making [EB/OL]. [extension://oikmahipjniocckomdcccplodldodja/pdf-viewer/web/viewer.html?file=https%3A%2F%2Fwww.aaai.org%2FPapers%2FAAAI%2F2008%2FAAAI08-203.pdf#=&zoom=130](https://oikmahipjniocckomdcccplodldodja/pdf-viewer/web/viewer.html?file=https%3A%2F%2Fwww.aaai.org%2FPapers%2FAAAI%2F2008%2FAAAI08-203.pdf#=&zoom=130), 2008/2019-02-05.
- [12] Moon A, Calisgan C, Operto F, et al. Open roboethics: Establishing an online community for accelerated

policy and design change [EB/OL].
http://robots.law.miami.edu/wp-content/uploads/2012/01/Moon_et_al_Open-Roboethics-2012.pdf,
2012/2019-01-05.

[13] Li Lun, Sun Baoxue. Enduing Artificial Intelligence "Moral Chips(Conscience)" : Four Dimensions of Ethical Research of Artificial Intelligence [J]. *Teaching and Research*, 2018(08):72-79. (in Chinese)

[14] Van Wynsberghe A. Designing robots for care: Care centered value-sensitive design[J]. *Science and engineering ethics*, 2013, 19(2): 407-433.

[15] Nagenborg M. Artificial moral agents: an intercultural perspective[J]. *International Review of Information Ethics*, 2007, 7(09): 129-133.

[16] Spiekermann S. *Ethical IT innovation: A value-based system design approach*[M]. New York: Auerbach Publications, 2015: 239-241.