

机器人和人工智能伦理丛书

中国机器人 伦理标准化前瞻 (2019)

北京大学

国家机器人标准化总体组——著

Chinese
Prospects for
the Standardization of
Robot Ethics 2019



北京大学出版社
PEKING UNIVERSITY PRESS



机器人和人工智能伦理丛书

中国机器人 伦理标准化前瞻 (2019)

Chinese
Prospects for
the Standardization
of Robot Ethics 2019

北京大学
国家机器人标准化总体组——著



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目 (CIP) 数据

中国机器人伦理标准化前瞻. 2019 / 北京大学国家机器人标准化总体组
著. —北京: 北京大学出版社, 2019.10
(机器人和人工智能伦理丛书)
ISBN 978-7-301-30842-4

I. ①中… II. ①北… III. ①机器人-技术伦理学-标准化-中国-
2019 IV. ①TP242-65 ②B82-057

中国版本图书馆 CIP 数据核字 (2019) 第 219441 号

书 名	中国机器人伦理标准化前瞻 (2019) ZHONGGUO JIQIREN LUNLI BIAOZHUNHUA QIANZHAN 2019
著作责任者	北京大学国家机器人标准化总体组 著
责任编辑	田 炜
标准书号	ISBN 978-7-301-30842-4
出版发行	北京大学出版社
地 址	北京市海淀区成府路 205 号 100871
网 址	http://www.pup.cn
电子信箱	pkuwsz@126.com
电 话	邮购部 010-62752015 发行部 010-62750672 编辑部 010-62750577
印 刷 者	北京中科印刷有限公司
经 销 者	新华书店
	880 毫米 × 1230 毫米 16 开本 11.25 印张 112 千字
	2019 年 10 月第 1 版 2019 年 10 月第 1 次印刷
定 价	50.00 元

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究

举报电话: 010-62752024 电子信箱: fd@pup.pku.edu.cn

图书如有印装质量问题, 请与出版部联系, 电话: 010-62756370



缩略语

BSI	英国标准委员会
COSDP	中国优化共生设计方案
EGE	欧盟科学与新技术伦理组
EURON	欧洲机器人学研究网
IEEE	电器与电子工程师协会
IFR	国际机器人学联合会
ISO	国际标准化组织
SAC	中国国家标准化管理委员会
UN	联合国

Abbreviations



BSI	British Standards Institution
COSDP	Chinese Optimizing Symbiosis Design Programme
EGE	European Group on Ethics in Science and New Technologies
EURON	European Robotics Research Network
IEEE	Institute of Electrical and Electronics Engineers
IFR	International Federation of Robotics
ISO	International Organization for Standardization
SAC	Standardization Administration of the People's Republic of China
UN	United Nations



序言

王天然

中国工程院院士

中国科学院沈阳自动化研究所

追随着机器人技术的进步和全球机器人产业的蓬勃发展，机器人的应用场景已经由早期的工业领域，扩展到了教育、医疗、家庭、安防等领域，机器人为社会带来了新的增长点，为家庭带来新的生活模式。

机器人与人工智能技术及其他科学技术的深度结合，使得机器人的智能化水平、适应环境的能力、与人交互和协作的能力不断提高。未来机器人会越来越地融入人类的生活中，与人紧密协调，人和机器人会逐步形成新的“人机共融关系”。例如在制造业中，工人和机器人打破物理区隔，互相配合，共同完成一项作业任务；在家庭生活中，服务机器人与人存在于同一空间，机器人能够与人聊天、护理病人、完成家务等。这些都意味着机器人将从“工具”变为人类的“合作伙伴”。



Foreword

Wang Tianran

Academician of Chinese Academy of Engineering
Shenyang Institute of Automation
Chinese Academy of Sciences

Consistent with the rapid progress in robotic technologies and the global boom of the robotics industry, the role of robots in society is being extended from the familiar locus of industry to areas such as education, healthcare, family life, and security. Robots are adding energy to economic growth, and are producing a new mode of life in our homes.

The productive intersection of robotics, AI, and other technologies is raising the level of robotic intelligence, and facilitating their adaptation to the environment and their interaction with humans. In the future, robots will become increasingly integrated into human life and will collaborate closely with us. A truly “harmonious relationship” between humans and robots is taking shape. For example, we are lowering the physical barrier against robots and are working together with them on the factory floor. At home, service robots share our living space, are capable of keeping us company, provide home care for us, and undertake a multitude of tasks. Such developments show clearly that the role of robots is changing from being mere “instruments” to becoming our “cooperative friends.”

机器人与人共融的紧密关系，给人类工作和生活带来极大便利，但也给我们原有的社会秩序带来了新的挑战。比如工业机器人与人共同工作，一旦发生事故，责任如何划分；家用服务机器人进入家庭，会不会泄露个人隐私，人类会不会对其产生“感情”；服务于医院的公共机器人决策过程中会不会带有偏见，例如分配急诊病人的优先次序时带有偏见等。人们担忧机器人不知何时会失控、在人类不知情的情况下侵犯人类的权利，更担忧这些行为无法追溯和问责。

这些问题并非是技术瓶颈，而是机器人如何“做决定”的问题，也就是机器人伦理问题。这些问题的解决方式直接决定了公众对于机器人及其技术的接受程度。因此如何对机器人的行为加以约束，如何引导机器人的行为和公共决策能够符合人类的基本价值观，成为机器人设计者、生产者、应用者乃至社会管理者都十分关注的课题。本书以此为出发点，从标准化的角度，来研究和探讨机器人的伦理问题，提出了一整套机器人伦理体系，用于引导机器人的设计、生产和应用，规避机器人的伦理风险。

我深信，本书必将引起社会的广泛关注，引发全社会对于机器人伦理的热烈讨论，吸引更多的有识之士投入到机器人的伦理研究中。同时希望这项研究能够为未来机器人伦理标准的制定乃至伦理法律治理奠定基础，推动机器人的设计和研发越来越有益于人类和自然的和谐发展。

While robots offer significant benefits in the workplace and in other life activities, at the same time the intimate relationship emerging between robots and human beings also challenges our existing societal order. For example, how to assign liability becomes unclear when accidents occur in the collaboration between industrial robots and humans. Again, service robots raise the risk of leaking private data when they enter our home space. It has also been discovered that human emotions function to bind us unilaterally to our robot friends. Furthermore, healthcare robots in the hospital environment could conceivably manifest biases in the processes of decision-making such as prioritizing certain groups of people in emergency services. And finally, there is real concern that we might lose control over robots to the extent of human rights being violated, and the difficulty of accountability and assigning liability as such cases occur.

Such problems are not technical in nature, but rather raise concerns about how robots “make decisions,” and thus are problems of robot ethics. Success in solving such problems will have an immediate impact on the public’s willingness to accept robots and robotic technologies. It is for this reason that how to restrain robotic “behaviours” and how to align robotic “behaviours” and “decisions” with our fundamental values becomes a major concern of robotic designers, manufacturers, users, and government agencies. Facing such challenging issues, this publication is a strong beginning in the research and discussion necessary to bring standardization to robot ethics, and offers a comprehensive ethical system to preempt ethical risks and hazards in the expanded use of robots and that can be used to orient efforts in robotic design, manufacture, and application.

I am certain that this present work will attract wide attention among the public, and stimulate robust discussions concerning robot ethics across the whole of society, and in so doing, will encourage more people to undertake research on robot ethics. At the same time, I am sure that this work will serve to ground our ethical standards and inform our legal governance over robotics for the immediate future, and will inspire the kind of robotic design and development that will foster a harmonious relationship between humans and nature.

王博

哲学系教授

北京大学副校长

从来没有任何一种生物像人类一样试图通过知识和技术改变这个世界，并取得了巨大的成功。这种成功对于人类生活的影响显而易见。从平均寿命的延长到日常生活的便利，从外太空的探索到精神世界的拓展，人类不断地赞美和惊叹于自己的力量，同时思考这种力量是否会让人类的未来充满了更多的不确定性。机器人革命时代的来临让人们更加兴奋，也让人们更加忧虑。

对于技术的担心贯穿人类精神生活的始终。古代中国的哲人庄子曾经说过：“有机械者必有机事，有机事者必有机心。”在直接的意义上，庄子忧虑技术给人心和人类生活带来的不可逆的负面影响。而在更深刻的意义上，他意识到的是知识和技术的有限性，它们固然可以解决一些问题，却也带来了更大的问题。《周易》同样指出单向度的“进”会导致“亢龙有悔”的局面，肯定“知进退存亡而不失其正者”，才是所谓的圣人。庄子和《周易》的思考当然无意也无法阻挡知识和技术的进步，但仍然有助于人们在热闹中保持清醒的头脑，兼顾科技和人文的平衡。

Wang Bo

Professor of Philosophy,
Vice-President of Peking University

There has never been a species like humankind that has attempted to change its world through knowledge and technology, and that has been so successful in doing so. The impacts of this knowledge and technology on our way of life is everywhere evident. From a much-extended longevity and the conveniences of everyday life to space exploration and expanding spiritual world, we human beings have accomplished much and have come to extol our own power in accruing such magnificent achievements. At the same time, we cannot help but wonder whether or not the unleashing of such power will bring with it real uncertainties for the human experience. The pending era of robotic revolution excites and worries us both at once.

Concern over the impact of new technologies is pervasive in human life. Zhuangzi, one of the most prominent sages in ancient China, once stated: “where there are machines, there are bound to be machine worries; where there are machine worries, there are bound to be machine hearts”. On the surface, Zhuangzi seems to be concerned about the irreversibility of the negative impacts that technology may have on the human life and spirit. Yet at a deeper level, he offers us a profound reflection on the limits of our knowledge and technology. For him, knowledge and technology create more problems than they solve. The *Yijing*, as the first among the Chinese classics, avers that progress might bring with it the disastrous consequences captured in the image of “the wilful dragon repents” and says further that only those persons who “while understanding progress and recession, life and death, can still maintain their integrity” can be called a sage. Neither of these canonical texts are recommending that we abandon intellectual and technological progress, but rather that in our enthusiasm to move ahead we remain clear-headed and maintain a balance between scientific knowledge and our humanistic concerns.

中国传统艺术的一个重要观念是留白。基于对有和无、实和虚之间关系的理解，艺术家们并不想占有全部的空间，而是把很大的一部分留给事物，也留给欣赏者。这不仅创造了美，也创造了事物的自由流动和欣赏者自由的想象。在这个意义上，留白是善，也是智慧。对于机器人、基因工程等技术来说，意识到留白的妙处，是非常必要的。

哲学家的特点是从整体上来理解世界，关注事物的多个面向，并寻求最恰当的生活。这个特点体现在《中国机器人伦理标准化前瞻（2019）》（以下简称《前瞻》）的工作中。以整体论的善为根基，围绕着多元、自然、正义和繁荣，倡导世界共生秩序，使这项工作充满了对伦理价值的尊重和社会责任的担当。而去人类中心主义的立意不仅让《前瞻》关心人类福祉，也关心自然存在物，引导机器人业通向一个共生和谐的良性世界秩序。我们可以看到这项工作的主持者们自觉地吸取传统中国的思想资源和现代世界的一般价值，力图使这份中国优化共生设计方案取得广泛的国际共识。希望这个方案可以引起科学家、人文学者、企业家、政治家和普通人群的关注，在广泛的讨论中得到进一步的完善。

In Chinese traditional painting, one of the most important concepts is to leave spaces empty. Based on their understanding of the relation between presence and absence and between fullness and emptiness, Chinese painters never occupy the entire space of a picture but leave most of it blank for its figures and for their observers. The art of leaving blank space not only expresses a kind of beauty, but also allows for the flow of the figures and for the free play of imagination on the part of the observers. In this sense, leaving blank space is felicitous, and manifests a kind of wisdom as well. On the same principle, it is thus indispensable to acknowledge the significance of leaving blank space in developing cutting-edge technologies of robotics and gene editing.

An important characteristic of Chinese philosophy is to understand the world holistically, and in thus seeing things from multitude of perspectives, to seek after the most appropriate form of life for all of us. Such an attitude is much in evidence in the production of the *Chinese Prospects for the Standardization of Robot Ethics 2019* in its respect for ethical values and social responsibility. The *Prospects* are rooted in a holistic concept of an optimal good from which it derives a formal framework that includes the fundamental principles of pluralism, nature, justice and well-doing. With its commitment to optimizing symbiosis, the *Prospects* eschew any kind of anthropocentrism in its approach to human welfare and the protection of the natural environment, and seek to incorporate the development of robotics into a well-ordered and integrated world. The authors of the *Prospects* have deliberately drawn upon both the Chinese traditions and modern values in developing Chinese Optimizing Symbiosis Design Program (COSDP) that has the capaciousness to attract a broad international consensus. It is our hope that the proposed program will draw the attention among scientists, humanities scholars, entrepreneurs, decision-makers and the public at large, and that it can be further improved upon by continuing discussion among all of the constituents of society.

北京大学一直关注变化世界中人类所面临的重大问题和挑战。学校在人文、社会科学和科学等领域的知识和人才储备，无疑是主导这项工作的基础。而不同学科和机构学者之间的合作，让这项工作的意义超出了工作本身。人文学者对于科技革命以及它所改变的世界的关注，会让人文学更具生命力和影响力。而科学家们的人文情怀和视角，也会让智能科技时代充满对事物和人类体贴的精神。

Peking University is unrelentingly attentive to the big questions and challenges that we face in an always changing world. Our rich intellectual sources in the areas of the humanities and in the social and natural sciences have provided a strong and certain foundation for producing the *Prospects*. The kind of cross-disciplinary collaboration among scholars from various institutes and disciplines makes such an undertaking more significant than the assemblage of any particular product. For the humanists, attention to the revolution in the sciences and to ongoing advances in technology can instill vigor and extend the reach and influence of our humanities research. And for the scientists, an awareness of the perspective brought by the humanists in this age of artificial intelligence will contribute importantly to appropriate concerns about both our world of things and our own species.



执行总结

致谢

在中国国家标准化管理委员会的精心组织和协调下，国家机器人标准化总体组与北京大学哲学系和外国哲学研究所紧密合作，正式制定并颁布首部《中国机器人伦理标准化前瞻（2019）》。

《中国机器人伦理标准化前瞻（2019）》的成功撰写依赖众多机构、学者和机器人业参与者的支持和帮助。我们首先诚挚感谢中国国家标准化管理委员会和国家机器人标准化总体组远见卓识地设立引导专项，及时启动研究并制定中国机器人伦理标准化体系。我们感谢国家机器人标准化总体组对机器人伦理标准化工作始终不渝的信任、鼓励和支持。

我们感谢北京大学，北京大学哲学系和北京大学外国哲学研究所从始至终对这部《中国机器人伦理标准化前瞻（2019）》研究和制定工作的鼓励与支持。

这份文件的撰写来自北京大学和浙江大学专业学者敬业、默契、合作和高效的集体工作。我们感谢工作组所有成员的忘我付出、出色行文、严谨分析和睿智思路。

Executive Summary



Acknowledgements

Under the guidance of the Standardization Administration of the People's Republic of China, The Department of Philosophy and the Institute of Foreign Philosophy, Peking University, have collaborated with the National Robotics Standardization Committee to produce this document *Chinese Prospects for the Standardization of Robot Ethics* (2019).

We would like to thank the Standardization Administration of the People's Republic of China (SAC) and the National Robotics Standardization General Working Group of China for taking the timely initiative to elaborate a Chinese robot ethical standard. We again thank the National Robotics Standardization General Working Group, for the persistent trust, encouragement and their support.

We would also like to thank the Department of Philosophy and the Institute of Foreign Philosophy at Peking University for their continuing encouragement and support for this *Prospects* project.

This document is the result of cooperative team work of academic people from both Peking University and ZheJiang University. We thank all the members of the team for their energetic engagement with this project, for their substantial contributions to it, for the rigor they brought to the analysis, and for the intelligent solutions they had on offer.