AI and Ethics (2023) 3:505–511
https://doi.org/10.1007/s43681-022-00180-6

ORIGINAL RESEARCH

Confucius, cyberpunk and Mr. Science: comparing AI ethics principles between China and the EU

Pascale Fung[1] · **Hubert Etienne**[2]

Received: 29 March 2022 / Accepted: 21 May 2022 / Published online: 20 June 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract
We propose a comparative analysis of the AI ethical guidelines endorsed by China (from the Chinese National New Generation Artificial Intelligence Governance Professional Committee) and by the EU (from the European High-level Expert Group on AI). We show that behind an apparent likeness in the concepts mobilized, the two documents largely differ in their normative approaches, which we explain by distinct ambitions resulting from different philosophical traditions, cultural heritages and historical contexts. In highlighting such differences, we show that it is erroneous to believe that a similarity in concepts necessarily translates into a similarity in ethics as even the same words may have different meanings from a country to another—as exemplified by that of "privacy". It would, therefore, be erroneous to believe that the world would have adopted a common set of ethical principles in only three years. China and the EU, however, share a common scientific method, inherited in the former from the "Chinese Enlightenment", which could contribute to better collaboration and understanding in the building of technical standards for the implementation of such ethics principles.

Keywords AI ethics · Europe · China · Cyberpunk · Regulation

1 Introduction

The exponential development of artificial intelligence triggered an unprecedented global concern for potential social and ethical issues. Stakeholders from different industries, international foundations, governmental organizations, and standards institutions quickly reacted in improvising codes of ethics for the purpose of establishing a first layer of control in the absence of existing State laws. This exercise is comparable to that of controlling the proliferation of nuclear weapons in the 1960's. Its objective was no less than to reach a global agreement on common ethical standards to regulate one of the most promising technologies whose crucial strategic implications on both business and political grounds are already well acknowledged. The resulting profusion of documents on AI ethical standards, as much as 84 identified by Jobin et al. [14] and 160 in Algorithm Watch's *AI Ethics guidelines Global Inventory* [1], however, deserve to be scrutinized.

A major concern is the large homogeneity and presumed consensualism around these principles. Jobin et al. [14] identified 11 clusters of ethical principles among 84 documents and Fjeld et al. [9] found 8 key themes across 36 of the most influential of these. They both noted a general convergence, which leads Fjeld et al. to conclude that **"**the conversation around principled AI is beginning to converge" and that "these themes may represent the 'normative core' of a principle-based approach to AI ethics and governance". However, we argue that ethics, by nature, is not consensual. While it is true that some ethical doctrines, such as Kantian deontologism, aspire to universalism, they are however not universal in practice. In fact, ethical pluralism is more about differences in which relevant questions to ask rather than different answers to a common question. When people abide by different moral doctrines, they tend to disagree on the very approach to an issue. Therefore, even when people from different cultures agree on a set of common principles, it does not necessarily mean that they share the same understanding of these concepts and what they entail.

✉ Hubert Etienne
hubert.etienne@sciencespo.fr

[1] Centre for Artificial Intelligence Research (CAiRE), The Hong Kong University of Science and Technology, Hong Kong, China

[2] Department of Philosophy, Ecole Normale Supérieure, Paris, France

To better understand the philosophical roots and cultural context underlying ethical principles in AI, we propose to analyze and compare the ethical principles endorsed by the Chinese National New Generation Artificial Intelligence Governance Professional Committee (CNNGAIGPC) and those elaborated by the European High-level Expert Group on AI (HLEGAI). China and the EU have very different political systems and diverge in their cultural heritages. In our analysis, we wish to highlight that principles which seem similar a priori may actually have different meanings, derived from different approaches and reflect distinct goals (Table 1).

2 Promotional vs prohibitive approaches

At first glance, the Chinese ethical principles seem similar to those of the EU in many aspects. Both notably promote fairness, robustness, privacy, safety and transparency. Their prescriptive approaches, however, reveal different cultural perspectives associated with different objectives.

2.1 A collective vs an individualist cultural heritage

Confucian philosophy has shaped the governing system in China and the rest of East Asia for centuries. It emphasizes the "rule for the people", rather than "rule by the people", and favors an elitist leadership, associating political mandates with competence and merit. The Chinese government's belief in "doing the right thing" for its citizens is informed by the Confucian ideas of virtuous authority and exemplary person, grounded in *ren* (humaneness), *yi* (appropriateness), *li* (rite), and *zhi* (wisdom). This philosophical tradition explains the community-focused and goal-oriented perspective, from which the Chinese guidelines derive, together with the promotions of principles, such as "harmony and friendship", "shared responsibilities", "tolerance and sharing", and "open collaboration".

Table 1 The ethical principles endorsed by the Chinese National New Generation Artificial Intelligence Governance Professional Committee (CNNGAIGPC) and those elaborated by the European High-level Expert Group on AI (HLEGAI)

Chinese ethical principles [17]	EU key requirements [10]
Harmony and friendship	Societal and environmental well-being
Fairness and justice	Diversity, non-discrimination and fairness
Tolerance and sharing	Human agency and oversight
Respect privacy	Privacy and data governance
Safe and controllable	Technical Robustness and safety
Share responsibilities	Transparency
Open collaboration	Accountability
Agile governance	

"A high sense of social responsibility and self-discipline" is also expected from individuals to harmoniously partake in a community while promoting shared responsibilities and open collaboration. The emphasis is explicitly informed by the Confucian value of "harmony" as an ideal balance to be achieved through the control of extreme passions to avoid conflicts. Other than a stern admonition against "illegal use of personal data", such value leaves little room for constraining rules. These principles are not paths to regulation, what would be detrimental to the development of research and business opportunities in a highly competitive environment where innovation is crucial. Rather, they are framed to guide AI developers in a way that would promote collective good for the Chinese society.

The European ethical principles, in contrast, emerge from a more individual-focused and rights-based approach. They express a different aspiration, rooted in the Enlightenment tradition, and colored by the European history. Their primary goal is to protect individuals against well-identified harms. Whereas the Chinese principles emphasize the promotion of good practices, the EU focuses on the prevention of evil consequences. The former draws a direction for the development of AI, so that it contributes to the improvement of society. The latter sets limitations to its uses, so that it does not happen at the expense of certain categories of people. This distinction is clearly illustrated by the presentation of fairness, diversity and inclusiveness. While the EU emphasizes fairness and diversity with regard to individuals from specific demographic groups (specifying gender, ethnicity, disability, etc.), Chinese guidelines urge for the upgrade of "all industries", reduction of "regional disparities" and prevention of data monopoly. While the EU insists on the protection of vulnerable *persons* and potential victims, China prescribes "inclusive development through better education and training, support".

The individualist perspective reflected by the European approach to AI ethics should, however, not be mistaken for a form of selfish moral individualism, but rather as a result from the European history of individual reasoning. It is worth noting that the first claim of the Enlightenment did not target political self-determination nor the possibility for people to partake in collective decision-making, but rather ontological autonomy, freeing them from the subjection to the king and the power to the State. This was famously defined by Kant [13] as "man's emergence from his self-incurred state of immaturity". This point is also illustrated in the *Declaration of Human and Citizen Rights* which gives prevalence of a citizen's protection against political power abuse — i.e., negative rights over positive rights. Finally, the repeated clashes of European nationalism that culminated in WWII and the trauma of totalitarianism acted as a powerful reminder to the Europeans of the dangers of political holism.

Consequently, European societies have shown a clear preference for individualist and rights-based approaches of governance.

2.2 Promotional vs prohibitive approaches

AI governance in both the EU and China are led by the transnational and national governments in consultation with industry stakeholders and academic experts. It is therefore pertinent to compare the actual and perceived roles of their governments in setting AI ethical guidelines. Philosophers have debated the compatibility of Confucian values with Western liberal democracies. There have also been debates on normative versus empirical legitimacy of a government, where scholars study the question of why "the observed level of regime legitimacy under non-democratic regimes has been substantially higher than either established or emerging democracies" [3]. Commenting on Shin's work [21] based on the Asian Barometer Survey, Chu [4] states that "the majority of East Asians in other countries with a Confucian legacy also tend to be attached to 'paternalistic meritocracy', prioritize economic well-being over freedom, and define democracy in substantive (rather than procedural) terms." China is more of an assumed authoritarian technocracy than an anti-democracy. Its political elite, composed of civil servants mostly with backgrounds in science, technology, engineering, and mathematics (STEM), has adopted a pragmatic approach to AI ethics, grounded in existing applications and driven by society's needs.

The Chinese leadership routinely holds workshops with scientists to keep up to date with the latest trends in advanced technologies. This proximity between political leaders and scientific research, together with the greater control exercised by the government on the development of this technology, are included in the centralized planning of the economy to serve national strategic objectives. It explains the great pragmatism in the Chinese governmental approach to AI ethics, focusing on foreseeable harms that may derive from the use of AI in the near future. The Chinese government is not averse to regulations, but in the case of AI governance, it is unlikely to regulate with a broad brush before AI has been widely applied and has found to be of serious negative impact in specific areas or posing danger to the society. Nevertheless, in areas of immediate societal impact and concern, such as data privacy, China has also been able to devise strict laws and regulations. Therefore, even though the ethical guideline call for mere "respect for privacy", it is understood that companies must exercise a great self-discipline in terms of user data protection.

In contrast, the traditional training of the European political elite does not always allow the same interest, nor understanding, for AI and new technologies in general. Furthermore, the experience with totalitarianism in Europe always serves as a reminder, calling for prudence. It also explains the greater skepticism from the European citizens about technologies, especially those that can be used for surveillance purposes. A recent example is the strong general reluctance of the French population to adopt the official COVID-tracking application out of fear of what it could be used for by the government. This is why the European approach to AI ethics was, since the beginning, conceived as both a way to regulate the private sector from foreseeable risks, and to prevent systematic distrust against AI by the public. It is intended to provide some sort of guarantee against potential abuses from public–private partnerships between governments and AI companies.

The differences between Chinese and European cultural heritage, their respective historical contexts and the background of their political elites translate into two different types of moral imperatives. The European requirements, centered on satisfying initial conditions, dictate a strict abidance by deontologist rules in the pure Kantian tradition. In contrast, the Chinese principles, referring to an ideal to aim for, express rather softer constraints at different levels, as part of a process to improve society. For the Europeans the development of AI *"must* be fair"; for the Chinese it should "eliminate prejudices and discriminations *as much as possible"*. The EU requires "processes to be transparent" while China requires to "continuously *improve* transparency". The EU principles aim to protect European citizens from vertical and horizontal abuses, conscient of the danger of nationalism. The Chinese governance system, in contrast, adopts a holistic approach, holding that the social group it forms is not to be reduced to the sum of its parts, but produces something more, namely the Chinese nation. Its ethical principles thus aim to benefit Chinese citizens through the service of the Chinese nation, considered as common good citizens are associated with.

3 A utopian vs a dystopian vision by populations

Other than roles of the governments, the two ethical sets of guidelines are informed by opposing views from the European and Chinese populations regarding AI. The main fears expressed by western society toward AI are related to privacy and surveillance, job automation [12], and the possibility of a loss of control resulting in existential risks for humanity [8]. These are greatly dependent on people's trust in political and technology leadership, on the narratives surrounding the development of AI in mainstream media, and on the representation of AI in science fiction.

3.1 The question of trust in the Government

Public opinion studies show that Chinese people are largely supportive of AI, which they associate with a great potential to benefit society, and as an engine of economic growth. Strong government support, a vibrant commercial market for AI, and media content favorable to AI all contribute to this positive perception [6]. A comparative study of German, Chinese and UK participants to assess the Attitude Towards Artificial Intelligence (ATAI scale) showed that the Chinese scored the highest on the ATAI Acceptance scale and lowest on the ATAI Fear scale [22]. Such findings are supported by another survey conducted by Ipsos, according to which 70 percent of Chinese respondents stated that they trust artificial intelligence [23]. Overall, Asian public opinion tends to be more favorable to AI. For example, a Pew Research Center survey in 2020 found that "majorities in most Asian public—Singapore (72%), South Korea (69%), India (67%), Taiwan (66%) and Japan (65%)—surveyed see AI as a good thing" for society, whereas more than half of the EU population view AI as negative. "In France, for example, views are particularly negative with only 37% of survey people considering AI as good for society versus 47% of them viewing it as bad. In the US and UK, about as many say it has been a good thing for society as a bad thing." [15]. The European historical context somewhat led to a general state of distrust in governments in many liberal democracies, which is described as the "counter-democracy" by Rosanvallon [20]. In France [16], as in the US [19], for instance, more than three quarters of the citizens think their political representatives behave unethically. The fear of AI being used by governments for mass surveillance is a major concern, and public–private collaborations are also regarded with high skepticism. From the private sector, multiple incidents and scandals related to user privacy, surveillance and nudging involving top technology companies in the past few years severely dampened consumer enthusiasm, together with the perception of these companies' intentions to do good or to operate responsibly [2].

This trust gap is particularly well illustrated by the perception of "privacy". Data privacy is promoted by both the European and the Chinese ethical guidelines, but with different meanings. The European promotion of privacy, as highlighted by General Data Protection Regulation (GDPR), encompasses the protection of individual data from both state and commercial entities. The Chinese privacy guidelines in contrast only target private companies and potential malicious agents. Whereas personal data are strictly protected both in the EU and in China from commercial entities, the State retains full access in China. Such a practice would be shocking in Western countries; it is however readily accepted by Chinese citizens, accustomed to living in a protected society and have consistently shown high trust in their government [7]. It is within the social norm in China where Chinese parents routinely have access to their children's personal information to provide guidance and protection. This difference goes back to the Confucian tradition of trusting and respecting the heads of State and family. The trust is nowadays strengthened by the great economic growth the country's leaders succeeded in achieving. A recent survey showed that Chinese government's successful domestic management of the COVID-19 crisis is likely to inspire more trust by its citizens [25]. This suggests that the trust gap would also be related to people's perception of government competency, and thus to the objectives these governments aim to achieve with AI. The most developed European countries are former global powers, which gave up on their past expansionist ambitions, and now focus on domestic policies to solve their social issues, while trying not to be left behind in the innovation race. In contrast, China has recently established itself as a world leading economy. This rapid ascend onto the world stage, together with the clear ambition to challenge the US leadership, has played a significant role in securing trust from the Chinese people in the actions of their government, including the strategic support given to AI.

3.2 The influence of the cyberpunk culture

The gap in the cultural representation of AI, perceived as a force for good in China, and as a menacing force in the dystopian technological future in the Western world, could be rooted in the influence of popular culture. Robots are assistants and companions in the Chinese vision of a technological future, whereas they tend to become insurrectional machines as portrayed by a Western media heavily influenced by the cyberpunk subgenre of science fiction and illustrated by success movies, such as *2001: A Space Odyssey* (Stanlet Kubrick, 1968), *Blade Runner* (Ridley Scott, 1982), *The Matrix* (Lana & Lilly Wachowski, 1999) or *Minority Report* (Steven Spielberg, 2002). Cyberpunk emerged in the 1960s in the West, as a subgenre of science fiction. It represents a view of a high-tech future where social orders are broken down and renegade rebel forces battle against a Big Brother government that uses technology to control the people. This vision, embodied in the works of Philip Dick and others, is a stark departure from a more positive vision of a technological future espoused by Isaac Asimov or Jules Verne in previous generations of science fiction. The influence of popular culture in shaping public opinion is well acknowledged. More particularly, Young and Carpenter [24] found that "consumption of frightening armed AI films is associated with greater opposition to autonomous weapons". Since lethal (fully) autonomous weapons systems have no official existence—or at least can we say that their use is not common yet—people's opinion about these is greatly influenced by their representations from science

fiction literature and movies: "Sci-fi as a genre, and certain iconic killer robot films in particular, appears most salient in rhetorical arguments against such weapons. […] [and robopocalyptic films themselves have been likelier to encourage a cautionary rather than techno-optimistic sentiment on armed AI, among at least sci-fi literate members of the American public" conclude Young and Carpenter.

Chinese science fiction in the early decades of the twentieth century was mainly translated from Soviet literature and designed for children. While in Chinese literature there is no tradition of describing a utopian future, there is not exactly a dystopian and cyberpunk influence either. Mainland China was mostly closed to the outside world before the 1980s, what prevented it from the influence of the cyberpunk culture and the dystopian visions of a technological future such as that conveyed by George Orwell's work. The influence of the Soviet Union stopped in 1960s when the diplomatic relationship between China and the USSR was severed, preventing the Chinese population from the dark visions of Stanisław Lem, for example. It is interesting to note an opposite trend to that of China in a country with a similar Buddhist/Confucian culture: Japan. The Japanese are found to have a relatively low level of trust in their government, in particular following the Fukushima nuclear plant crisis in 2011. Their trust in the government ranks below that of many EU countries including Germany [18]. Despite its world leading pioneer position in robotics and technophile population, there is a general doomsday malaise from a collective memory of the only atomic bomb detonation in history. The cyberpunk animation classic *Akira* (Katsuhiro Otomo, 1988) foretold a post-apocalyptic dystopian future in 2019 rife with anti-government protests and gang violence, superpowers and government sponsored assassination attempts, all in the shadow of an impending Olympic game. Akira inspired a cult following and had a strong influence on Western science fiction culture that followed, including the *Matrix* series. Nevertheless, whereas the Japanese have suffered a number of data breaches prompting their government to amend the Act on the Protection of Personal Information (APPI) in 2020, they are still relatively optimistic about AI. This is likely due to the familiarity of most Japanese with the long-standing use of AI and robotics in their manufacturing and health care sectors.

4 A scientific common ground

These gaps in both cultural representations of technology and levels of trust toward governments constitute valuable signals to explain why the Chinese principles work as paternalistic guidelines where trust is not an objective, because mistrust is not an issue, while the European principles establish the conditions for AI to be "trustworthy", as distrust has become the norm. Despite the seemingly different, though not contradictory, approaches on AI ethics from China and the EU, the presence of major commonalities between them points to a more promising and collaborative future in the implementation of these standards.

Much of operationalization and implementation of ethical standards in AI is in organizational governance, that is to say the process and application choices we make. In addition to governance, ethical principles need to be incorporated into the design of AI systems and a significant part of operationalizing these standards lies in improvements and modifications to the methodology and the architecture of modern AI software systems. AI systems research and development is an open and collaborative process across nations. Their designers from China, the US or the EU are all trained in a similar computer science and engineering curriculum based on the "scientific method". This latter paradigm—which consists in formulating hypothesis and devising empirical experiments to verify these to arrive at a claim or thesis—has underpinned research areas from statistics, signal processing, optimization, machine learning, and pattern recognition, all forming the multidisciplinary area that is modern AI today. The scientific method was first adopted by China among other Enlightenment values during the May Fourth Movement in 1919. Coined the "Chinese Enlightenment", this movement resulted in the first repudiation of traditional Confucian values, and it was then believed that only by adopting Western ideas of "Mr. Science" and "Mr. Democracy" in place of "Mr. Confucius" could the nation be strengthened. In the years since the third generation of Chinese leaders, the Confucian value of the "harmonious society" is again promoted as a cultural identity of the Chinese nation. Nevertheless, "Mr. Science" and "technological development" continue to be seen as a major engine for economic growth and livelihood improvement, hence leading to the betterment of the "harmonious society". For both governance and design, two leading international standards bodies, namely the International Standards Organization (ISO) and the Institute of Electrical and Electronic Engineers (IEEE) are working on and publishing governance and best practice guidelines for the industry. Since ISO and IEEE standards lend credibility to products and services, they are widely accepted and recognized by countries. Chinese as well as European representatives are also actively involved in these standards organizations ensuring that such standards and best practice guidelines take into account cultural norms and differences. It results that "ISO data security standards have been widely adopted by cloud computing providers, e.g., Alibaba, Amazon, Apple, Google, Microsoft, and Tencent." [5] while the working group on IEEE Guidelines for Ethically Aligned Design keeps exploring "established ethics systems, including both philosophical traditions (utilitarianism, virtue ethics, and deontological ethics) and religious

and culture-based ethical systems (Buddhism, Confucianism, African Ubuntu traditions, and Japanese Shinto) and their stance on human morality in the digital age. In doing so, […] [they] critique [ethical] assumptions [… and they] attempt[ed] to carry these inquiries into artificial systems' decision-making processes." [11].

Another reason for China to acknowledge this common scientific ground in its ethical principle of "open collaboration" relates to the hundreds of thousands of Chinese students who have gone to study in the US and the EU since the 1980's, most of them in the STEM fields. American technology companies, such as Microsoft, Amazon, Google, have all established research centers in the PRC where Chinese researchers are recruited to work with their counterparts in the US headquarters. Chinese graduate students in AI have one time or another worked as interns in these companies in China. A sampled study from the authors of one NeurIPS conference showed that nearly 30% of the authors received their undergraduate degrees in China, more than from any other country, meanwhile over 50% get their graduate degrees from the US and 16% from the EU. A significant number of Chinese AI researchers do not return to China within five years of completing their graduate studies abroad. In recent years, top Chinese AI companies, such as Tencent, Baidu, Huawei, and latecomers, such as Didi and Bytedance, also have established research labs in the US and EU to attract AI talents. Such a collaborative melting between AI researchers from various parts of the world should encourage discussion around AI ethics to help us achieve more consensus around ethical principles.

5 Conclusion

We analyzed and compared AI ethical guidelines from China and the EU both from the perspective of governmental roles, of public opinion and popular cultures, as well as from the scientific common ground for the research and development of AI in China and the West. Whereas the EU framework is rooted in the core Enlightenment values of individual freedom, equal rights and serves to protect against State abuses, the Chinese guidelines are based on the Confucian values of virtuous government, harmonious society, and targets to protect against commercial exploitation. The EU ethical framework is also built as a dialectic system between users on one side, and AI developers and service providers on the other side. These normative rules are perceived as necessary to enable trust from users, as well as that of positive interactions between these two poles. This system is dynamic and includes effective feedback loops, allowing people to keep control and improve the system via their ability to "contest and seek effective redress against decisions made by AI systems and by the humans operating them" [10]. In other words, the transparency and explicability of AI systems are required for decisions to "be duly contested". The EU principles assume skepticism from users and attempts to assuage such negative sentiment with protective rules. Although the Chinese AI ethical principles seem similar to those of the EU in many ways, they however largely differ in the overall approach. The Chinese principles start with the assumption that Chinese citizens trust the state to guide and protect them against commercial and third-party abuses. They are pointing a future direction for the development of AI, rather than its limitations. Finally, the EU principles mostly refer to deontologist normative rules—mainly negative obligations–, whereas the Chinese principles, stemming from Confucian values, tend to combine some strict deontologist normative rules (e.g., prohibiting evil uses and illegal activities) with softer constraints that could be satisfied on different levels (e.g., promote shared and inclusive development) and even some aspects of virtue ethics, referring to "vigilance" and "self-discipline. Whereas the Chinese principles tend to suggest directions to shape how AI should be developed and applied, the EU principles aim to precisely define what it should not be allowed to do.

References

1. Algorithm Watch. AI ethics guidelines global inventory. https://inventory.algorithmwatch.org/ (2020). Accessed 22 June 2021
2. Buchholz, K.: Americans fear the AI apocalypse. Statista Infographics. https://www.statista.com/chart/16623/attitudes-of-americans-towards-ai/ (2019)
3. Chu, Y.-H.: Sources of regime legitimacy in Confucian societies. J. Chin. Gov, **1**(2), 195–213 (2016)
4. Chu, Y.-H.: Sources of regime legitimacy and the debate over the Chinese model. China Rev. **13**, 1–42 (2013)
5. Cihon, P.: Standards for AI governance: international standards to enable global coordination in AI research & development. Future of Humanity Institute's technical report (2019)
6. Cui, Di., Wu, F.: The influence of media use on public perceptions of artificial intelligence in China: evidence from an online survey. Inf. Dev. **37**(1), 45–57 (2019)
7. Cunningham, E., Saich, T., Turiel, J.: Understanding CCP resilience: surveying Chinese public opinion through time. Ash Center for Democratic Governance and Innovation. https://ash.harvard.edu/publications/understanding-ccp-resilience-surveying-chinese-public-opinion-through-time (2020)
8. Gherheş, V.: Why are we afraid of artificial intelligence (AI)? Eur. Rev. Appl. Sociol. **11**(17), 6–15 (2018)
9. Fjeld, J., Achten, N., Hilligoss, H., Nagy, A>, Srikumar, M.: Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center Research Publication n° 2020–1, (2020)
10. High Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI, European Commission. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html (2021). Accessed 22 June 2021

11. IEEE Standards Association. IEEE ethics in action in autonomous and intelligent systems, IEEE. https://ethicsinaction.ieee.org/ (2021)
12. Johnson, C., Tyson, A.: People globally offer mixed views of the impact of artificial intelligence, job automation on society. Pew Research Center (2020)
13. Kant, I.: Beantwortung der Frage: Was ist Aufklärung?. English trad. from http://donelan.faculty.writing.ucsb.edu/enlight.html (1784)
14. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. Nat. Mach. Intell. **1**, 389–399 (2019)
15. Johnson, C., Tyson, A,: People globally offer mixed views of the impact of artificial intelligence, job automation on society. Pew Research Center (2020)
16. Lévy, J.-D., Bartoli, P.-H., Hauser, M.: Les perceptions de la corruption en France. Harris interactive (2019)
17. National Governance Committee for the New Generation Artificial Intelligence. Governance principles for the new generation artificial intelligence--developing responsible artificial intelligence, China Daily. http://www.chinadaily.com.cn/a/201906/17/WS5d07486ba3103dbf14328ab7.html (2021). Retrieved 22 June 2021
18. OECD. "Trust in government", The OECD data. https://data.oecd.org/gga/trust-in-government.htm (2021)
19. Pew Research Center. Why Americans don't fully trust many who hold positions of power and responsibility (2019)
20. Rosanvallon, P.: La Contre-Démocratie. La politique à l'âge de la défiance, Paris, Seuil
21. Shin, D.C.: Confucian Legacies and the Making of Democratic Citizens: Civic Engagement and Democratic Commitment in Six East Asian Countries. Cambridge University Press, Cambridge (2011)
22. Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H.S., Li, M., Sariyska, R., Stavrou, M., Becker, B., Montag, C.: Assessing the attitude towards artificial intelligence: introduction of a short measure in German, Chinese, and English language. KI-Künstliche Intelligenz **35**(1), 109–118 (2021)
23. Slotta, D.: Share of people trusting AI in China 2018. Statistica. https://www.statista.com/statistics/947661/china-share-of-people-trusting-artificial-intelligence/ (2018)
24. Young, K.L., Carpenter, C.: Does science fiction affect political fact? Yes and no: a survey experiment on 'Killer Robots.' Int. Stud. Quart. **62**(3), 562–576 (2018)
25. Wu, C.: How Chinese citizens view their government's coronavirus response. *The Conversation*. https://theconversation.com/how-chinese-citizens-view-their-governments-coronavirus-response-139176 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.