

DOI: 10.35643/Info.26.1.3

**Dossier temático: Ética de la Información**

**AI for Social Good, AI for Datong**

**IA para el bien común, IA para Datong**

**IA para o Bem Social, IA para Datong**

**Pak-Hang Wong<sup>a</sup>**

<sup>a</sup> Senior Responsible AI & Data Specialist at H&M Group, Stockholm, Sweden. ORCID: [0000-0001-8720-9912](https://orcid.org/0000-0001-8720-9912). Correo electrónico: [pak-hang.wong@hm.com](mailto:pak-hang.wong@hm.com).

**Abstract**

The Chinese government and technology companies assume a proactive stance towards digital technologies and AI and their roles in users'—and more generally, people's—lives. This vision of 'Tech for Good', i.e., the development of good digital technologies and AI or the application of them for good, is also shared by major technology companies in the globe, e.g., Google, Microsoft, and Facebook. Interestingly, these initiatives have invited a number of critiques for their feasibility and desirability, particularly in relation to the social and political conditions of liberal democratic societies. In this article, I discuss whether these critiques also apply to the Chinese context and contend that Confucian philosophy provides the normative resources to answer these critiques. This cross-cultural analysis, therefore, allows us to formulate a different account of AI4SG, which I shall call 'AI for Datong', and helps us in our reimagining of the normative vision for AI.

**Keywords:** AI FOR SOCIAL GOOD; CONFUCIANISM; TECH FOR GOOD; PLURALISM.

**Resumen**

El gobierno chino y las compañías tecnológicas asumen una postura proactiva hacia las tecnologías digitales y la inteligencia artificial (AI), y su papel en la vida de los usuarios y las personas en general. Esta visión de 'Tech for Good' (por ejemplo, el desarrollo de buenas tecnologías digitales y de AI, o su aplicación para cosas buenas), es compartida también por las mayores compañías tecnológicas del mundo (Google, Microsoft y Facebook). Sin embargo, estas iniciativas han recibido numerosas críticas por su viabilidad y conveniencia, en particular en relación con las condiciones sociales y políticas de las sociedades democráticas liberales. En este artículo se discute si las críticas también aplican al

contexto chino y sostiene que la filosofía de Confucio proporciona los recursos normativos para responder a esas críticas. Este análisis de cruces culturales, permite formular, por lo tanto, un relato diferente de AI4SG, que aquí es denominada ‘AI for Datong’, y ayuda a re-imaginar la visión normativa para la AI.

**Palabras clave:** INTELIGENCIA ARTIFICIAL; CONFUCIANISMO; TECH FOR GOOD; PLURALISMO.

### **Resumo**

O governo chinês e as empresas de tecnologia assumem uma postura pró-ativa em relação às tecnologias digitais e à IA e aos seus papéis na vida dos utilizadores - e, de um modo mais geral, das pessoas. Esta visão de "Tecnologia para o Bem", ou seja, o desenvolvimento de boas tecnologias digitais e IA, ou a sua aplicação para o bem, também é compartilhada pelas principais empresas de tecnologia do mundo, por exemplo, Google, Microsoft e Facebook. Curiosamente, estas iniciativas convidaram uma série de críticas pela sua viabilidade e conveniência, em particular no que se refere às condições sociais e políticas das sociedades democráticas liberais. Neste artigo, discuto se essas críticas também se aplicam ao contexto chinês e defendo que a filosofia confuciana fornece os recursos normativos para responder a essas críticas. Esta análise intercultural, portanto, nos permite formular um relato diferente da AI4SG, que chamarei de "IA para Datong", e nos ajuda em nossa reimaginação da visão normativa para a IA.

**Palavras-chave:** INTELIGÊNCIA ARTIFICIAL; CONFUCIONISMO; TÉCNICA PARA O BEM; PLURALISMO.

---

Received: 20/08/2020

Accepted: 02/05/2021

---

## **1. Introduction**

On November 11, 2019, Tencent, one of the largest Internet and technology companies of China and in the world, has revised its mission and vision to ‘Value for Users, Tech for Good’, and elaborated the new mission and vision. On the company’s website, it states:

Technology is powerful and evolving rapidly. The appropriate use of technology can have a significant impact on our social welfare. Technology is a tool, but the use of technology for good is a conscious choice. To us, the choice is to provide better products and services to users, to continually enhance their productivity and quality of life. We have strong convictions of right or wrong. To live up to our promise, we prioritize the needs of our

users and incorporate the consideration of social responsibility in our products and services [...]. (Tencent 2019)

There are multiple ways to interpret Tencent's pronouncement of 'Tech for Good'. However, the new mission and vision clearly express a *proactive* stance of the company to choose (or, make available) *good* and *right* products and services that enhance productivity and quality of life and account for user needs and social responsibility *for its users* and more generally *for the public*. In China, Tencent is not alone in advocating the view of 'Tech for Good', as the view has been echoed by other major Chinese technology companies such as Alibaba Group and Huawei (Cheng 2019).

This proactive stance is also evident in China's digital and AI strategies, as digital technologies and AI are being developed and applied in the public in areas such as education, healthcare, urban operations, judicial services, social governance, national security, etc. in efforts to resolve problems China faces in those areas and improve on existing practices (see, Nesta 2020). Indeed, the three sets of AI governance principles published by a key research institute for AI in Beijing, i.e., Beijing AI Principles (BAAI 2019), by the industry, i.e., the Joint Pledge on Artificial Intelligence Industry Self-Discipline (Webster 2019), and by the Ministry of Science and Technology of China, i.e., the Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence (Laskai & Webster 2019), have included "Do good", "Enhance well-being", and "Harmony and friendliness" (which amounts to "the objective of enhancing the common well-being of humanity"). While there is justified concern over whether the inclusion of 'the good' and 'well-being' is merely rhetorical, but, in fact, is an attempt to legitimize the use of technologies and AI that strengthen state control (Lucero 2019), the *explicit* mentioning of 'the good' and 'well-being', I think, demonstrates a marked difference between the Chinese perspective and the European approach to AI governance, which focuses primarily on rights (and what is right).

This is, however, not to suggest the vision of developing *good* digital technologies and AI or using them *for good* is absent outside China. Major technology companies, e.g., Google, Microsoft, and Facebook, have also introduced similar

initiatives for AI (Data) for Social Good (henceforth, AI4SG). Nevertheless, these initiatives have invited a number of critiques for their *desirability*, particularly in relation to the social and political conditions of liberal democratic societies (Latonero 2019; Green 2019; Moore 2019; also, see Berendt 2019). In this article, I discuss whether these critiques of AI4SG also apply to the Chinese context. I contend that Confucian philosophy provides the normative resources to answer these critiques. This cross-cultural analysis, therefore, allows us to formulate a different account of AI4SG, which I shall call ‘AI for *Datong*’, and helps us in our reimagining of the normative vision for AI.

Before I begin my discussion, a clarification of the *type* of argument I am making is in order. More specifically, the nature of my argument is *philosophical* but not *empirical*. Hence, I am *not* arguing that the Chinese government and technology companies’ vision of ‘Tech for Good’ is *in fact* grounded on Confucian values, nor am I arguing that the Chinese people support this vision because they *in fact* hold Confucian values. These are empirical questions that fall outside the scope of this article. What I argue instead is that Confucian values can address the critiques of AI4SG, which are based on liberal democratic values.

## **2. AI for Social Good: Definition, Critiques, and Responses**

Luciano Floridi and colleagues define AI4SG broadly as “the design, development, and deployment of AI systems in ways that (i) prevent, mitigate or resolve problems adversely affecting human life and/or the wellbeing of the natural world, and/or (ii) enable socially preferable and/or environmentally sustainable developments” (2020, 1773-4).<sup>[1]</sup> As they acknowledge, however, what constitutes a ‘socially good outcome’ will remain deeply contested (Floridi *et al.* 2020, 1774). Indeed, the contested nature of ‘the (social) good’ is an inherent feature of liberal democratic societies, which John Rawls calls “the fact of reasonable pluralism”, i.e., “a pluralism of comprehensive religious, philosophical, and moral doctrines [and, more importantly,] a pluralism of incompatible yet reasonable comprehensive doctrines” (1993, xvi); and, a comprehensive doctrine includes, among other things, “conceptions of what is of value in human life” (Rawls 1993, 13). It is the contested nature of the (social)

good that leads Ben Green (2019) and Jared Moore (2019) to caution against AI4SG.

Green (2019) argues that computer science (or artificial intelligence, data science) lacks a nuanced understanding of the (social) good, and thus the proponents and practitioners of AI4SG are often ignorant of the complexities involve in their choices of what constitutes the social good and their visions of what is socially desirable, thereby neglecting important questions such as who will benefit from the AI4SG project, who will be harmed by it, whether and how power will be shifted after the implementation, who should make these decisions, etc. Moreover, Moore notes that the language of ‘AI for the good’ can be “*strategically vague* [to leave out] the intensely political nature of any one of the areas associated with [AI4SG]” (2019, 5). So construed, the lack of a substantive understanding of the good and/or robust ways to specify the good calls into question the desirability of AI4SG. It should be acknowledged that this challenge from the multiplicity and/or vagueness of the good is not merely a problem in computer science and related fields, but it is *the* background condition of liberal democratic societies. AI4SG, therefore, requires a clearer articulation of the good in use, and to ensure it is not arbitrary or idiosyncratic but shared by the society.

In addition, Moore (2019) rightly observes that calling some AI technologies ‘for the good’ seems to imply the technologies and technological solutions are intrinsically better than the social systems, which presumably are the sources of the problems which AI is supposed to solve. Green (2019) also points out that computer science and related fields often do not have the appropriate methods and toolkits to assess the long-term social impact of technological interventions, which, in turn, results in the proponents and practitioners of AI4SG prioritizing the immediate results of technologies and focusing on optimizing the existing technological solutions. In this way, AI4SG encourages technological solutionism—to reduce social and political problems into technical problems to be solved *exclusively* with technologies and engineering (Morozov 2013). In other words, AI4SG needs to include more vigorous methods and tools to examine the broader and long-term impact of AI technologies and AI-based solutions beyond

the immediate results and the technologies themselves and realize the objective of *for the good* is fundamentally *social* and *political*.

Floridi and colleagues (2020) recognize the above difficulties for AI4SG and propose seven essential factors for AI4SG projects to be considered as genuinely advancing the good, and these factors are: (1) falsifiability and incremental deployment; (2) safeguards against the manipulation of predictors; (3) receiver-contextualized intervention; (4) receiver-contextualized explanation and transparent purposes; (5) privacy protection and data subject consent; (6) situational fairness; and (7) human-friendly semanticization. Some of these factors are intended to ensure AI4SG projects in question to satisfy the *basic moral requirement of doing no harm*, e.g., (2), (5), (6), (7), thereby *being just* (or *morally right*); and, the other factors address more directly the nature of the good in AI4SG projects, i.e., (1), (3), (4).

For instance, the falsifiability (in (1)) of AI4SG projects allows us to assess whether AI technologies and AI-based solutions in use realize the values (and the good) they are designed for and intended. Accordingly, falsifiability prevents vague claims for the good made by AI4SG projects. Relatedly, the receiver-contextualization (in (3) and (4)) aims at aligning the values and the understanding of the good in AI4SG projects to the values and the understanding of the good of those who will be affected by the projects, which is to be achieved through *consultation* in the process of design and implementation and through *explanation* and *rational persuasion* of the purpose of AI technologies and AI-based solutions. Receiver-contextualization, therefore, enables AI4SG projects to be grounded on a shared understanding of the good. More importantly, when receiver-contextualization is done right, the inputs from the users should broaden the scope of assessment for AI4SG to include the social and political dimension of the problems AI technologies and AI-based solutions are introduced to resolve, thereby preventing the AI4SG from focusing only on immediate results and/or technological details.

The seven essential factors identified by Floridi and colleagues helpfully provide a *procedural* account of the good for AI4SG (Mansbridge 1998). However, insofar as a *substantive* account of the good is missing, and AI4SG projects are

implemented at a large-scale affecting people across different groups, receiver-contextualization is unlikely to supply AI4SG projects with a *coherent* or *stable* understanding of the good, as people in liberal democratic societies have diverse interests and different visions of the good. Recently, proponents and practitioners of AI4SG have turned to the United Nations Sustainable Development Goals (SDGs) as the substantive account of the good for AI4SG (see, e.g., Vinuesa *et al.* 2020; Tomašev *et al.* 2020; Ryan *et al.* 2020). The SDGs are a set of 17 goals with 169 targets *all members of the United Nations* agreed upon to work towards for a better future for *all* and *our planet* (United Nations n.d.). Given the widespread acceptance of SDGs by different nations and various sectors of the society, including the governments, NGOs, corporations, the SDGs seem to suitably serve as the substantive account of the good for AI4SG.

Despite offering a *universal* vision of a better future for all, the *ethical base* of the SDGs remains an open question, especially regarding whether the SDGs are *instrumentally* or *intrinsically* normative—that their normative force is derived either from the belief that achieving them leads to a win-win situation for all or from other more fundamental *ethical* reasons (Gasper 2019). The answer to this question, in turn, decides how people can be persuaded into conceiving the SDGs as the substantive account of the good and be motivated to support them, thereby offering a normative ground for AI4SG projects based on the SDGs. So construed, referring to the SDGs alone does *not* resolve the questions about the good in AI4SG. In the remaining of this article, I shall articulate a different normative ground for AI4SG through Confucianism, thereby formulating an alternative account of AI4SG.

### 3. Confucian *Dao*, Harmony, and AI for Social Good

It is well to state in the outset that Confucianism does not distinguish sharply between *the right* and *the good*. Unlike the liberal view that looks to confine the good as *personal* matters, i.e., something to be determined by the individuals themselves, both the right and the good are subjected to a substantive normative ideal in Confucian philosophy. Accordingly, Confucians would reject both the

claim that the (social) good is deeply contested and that there is no substantive account of the good for AI4SG projects.

In my earlier work, I have argued that Confucian *dao* is constitutive of both the right and the good, and it is the ultimate source of normativity in Confucianism (Wong 2012). In particular, Confucianism holds that the *oneness of heaven*<sup>[2]</sup> and *humanity* as an ideal for human beings, which is to be attained through the realization of *dao*. As such, the realization of *dao* is equivalent to human flourishing, i.e., the good. Also, *dao* specifies the right ways of its realization, e.g., *the Analects* states: “Riches and honors are what men desire. If it cannot be obtained in the proper way [*dao*], they should not be held. Poverty and meanness are what men dislike. If it cannot be avoided in the proper way [*dao*], they should not be avoided” (*The Analects* 4.5, in Legge 1861). For Confucians, even if something is good, when pursuing it in a way that goes against *dao*, we ought not to do so. Hence, *dao* would provide a *definitive* normative ground for AI4SG, thereby responding to the concerns over the arbitrariness or idiosyncrasy of the meaning of the good.

Moreover, I have defended *harmony* as the *normative standard* in Confucianism and elaborated on Kam-por Yu’s (2010) account of *harmony*, in which he characterizes the notion as: (1) Harmony is not complete agreement; (2) Harmony is not unprincipled compromise; (3) Harmony is balancing one thing with another one; and, (4) Harmony is the mutual complementation of acceptance and rejection (Yu 2010, 21-25). Based on Yu (2010), I argued that harmony calls for mutual enrichment for *all* parties involved in the situations requiring ethical decisions, and aims at optimization in and of *those* concrete situations, which takes into account *both* what is right (and wrong) and what is good (and bad). I contended that this understanding of harmony is *processual*, which calls for a *continuous* negotiation and adjustment of relationships between human beings, society, and technology (Wong 2012). Recently, Berberich and colleagues (2020) have further explored the role of harmony in the design and implementation of AI4SG and proposed harmony as an additional core principle of AI ethics. In particular, they argue that the idea of harmony requires us to consider “[f]or which tasks should the [AI] systems not be used? When should it remain silent?” (tactful restraint)



“How should the system interact with humans to achieve smooth interactions, avoid causing offense and positively mediate human-human interactions?” (tactful interaction and mediation) “Which information should the system not ask for, record, extract or share?” (tactful privacy)” (Berberich *et al.* 2020, 22). The inclusion of harmony broadens the scope of assessment for AI4SG through the emphasis on various situational factors and the primacy to enrich *all* parties affected by AI4SG projects, thereby addressing the concerns about the myopic vision of the good in computer science and related fields as well as the danger of technological solutionism in AI4SG projects.

So far, I have argued that Confucian *dao*, which encompasses both *the good* and *the right*, would provide the normative ground for AI4SG. I have also argued that *harmony*, as a normative standard, would broaden the scope of assessment for AI4SG projects and require them to optimize for *all* parties by balancing their interests (and needs). *Dao* and *harmony*, therefore, affords a different way to conceive AI4SG. What is still missing in this Confucian reconceptualization of AI4SG, however, is a more specific vision of *good society* for AI4SG projects to aim for.<sup>[3]</sup> There are many ways to characterize the Confucian vision of the good society, but I shall introduce the idea of *datong* (Grand Union) for our purpose. As Albert H. Y. Chen (2015) shows, *datong* is the expression of common good in Confucian philosophy.

*Datong* expresses the ideal society of Confucianism, and a *datong* society is described in *Li Yun* as follows:

When the Grand course was pursued, a public and common spirit ruled all under the sky [*tianxia wei gong*]; they chose men of talents, virtue, and ability; their words were sincere, and what they cultivated was harmony. Thus men did not love their parents only, nor treat as children only their own sons. A competent provision was secured for the aged till their death, employment for the able-bodied, and the means of growing up to the young. They showed kindness and compassion to widows, orphans, childless men, and those who were disabled by disease, so that they were all sufficiently maintained. Males had their proper work, and females had their homes. (They accumulated) articles (of value), disliking that they should be thrown away upon the ground, but not wishing to keep them for their own gratification. (They laboured) with their strength, disliking that it should not be exerted, but not exerting it (only) with a view to their own advantage. In this way (selfish) schemings were repressed and found no development. Robbers, filchers, and rebellious traitors did not show themselves, and hence

the outer doors remained open, and were not shut. This was (the period of) what we call the Grand Union [*datong*].” (*Li Yun*, in Legge 1885)

From the description of the *datong* society in *Li Yun*, several observations can be made. Firstly, the *datong* society is characterized by a world that belongs to and shared by the general public (*gong*), where the ruling authority and ruled subjects work to contribute to the (social) good and not for individual gain. Relatedly, the people of the *datong* society are characterized by their altruistic care for others and their aversion to self-centeredness. As Julia Tao (2000) and Li-Hsiang Lisa Rosenlee (2017) summarize, a *datong* society is one that all are cared for. Finally, the *datong* society induces a condition of moral (self-)transformation, where occasions for selfish desires and other vices do not arise.

We can now rethink AI4SG with the idea of *datong*. The idea of *datong* demands AI4SG to be public-centered, i.e., AI4SG projects ought to be motivated and justified by *the good of the general public* but not the interests of specific groups of individuals. Moreover, AI4SG should be based on *altruistic care for all*, and do not aim for any personal (or company) advantages. Most importantly, AI4SG should not merely attempt to prevent, mitigate or resolve problems adversely affecting human beings and the environment, but to more fundamentally *transform* the *individuals* and *social conditions* such that the problems do not arise. Here, the idea of *datong* helps to refocus on the causes of the problems but not the symptoms of the problems; and, it also helps to counter technological solutionism as the transformation is ultimately about *people* and *society* but not technology *per se*.

It is instructive to note that the Confucian reconceptualization of AI4SG, which I shall call *AI for Datong*, is (will be) compatible with the AI4SG projects aiming to achieve SDGs. However, *AI for Datong* differs from them by being (i) *public-centered*, (ii) *care-centric*, and (iii) *transformative*. In other words, *AI for Datong* takes the general public as the basis of normative assessment (or, the source of the good), is motivated by *care*, and aims to *transform* the situations. Hence, I contend that *AI for Datong* offers a formulation of AI4SG with a more explicit normative ground than merely referring to SDGs.

#### 4. Concluding Remarks

The above discussion demonstrates how Confucian philosophy can contribute to rethink AI4SG through the idea of *datong*. However, the discussion remains at an abstract level without reference to examples. By way of concluding remarks, I briefly look at the case of China's Health Code system through AI for *Datong*.

In the midst of COVID-19 pandemic, the Chinese government have partnered with technology companies, such as Tencent and Alibaba, to implement the Health Code systems in order to manage the spread of the virus. Individuals are requested to install an app on their smartphone, which allows the systems to evaluate whether they are at risk of COVID-19 based on the personal data collected by the app, and then their movement can be restricted based on the evaluation (presented by Green-, Amber-, and Red- QR code). While the systems have been viewed as authoritarian strategies to tighten control over the Chinese public, and there are complaints and misgivings over privacy invasion, technical errors, and discrimination from within and outside China, the Health Code apps are widely used by the people in China (see, e.g., Mozur *et al.* 2020; Davidson 2020; Grover 2020).

One may be tempted to explain the 'success' of the Health Code systems exclusive in terms of state coercion or a lack of concern for privacy in China. In doing so, however, one ignores the fact that the people in China do have control over what they choose to use and that they do care about privacy, as shown by the backlash caused by various attempts to extend the application of the Health Code systems for other purposes. Another explanation of the widespread use of the apps, therefore, is that they are considered by the Chinese public as *a technology for the good*; and, indeed, they have been promoted as technology for the good in the People's Daily (see, e.g., Zhu 2020) and the technology companies. However, do the Health Code systems genuinely qualify as AI4SG, in particular, from the perspective of AI for *Datong*?

Ideally, the Health Code systems are to enable *the general public* to have a safe and healthy environment to return to their normal lives. In other words, the systems aim at the good of the general public, and they are not based on the interests of particular groups of individuals. If they function *exclusively* to ensure a safe and healthy environment for the public, then the systems do satisfy the

*public-centeredness* requirement in AI for *Datong*. Unfortunately, the systems are open to mission creep, where the application of the systems can easily go *beyond* the provision and maintenance of a safe and healthy environment. As such, the systems can satisfy the public-centeredness requirement only if there are sufficient measures to avoid mission creep. Relatedly, the systems *should not* only aim to detect and restrict individuals, who are considered to be at risk of infection, but also to provide suitable assistance to them. In this respect, the systems are accompanied by other tools to assist the people (see, Grover 2020). Nonetheless, we may still question whether the systems are designed and implemented out of the duty of care (for all), or they are driven by other political and/or commercial considerations, and so whether they satisfy the requirement of *being care-centric* in AI for *Datong*. In order to demonstrate the systems are operated on the basis of care, the *reasons* for and *functioning* of the systems ought to be transparent. Finally, AI for *Datong* requires the systems to be *transformative*; and, in the context of pandemic, I take the (self-) transformative dimension to be about *responsibility* and *trust*. Here, it is an open question whether the systems will make individuals *more responsible* and *trustworthy*, as the systems *remind* them the risks about the virus, or whether the systems will make them *less responsible* and *less trusting*, as the systems impose boundaries between people from the outside.

In short, from the perspective of AI for *Datong* the Health Code systems could still be used for the good if the design and implementation of the systems work towards answering the issues raised above, and thus being *public-centered*, *care-centric*, and *transformative*.<sup>[4]</sup>

## Bibliography

Ames, R., Rosemont, H. (1999). *The Analects of Confucius: a Philosophical Translation*. New York: Ballantine Books (trans.).

Beijing Academy of Artificial Intelligence [BAAI] (2019, May 28). *Beijing AI Principles*. Available at: <https://www.baai.ac.cn/news/beijing-ai-principles-en.html>

- Berendt, B. (2019). AI for the Common Good?! Pitfalls, Challenges, and Ethics Pen-Testing. *Paladyn, Journal of Behavioral Robotics*, 10, 1, 44-65.  
<https://doi.org/10.1515/pjbr-2019-0004>
- Berberich, N., Nishida, T., Suzuki, S. (2020). Harmonizing Artificial Intelligence for Social Good. *Philosophy & Technology*. Available at:  
<https://doi.org/10.1007/s13347-020-00421-8>
- Chen, A. H. Y. (2014). The Concept of “ Datong ” in Chinese Philosophy as an Expression of the Idea of the Common Good. In: Solomon, D., Lo, P. C. (Eds.). *The Common Good: Chinese and American Perspective*. Dordrecht: Springer. p. 85-102.
- Cheng, M. (2019, November 20). Tech for Good: The Second Half of a Large Social Experiment? *Internet Frontiers*. Available at:  
[https://www.tisi.org/?page\\_id=7043&post\\_id=839](https://www.tisi.org/?page_id=7043&post_id=839)
- European Commission High-Level Expert Group on Artificial Intelligence [AI HLEG] (2019). *Ethics Guidelines for Trustworthy AI*. Available at:  
<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- Floridi, L., Cows, J., King, T. C., Taddeo, M. (2020). How to Design AI for Social Good: Seven Essential Factors. *Science and Engineering Ethics*, 26, 1771-1796.
- Davidson, H. (2020, April 1). China's Coronavirus Health Code Apps Raise Concerns over Privacy. *The Guardian*. Available at:  
<https://www.theguardian.com/world/2020/apr/01/chinas-coronavirus-health-code-apps-raise-concerns-over-privacy>
- Gasper, D. (2019) The Road to the Sustainable Development Goals: Building Global alliances and norms. *Journal of Global Ethics*, 15, 2, 118-137.
- Green, B. (2019). “Good” isn’t Good Enough. Paper presented at *AI for Social Good Workshop, NeurIPS 2019*. Vancouver, Canada. Available at:  
<https://www.benzevgreen.com/wp-content/uploads/2019/11/19-ai4sg.pdf>
- Grover, D. (2020, April 5). *How Chinese Apps Handled Covid-19*. Available at:  
<http://dangrover.com/blog/2020/04/05/covid-in-ui.html>
- Laskai, L., Webster, G. (2019, June 17). Translation: Chinese Expert Group Offers ‘Governance Principles’ for ‘Responsible AI’. *New America*. Available at: <https://www.newamerica.org/cybersecurity->

[initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/](#)

Latonero, M. (2019, November 18). AI for Good is Often Bad. *Wired*. Available at: <https://www.wired.com/story/opinion-ai-for-good-is-often-bad/>

Legge, J. (1861). *The Analects* (trans.). Available at: <https://ctext.org/analects>

Legge, J. (1885). *The Li Ki* (trans.). Available at: <https://ctext.org/liji>

Lucero, K. (2019). Artificial Intelligence Regulation and China's Future. *Columbia Journal of Asian Law*, 33, 1, 94-171.

Mansbridge, J. (1998). On the Contested Nature of the Public Good. In W. W. Powell and E. S. Clemens (Eds.). *Private Action and the Public Good*. New Haven: Yale University Press. p. 3-19.

Moore, J. (2019). AI for Not Bad. *Frontiers in Big Data*, 2, 32. Available at: <https://doi.org/10.3389/fdata.2019.00032>

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: PublicAffairs.

Mozur, P., Zhong, R., Krolik, A. (2020, March 1). In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags, *New York Times*. Available at: <https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html>

Nesta (2020). *The AI Powered State: China's approach to public sector innovation*. Available at: [https://www.nesta.org.uk/documents/1860/Nesta\\_TheAIPoweredState\\_2020.pdf](https://www.nesta.org.uk/documents/1860/Nesta_TheAIPoweredState_2020.pdf)

Rawls, J. (1993). *Political Liberalism*. New York: Columbia University Press.

Rosenlee, L. L. (2017). Ritual, Dependency Care and Confucian Political Authority. *International Communication of Chinese Culture*, 4, 493-513.

Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., Stahl, B. (2020). The Ethical Balance of Using Smart Information Systems for Promoting the United Nations' Sustainable Development Goals. *Sustainability*, 12, 12, 4826. Available at: <https://doi.org/10.3390/su12124826>

Tao, J. P.-W. L. (2000). Two Perspectives of Care: Confucian Ren and Feminist Care. *Journal of Chinese Philosophy*, 27, 2, 215-240.

- Tencent (2019, November 11). *Value for Users, Tech for Good*. Available at: <https://www.tencent.com/en-us/articles/2200913.html>
- Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, et. al (2020). AI for Social Good: Unlocking the Opportunity for Positive Impact. *Nature Communications*, 11, 1, 2468. Available at: <https://doi.org/10.1038/s41467-020-15871-z>
- United Nations (n.d.). *Sustainable development Goals—United Nations*. Available at <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>
- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, et. al. (2020). The Role of Artificial Intelligence in Achieving the Sustainable Development Goals. *Nature Communications*, 11, 1, 233. Available at: <https://doi.org/10.1038/s41467-019-14108-y>
- Webster, G. (2019, June 17). Translation: Chinese AI Alliance Drafts Self-Discipline 'Joint Pledge'. *New America*. Available at: <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-ai-alliance-drafts-self-discipline-joint-pledge/>
- Wong, P.-H. (2012). Dao, Harmony and Personhood: Towards a Confucian Ethics of Technology. *Philosophy & Technology*, 25, 67-86.
- Wong, P.-H. (2019). Rituals and Machines: A Confucian Response to Technology-Driven Moral Deskillling. *Philosophies*, 4, 4, 59. Available at: <https://doi.org/10.3390/philosophies4040059>
- Yu, K. P. (2010). The Confucian Conception of Harmony. In: Tao, J., Cheung, A., Painter, M., Li, C. (Eds.), *Governance for Harmony in Asia and Beyond*. London: Routledge. p. 15-36.
- Zhu, Y. (2020, February 17). Making Good Use of Science and Technology to Solve the “Pandemic” Emergency. *People’s Daily*. Available at: <http://opinion.people.com.cn/n1/2020/0217/c1003-31591258.html>

---

## Footnotes

[1] There are different understandings of the term ‘AI’, which foreground different concerns. In this article, I shall refer to the understanding of AI by the European Commission High-Level Expert Group on Artificial Intelligence, which defines AI as “software (and possibly also hardware) systems designed by humans

that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected [...] data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions” (AI HLEG 2019).

[2] It should be noted that the translation of Tian as ‘Heaven’ is not uncontroversial. Notably, Ames and Rosemont have argued that ‘Heaven’ has an unnecessary and undesirable connotation to the transcendental realm in Judeo-Christian tradition, which is not apparent in the Confucian tradition; and, the translation of Tian as ‘Heaven’ also downplays and the moral (-political) connotation in the concept of Tian. Accordingly, they think that it will be misleading to translate Tian as ‘Heaven’ (Ames and Rosemont 1999, 46ff). So, I use this translation only because of lacking a better terminology to convey the multifarious meanings of Tian.

[3] Alternatively, we can also look at the Confucian vision of the good life and examine how to design and implement AI to achieve this vision, see Wong (2019).

[4] The author wrote this article in his personal capacity. The views expressed in this article are his own and do not represent the views of his employer.

---

#### **Author contribution**

The entirety of this manuscript was prepared by Pak-Hang Wong.

---

#### **Editor’s notes**

The editor responsible for the publication of this article was Rafael Capurro.

Style editing and linguistic revision to the wording in this text has been performed by Prof. Adj. Hugo E. Valanzano (State University, Uruguay).

Nilzete Ferreira Gomes (Universidade Federal Rural da Amazõnia (UFRA), Pará, Brazil), was in charge of translating from Portuguese to Spanish.