



The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation

Huw Roberts¹ · Josh Cowlis^{1,2} · Jessica Morley¹ · Mariarosaria Taddeo^{1,2} · Vincent Wang³ · Luciano Floridi^{1,2,3}

Received: 21 September 2019 / Accepted: 17 May 2020 / Published online: 17 June 2020
© The Author(s) 2020

Abstract

In July 2017, China’s State Council released the country’s strategy for developing artificial intelligence (AI), entitled ‘New Generation Artificial Intelligence Development Plan’ (新一代人工智能发展规划). This strategy outlined China’s aims to become the world leader in AI by 2030, to monetise AI into a trillion-yuan (ca. 150 billion dollars) industry, and to emerge as the driving force in defining ethical norms and standards for AI. Several reports have analysed specific aspects of China’s AI policies or have assessed the country’s technical capabilities. Instead, in this article, we focus on the socio-political background and policy debates that are shaping China’s AI strategy. In particular, we analyse the main strategic areas in which China is investing in AI and the concurrent ethical debates that are delimiting its use. By focusing on the policy backdrop, we seek to provide a more comprehensive and critical understanding of China’s AI policy by bringing together debates and analyses of a wide array of policy documents.

Keywords Artificial intelligence · China · Cyber warfare · Digital ethics · Economic growth · Governance · Innovation · International competition · New Generation Artificial Intelligence Development Plan · Policy · Privacy · Social governance

1 Introduction

In March 2016, a Google DeepMind artificial intelligence (AI) designed for playing the board game Go (AlphaGo) defeated Lee Sedol, a South Korean professional Go player. At the time, Sedol had the second-highest number of Go international championship victories, yet lost against AlphaGo by four games to one (Borowiec 2016). While the match received some coverage in the West, it was a major event in China, where over 280 million people watched it live. Two government insiders described this match as a ‘Sputnik moment’ for the development of AI within China (Lee 2018, p. 3). Although there had been AI policy initiatives in the country previously, the victory for AlphaGo contributed to an increase in focus, as indicated by the 2017

‘New Generation Artificial Intelligence Development Plan’ (AIDP). The AIDP set out strategic aims and delineated the overarching goal of making China the world leader in AI by 2030.¹

A limited number of reports have attempted to assess the plausibility of China’s AI strategy given China’s current technical capabilities (Ding 2018; ‘China AI Development Report’ 2018). Others have sought to understand specific areas of development, for instance, security or economic growth (Barton et al. 2017; ‘Net Impact of AI on jobs in China’ 2018; Allen 2019). However, to grasp the ramified implications and direction of the AIDP, it is insufficient to analyse specific elements in isolation or to consider only technical capabilities. Instead, a more comprehensive and critical analysis of the driving forces behind China’s AI strategy, its political economy, cultural specificities, and the current relevant policy debates, is required to understand China’s AI strategy. This is the task we undertake in this article.

✉ Luciano Floridi
luciano.floridi@oii.ox.ac.uk

¹ Oxford Internet Institute, University of Oxford, 1 St Giles’, Oxford OX1 3JS, UK

² Alan Turing Institute, British Library, 96 Euston Rd, London NW1 2DB, UK

³ Department of Computer Science, University of Oxford, 15 Parks Rd, Oxford OX1 3QD, UK

¹ In the rest of this article, we shall use ‘China’ or ‘Chinese’ to refer to the political, regulatory, and governance approach decided by the Chinese national government concerning the development and use of AI capabilities.

To provide this contextualised understanding, Sect. 2 maps relevant AI legislation in China. We argue that, although previous policy initiatives have stated an intent to develop AI, these efforts have been fractious and viewed AI as one of many tools in achieving a different set goal. In contrast, the AIDP is the first national-level legislative effort that focuses explicitly on the development of AI as a unified strategy. Following this, Sect. 3 analyses the interventions and impact of the AIDP on three strategic areas identified in the document, namely: international competition, economic growth, and social governance. Section 4 focuses on China's aim to develop ethical norms and standards for AI. There we argue that, although the debate is in its early stages, the desire to define normative boundaries for acceptable uses of AI is present and pressing. Altogether, this article seeks to provide a detailed and critical understanding of the reasons behind, and the current trajectory of, China's AI strategy. It emphasises that the Chinese government is aware of the potential benefits, practical risks, and the ethical challenges that AI presents, and that the direction of China's AI strategy will largely be determined by the interplay of these factors and by the extent to which government's interests may outweigh ethical concerns. Section 5 concludes the paper by summarising the key findings of our analysis.

2 AI legislation in China

Since 2013, China has published several national-level policy documents, which reflect the intention to develop and deploy AI in a variety of sectors. For example, in 2015, the State Council released guidelines on China's 'Internet+' action. It sought to integrate the internet into all elements of the economy and society. The document clearly stated the importance of cultivating emerging AI industries and investing in research and development. In the same year, the 10-year plan 'Made in China 2025' was released, with the aim to transform China into the dominant player in global high-tech manufacturing, including AI (McBride and Chatzky 2019). Another notable example is the Central Committee of the Communist Party of China's (CCP) 13th 5-year plan,² published in March 2016. The document mentioned AI as one of the six critical areas for developing the country's emerging industries (CCP 2016), and as an important factor in stimulating economic growth. When read together, these documents indicate that there has been a conscious effort to develop and use AI in China for some time, even before 'the Sputnik moment'. However, prior to 2016, AI was presented merely as one technology among

many others, which could be useful in achieving a range of policy goals. This changed with the release of the AIDP.

2.1 The New generation artificial intelligence development plan (AIDP)

Released in July 2017 by the State Council (which is the chief administrative body within China), the 'New Generation Artificial Intelligence Development Plan' (AIDP) acts as a unified document that outlines China's AI policy objectives. Chinese media have referred to it as 'year one of China's AI development strategy' ('China AI Development Report' 2018, p. 63). The overarching aim of the policy, as articulated by the AIDP, is to make China the world centre of AI innovation by 2030, and make AI 'the main driving force for China's industrial upgrading and economic transformation' (AIDP 2017). The AIDP also indicates the importance of using AI in a broader range of sectors, including defence and social welfare, and focuses on the need to develop standards and ethical norms for the use of AI. Altogether, the Plan provides a comprehensive AI strategy and challenges other leading powers in many key areas.

The AIDP delineates three key steps, each of which contains a series of goals, some of which are tightly defined, while others are vaguer. They are summarised as follows and in Fig. 1 below:

1. By 2020, China aims to maintain competitiveness with other major powers and optimise its AI development environment. In monetary terms, China intends to create an AI industry worth more than 150 billion yuan (ca. 21 billion dollars). Lastly, it seeks to establish initial ethical norms, policies, and regulations for vital areas of AI.
2. By 2025, China aims to have achieved a 'major breakthrough' (as stated in the document) in basic AI theory and to be world-leading in some applications ('some technologies and applications achieve a world-leading level'). China also targets an increase in the worth of its core AI industry to over 400 billion yuan (ca. 58 billion dollars), and plans to expand upon, and codify in law, ethical standards for AI.
3. By 2030, China seeks to become the world's innovation centre for AI. By then, growth in the core AI industry is expected to more than double again and be valued at 1 trillion yuan (ca 147 billion dollars), and further upgrades in the laws and standards are also to be expected, to deal with newly emerging challenges.

2.2 Implementing the AIDP

The Plan will be guided by a new AI Strategy Advisory Committee, established in November 2017, and will be coordinated by the Ministry of Science and Technology (MIST),

² The 5-year plans are a central pillar in China's economic growth policy (Heilmann and Melton 2013; Hu 2013).

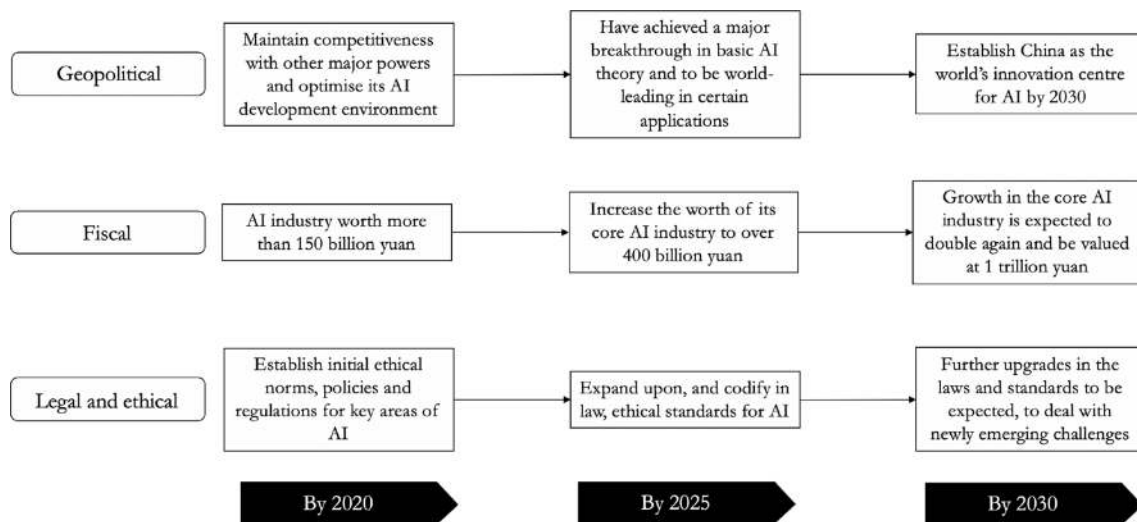


Fig. 1 Visualising China's AIDP

alongside the AI Plan Promotion Office, and other relevant bodies (“AI Policy-China” n.d.).³ Although these bodies will provide central guidance, the Plan is not meant to act as a centrally enacted initiative. The AIDP instead functions as a stamp of approval for de-risking and actively incentivising local projects that make use of AI. Recognising this point is important: the AIDP is an ambitious strategy set by the central government, but the actual innovation and transformation is expected to be driven by the private sector and local governments. In other words, it is more appropriate to view the AIDP as a highly incentivised ‘wish list’, to nudge, and coordinate other relevant stakeholders, rather than a central directive (Sheehan 2018). This is why the 3-year plan promoting the AIDP (2018–2020) emphasises coordination between provinces and local governments.

With regard to the private sector, China has selected ‘AI national champions’: businesses endorsed by the government to focus on developing specific sectors of AI. For example, Baidu has been tasked with the development of autonomous driving, Alibaba with the development of smart cities, and Tencent with computer vision for medical diagnoses (Jing and Dai 2017). Being endorsed as a national champion involves a deal whereby private companies agree to focus on the government’s strategic aims. In return, these companies receive preferential contract bidding, easier access to finance, and sometimes market share protection. Although other companies can compete in these fields, historically the

status of ‘national champion’ has helped larger companies dominate their respective sectors (Graceffo 2017).

With this said, the new AI ‘national team’ differs from previous state-sponsored national champions in that they are already internationally successful in their respective fields, independently of this preferential treatment. Furthermore, there is extensive domestic competition in the areas where national champions have been selected. This suggests that competition may not be stymied in the traditional manner. For instance, all the companies selected as AI national champions are developing technologies in Alibaba’s designated area of smart cities (Ding 2019). In parallel with this, patronage does not prohibit smaller companies benefiting from the financial incentive structure. Technology start-ups within China often receive government support and subsidies for developing AI technologies. As an example, Zhongguancun Innovation Town is a purpose-built, government subsidised, incubator workspace that provides a suite of services to help Chinese technology start-ups succeed, often in the sectors where national champions have been selected. Finally, there are also cases where there is no specific endorsement. For example, while the AIDP promotes smart courts, with a stated desire to develop AI for evidence collection, case analysis, and legal document reading, as of April 2020 there is no national champion selected for developing AI applications for the administration of justice.

Concerning local governments, the political structure within China creates a system of incentives for fulfilling national government policy aims. Short term limits for provincial politicians and promotions based on economic performance provide strong incentives for following centrally-defined government initiatives (Li and Zhou 2005; Persson and Zhuravskaya 2016). Thus, local governments become hotbeds for testing and developing central government

³ It should be noted that, although MIST has been tasked with coordinating the AIDP, it was the Ministry of Industry and Information Technology (MIIT) that released the guidance for the implementing the first step of the AIDP.

policy. The strength of this incentive system can be seen in the decision made by the administration of the city of Tianjin to establish a \$5 billion fund for the development of AI, around the same time as the publication of the AIDP (Mozur 2017). At the same time, it is important to recognise how the absence of an effective accountability review of local government spending creates problems within this system. Notably, it has facilitated a mindset in which local politicians know that the central government will bail them out for failed projects, leading to poor budget management (Ji 2014). A clear example of this is the large-scale port building initiatives developed by provincial governments in East coast provinces that were based more on prestige than any economic rationale, and which led to overcapacity and disorderly competition (Zhu 2019).

These incentive structures contain a subtle distinction. A national team has been selected to lead the research and development in a handful of designated strategic areas. Beyond these selected companies, there are few specific guidelines provided to industry and local state agents as to which items to pursue on the AIDP's 'wish list'. This enables companies to cherry-pick the technologies they want to develop and provides local governments with a choice of private sector partners for integrating AI into city infrastructure or governance (Sheehan 2018). Subsequent documentation has emphasised the importance of strengthening organisation and implementation,⁴ including between provinces and ministries, yet it is unclear how this coordination would function in practice. Thus, the AIDP may work as a 'wish list', but the exact guidance, incentivisation and risk differ depending on the type of stakeholder.

The AIDP should not be read in isolation when considering China's AI strategy (Ding 2018), but it does provide the most transparent and influential indication of the driving forces behind China's AI strategy. Because of the AIDP's significance (in terms of policy) and importance (in terms of strategy), in the rest of this article, we shall use it as the organisational skeleton for explaining the drivers and ethical boundaries shaping China's approach to AI.

3 China's AI strategic focus

The AIDP provides a longitudinal perspective on China's strategic situation regarding AI, including its comparative capabilities, the opportunities offered, and the potential

risks. Following a technology-first approach, it may be tempting to concentrate one's attention on the stated capabilities of AI, to gain an insight into the types of technologies in which China is investing. However, this would likely offer only a short-term perspective and would soon be out of date as technological innovation advances rapidly. Furthermore, it would do little to explain why China is seeking to develop a strong AI sector in the decades to come. To this end, it is more useful to try to understand China's strategic focus from a policy-first approach, by analysing the areas where China considers that AI presents opportunities. In this section, we focus on these areas of particular importance to China, on how and what China expects to gain from developing AI in each of them, and on some of the perceived risks present in each of these areas. The AIDP highlights three areas where AI can make a substantial difference within China: international competition, economic development, and social governance. They are strictly interrelated but, for the sake of clarity, we shall analyse them separately, and contextualise each of them by discussing the relevant literature surrounding the broader political backdrop and contemporary policy debates.

3.1 International competition

The AIDP states that AI has become a new focus of international competition and that 'the development of AI [is] [...] a major strategy to enhance national competitiveness and protect national security' (AIDP 2017). It emphasises that China should take the strategic opportunity afforded by AI to make 'leapfrog developments'⁵ in military capabilities. Although China and the US are regularly portrayed as geopolitical rivals (Mearsheimer 2010; Zhao 2015), the military budgets of the two powers remain significantly different. China has the world's second-largest military budget, with \$175 billion allocated in 2019 (Chan and Zhen 2019), but its spending is still only a third of the US budget (Martina and Blanchard 2019). Rather than outspending the US in conventional weaponry, China considers investing in AI as an opportunity to make radical breakthroughs in military technologies and thus overtake the US.

Attempts to use technologies to challenge US hegemony are nothing new within China's military strategy. Since the late 1990s, the country has been following a policy of 'shashoujian' (杀手锏), which roughly translates as 'trump-card' (Bruzdzinski 2004). Rather than directly competing with the US, China has sought to develop asymmetric capabilities, which could provide a critical advantage in warfare

⁴ To accompany the three steps outlined earlier, the Ministry of Industry and Information Technology (MIIT) provides documents to flesh out these aims. The first of these, 'Three-Year Action Plan for Promoting Development of a New Generation Artificial Intelligence Industry (2018–2020)', has already been released.

⁵ 'This term refers to 'an actor, which lags behind its competitors in terms of development, coming up with a radical innovation that will allow it to overtake its rivals' (Brezis et al. 1993).'

and credible deterrence in peacetime (Blasko 2011). This trump-card strategy seeks to use unorthodox technologies against enemies' weaknesses to gain the initiative in war (Peng and Yao 2005). The trump-card approach was echoed by the former Party Chairman, Jiang Zemin, who emphasised that technology should be the foremost focus of the military, especially the technology that the 'enemy fears [the] most' (Cheung et al. 2016).

One area in which China has been developing these asymmetric tactics is cyber warfare, where capabilities have been developed for targeting the US military's battle-critical networks if needed (Kania 2017a). Alongside this, evidence points to the persistent use of cyberattacks to collect scientific, technological and commercial intelligence (Inkster 2010). The Chinese position on these capabilities is ambivalent. On the one hand, China has officially promoted international initiatives for regulating hostile state-run activities in cyberspace, and to fill the existing regulatory gap for state behaviour in this domain (Ku 2017; Austin 2016; Taddeo 2012; Taddeo 2016). For example, China co-sponsored the International Code of Conduct for Information Security at the UN General Assembly in September 2011, which sought a commitment against using information technologies in acts of aggression and has provided continued support for dialogue by the UN Group of Government Experts in preventing cyber conflicts (Meyer 2020). On the other hand, China has also run cyber operations targeting US infrastructure and aiming at extracting commercial and scientific information as well as acquiring relevant intelligence against several countries, including Australia, Philippines, Hong Kong, and the US.⁶

The desire to leapfrog the US is echoed in statements from China's political and military leadership. For instance, President Xi Jinping stated in 2017 that 'under a situation of increasingly fierce international military competition, only the innovators win' (Kania 2020, p. 2). This sentiment is shared by Lieutenant General Liu Guozhi, deputy of the 19th National Congress and director of the Science and Technology Committee of the Central Military Commission, who stated in an interview that AI presented a rare opportunity for taking shortcuts to achieve innovation and surpass rivals ("AI military reform" 2017). In parallel, academics affiliated with the People's Liberation Army (PLA) highlight that AI will be used to predict battlefield situations and identify optimal approaches, facilitating 'winning before the war' (Li 2019). Some members of the PLA go further than this in anticipating a battlefield 'singularity', where AI outpaces human decision-making (Kania 2017a). These statements emphasise the belief, which is widespread throughout

China's military and defence circles, in the importance of utilising emergent technologies including AI to achieve a competitive military advantage.

As China has developed economically and militarily, the focus of the country's military strategy has also matured. Over the past few years, China's strategy has coalesced around efforts to develop 'new concept weapons' to surpass the US's military capabilities. These are not limited to AI alone, and are applicable to China's investments in other fields of emerging military technologies, like hypersonic weaponry (Kania 2017b). Therefore, China's efforts to use technology to gain an advantage in military affairs should not be seen as something new, but instead understood within a broader historical context of finding innovative ways to challenge the hegemony of the US.

Although the push for leapfrog developments marks a continuation of previous policy, there are strong concurrent indications that Chinese officials are also concerned about AI causing an arms race and potential military escalation. Statements of senior officials seem to suggest a belief in cooperation and arms control to mitigate the risks that AI's military development poses. In particular, three major risks are central to the debate:

- (i) human involvement and control once AI-based weapons are deployed;
- (ii) the absence of well-defined norms for state behaviour and use of AI weapons, which in turn increases;
- (iii) the likelihood of misperceptions or unintentional conflict escalation (Taddeo and Floridi 2018; Allen 2019).

These concerns underpin China's support to restrict the use of autonomous weapons, as expressed at the 5th Convention on Certain Conventional Weapons ("Chinese Position Paper" 2016) and, more recently, the desire to ban autonomous lethal weapons (Kania 2018a). Despite concerns (i)–(iii), it is crucial to stress that China is the actor pursuing the most aggressive strategy for developing AI for military uses among the major military powers (Pecotic 2019).

Digging more deeply into China's actions on the international stage is revealing. The ban that China advocated encompassed only usage and not the development or production of autonomous lethal weapon systems. Thus, it would not prevent the existence of autonomous lethal weapons serving as a deterrent, in much the same way that China has a putative 'no first use' (NFU) doctrine for nuclear weapons. Furthermore, the definition of autonomy embraced by China is extremely narrow, including only fully autonomous weapons ("UN Seeks Human Control Over Force" 2018). Some commentators argue that this juxtaposition of cautious concerns about deployment, on the one hand, and an aggressive approach to development, on the other, can be explained

⁶ <https://www.csis.org/programs/technology-policy-program/significant-cyber-incidents>.

by the Chinese efforts to exert pressure on other militaries whose democratic societies are more sensitive to the controversies of using automated weapons (Kania 2018a). This is a reasonable claim: a continuation of propaganda may be part of the explanation. For instance, China was the first nuclear power to pledge ‘no first use’ of nuclear weapons (so far only India has a similar pledge; other countries, including the US and the UK, have pledged to use nuclear weapons only defensively). But rather than offering a genuine commitment to NFU, this pledge was meant as internal and external propaganda tool, which would be circumvented by semantics if needed (Schneider 2009).

Taken together, China’s focus on military AI can be considered as a continuation of a longer-term strategy, which privileges developing (with the threat of deploying) technology to gain a military advantage. There remains a conscious recognition, by several actors in China, that developing AI presents an especially fraught risk of igniting an arms race or causing unintentional escalation due to the autonomy of these technologies (Taddeo and Floridi 2018; Allen 2019). But at the political level, efforts to curtail the use of military AI internationally may also be seen as part of a propaganda strategy.

3.2 Economic development

Economic development is the second strategic opportunity explicitly mentioned in the AIDP. It is stated that AI will be the driving force behind a new round of industrial transformation, which will ‘inject new kinetic energy into China’s economic growth’ (AIDP 2017). The reconstruction of economic activity is targeted in all sectors, with manufacturing, agriculture, logistics, and finance being the examples promoted in the AIDP.

China’s rapid growth has frequently been referred to as an ‘economic miracle’, due to the country’s shift from having a slow-growth economy to enjoying some of the world’s highest growth rates for almost three decades (Ray 2002; Naughton and Tsai 2015). A number of factors facilitated this economic growth, of which the demographic dividend is one. A large workforce, in combination with a small dependent population, fostered high levels of savings and heavy investment (Cai and Lu 2013). Structural changes, including a conscious shift from a predominantly agricultural to a manufacturing economy, and the opening up of markets, are additional, critical factors. By 2012, China’s labour force growth dropped to around zero, and its shift from an agricultural to a manufacturing economy had largely matured. These trends have led Chinese policymakers to the realisation that an alternative development model is necessary for maintaining high rates of growth. This model rests on the shift from heavy investment in the industry to growth

stimulated by an innovative society (Naughton and Tsai 2015). Recently, science and technology have been put forward as a crucial means for achieving this type of innovative growth (Zhang 2018).

Some commentators have argued that maintaining these high levels of growth is particularly important for China due to the implicit trading by citizens of political freedoms for economic growth and embourgeoisement (Balding 2019). Research has highlighted that support for the party and a relatively lackluster desire for democracy stems from satisfaction with employment and material aspects of life, particularly within the middle classes (Chen 2013). Slowing economic growth would likely sow dissatisfaction within the populace and make inherent features within the Chinese political system, such as corruption, less tolerable (Diamond 2003; Pei 2015). A lack of a democratic outlet for this frustration could lower the overall support that the government currently receives. Some maintain that this creates a ‘democratise or die’ dynamic (Huang 2013), however, this may be unfeasible, given China’s political control (Dickson 2003; Chin 2018).

Against this backdrop, a report by PwC suggested that China is the country that has the most to gain from AI, with a boost in GDP of up to 26% by 2030 (“Sizing the Prize” 2017). Estimates also suggest that AI could facilitate an increase in employment by 12% over the next two decades (“Net Impact of AI on Jobs in China?” 2018). Because of these potential benefits, President Xi has frequently spoken of the centrality of AI to the country’s overall economic development (Hickert and Ding 2018; Kania 2018b). China has been pursuing the potential economic benefits of AI concretely and proactively for some time. For example, there has been a 500% increase in the annual installation of robotic upgrades since 2012. This rate is staggering, especially when compared to a rate of just over 100% in Europe (Shoham et al. 2018), equating to over double the number of robot installations in China than Europe.

AI can be a double-edged sword, because the benefits and improvements brought about by AI come with the risk, amongst others, of labour market disruptions. This is a concern explicitly stated in the AIDP. Although the aforementioned PwC report predicts that automation will increase the net number of jobs in China, disruption will likely be unevenly spread (“Net Impact of AI on Jobs in China?” 2018). Smarter automation will most immediately affect low- and medium-skilled jobs while creating opportunities for higher-skilled technical roles (Barton et al. 2017). China has been active in its efforts to adapt to such AI-related risks, especially with an education overhaul promoted by the ‘National Medium- and Long-term Education Reform and Development Plan (2010–2020)’. This plan has the goal of supporting the skilled labour required in the information age (“Is China Ready for Intelligent Automation” 2018). In

the same vein, China is addressing the shortage in AI skills specifically by offering higher education courses on the subject (Fang 2019). Accordingly, China seems to be preparing better than other middle-income countries to deal with the longer-term challenges of automation (“Who is Ready for Automation?” 2018).

Although these efforts will help to develop the skill set required in the medium and long term, they do little to ease the short-term structural changes. Estimates show that, by 2030, automation in manufacturing might have displaced a fifth of all jobs in the sector, equating to 100 million workers (“Is China Ready for Intelligent Automation” 2018). These changes are already underway, with robots having replaced up to 40% of workers in several companies in China’s export-manufacturing provinces of Zhejiang, Jiangsu and Guangdong (Yang and Liu 2018). In the southern city of Dongguan alone, reports suggest that 200,000 workers have been replaced with robots (“Is China Ready for Intelligent Automation” 2018). When this is combined with China’s low international ranking in workforce transition programmes for vocational training (“Is China Ready for Intelligent Automation” 2018), it can be suggested that the short-term consequence of an AI-led transformation is likely to be significant disruptions to the workforce, potentially exacerbating China’s growing inequality (Barton et al. 2017).

3.3 Social governance

Social governance, or more literally in Chinese ‘social construction’,⁷ is the third area in which AI is promoted as a strategic opportunity for China. Alongside an economic slowdown, China is facing emerging social challenges, hindering its pursuit of becoming a ‘moderately prosperous society’ (AIDP 2017). An aging population and constraints on the environment and other resources are explicit examples provided in the AIDP of the societal problems that China is facing. Thus, the AIDP outlines the goal of using AI within a variety of public services to make the governance of social services more precise and, in doing so, mitigate these challenges and improve people’s lives.

China has experienced some of the most rapid structural changes of any country in the past 40 years. It has been shifting from a planned to a market economy and from a rural to an urban society (Naughton 2007). These changes have helped facilitate economic development, but also introduced a number of social issues. One of the most pressing social challenges China is facing is the absence of a well-established welfare system (Wong 2005). Under the planned

economy, workers were guaranteed cradle-to-grave benefits, including employment security and welfare benefits, which were provided through local state enterprises or rural collectives (Selden and You 1997). China’s move towards a socialist market economy since the 1990s has accelerated a shift of these provisions from enterprises and local collectives to state and societal agencies (Ringen and Ngok 2017). In practice, China has struggled to develop mature pension and health insurance programmes, creating gaps in the social safety net (Naughton 2007). Although several initiatives have been introduced to alleviate these issues (Li et al. 2013), the country has found it difficult to implement them (Ringen and Ngok 2017).

The serious environmental degradation that has taken place in the course of China’s rapid development is another element of concern. For most of China’s development period, the focus has been on economic growth, with little or no incentive provided for environmental protection (Rozelle et al. 1997). As a result, significant, negative externalities and several human-induced natural disasters have occurred that have proven detrimental for society. One of the most notable is very poor air quality, which has been linked to an increased chance of illness and is now the fourth leading cause of death in China (Delang 2016). In parallel, 40% of China’s rivers are polluted by industry, causing 60,000 premature deaths per year (Economy 2013). Environmental degradation of this magnitude damages the health of the population, lowers the quality of life, and places further strain on existing welfare infrastructure.

The centrality of these concerns could be seen at the 19th National Party Congress in 2017, where President Xi declared that the ‘principal contradiction’ in China had changed. Although the previous ‘contradiction’ focused on ‘the ever-growing material and cultural needs of the people and backward social production,’ Xi stated ‘what we now face is the contradiction between unbalanced and inadequate development and the people’s ever-growing needs for a better life’ (Meng 2017). After years of focusing on untempered economic growth, President Xi’s remarks emphasise a broader shift in China’s approach to dealing with the consequences of economic liberalisation.

These statements are mirrored in several government plans, including the State Council Initiative, ‘Healthy China 2030’, which seeks to overhaul the healthcare system. Similar trends can be seen in China’s efforts to clean up its environment, with a new 3-year plan building on previous relevant initiatives (Leng 2018). China has recently focused on AI as a way of overcoming these problems and improving the welfare of citizens. It has been pointed out that China’s major development strategies rely on solutions driven by big data (Heilmann 2017). For example, ‘Healthy China 2030’ explicitly stresses the importance of technology in achieving China’s healthcare reform strategy, and emphasises a

⁷ The Chinese text (社会建设) directly translates to ‘society/community’ and ‘build/construction’.

switch from treatment to prevention, with AI development as a means to achieve the goal (Ho 2018). This approach also shapes environmental protection, where President Xi has been promoting ‘digital environmental protection’ (数字环保) (Kostka and Zhang 2018). Within this, AI is being used to predict and mitigate air pollution levels (Knight 2015), and to improve waste management and sorting (“AI-powered waste management underway in China” 2019).

The administration of justice is another area where the Chinese government has been advancing using AI to improve social governance. Under Xi Jinping, there has been an explicit aim to professionalise the legal system, which suffers from a lack of transparency, issues of local protectionism, and interference in court cases by local officials (Finder 2015). A variety of reforms have been introduced in an attempt to curtail these practices including, transferring responsibility for the management of local courts from local to provincial governments, the creation of a database where judges can report attempts at interference by local politicians, and a case registration system that makes it more difficult for courts to reject complex or contentious cases (Li 2016).

Of particular interest, when focusing on AI, is the ‘Several Opinions of the Supreme People’s Court on Improving the Judicial Accountability System’ (2015), which requires judges to reference similar cases in their judicial reasoning. Furthermore, it stipulates that decisions conflicting with previous similar cases should trigger a supervision mechanism with more senior judges. To help judges minimise inconsistencies, an effort has been made to introduce AI technologies that facilitate making ‘similar judgements in similar cases’ (Yu and Du 2019). In terms of technology, two overarching types of systems have emerged. The first is a ‘similar cases pushing system’, where AI is used to identify judgements from similar cases and provide judges these for reference. This type of system has been introduced by, amongst others, Hainan’s High People’s Court who has also encouraged the use of AI systems in lower-level courts across the province (Yuan 2019). The second type uses AI techniques to provide an ‘abnormal judgement warning’ that would detect if a judgement made differs from similar cases. If an inconsistent judgement does occur, the system alerts the judge’s superiors, prompting an intervention. Despite the potential that AI-assisted sentencing holds for reducing the workload of judges and lessening corruption, feedback from those who have used the systems suggests that the technology is currently too imprecise (Yu and Du 2019). Some legal theorists have gone further in their criticism, highlighting the inhumane effects of using technology in sentencing and the detriment that it could cause for ‘legal hermeneutics, legal reasoning techniques, professional training and the ethical personality of the adjudicator’ (Ji 2013, p. 205).

Looking forward, the focus on China’s use of AI in governance seems most likely to centre on the widely reported ‘Social Credit’ System, which is premised upon developing the tools required to address China’s pressing social problems (Chorzempa et al. 2018). To do this, the system broadly aims at increasing the state’s governance capacity, promoting the credibility of state institutions, and building a viable financial credit base (Chai 2018). Currently, ‘the’ Social Credit System is not one unified nationwide system but rather comprises national blacklists that collate data from different government agencies, individual social credit systems run by local governments, and private company initiatives (Liu 2019). These systems are fractious and, in many cases, the local trials lack technical sophistication, with some versions relying on little more than paper and pen (Gan 2019). Nonetheless, the ambitious targets of the Social Credit System provide a compelling example of the government’s intent to rely on digital technology for social governance and also for more fine-grained regulation of the behaviour of its citizens.

3.4 Moral governance

Social governance/construction in China does not just encompass material and environmental features, but also the behaviour of citizens. Scholars have argued that the disruption of the Maoist period followed by an ‘opening up’ has created a moral vacuum within China (Yan 2009; Lazarus 2016). These concerns are echoed by the Chinese public, with Ipsos Mori finding that concerns over ‘moral decline’ in China were twice as high as the global average (Atkinson and Skinner 2019).⁸ This is something that has been recognised by the Chinese government, with high-level officials, including President Xi, forwarding the idea of a ‘minimum moral standard’ within society (He 2015). This goal is not limited to ensuring ‘good’ governance in the traditional sense; it extends to the regulating the behaviour of citizens and enhancing their moral integrity, which is considered a task within the government’s remit (“Xi Jinping’s report at 19th CPC National Congress” 2017). In view of the government, AI can be used to this end.

The AIDP highlights AI’s potential for understanding group cognition and psychology (2017). The intention to rely on AI for moral governance can be seen in further legislation, with perhaps the clearest example being the State Council’s ‘Outline for the Establishment of a Social Credit System’, released in 2014. This document underscored that

⁸ It is worth highlighting, however, that the Chinese are more than double the world average, and ranked first, when it comes to answering the question “whether the country is going in the right direction”, with 94% of the respondents in agreement.

the Social Credit System did not just aim to regulate financial and corporate actions of business and citizens, but also the social behaviour of individuals. This document outlines several social challenges that the plan seeks to alleviate, including tax evasion, food safety scares, and academic dishonesty (Chorzempa et al. 2018). As highlighted, current efforts to implement these systems have been fractious, yet a number have already included moral elements, such as publicly shaming bad debtors (Hornby 2019).

Further concrete examples of how China has been utilising AI in social governance can be seen in the sphere of internal security and policing. China has been at the forefront of the development of smart cities, with approximately half of the world's smart cities located within China. The majority of resources that have gone into developing these cities have focused on surveillance technologies, such as facial recognition and cloud computing for ordinary policing (Anderlini 2019). The use of advanced 'counterterrorism'⁹ surveillance programmes in the autonomous region of Xinjiang offers clearer and more problematic evidence of governmental efforts to use AI for internal surveillance. This technology is not limited to facial recognition, but also includes mobile phone applications to track the local Uyghur population, who are portrayed by the government as potential dissidents or terrorists ("China's Algorithms of Repression" 2019). When government statements are read in parallel with these developments, it seems likely that some form of the Social Credit System(s) will play a central role in the future of China's AI-enabled governance (Ding 2018), putting the rights of citizens under a sharp devaluative pressure. For example, most citizens generate large data footprints, and nearly all day-to-day transactions in cities are cashless and done with mobile apps (Morris 2019), internet providers enact 'real-name registration', linking all online activity to the individual (Sonnad 2017), enabling the government to identify and have access to the digital profile of all citizens using mobile-internet services.

The significant and likely risks related to implementing AI for governance stem from the intertwining of the material aspects of social governance with surveillance and moral control. Articles in the Western media often emphasise the problematic nature of 'the' Social Credit System, due to the authoritarian undertones of this pervasive control (Clover 2016; Botsman 2017). Examples of public dissatisfaction with specific features of locally run social credit systems appear to support this viewpoint (Zhang and Han 2019). In some cases, there have even been cases of public backlash leading to revisions in the rating criteria for local social

credit systems. In contrast, some commentators have emphasised that, domestically, a national social credit system may be positively received as a response to the perception of moral decline in China, and a concomitant desire to build greater trust; indeed, it has been suggested that the system may be better named the 'Social Trust' system (Song 2018; Kobie 2019). When looking at the punishments distributed by social credit systems, some measures, including blacklisting citizens from travelling due to poor behaviour on trains, have received a positive response on Chinese social media. Government censorship and a chilling effect could account for this support, but there is currently no evidence of censors specifically targeting posts concerning social credit systems (Koetse 2018).

Efforts have also been made to understand public opinion on the systems as a whole, rather than just specific controversies or cases of blacklisting. A nationwide survey by a Western academic on China's social credit systems found high levels of approval within the population (Kostka 2019). With this said, problems with the methodology of the paper, in particular with the translation of 'social credit system', indicate that it may be more appropriate to consider a general lack of awareness, rather than a widespread sentiment of support ('Beyond Black Mirror' 2019). These points indicate that it is too early to measure public sentiment in China surrounding the development of the Social Credit System(s).

It is important to recognise that despite the relative mundanity of current applications of the Social Credit System (Daum 2019; Lewis 2019), looking forward, substantial ethical risks and challenges remain in relation to the criteria for inclusion on a blacklist or receiving a low score and the exclusion that this could cause. In terms of the former, national blacklists are comprised of those who have broken existing laws and regulations, with a clear rationale for inclusion provided (Engelmann et al. 2019). However, the legal documents on which these lists are built are often ill-defined and function within a legal system that is subordinate to the Chinese Communist Party (Whiting 2017). As a result, legislation, like the one prohibiting the spread of information that seriously disturbs the social order, could be used to punish individuals for politically undesirable actions, including free speech (Arsène 2019). Still, it is more appropriate to consider this a problem of the political-legal structure and not a social credit system per se. The fundamental, ethical issue of an unacceptable approach to surveillance remains unaddressed.

In relation to local score-based systems that do not solely rely on illegality, assessment criteria can be even vaguer. For instance, social credit scores in Fuzhou account for 'employment strength', which is based on the loosely defined 'hard-working/conscientious and meticulous' (Lewis 2019). This is ethically problematic because of the opaque and arbitrary inclusion standards that are introduced for providing people

⁹ The word 'counterterrorism' started to be used after 9/11, with the phrase 'cultural integration' favoured before this ("Devastating Blows 2005).

certain benefits. In tandem with inclusion is the exclusion that punishments from these systems can cause. At present, most social credit systems are controlled by separate entities and do not connect with each other (Liu 2019), limiting excessive punishment and social exclusion. Nonetheless, memorandums of understanding are emerging between social credit systems and private companies for excluding those blacklisted from activities such as flying (Arsène 2019). As a result, it is important to emphasise that whilst the Social Credit System(s) is still evolving, the inclusion criteria and potential exclusion caused raise serious ethical questions.

4 The debate on digital ethics and AI in China

Alongside establishing material goals, the AIDP outlines a specific desire for China to become a world leader in defining ethical norms and standards for AI. Following the release of the AIDP, the government, public bodies, and industry within China were relatively slow to develop AI ethics frameworks (Lee 2018; Hickert and Ding 2018). However, there has been a recent surge in attempts to define ethical principles. In March 2019, China's Ministry of Science and Technology established the National New Generation Artificial Intelligence Governance Expert Committee. In June 2019, this body released eight principles for the governance of AI. The principles emphasised that, above all else, AI development should begin from enhancing the common well-being of humanity. Respect for human rights, privacy and fairness were also underscored within the principles. Finally, they highlighted the importance of transparency, responsibility, collaboration, and agility to deal with new and emerging risks (Laskai and Webster 2019).

In line with this publication, the Standardization Administration of the People's Republic of China, the national-level body responsible for developing technical standards, released a white paper on AI standards. The paper contains a discussion of the safety and ethical issues related to technology (Ding and Triolo 2018). Three key principles for setting the ethical requirements of AI technologies are outlined. First, the principle of human interest states that the ultimate goal of AI is to benefit human welfare. Second, the principle of liability emphasises the need to establish accountability as a requirement for both the development and the deployment of AI systems and solutions. Subsumed within this principle is transparency, which supports the requirement of understanding what the operating principles of an AI system are. Third, the principle of consistency of [sic] rights and responsibilities emphasised that, on the one hand, data should be properly recorded and oversight present but, on

the other hand, that commercial entities should be able to protect their intellectual property (Ding and Triolo 2018).

Government affiliated bodies and private companies have also developed their own AI ethics principles. For example, the Beijing Academy of Artificial Intelligence, a research and development body including China's leading companies and Beijing universities, was established in November 2018 (Knight 2019). This body then released the 'Beijing AI Principles' to be followed for the research and development, use, and governance of AI ("Beijing AI Principles" 2019). Similar to the principles forwarded by the AIDP Expert Committee, the Beijing Principles focus on doing good for humanity, using AI 'properly', and having the foresight to predict and adapt to future threats. In the private sector, one of the most high-profile ethical framework has come from the CEO of Tencent, Pony Ma. This framework emphasises the importance of AI being available, reliable, comprehensible, and controllable (Si 2019). Finally, the Chinese Association for Artificial Intelligence (CAAI)¹⁰ has yet to establish ethical principles but did form an AI ethics committee in mid-2018 with this purpose in mind ("AI association to draft ethics guidelines" 2019).

The aforementioned principles bear some similarity to those supported in the Global North (Floridi and Cowls 2019), yet institutional and cultural differences mean that the outcome is likely to be significantly different. China's AI ethics needs to be understood in terms of the country's culture, ideology, and public opinion (Triolo and Webster 2017). Although a full comparative analysis is beyond the scope of this article, it might be anticipated, for example, that the principles which emerge from China place a greater emphasis on social responsibility and group and community relations, with relatively less focus on individualistic rights, thus echoing earlier discussions about Confucian ethics on social media (Wong 2013).

In the following sections, we shall focus on the debate about AI ethics as it is emerging in connection with privacy and medical ethics because these are two of the most mature areas where one may grasp a more general sense of the current 'Chinese approach' to digital ethics. The analysis of the two areas is not meant to provide an exhaustive map of all the debates about ethical concerns over AI in China. Instead, it may serve to highlight some of the contentious issues that are emerging, and inform a wider understanding of the type of boundaries which may be drawn in China when a normative agenda in the country is set.

¹⁰ The Chinese Association for Artificial Intelligence (CAAI) is the only state-level science and technology organization in the field of artificial intelligence under the Ministry of Civil Affairs.

4.1 Privacy

All of the sets of principles for ethical AI outlined above mention the importance of protecting privacy. However, there is a contentious debate within China over exactly what types of data should be protected. China has historically had weak data protection regulations—which has allowed for the collection and sharing of enormous amounts of personal information by public and private actors—and little protection for individual privacy. In 2018, Robin Li, co-founder of Baidu, stated that ‘the Chinese people are more open or less sensitive about the privacy issue. If they are able to trade privacy for convenience, safety and efficiency, in a lot of cases, they are willing to do that’ (Liang 2018). This viewpoint—which is compatible with the apparently not too negative responses to the Social Credit System—has led some Western commentators to misconstrue public perceptions of privacy in their evaluations of China’s AI strategy (Webb 2019). However, Li’s understanding of privacy is not one that is widely shared, and his remarks sparked a fierce backlash on Chinese social media (Liang 2018). This concern for privacy is reflective of survey data from the Internet Society of China, with 54% of respondents stating that they considered the problem of personal data breaches as ‘severe’ (Sun 2018). When considering some cases of data misuse, this number is unsurprising. For example, a China Consumers Association survey revealed that 85% of people had experienced a data leak of some kind (Yang 2018). Thus, contrary to what may be inferred from some high-profile statements, there is a general sentiment of concern within the Chinese public over the misuse of personal information.

As a response to these serious concerns, China has been implementing privacy protection measures, leading one commentator to refer to the country as ‘Asia’s surprise leader on data protection’ (Lucas 2018). At the heart of this effort has been the Personal Information Security Specification (the Specification), a privacy standard released in May 2018. This standard was meant to elaborate on the broader privacy rules, which were established in the 2017 Cyber Security Law. In particular, it focused on both protecting personal data and ensuring that people are empowered to control their own information (Hong 2018). A number of the provisions within the standard were particularly all-encompassing, including a broad definition of sensitive personal information, which includes features such as reputational damage. The language used in the standard led one commentator to argue that some of the provisions were more onerous than those of the European Union’s (EU) General Data Protection Regulation (GDPR) (Sacks 2018).

Despite the previous evidence, the nature of the standard means that it is not really comparable to the GDPR. On the one hand, rather than being a piece of formally enforceable legislation, the Specification is merely a ‘voluntary’ national

standard created by the China National Information Security Standardization Technical Committee (TC260). It is on this basis that one of the drafters stated that this standard was not comparable to the GDPR, as it is only meant as a guiding accompaniment to previous data protection legislation, such as the 2017 Cyber Security law (Hong 2018). On the other hand, there remains a tension that is difficult to resolve because, although it is true that standards are only voluntary, standards in China hold substantive clout for enforcing government policy aims, often through certification schemes (Sacks and Li 2018). In June 2018, a certification standard for privacy measures was established, with companies such as Alipay and Tencent Cloud receiving certification (Zhang and Yin 2019). Further, the Specification stipulates the specificities of the enforceable Cybersecurity Law, with Baidu and AliPay both forced to overhaul their data policies due to not ‘complying with the spirit of the Personal Information Security Standard’ (Yang 2019).

In reality, the weakness in China’s privacy legislation is due less to its ‘non-legally binding’ status and more to the many loopholes in it, the weakness of China’s judicial system, and the influential power of the government, which is often the last authority, not held accountable through democratic mechanisms. In particular, significant and problematic exemptions are present for the collection and use of data, including when related to security, health, or the vague and flexibly interpretable ‘significant public interests’. It is these large loopholes that are most revealing of China’s data policy. It may be argued that some broad consumer protections are present, but actually this is not extended to the government (Sacks and Laskai 2019). Thus, the strength of privacy protection is likely to be determined by the government’s decisions surrounding data collection and usage, rather than legal and practical constraints. This is alarming.

It is important to recognise that the GDPR contains a similar ‘public interest’ basis for lawfully processing personal data where consent or anonymisation are impractical and that these conditions are poorly defined in legislation and often neglected in practice (Stevens 2017). But the crucial and stark difference between the Chinese and EU examples concerns the legal systems underpinning the two approaches. The EU’s judicial branch has substantive influence, including the capacity to interpret legislation and to use judicial review mechanisms to determine the permissibility of legislation more broadly.¹¹ In contrast, according to the Chinese legal system, the judiciary is subject to supervision and interference from the legislature, which has de jure legislative supremacy (Ji 2014); this gives de facto

¹¹ As a practical example of this, the Court of Justice of the European Union gave judgment in *Rīgas Case* (2017) that has been used in defining what is meant by ‘legitimate interest.’

control to the Party (Horsley 2019). Thus, the strength of privacy protections in China may be and often is determined by the government's decisions surrounding data collection and usage rather than legal and practical constraints. As it has been remarked, 'The function of law in governing society has been acknowledged since 2002, but it has not been regarded as essential for the CCP. Rather, morality and public opinion concurrently serve as two alternatives to the law for the purpose of governance. As a result, administrative agencies may ignore the law on the basis of party policy, morality, public opinion, or other political considerations' (Wang and Liu 2019, p. 6).

When relating this back to AI policy, China has benefited from the abundance of data that historically lax privacy protections have facilitated (Ding 2018). On the surface, China's privacy legislation seems to contradict other development commitments, such as the Social Credit System, which requires extensive personal data. This situation creates a dual ecosystem whereby the government is increasingly willing to collect masses of data, respecting no privacy, while simultaneously admonishing tech companies for the measures they employ (Sacks and Laskai 2019). Recall that private companies, such as the AI National Team, are relied upon for governance at both a national and local level, and therefore may receive tacit endorsement rather than admonishment in cases where the government's interests are directly served. As a result, the 'privacy strategy' within China appears to aim to protect the privacy of a specific type of consumer, rather than that of citizens as a whole, allowing the government to collect personal data wherever and whenever it may be merely useful (not even strictly necessary) for its policies. From an internal perspective, one may remark that, when viewed against a backdrop of high levels of trust in the government and frequent private sector leaks and misuses, this trade-off seems more intelligible to the Chinese population. The Specification has substantial scope for revising this duality, with a number of loopholes being closed since the initial release (Zhang and Yin 2019), but it seems unlikely that privacy protections from government intrusion will be codified in the near future. The ethical problem remains unresolved.

4.2 Medical ethics

Medical ethics is another significant area impacted by the Chinese approach to AI ethics. China's National Health Guiding Principles have been central to the strategic development and governance of its national healthcare system for the past 60 years (Zhang and Liang 2018). They have been re-written several times, as the healthcare system has transitioned from being a single-tier system, prior to 1978, to a two-tier system that was reinforced by healthcare reform in 2009 (Wu and Mao 2017). The last re-write of the Guiding

Principles was in 1996 and the following principles still stand (Zhang and Liang 2018):

- (a) People in rural areas are the top priority
- (b) Disease prevention must be placed first
- (c) Chinese traditional medicine and Western medicine must work together
- (d) Health affairs must depend on science and education
- (e) Society as a whole should be mobilised to participate in health affairs, thus contributing to the people's health and the country's overall development.

All five principles are relevant for understanding China's healthcare system as a whole but, from the perspective of analysing the ethics of China's use of AI in the medical domain, principles (a), (b), and (e) are the most important. They highlight that—in contrast to the West, where electronic healthcare data are predominantly focused on individual health, and thus AI techniques are considered crucial to unlock 'personalised medicine' (Nittas et al. 2018)—in China, healthcare is predominantly focused on the health of the population. In this context, the ultimate ambition of AI is to liberate data for public health purposes¹² (Li et al. 2019a, b). This is evident from the AIDP, which outlines the ambition to use AI to 'strengthen epidemic intelligence monitoring, prevention and control,' and to 'achieve breakthroughs in big data analysis, Internet of Things, and other key technologies' for the purpose of strengthening community intelligent health management. The same aspect is even clearer in the State Council's 2016 official notice on the development and use of big data in the healthcare sector, which explicitly states that health and medical big data sets are a national resource and that their development should be seen as a national priority to improve the nation's health (Zhang et al. 2018).¹³

From an ethical analysis perspective, the promotion of healthcare data as a public good throughout public policy—including documents such as Measures on Population Health Information and the Guiding Opinions on Promoting and Regulating the Application of Big Medical and Health Data (Chen and Song 2018)—is crucial. This approach,

¹² This is not to imply that the West is not interested in using AI for population health management purposes, or that China is not interested in using AI for personalised health purposes. China is, for example, also developing an integrated data platform for research into precision medicine (Zhang et al. 2018). We simply mean to highlight that the order of priority between these two goals seems to differ.

¹³ The challenges section outlines some concrete benefits of implementing AI, illustrating some perceived gains to China. A separate (though more technological than ethical) point substantiated by the article is there is a lot of medical data which could potentially be beneficial, but the data are spread out among hospitals, not used for research, and largely unstructured.

combined with lax rules about data sharing within China (Liao 2019; Simonite 2019), and the encouragement of the open sharing of public data between government bodies (“Outline for the Promotion of Big Data Development” 2015), promotes the collection and aggregation of health data without the need for individual consent, by positioning group beneficence above individual autonomy. This is best illustrated with an example. As part of China’s ‘Made in 2025’ plan, 130 companies, including ‘WeDoctor’ (backed by Tencent, one of China’s AI national champions) signed co-operation agreements with local governments to provide medical check-ups comprised of blood pressure, electrocardiogram (ECG), urine and blood tests, free of charge to rural citizens (Hawkins 2019). The data generated by these tests were automatically (that is with no consent from the individual) linked to a personal identification number and then uploaded to the WeDoctor cloud, where they were used to train WeDoctor’s AI products. These products include the ‘auxiliary treatment system for general practice’, which is used by village doctors to provide suggested diagnosis and treatments from a database of over 5000 symptoms and 2000 diseases. Arguably, the sensitive nature of the data can make ‘companies—and regulators—wary of overseas listings, which would entail greater disclosure and scrutiny’ (Lucas 2019). Although this, and other similar practices, do involve anonymisation, they are in stark contrast with the European and US approaches to the use of medical data, which prioritise individual autonomy and privacy, rather than social welfare. A fair balance between individual and societal needs is essential for an ethical approach to personal data, but there is an asymmetry whereby an excessive emphasis on an individualistic approach may be easily rectified with the consensus of the individuals, whereas a purely societal approach remains unethical insofar as it overrides too easily individual rights and cannot be rectified easily.

Societal welfare may end up justifying the sacrifice of individual rights as a means. This remains unethical. However, how this is perceived within China remains a more open question. One needs to recall that China has very poor primary care provision (Wu and Mao 2017), that it achieved 95% health coverage (via a national insurance scheme) only in 2015 (Zhang et al. 2018), it has approximately 1.8 doctors per 1000 citizens compared to the OECD average of 3.4 (Liao 2019), and is founded on Confucian values that promote group-level equality. It is within this context that the ethical principle of the ‘duty of easy rescue’ may be interpreted more insightfully. This principle prescribes that, if an action can benefit others and poses little threat to the individual, then the ethical option is to complete the action (Mann et al. 2016). In this case, from a Chinese perspective, one may argue that sharing of the healthcare data may pose little immediate threat to the individual, especially as Article 6 of the Regulations on the Management of Medical Records

of Medical Institutions, Article 8 of the Management Regulations on Application of Electronic Medical Records, Article 6 of the Measures for the Management of Health Information, the Cybersecurity Law of the People’s Republic of China, and the new Personal Information Security Specification all provide specific and detailed instructions to ensure data security and confidentiality (Wang 2019). However, it could potentially deliver significant benefit to the wider population.

The previous ‘interpretation from within’ does not imply that China’s approach to the use of AI in healthcare is acceptable or raises no ethical concerns. The opposite is actually true. In particular, the Chinese approach is undermined by at least three main risks.

First, there is a risk of creating a market for human care. China’s two-tiered medical system provides state-insured care for all, and the option for individuals to pay privately for quicker or higher quality treatment. This is in keeping with Confucian thought, which encourages the use of private resources to benefit oneself and one’s family (Wu and Mao 2017). With the introduction of Ping An [sic] Good Doctor’s unmanned ‘1-min clinics’ across China (of which there are now as many as 1,000 in place), patients can walk in, provide symptoms and medical history, and receive an automated diagnosis and treatment plans (which are only followed up by human clinical advice for new customers), it is entirely possible to foresee a scenario in which only those who are able to pay will be able to access human clinicians. In a field where emotional care, and involvement in decision making, are often as important as the logical deduction of a ‘diagnosis,’ this could have a significantly negative impact on the level and quality of care accessed across the population and on the integrity of the self (Andorno 2004; Pasquale 2015),¹⁴ at least for those who are unable to afford human care.

Second, in the context of a population that is still rapidly expanding yet also ageing, China is investing significantly in the social informatisation of healthcare and has, since at least 2015, been linking emotional and behavioural data extrapolated from social media and daily healthcare data (generated from ingestibles, implantables, wearables, carebots, and Internet of Things devices) to Electronic Health Records (Li et al. 2019a, b), with the goal of enabling community care of the elderly. This further adds to China’s culture of state-run, mass-surveillance and, in the age of the Social Credit System, suggests that the same technologies designed to enable people to remain independent in the community as they age may one day be used as a means of social control (The Medical Futurist 2019), to reduce the incidence of ‘social diseases’—such as obesity and type II diabetes

¹⁴ Note that the emphasis on individual wellbeing must also be contextualised culturally.

(Hawkins 2019)—under the guise of ‘improving peoples lives’ through the use of AI to improve the governance of social services (as stated in the AIDP).

The third ethical risk is associated with CRISPR gene modification and AI. CRISPR is a controversial gene modification technique that can be used to alter the presentation of genes in living organisms, for example for the purpose of curing or preventing genetic diseases. It is closely related to AI, as Machine Learning techniques can be used to identify which gene or genes need to be altered with the CRISPR method. The controversies, and potential significant ethical issues, associated with research in this area are related to the fact that it is not always possible to tell where the line is between unmet clinical need and human enhancement or genetic control (Cohen 2019). This became clear when, in November 2018, biophysics researcher He Jiankui revealed that he had successfully genetically modified babies using the CRISPR method to limit their chances of ever contracting HIV (Cohen 2019). The announcement was met by international outcry and He’s experiment was condemned by the Chinese government at the time (Belluz 2019). However, the drive to be seen as a world leader in medical care (Cheng 2018), combined with the promise gene editing offers for the treatment of diseases, suggest that a different response may be possible in the future (Cyranski 2019; “China opens a Pandora’s Box”, 2018). Such a change in government policy is especially likely as global competition in this field heats up. The US has announced that it is enrolling patients in a trial to cure an inherited form of blindness (Marchione 2019); and the UK has launched the Accelerating Detection of Disease challenge to create a five-million patient cohort whose data will be used to develop new AI approaches to early diagnosis and biomarker discovery (UK Research and Innovation 2019). These announcements create strong incentives for researchers in China to push regulatory boundaries to achieve quick successes (Tatlow 2015; Lei et al. 2019). China has filed the largest number of patents for gene-editing on animals in the world (Martin-Laffon et al. 2019). Close monitoring will be essential if further ethical misdemeanours are to be avoided.

5 Conclusion

In this article, we analysed the nature of AI policy within China and the context within which it has emerged, by mapping the major national-level policy initiatives that express the intention to utilise AI. We identified three areas of particular relevance: international competitiveness, economic growth, and social governance (construction). The development and deployment of AI in each of these areas have implications for China and for the international community. For example, although the ‘trump-card’ policy to

gain a military advantage may not be something new, its application to AI technologies risks igniting an arms race and undermining international stability (Taddeo and Floridi 2018). Efforts to counteract this trend seem largely hollow. Our analysis indicates that China has some of the greatest opportunities for economic benefit in areas like automation and that the country is pushing forward in AI-related areas substantially. Nonetheless, efforts to cushion the disruptions that emerge from using AI in the industry are currently lacking. Thus, AI can help foster increased productivity and high levels of growth, but its use is likely to intensify the inequalities present within society and even decrease support for the government and its policies. The AIDP also promotes AI as a way to help deal with some of the major social problems, ranging from pollution to standards of living. However, positive impact in this area seem to come with increased control over individuals’ behaviour, with governance extending into the realm of moral behaviour and further erosion of privacy.

Ethics also plays a central role in the Chinese policy effort on AI. The AIDP outlines a clear intention to define ethical norms and standards, yet efforts to do so are at a fledgling stage, being broadly limited to high-level principles, lacking implementation. Analyses of existing Chinese approaches and emerging debates in the areas of privacy and medical ethics provide an insight into the types of frameworks that may emerge. With respect to privacy, on the surface, recently introduced protections may seem robust, with definitions of sensitive personal information even broader than that used within the GDPR. However, a closer look exposes the many loopholes and exceptions that enable the government (and companies implicitly endorsed by the government) to bypass privacy protection and fundamental issues concerning lack of accountability and government’s unrestrained decisional power about mass-surveillance.

In the same vein, when focusing on medical ethics, it is clear that, although China may agree with the West on the bioethical principles, its focus on the health of the population, in contrast to the West’s focus on the health of the individual, may easily lead to unethical outcomes (the sacrifice imposed on one for the benefit of many) and is creating a number of risks, as AI encroaches on the medical space. These are likely to evolve over time, but the risks of unequal care between those who can afford a human clinician and those who cannot, control of social diseases, and of unethical medical research are currently the most significant.

China is a central actor in the international debate on the development and governance of AI. It is important to understand China’s internal needs, ambitions in the international arena, and ethical concerns, all of which are shaping the development of China’s AI policies. It is also important to understand all this not just externally, from a Western perspective, but also internally, from a Chinese perspective. However, some ethical safeguards, constraints and desiderata

are universal and are universally accepted and cherished, such as the nature and scope of human rights.¹⁵ They enable one to evaluate, after having understood, China's approach to the development of AI. This is why in this article we have sought to contribute to a more comprehensive and nuanced analysis of the structural, cultural and political factors that ground China's stance on AI, as well as an indication of its possible trajectory, while also highlighting where ethical problems remain, arise, or are likely to be exacerbated. They should be addressed as early as it is contextually possible.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- AI association to draft ethics guidelines (2019). Xinhua. https://www.xinhuanet.com/english/2019-01/09/c_137731216.htm. Accessed 1 May 2019
- AI Policy-China (n.d.) Future of Life Institute. <https://futureoflife.org/ai-policy-china/>. Accessed 14 Mar 2019
- AI-powered waste management underway in China (2019) People's Daily Online. <https://en.people.cn/n3/2019/0226/c98649-9549956.html>. Accessed 13 Apr 2019
- Allen G (2019) Understanding China's AI Strategy. Center for a New American Security. <https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy>. Accessed 13 Mar 2019
- Anderlini J (2019) How China's smart-city tech focuses on its own citizens. Financial Times. <https://www.ft.com/content/46bc137a-5d27-11e9-840c-530737425559>. Accessed 25 Apr 2019
- Arsène S (2019) China's Social Credit System A Chimera with Real Claws. *Asie.Visions*, No. 110. https://www.ifri.org/sites/default/files/atoms/files/arsene_china_social_credit_system_2019.pdf. Accessed 12 Mar 2020
- Atkinson S, Skinner G (2019) What worries the world?. IPSOS Public Affairs, Paris
- Austin G (2016) International legal norms in cyberspace: evolution of China's National Security Motivations. In: Osula A, Rõigas H (eds) *International cyber norms: legal, policy and industry perspectives*. NATO CCD COE Publications, Tallinn
- Balding C (2019) What's causing China's economic slowdown. *Foreign Affairs*. <https://www.foreignaffairs.com/articles/china/2019-03-11/whats-causing-chinas-economic-slowdown>. Accessed 20 Mar 2019
- Barton D, Woetzel J, Seong J, Tian Q (2017) *Artificial intelligence: implications for China*. McKinsey Global Institute, San Francisco
- Beijing AI Principles (2019) Beijing Academy of Artificial Intelligence. <https://www.baai.ac.cn/blog/beijing-ai-principles>. Accessed 05 June 2019
- Belluz J (2019) Is the CRISPR baby controversy the start of a terrifying new chapter in gene editing? *VOX*. <https://www.vox.com/science-and-health/2018/11/30/18119589/crispr-gene-editing-he-jianku>. Accessed 12 Apr 2019
- Beyond Black Mirror-China's Social Credit System (2019) *Re:publica*. <https://19.re-publica.com/de/session/beyond-black-mirror-china-s-social-credit-system>. Accessed 30 June 2019
- Blasko DJ (2011) 'Technology Determines Tactics': the relationship between technology and doctrine in Chinese military thinking. *J Strat Stud* 34(3):355–381. <https://doi.org/10.1080/01402390.2011.574979>
- Borowiec S (2016) Google's AI machine v world champion of 'Go': everything you need to know. *The Guardian*. <https://www.theguardian.com/technology/2016/mar/09/googles-ai-machine-v-world-champion-of-go-everything-you-need-to-know>. Accessed 29 Mar 2019
- Botsman R (2017) Big data meets Big Brother as China moves to rate its citizens. *Wired UK*. <https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>. Accessed 10 Apr 2019
- Brezis ES, Krugman PR, Tsiddon D (1993) Leapfrogging in international competition: a theory of cycles in National Technological Leadership. *Am Econ Rev* 83(5):1211–1219
- Bruzdzinski JE (2004) Demystifying Shashoujian: "China's Assassin's Mace" Concept. In: Scobell A, Wortzel L (eds) *Civil-Military Change in China Elites, Institutes, and Ideas after the 16th Party Congress*. Diane Publishing Co, Darby
- Cai F, Lu Y (2013) Population change and resulting slowdown in potential GDP growth in China. *China World Econ* 21(2):1–14. <https://doi.org/10.1111/j.1749-124X.2013.12012.x>
- Chai S (2018) The social credit system: is technology helping china legitimate and augment authoritarian rule? IPP2018 Conference Paper. <https://oxis.oii.ox.ac.uk/wp-content/uploads/sites/77/2018/08/IPP2018-Chai.pdf>. Accessed 03 June
- Chan J (1999) A confucian perspective on human rights for contemporary China. In: Bauer JR, Bell DA (eds) *The East Asian challenge for human rights*. Cambridge University Press, Cambridge
- Chan M, Zhen L (2019) Modern military remains top priority as China boosts defence spending. *South China Morning Post*. <https://www.scmp.com/news/china/military/article/2188771/modernising-military-remains-top-priority-china-boosts-defence>. Accessed 06 Apr 2019
- Chen J (2013) *A middle class without democracy: economic growth and the prospects for democratization in China*. Oxford University Press, Oxford
- Chen Y, Song L (2018) China: concurring regulation of cross-border genomic data sharing for statist control and individual protection. *Hum Genet* 137(8):605–615. <https://doi.org/10.1007/s00439-018-1903-2>
- Cheng Y (2018) China will always be bad at bioethics. *Foreign Policy*. <https://foreignpolicy.com/2018/04/13/china-will-always-be-bad-at-bioethics/>. Accessed 10 Apr 2019
- Cheung TM, Mahnken T, Seligsohn D, Pollpeter K, Anderson E, Yang F (2016) *Planning for innovation: understanding China's plans for technological, energy, Industrial, and Defense Development*. US–China Economic and Security Review Commission
- Chin JJ (2018) The longest March: why China's democratization is not imminent. *J Chin Polit Sci* 23(1):63–82. <https://doi.org/10.1007/s11366-017-9474-y>

¹⁵ For arguments on the universality of human rights coming from *within* cultural perspectives, see Chan (1999) on Confucianism and human rights.

- China AI Development Report (2018) Tsinghua University. https://www.sppm.tsinghua.edu.cn/eWebEditor/UploadFile/China_AI_development_report_2018.pdf. Accessed 21 Mar 2019
- China Opens a ‘Pandora’s Box’ of Human Genetic Engineering (2018) Bloomberg News. <https://www.bloomberg.com/news/articles/2018-11-27/china-opens-a-pandora-s-box-of-human-genetic-engineering>. Accessed 30 July 2019
- China’s Algorithms of Repression: Reverse Engineering a Xinjiang Police Mass Surveillance App (2019) Human Rights Watch. <https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance>. Accessed 14 May 2019
- Chorzempa M, Triolo P, Sacks S (2018) China’s social credit system: a mark of progress or a threat to privacy? Peterson Institute for International Economics. <https://piie.com/publications/policy-briefs/chinas-social-credit-system-mark-progress-or-threat-privacy>. Accessed 29 Apr 2019
- Clover C (2016) China: when big data meets big brother. Financial Times. <https://www.ft.com/content/b5b13a5e-b847-11e5-b151-8e15c9a029fb>. Accessed 30 Apr 2019
- Cohen J (2019) The untold story of the ‘circle of trust’ behind the world’s first gene-edited babies. Science. <https://doi.org/10.1126/science.aay9400>
- Cyranoski D (2019) The CRISPR-baby scandal: what’s next for human gene-editing. Nature 566:440. <https://doi.org/10.1038/d41586-019-00673-1>
- Daum J (2019) Untrustworthy: social Credit Isn’t what you think it is. Verfassungsblog. <https://verfassungsblog.de/untrustworthy-social-credit-isnt-what-you-think-it-is/>. Accessed 17 Mar 2020
- Delang CO (2016) China’s air pollution problems. Routledge, Abingdon
- Devastating Blows: Religious Repression of Uighurs in Xinjiang (2005) Human Rights Watch. <https://www.hrw.org/report/2005/04/11/devastating-blows/religious-repression-uighurs-xinjiang>. Accessed 04 Apr 2020
- Diamond L (2003) The rule of law as transition to democracy in China. J Contemp China 12(35):319–331. <https://doi.org/10.1080/1067056022000054632>
- Dickson BJ (2003) Red capitalists in China: the party, private entrepreneurs, and prospects for political change. Cambridge University Press, Cambridge
- Ding J (2018) Deciphering China’s AI dream. Future of Humanity Institute, University of Oxford, Oxford
- Ding J (2019) ChinAI #51: China’s AI “National Team”, ChinAI Newsletter. <https://chinai.substack.com/p/chinai-51-chinas-ai-national-team>. Accessed 27 July 2019
- Ding J, Triolo P (2018) Translation: excerpts from China’s ‘White Paper on Artificial Intelligence Standardization’. New America. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-excerpts-chinas-white-paper-artificial-intelligence-standardization/>. Accessed 25 Apr 2019
- Economy EC (2013) China’s water pollution crisis. The Diplomat. <https://thediplomat.com/2013/01/forget-air-pollution-chinas-has-a-water-problem/>. Accessed 25 Apr 2019
- Engelmann S, Chen M, Fischer F, Kao CY, Grossklags J (2019) Clear sanctions, vague rewards: how China’s social credit system currently defines “good” and “bad” Behavior. In: Proceedings of the conference on fairness, accountability, and transparency. pp 69–78. <https://doi.org/10.1145/3287560.3287585>
- Fang A (2019) Chinese colleges to offer AI major in challenge to US. Nikkei Asian Review. <https://asia.nikkei.com/Business/China-tech/Chinese-colleges-to-offer-AI-major-in-challenge-to-US>. Accessed 14 Apr 2019
- Finder S (2015) China’s master plan for remaking its courts. The Diplomat. <https://thediplomat.com/2015/03/chinas-master-plan-for-remaking-its-courts/>. Accessed 13 Mar 2020
- Floridi L, Cowls J (2019) A unified framework of five principles for AI in society. Harvard Data Sci Rev. <https://doi.org/10.1162/99608f92.8cd550d1>
- Gan N (2019) China’s social credit horrifies the West. But do the Chinese even notice? South China Morning Post. <https://www.scmp.com/news/china/politics/article/2185303/hi-tech-dystopia-or-low-key-incentive-scheme-complex-reality>. Accessed 30 Apr 2019
- Graceffo A (2017) China’s National Champions: State Support Makes Chinese Companies Dominant. Foreign Policy Journal. <https://www.foreignpolicyjournal.com/2017/05/15/chinas-national-champions-state-support-makes-chinese-companies-dominant/>. Accessed 10 June 2019
- Guidance of the State Council on the Active Promotion of the “Internet Plus” (2015) State Council. [国务院关于积极推进“互联网+”行动的指导意见, 国务院]. https://www.gov.cn/zhengce/content/2015-07/04/content_10002.htm. Accessed 14 Apr 2019
- Hawkins A (2019) How elderly, sickly farmers are quenching China’s thirst for data. Wired. <https://www.wired.co.uk/article/china-ai-healthcare>. Accessed 01 June
- He H (2015) Social ethics in a changing china: moral decay or ethical awakening?. Brookings Institution Press, Washington
- Healthy China 2030 (from vision to action) (2016) WHO. <https://www.who.int/healthpromotion/conferences/9gchp/healthy-china/en/>. Accessed 13 Apr 2019
- Heilmann S (2017) Big data reshapes China’s approach to governance. Financial Times. <https://www.ft.com/content/43170fd2-a46d-11e7-b797-b61809486fe2>. Accessed 29 Apr 2019
- Heilmann S, Melton O (2013) The reinvention of development planning in China, 1993–2012. Mod China 39(6):580–628. <https://doi.org/10.1177/0097700413497551>
- Hickert C, Ding J (2018) Read what top Chinese officials are hearing about AI competition and policy. New America. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/read-what-top-chinese-officials-are-hearing-about-ai-competition-and-policy/>. Accessed 31 Mar 2019
- Ho A (2018) AI can solve China’s doctor shortage. Here’s how. World Economic Forum. <https://www.weforum.org/agenda/2018/09/ai-can-solve-china-s-doctor-shortage-here-s-how/>. Accessed 10 May 2020
- Hong Y (2018) Responses and explanations to the five major concerns of the Personal Information Security Code. Network. [洪延青. 对《个人信息安全规范》五大重点关切的回应和解释. 网安寻路人]. <https://mp.weixin.qq.com/s/rSW-Ayu6zNXw87itYHcPYA>. Accessed 01 June 2019
- Hornby L (2019) Chinese app names and shames bad debtors. Financial Times. <https://www.ft.com/content/2ad7feea-278e-11e9-a5ab-ff8ef2b976c7>. Accessed 01 June 2019
- Horsley J (2019) Party Leadership and Rule of Law in the Xi Jinping Era. Global China. https://law.yale.edu/sites/default/files/area-center/china/document/horsley_china_party_legal_development.pdf. Accessed 14 Mar 2020
- Hu A (2013) The distinctive transition of China’s five-year plans. Mod China 39(6):629–639. <https://doi.org/10.1177/0097700413499129>
- Huang Y (2013) Democratize or Die. <https://www.foreignaffairs.com/articles/china/2012-12-03/democratize-or-die>. Accessed 01 Apr 2019
- Inkster N (2010) China in cyberspace. Survival 52(4):55–66. <https://doi.org/10.1080/00396338.2010.506820>
- Is China ready for intelligent automation? (2018) China Power, Center for Strategic and International Studies. <https://chinapower.csis.org/china-intelligent-automation/>. Accessed 06 Apr 2019
- Ji W (2013) The judicial reform in China: the status quo and future directions. Indiana J Glob Legal Stud 20(1):185–220

- Ji W (2014) The rule of law in a Chinese way: social diversification and reconstructing the system of authority. *Asian J Law Soc* 1(2):305–338. <https://doi.org/10.1017/als.2014.9>
- Jing M, Dai S (2017) China recruits Baidu, Alibaba and Tencent to AI ‘national team’. *South China Morning Post*. <https://www.scmp.com/tech/china-tech/article/2120913/china-recruits-baidu-alibaba-and-tencent-ai-national-team>. Accessed 14 Apr 2019
- Kania EB (2020) “AI weapons” in China’s military innovation. *Brookings: Global China*. https://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf. Accessed 15 May 2020
- Kania EB (2017a) 杀手锏和跨越发展: trump cards and leapfrogging. *Strategy Bridge*. <https://thestrategybridge.org/the-bridge/2017/9/5/-and-trump-cards-and-leapfrogging>. Accessed 24 Apr 2019
- Kania EB (2017b) Battlefield singularity: artificial intelligence, military revolution, and China’s future military power, Center for a New American Security
- Kania EB (2018a) China’s strategic ambiguity and shifting approach to lethal autonomous weapons systems. *Lawfare*. <https://www.lawfareblog.com/chinas-strategic-ambiguity-and-shifting-approach-lethal-autonomous-weapons-systems>. Accessed 06 Apr 2019
- Kania EB (2018b) China’s embrace of AI: Enthusiasm and challenges. *European council on foreign relations*. https://www.ecfr.eu/article/commentary_chinas_embrace_of_ai_enthusiasm_and_challenges. Accessed 30 Mar 2019
- Knight W (2015) How artificial intelligence can be used to fight air pollution in China. *MIT Technology Review*. <https://www.technologyreview.com/s/540806/how-artificial-intelligence-can-fight-air-pollution-in-china/>. Accessed 24 Mar 2019
- Knight W (2019) Why does Beijing suddenly care about AI ethics? *MIT Technology Review*. <https://www.technologyreview.com/s/613610/why-does-china-suddenly-care-about-ai-ethics-and-privacy/>. Accessed 01 June 2019
- Kobie N (2019) The complicated truth about China’s social credit system. *Wired UK*. <https://www.wired.co.uk/article/china-social-credit-system-explained>. Accessed 03 Mar 2019
- Koetse M (2018) Insights into the social credit system on Chinese online media vs its portrayal in Western Media. What’s on Weibo. <https://www.whatsonweibo.com/insights-into-the-social-credit-system-on-chinese-online-media-and-stark-contrasts-to-western-media-approaches/>. Accessed 03 June 2019
- Kostka G (2019) China’s social credit systems and public opinion: explaining high levels of approval. *New Media Soc*. <https://doi.org/10.1177/1461444819826402>
- Kostka G, Zhang C (2018) Tightening the grip: environmental governance under Xi Jinping. *Environ Politics* 27(5):769–781. <https://doi.org/10.1080/09644016.2018.1491116>
- Ku J (2017) Tentative observations on China’s views on international law and cyber warfare. *Lawfare*. <https://www.lawfareblog.com/tentative-observations-chinas-views-international-law-and-cyber-warfare>. Accessed 17 Mar 2020
- Laskai L, Webster G (2019) Translation: Chinese expert group offers ‘governance principles’ for ‘responsible AI’. *New America Foundation*. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/>. Accessed 05 July 2019
- Lazarus LM (2016) China’s quest for a moral compass. *The Diplomat*. <https://thediplomat.com/2016/09/chinas-quest-for-a-moral-compass/>. Accessed 01 May 2019
- Lee KF (2018) AI superpowers: China, Silicon Valley, and the new world order. *Houghton Mifflin Harcourt*, Boston
- Lei R, Zhai X, Zhu W, Qiu R (2019) Reboot ethics governance in China. *Nature*. 569:184–186. <https://doi.org/10.1038/d41586-019-01408-y>
- Leng S (2018) China has new three-year plan to clean up environment. *South China Morning Post*. <https://www.scmp.com/news/china/policies-politics/article/2137666/china-has-new-three-year-plan-clean-environment>. Accessed 25 Apr 2019
- Lewis D (2019) All carrots and no sticks: a case study on social credit scores in Xiamen and Fuzhou. *Berkman Klein Centre Collection*. <https://medium.com/berkman-klein-center/social-credit-case-study-city-citizen-scores-in-xiamen-and-fuzhou-2a65feb2bb3>. Accessed 17 Mar 2020
- Li A (2016) Centralisation of power in the pursuit of law-based governance: legal reform in China under the Xi Administration. *China Perspect* 2016:2
- Li B, Li J, Jiang Y, Lan X (2019) Experience and reflection from China’s Xiangya medical big data project. *J Biomed Inform*. <https://doi.org/10.1016/j.jbi.2019.103149>
- Li H, Zhou LA (2005) Political turnover and economic performance: the incentive role of personnel control in China. *J Public Econ* 89(9–10):1743–1762
- Li M (2019) Where does the winning mechanism for intelligent wars change? *China Military Network*. [李明海. (2019). 智能化战争的制胜机理变在哪里. 中国军网]
- Li Q, Lan L, Zeng N, You L, Yin J, Zhou X, Meng Q (2019) A framework for big data governance to advance RHINs: a case study of China. *IEEE Access* 7:50330–50338. <https://doi.org/10.1109/ACCESS.2019.2910838>
- Li S, Sato H, Sicular T (eds) (2013) *Rising inequality in China: challenges to a harmonious society*. Cambridge University Press, Cambridge
- Liang J (2018) Baidu chief under fire for privacy comments. *People’s Daily Online*. <https://en.people.cn/n3/2018/0328/c90000-9442509.html>. Accessed 01 May 2019
- Liao R (2019) Two former Qualcomm engineers are using AI to fix china’s healthcare problem. *TechCrunch*. https://techcrunch.com/2019/02/10/12-sigma-profile/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLnVnbS8&guce_referrer_cs=JrhrZU0nYiUY9D_0Z1SQW. Accessed 01 June 2019
- Liu C (2019) Multiple social credit systems in China. *Econ Sociol Europ Electron News* 21:1
- Lieutenant General Liu Guozhi, deputy to the National People’s Congress and director of the Military Commission’s Science and Technology Commission: Artificial intelligence will accelerate the process of military reform (2017) *Xinhua*. [人大代表、军委科技委主任刘国治中将:人工智能将加速军事变革进程. 新华.]. https://www.xinhuanet.com/mil/2017-03/08/c_129504550.htm. Accessed 05 May 2020
- Lucas L (2018) China’s artificial intelligence ambitions hit hurdles. *Financial Times*. <https://www.ft.com/content/8620933a-e0c5-11e8-a6e5-792428919cee>. Accessed 24 Mar 2019
- Lucas L (2019) China’s WeDoctor shelves overseas listing over data concerns. *Financial Times*. <https://www.ft.com/content/4993fbd8-91a4-11e9-b7ea-60e35ef678d2>. Accessed 24 Mar 2019
- Mann SP, Savulescu J, Sahakian BJ (2016) Facilitating the ethical use of health data for the benefit of society: electronic health records, consent and the duty of easy rescue. *Philos Trans R Soc A Math Phys Eng Sci*. <https://doi.org/10.1098/rsta.2016.0130>
- Marchione M (2019) Researchers Think They Can Use CRISPR Gene Editing Inside People to Cure a Form of Blindness. *Time*. <https://time.com/5635177/crispr-gene-editing-blindness/>. Accessed 01 Aug 2019
- Martina M, Blanchard B (2019) Rise in China’s defence budget to outpace economic growth target. *Reuters*. <https://uk.reuters.com/article/uk-china-parliament-defence-idUKKCN1QM036>. Accessed 11 Apr
- Martin-Laffon J, Kuntz M, Ricroch AE (2019) Worldwide CRISPR patent landscape shows strong geographical biases. *Nat Biotechnol* 37(6):613–620. <https://doi.org/10.1038/s41587-019-0138-7>

- McBride J, Chatzky A (2019) Is ‘Made in China 2025’ a Threat to Global Trade? Council on Foreign Relations. <https://www.cfr.org/backgrounder/made-china-2025-threat-global-trade>. Accessed 04 Apr 2019
- Mearsheimer JJ (2010) The gathering storm: China’s challenge to US power in Asia. *Chin J Int Politics* 3(4):381–396. <https://doi.org/10.1093/cjip/poq016>
- Meng J (2017) Principal contradiction facing Chinese society has evolved in new era: Xi. *Xinhua*. https://www.xinhuanet.com/english/2017-10/18/c_136688132.htm. Accessed 25 Apr 2019
- Meyer P (2020) Norms of responsible state behaviour in cyberspace. In: Christen M, Gordijn B, Loi M (eds) *The ethics of cybersecurity*. The International library of ethics, law and technology, vol 21. Springer, Cham
- Morris H (2019) China’s march to be the world’s first cashless society. *The Straits Times*. <https://www.straitstimes.com/asia/east-asia/>. Accessed 19 Mar 2020
- Mozur P (2017) Beijing Wants A.I. to Be Made in China by 2030. *The New York Times*. <https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html>. Accessed 25 Mar 2019
- Naughton B (2007) *The Chinese economy: transitions and growth*. MIT press, Cambridge
- Naughton B, Tsai KS (2015) *State capitalism, institutional adaptation, and the Chinese miracle*. Cambridge University Press, Cambridge
- Nittas V, Mütsch M, Ehrler F, Puhana MA (2018) Electronic patient-generated health data to facilitate prevention and health promotion: a scoping review protocol. *BMJ Open*. <https://doi.org/10.1136/bmjopen-2017-021245>
- Notice of the State Council on Printing and Distributing the Action Plan for Promoting Big Data Development (2015) State Council. [国务院关于印发促进大数据发展行动纲要的通知. 国务院]. https://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm. Accessed 02 June 2019
- Pasquale F (2015) *The black box society: the secret algorithms that control money and information*. Harvard University Press, Cambridge
- Pecotic A (2019) Whoever Predicts the Future Will Win the AI Arms Race. *Foreign Policy*. <https://foreignpolicy.com/2019/03/05/whoever-predicts-the-future-correctly-will-win-the-ai-arms-race-russia-china-united-states-artificial-intelligence-defense/>. Accessed 29 Feb 2020
- Pei M (2015) The Twilight of Communist Party Rule in China. *The American Interest*. <https://www.the-american-interest.com/2015/11/12/the-twilight-of-communist-party-rule-in-china/>. Accessed 24 Apr 2019
- Peng G, Yao Y (eds) (2005) *The science of military strategy*. Military Science Publishing House, Beijing
- Persson P, Zhuravskaya E (2016) The limits of career concerns in federalism. *J Eur Econ Assoc* 14(2):338–374. <https://doi.org/10.1111/jeea.12142>
- Ray A (2002) The Chinese economic miracle: lessons to be learnt. *Econ Polit Wkly* 37(37):3835–3848
- Ringgen S, Ngok K (2017) What kind of welfare state is emerging in China? In: Yi I (ed) *Towards universal health care in emerging economies*. Palgrave Macmillan, pp 213–237. https://doi.org/10.1057/978-1-137-53377-7_8
- Rozelle S, Huang J, Zhang L (1997) Poverty, population and environmental degradation in China. *Food Policy* 22(3):229–251. [https://doi.org/10.1016/S0306-9192\(97\)00011-0](https://doi.org/10.1016/S0306-9192(97)00011-0)
- Sacks S (2018) New China Data Privacy Standard Looks More Far-Reaching than GDPR. Center for Strategic and International Studies. <https://www.csis.org/analysis/new-china-data-privacy-standard-looks-more-far-reaching-gdpr>. Accessed 31 Mar 2019
- Sacks S, Laskai L (2019) China is having an unexpected privacy awakening. *Slate*. <https://slate.com/technology/2019/02/china-consu-mer-data-protection-privacy-surveillance.html>. Accessed 31 Mar 2019
- Sacks S, Li MK (2018) How Chinese Cybersecurity Standards Impact Doing Business in China. Center for Strategic and International Studies. <https://www.csis.org/analysis/how-chinese-cybersecurity-standards-impact-doing-business-china>. Accessed 31 Mar 2019
- Schneider M (2009) The nuclear doctrine and forces of the People’s Republic of China. *Comp Strat* 28(3):244–270. <https://doi.org/10.1080/01495930903025276>
- Selden M, You L (1997) The reform of social welfare in China. *World Dev* 25(10):1657–1668. [https://doi.org/10.1016/S0305-750X\(97\)00055-7](https://doi.org/10.1016/S0305-750X(97)00055-7)
- Sheehan M (2018) How China’s Massive AI plan actually works. *Marco Polo*. <https://macropolo.org/chinas-massive-ai-plan-actually-works/>. Accessed 04 Apr 2019
- Shoham Y, Perrault R, Brynjolfsson E, Clark J, Manyika K, Niebles JC, Lyons T, Etchemendy J, Grosz B, Bauer Z (2018) *The AI Index 2018 Annual Report*. AI Index Steering Committee, Human-Centered AI Initiative
- Si J (2019) These rules could save humanity from the threat of rogue AI. *World Economic Forum*. <https://www.weforum.org/agenda/2019/05/these-rules-could-save-humanity-from-the-threat-of-rogue-ai/>. Accessed 25 May 2019
- Simonite T (2019) How Health Care Data and Lax Rules Help China Prosper in AI. *Wired*. <https://www.wired.com/story/health-care-data-lax-rules-help-china-prosper-ai/>. Accessed 03 May 2019
- Sizing the prize: What’s the real value of AI for your business and how can you capitalise? (2017). *PwC*. <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>. Accessed 01 May 2019
- Song B (2018) The West may be wrong about China’s social credit system. *The Washington Post*. https://www.washingtonpost.com/news/worldpost/wp/2018/11/29/social-credit/?utm_term=.200cea3c05ad. Accessed 07 June 2019
- Sonnad N (2017) In China you now have to provide your real identity if you want to comment online. *Quartz*. <https://qz.com/1063073/in-china-you-now-have-to-provide-your-real-identity-if-you-want-to-comment-online>. Accessed 03 Mar 2020
- Stevens LA (2017) Public interest approach to data protection law: the meaning, value and utility of the public interest for research uses of data. *Edinburgh Research Archive*.
- Sun Y (2018) China’s citizens do care about their data privacy, actually. *MIT Technology Review*. <https://www.technologyreview.com/the-download/610708/chinas-citizens-do-care-about-their-data-privacy-actually/>. Accessed 31 March 2019
- Taddeo M (2012) Information warfare: a philosophical perspective. *Philos Technol* 25(1):105–120
- Taddeo M (2016) On the risks of relying on analogies to understand cyber conflicts. *Mind Mach* 26(4):317–321. <https://doi.org/10.1007/s11023-016-9408-z>
- Taddeo M, Floridi L (2018) Regulate artificial intelligence to avert cyber arms race. *Nature* 556:296–298. <https://doi.org/10.1038/d41586-018-04602-6>
- Tatlow DK (2015) A scientific ethical divide between China and West. *The New York Times*. <https://www.nytimes.com/2015/06/30/science/a-scientific-ethical-divide-between-china-and-west.html>. Accessed 23 Apr 2019
- The 13th Five Year Plan for Economic and Social Development of the People’s Republic of China (2016) Communist Party of China. (trans. Compilation and Translation Bureau). <https://en.ndrc.gov.cn/newsrelease/201612/P020161207645765233498.pdf>. Accessed 03 Apr 2019

- The Automation Readiness Index: Who is Ready for the Coming Wave of Automation? (2018) The Economist Intelligence Unit. <https://automationreadiness.eiu.com/>. Accessed 24 Apr 2019
- The Medical Futurist (2019) China is building the ultimate technological health paradise. Or is it? The Medical Futurist. <https://medicalfuturist.com/china-digital-health>. Accessed 01 Mar 2019
- The position paper submitted by the Chinese delegation to CCW 5th Review Conference (2016) [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/DD1551E60648CEBBC125808A005954FA/\\$file/China's+Position+Paper.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/DD1551E60648CEBBC125808A005954FA/$file/China's+Position+Paper.pdf). Accessed 01 Apr 2019
- UK Research and Innovation (2019) Accelerating detection of disease. <https://www.ukri.org/innovation/industrial-strategy-challenge-fund/accelerating-detection-of-disease/>. Accessed 02 Aug 2019
- UN seeks to retain human control over force (2018) Campaign to Stop Killer Robots. <https://www.stopkillerrobots.org/2018/08/uns/>. Accessed 24 Apr 2019
- Wang J, Liu S (2019) Ordering power under the party: a relational approach to law and politics in China. *Asian J Law Soc* 6(1):1–18. <https://doi.org/10.1017/als.2018.40>
- Wang Z (2019) Data integration of electronic medical record under administrative decentralization of medical insurance and healthcare in China: a case study. *Israel J Health Policy Res.* <https://doi.org/10.1186/s13584-019-0293-9>
- Webb A (2019) China and the AI Edge. *Nieman Reports.* <https://niemanreports.org/articles/china-and-the-ai-edge/>. Accessed 03 June 2019
- Triolo P, Webster G (2017) China's Plan to 'Lead'. In: Webster G, Creemers R, Triolo P, Kania E (eds) *AI: purpose, prospects, and problems.* New America. <https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/>. Accessed 27 Mar 2019
- What will be the net impact of AI and related technologies on jobs in China? (2018) PwC. <https://www.pwc.com/gx/en/issues/artificial-intelligence/impact-of-ai-on-jobs-in-china.pdf>. Accessed 15 May 2019
- Whiting SH (2017) Authoritarian "Rule of Law" and Regime Legitimacy. *Comp Polit Stud* 50(14):1907–1940. <https://doi.org/10.1177/0010414016688008>
- Wong L (2005) *Marginalization and Social Welfare in China.* Routledge, Abingdon
- Wong PH (2013) Confucian social media: an oxymoron. *Dao* 12(3):283–296
- Wu J, Mao Y (2017) Liberty in health care: a comparative study between Hong Kong and Mainland China. *J Med Philo* 42(6):690–719. <https://doi.org/10.1093/jmp/jhx026>
- Yan Y (2009) The Good Samaritan's new trouble: a study of the changing moral landscape in contemporary China. *Soc Anthropol* 17(1):9–24. <https://doi.org/10.1111/j.1469-8676.2008.00055.x>
- Yang H (2019) China—the privacy, data protection and cybersecurity law review—Edition 6. *The Law Reviews.* <https://thelawreviews.co.uk/edition/the-privacy-data-protection-and-cybersecurity-law-review-edition-6/1210009/china>. Accessed 9 Mar 2020
- Yang Y (2018) China's data privacy outcry fuels case for tighter rules. *Financial Times.* <https://www.ft.com/content/fdeaf22a-c09a-11e8-95b1-d36dfef1b89a>. Accessed 31 Mar 2019
- Yang Y, Liu X (2018) China's AI push raises fears over widespread job cuts. *Financial Times.* <https://www.ft.com/content/1e2db400-ac2d-11e8-94bd-cba20d67390c>. Accessed 01 Apr 2019
- Yu M, Du G (2019) Why are Chinese courts turning to AI? *The Diplomat.* <https://thediplomat.com/2019/01/why-are-chinese-courts-turning-to-ai/>. Accessed 13 Feb 2020
- Yuan S (2019) AI-assisted sentencing speeds up cases in judicial system. *China Daily.* https://www.chinadaily.com.cn/cndy/2019-04/18/content_37459601.htm. Accessed 13 Mar
- Zhang G, Yin K (2019) More updates on the China data protection regime in 2019. *IAPP.* <https://iapp.org/news/a/more-positive-progress-on-chinese-data-protection-regime-in-2019/>. Accessed 01 May 2019
- Zhang L, Wang H, Li Q, Zhao M-H, Zhan Q-M (2018) Big data and medical research in China. *BMJ.* <https://doi.org/10.1136/bmj.j5910>
- Zhang P (2019) China's top AI scientist drives development of ethical guidelines. *South China Morning Post.* <https://www.scmp.com/news/china/science/article/2181573/chinas-top-ai-scientist-drive-s-development-ethical-guidelines>. Accessed 12 Apr 2019
- Zhang P, Liang Y (2018) China's National Health Guiding Principles: a perspective worthy of healthcare reform. *Prim Health Care Res Dev* 19(1):99–104. <https://doi.org/10.1017/S1463423617000421>
- Zhang Y, Han W (2019) In Depth: China's Burgeoning Social Credit System Stirs Controversy. *Caixin.* <https://www.caixinglobal.com/2019-04-01/in-depth-chinas-burgeoning-social-credit-system-stirs-controversy-101399430.html>. Accessed 30 Apr 2019
- Zhang Z (2018) Scientist: China to focus on innovation to boost economy, lives. *China Daily.* <https://www.chinadaily.com.cn/a/201801/09/WS5a543bd5a31008cf16da5fa9.html>. Accessed 01 June 2019
- Zhao S (2015) A new model of big power relations? China–US strategic rivalry and balance of power in the Asia-Pacific. *J Contemp China* 24(93):377–397. <https://doi.org/10.1080/10670564.2014.953808>
- Zhu V (2019) China Trends #2—large but not strong: the challenges for China's Domestic Ports. *Institut Montaigne.* <https://www.institutmontaigne.org/en/blog/china-trends-2-large-not-strong-challenges-chinas-domestic-ports>. Accessed 13 Mar 2019

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