

LOFTY PRINCIPLES, CONFLICTING INTERESTS

AI ethics and governance in China

Rebecca Arcesati

KEY FINDINGS

- **Rather than being driven entirely from the top**, China's AI ethics and governance landscape is shaped by multiple actors and their varied approaches, ranging from central and local governments to private companies, academia and the public. China's regulatory approach to AI will emerge from the complex interactions of these stakeholders and their diverse interests.
- **Despite notable advances in tackling ethics issues in specific AI sectors and application areas**, a large gap remains between defining broad ethical principles and norms to guide AI development and putting these into practice through standards, laws and government or corporate regulation.
- **This gap is not unique to China**, but particularly pronounced in the Chinese context since AI is seen as a core means for fully achieving the governance vision of the Chinese Communist Party, which prioritizes state control and political security over individual rights. Genuine concern for AI ethics coexists with Beijing's use of AI for mass surveillance and ethnic profiling.
- **Given China's rapid AI advancements**, its expanding presence in global standards bodies and Chinese tech companies' growing global reach, it will be critical for the EU to engage with Chinese actors. However, European policymakers must take the government's rhetoric on AI ethics with great caution and push back against China's use (and export) of AI for surveillance and other applications that threaten human rights and fundamental freedoms.



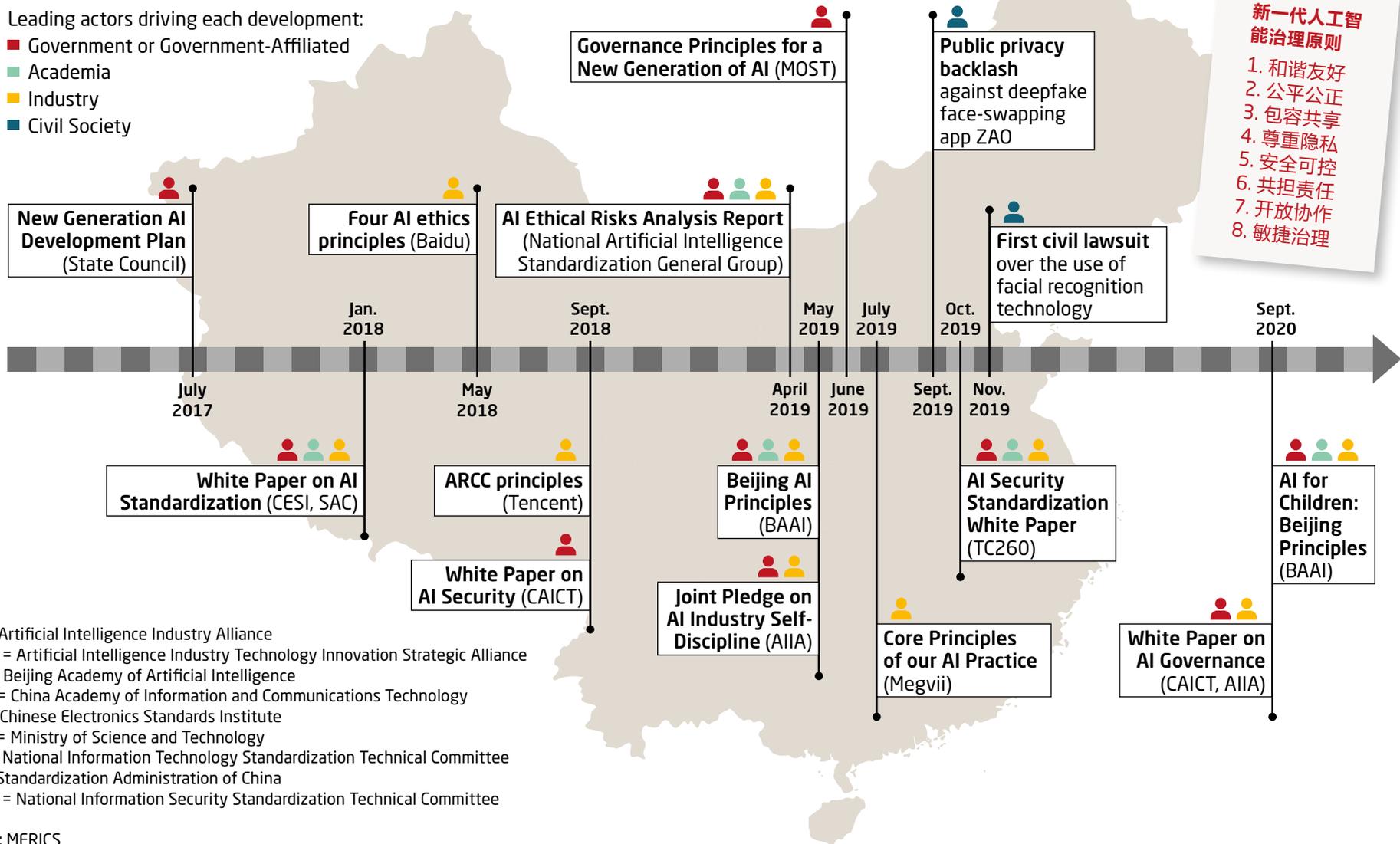
Various Chinese actors have tackled AI ethics issues

A timeline of seminal developments since 2017



Leading actors driving each development:

- Government or Government-Affiliated
- Academia
- Industry
- Civil Society



新一代人工智能治理原则

1. 和谐友好
2. 公平公正
3. 包容共享
4. 尊重隐私
5. 安全可控
6. 共担责任
7. 开放协作
8. 敏捷治理

AIIA = Artificial Intelligence Industry Alliance
 AITISA = Artificial Intelligence Industry Technology Innovation Strategic Alliance
 BAAI = Beijing Academy of Artificial Intelligence
 CAICT = China Academy of Information and Communications Technology
 CESI = Chinese Electronics Standards Institute
 MOST = Ministry of Science and Technology
 NITS = National Information Technology Standardization Technical Committee
 SAC = Standardization Administration of China
 TC260 = National Information Security Standardization Technical Committee

Source: MERICS

© MERICS

1. INTRODUCTION

Countries around the world are harnessing the transformative impact of artificial intelligence (AI) on their economies and societies. There has been much focus on the competition and rivalry between countries with advanced AI research and development (R&D) capabilities, with talk of an “AI race” between the United States and China – and to a lesser extent Europe. However, the ethical and safety risks of not getting AI right are as great as its beneficial potential. From facial recognition and recruitment algorithms carrying biases to self-driving cars endangering lives, the challenges associated with AI governance failures are enormous and require joint solutions.

Artificial Intelligence (AI) refers to both a scientific field and a broad suite of technologies that accomplish tasks generally believed to require human intelligence, such as making decisions through the collection, processing and interpretation of data. The EU Commission defines AI as “systems that display intelligent behavior by analyzing their environment and taking actions - with some degree of autonomy - to achieve specific goals.”

Under the umbrella term “AI ethics,” experts are discussing questions such as what role AI systems should play in our societies, what risks they involve and how we should control them. In recent years, professional associations, companies, governments and international organizations have published a plethora of AI ethics principles and guidelines. Several European countries and organizations have played a pivotal role in these efforts, with the EU strongly advocating for the development of risk frameworks and legislation to ensure “trustworthy” AI, cemented in April 2021 in a proposal for the world’s first dedicated AI regulations.

Understanding Chinese approaches to AI ethics and governance is vitally important for European stakeholders. China will be a fundamental force in shaping the trajectory of AI innovation and adoption as well as the way in which AI will be governed. It has embraced AI and aims to become the world’s primary AI innovation center by 2030. Chinese policymakers are paying increasing attention to ethics in the context of AI governance, having issued multiple related principles. Behind such initiatives there is a web of public and private players, interests and voices.

China will be a fundamental force in shaping the trajectory of AI innovation and adoption as well as the way in which AI will be governed

This MERICS Monitor provides an analysis of China's emerging AI ethics and governance landscape. **It examines three issues:**

- The various approaches to AI ethics taken by government, corporate, academic and civil society actors in China
- Ethical issues related to specific applications (healthcare, autonomous driving and public security) and how they are being addressed
- China’s role in global AI ethics and governance efforts and its implications for European stakeholders

2. BEIJING'S STRATEGIC CONSIDERATIONS AND APPROACH TO AI ETHICS AND GOVERNANCE

The government's ambition to lead the world in AI is accompanied by its growing attention to the technology's governance. In 2018, President Xi Jinping called for the "healthy development" of AI through the establishment of laws, ethics, institutional mechanisms and regulations.¹ In the leadership's view, researching and preventing the short-term risks, such as privacy and intellectual property infringements, and long-term challenges AI systems could pose to the economy, social stability and national security, such as unemployment and changes to social ethics, is of utmost importance.

2.1 China's policymakers pay increasing attention to ethics in the context of AI governance

Starting with the publication in 2017 of the State Council's New Generation Artificial Intelligence Development Plan (AIDP), the government expressed its intention to tackle ethical issues arising from AI systems. The plan states that by 2025 China will set up an initial system of laws, regulations, ethical norms and policies as well as a security assessment framework to "ensure the safe, reliable and controllable development of AI." A comprehensive system should be established by 2030. The AIDP calls for strengthening research on legal, ethical and social issues. It also urges measures like an ethical framework for human-machine collaboration and codes of conduct for personnel in AI product R&D and design.²

The Chinese government has expressed its intention to tackle ethical issues arising from AI systems

Since then, several principles and white papers have been issued to guide AI governance (see Exhibit 1). In a 2018 AI Standardization White Paper, the Chinese Electronics Standards Institute (CESI) recommended three overarching ethical considerations for AI: "human interest," "liability" and "consistency of rights and responsibilities." The document discusses safety, ethical and privacy issues and reflects the government's wish to use technical standardization as a tool in domestic and global AI governance efforts.³

In 2019, the Ministry of Science and Technology (MOST) issued the Governance Principles for a New Generation of AI, which put forward eight principles for developing "responsible AI."⁴ Drafted by a dedicated expert group, the Governance Principles are the most official formulation of China's approach to AI ethics to date.

Understanding the terms the government uses is necessary to gauge its vision for AI governance. Reference to "human rights" in the Governance Principles does not imply endorsement of liberal democratic values, while "societal security" implies maintaining stability by prioritizing collective wellbeing, as defined by the Chinese Communist Party (CCP), over individual freedoms. Additionally, the concept of human-machine harmony, read alongside the AIDP's call for strengthening "public opinion guidance," may indicate the intent to prepare society for greater data-driven monitoring and governance *through* AI.

2.2 The government directs a multi-stakeholder conversation on AI ethics

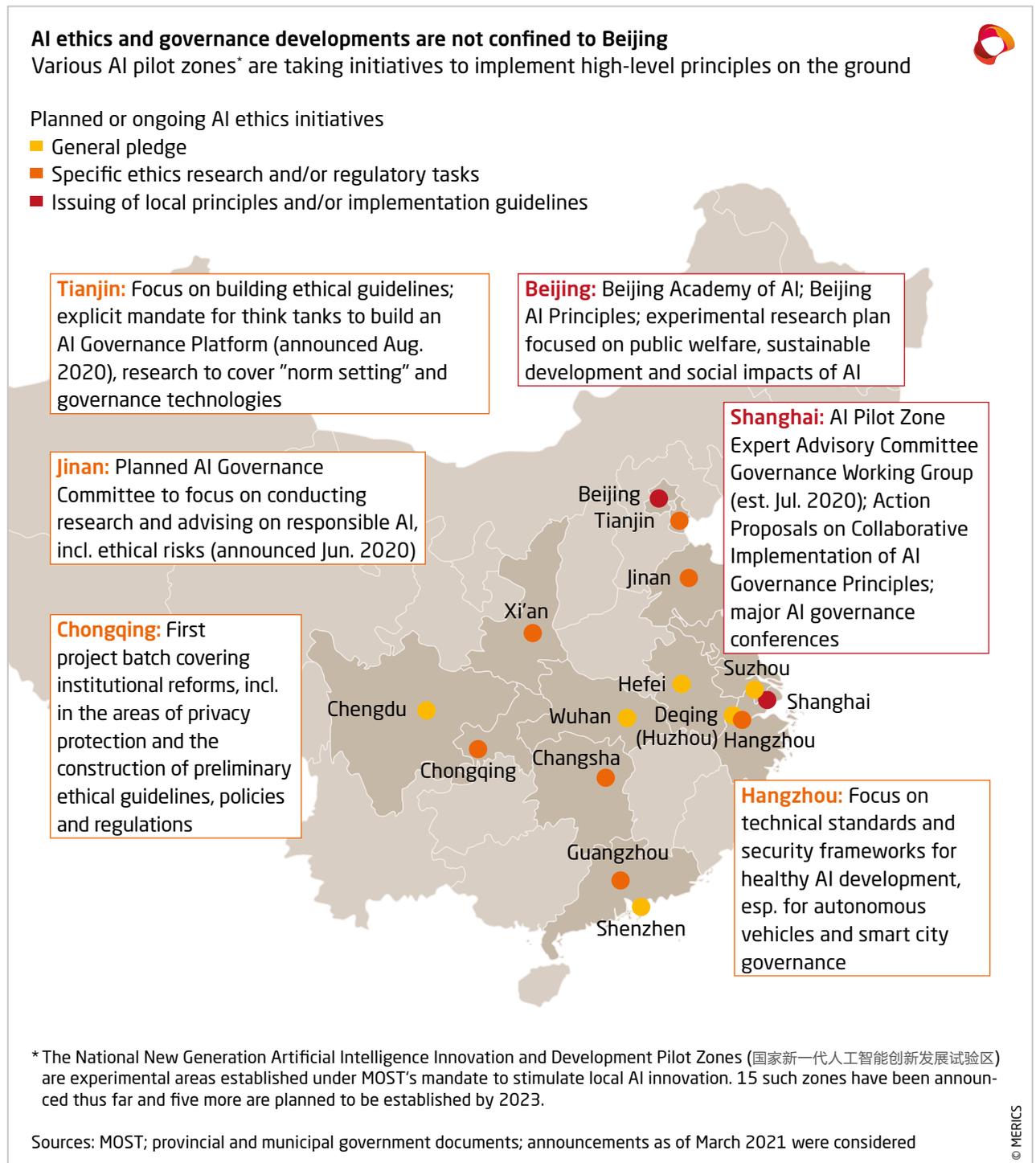
While the debate on AI ethics is overseen by Beijing and takes place within the strict limits of the party-state's goals and interests, it is a multi-stakeholder conversation.

MOST's AI governance committee comprises experts from leading universities, the Chinese Academy of Sciences and private AI companies. The Beijing AI Principles, a key document that

preceded the Governance Principles, also resulted from a deliberation involving universities and companies under the leadership of the Beijing Academy of Artificial Intelligence (BAAI), a leading AI research institute backed by MOST and the Beijing municipal government.⁵

The third seminal set of principles, the Joint Pledge on AI Industry Self-Discipline, similarly emerged from a consultation between different players.⁶ Its process was launched by the Artificial Intelligence Industry Alliance, an association of universities and tech firms led by the China Academy of Information and Communications Technology (CAICT) of the Ministry of Industry and Information Technology (MIIT), the top government-affiliated think tank for tech policy issues.

Exhibit 2



A central feature of all these discussions is their applied approach. To drive implementation at the local level, MOST is encouraging municipal governments to step up relevant work in AI pilot zones (see Exhibit 2). Additionally, both the Beijing AI Principles and the Joint Pledge focus on applicable and action-oriented goals and measures to ensure that the trajectory of AI development throughout the lifecycle of systems, from R&D to commercialization, is beneficial for society.⁷

2.3 Safeguarding stability is a key objective of China's AI strategy

The government's rhetoric and attention to ethics can appear hypocritical given its use of AI for mass surveillance, repression and ethnic profiling (see section 4.3). However, from the perspective of China's leadership and of its moral and ethical frameworks, this poses no contradiction. National security and stability are the highest collective goods, taking priority over personal privacy, transparency, accountability and individual human rights.

The CCP sees security and stability as preconditions as well as products of economic development, a key objective of China's AI strategy. A major goal of the AIDP is the modernization of social governance, which entails not only the optimized provision of public services but also the construction of a modernized socialist society through, for example, the use of AI to "grasp group cognition and psychological changes."⁸

Additionally, the party justifies its control over the legal system by arguing for the need to ward off threats from internal and external enemies to meet the superior goal of preserving political security.⁹ Thus, from the CCP's perspective, the use of AI against a part of the population it sees as a terrorist threat to society, as is the case with Uighurs in Xinjiang, can coexist with efforts to ensure that AI systems do not cause harm to the majority.

Ethical questions about algorithmic decision-making are framed around the interests of the collective – of which the party-state claims to be the sole legitimate representative – rather than the individual.¹⁰ This logic also explains why the emerging data protection regime aims to impose restrictions on companies' ability to collect personal information but leaves the government with nearly unrestrained power to harvest and use citizens' data for public security and law enforcement.¹¹

Ethical questions about algorithmic decision-making are framed around the interests of the collective rather than the individual

3. HOW INDUSTRY, ACADEMIA AND CIVIL SOCIETY DRIVE FORWARD ETHICAL AI

3.1 Industry plays a pivotal role in shaping Chinese discussions

China's leadership sees industry as a key driver in coordinating self-regulation, research and education on AI ethics, though regulators ultimately set governance rules. It has highlighted the importance of corporate self-regulation, with a recent white paper published by the CAICT identifying companies as the main AI governance entities in the near-term.¹² Many leading tech companies and startups have issued calls to address governance and ethics issues related to the development and commercialization of AI applications. They are also joining multi-stakeholder efforts to develop ethics principles and industry standards for responsible AI development, while initiating their own research and principles to tackle ethics issues.

Many companies were directly or indirectly involved in each of China's three seminal AI documents, of which the joint pledge is an industry commitment to self-regulation. The seven members of MOST's AI governance committee, for example, include two executives

from e-commerce giant JD.com and facial recognition unicorn Megvii, demonstrating that companies are directly involved in the formulation of policy recommendations and guiding documents such as the Governance Principles.

Tech giants and AI startups are founding members of the previously mentioned BAAI and other key industry alliances behind AI principles and white papers.¹³ Baidu and Tencent have also submitted proposals on AI ethics directly to China's leadership.¹⁴ Many companies are meanwhile active participants in domestic standard-setting activities related to AI.¹⁵

Corporate self-regulation has thus far primarily taken the shape of high-level ethics codes. Most notably, Baidu, Tencent and Megvii have issued documents that put forward ethics principles to guide their own and the industry's development of AI. All three highlight similar notions such as the importance of technical robustness and safety, human oversight, data privacy and accountability. Tencent's AI principles are the most detailed principles developed by a Chinese company so far. Issued in 2018, they urge for AI to be available, reliable, comprehensible and controllable, and highlight specific issues such as algorithmic transparency.¹⁶

Tencent's AI principles urge for AI to be available, reliable, comprehensible and controllable

Companies also conduct extensive research into governance and ethics issues through dedicated departments. Their research, much of which predates the government's increased attention to AI ethics, ranges from techniques for preserving privacy in machine learning to methods for protecting against adversarial attacks on deep-learning systems.¹⁷

CEOs and AI executives also advocate for interdisciplinary exchanges and collaborative action on AI ethics, while positioning themselves as thought leaders on AI governance issues at key industry forums such as Shanghai's annual World AI Conference. Some also raise public awareness of the risks of AI applications in everyday life, through campaigns such as AI for Good.¹⁸

While many tech companies and AI clearly recognize the importance of governing the societal and ethical impact of AI, few have institutionalized steps that turn high-level commitments into concrete procedures. Their AI ethics research and principles, while representing good-faith intentions, mostly lack concrete implementation measures that address the specific issues they identify, from algorithmic fairness to data privacy.

Megvii is one of the few companies to create internal structures such as an AI Ethics Committee to oversee the implementation of its AI principles. This committee is said to make recommendations to the board based on internal investigations and a whistleblowing procedure. However, one listed international member says he never joined the committee and it remains unclear what kinds of changes – if any – it has effected.¹⁹

It seems logical for companies to be at the forefront of identifying and addressing the harmful impacts of AI applications, given that they research, develop and deploy AI in real-life situations. They are also incentivized to anticipate and address the risks of their AI products and services to avoid backlash from regulators or the general public.

However, for now it is still unclear whether corporate AI ethics declarations are leading to meaningful changes in internal research and development processes, or whether they are ultimately empty commitments that serve only to enhance companies' reputation. Companies are also commonly reluctant to implement potentially costly and time-intensive mechanisms to ensure safe and ethical AI products.

The close relationship of tech and AI companies with the government adds an additional layer of complication since the government not only provides extensive policy support but is also often a major client for corporates. Companies' pledges on AI ethics thus often stand in stark contrast to their sale of AI products such as facial recognition or ethnic minority analytics tools to the public security apparatus (see section 4.3).

3.2 Chinese academic research also shapes AI ethics discussions

Academic research on the social and ethical implications of AI is increasingly informing discussions about AI governance in China. A review of relevant publications since 2017 reveals that although research efforts approach the issue from various angles, most are still limited to conceptualizing the changes brought about by AI and suggesting normative and regulatory frameworks. Critical research on specific applications is mostly lacking, although there are notable exceptions.²⁰

Ethics research is conducted through state-sponsored projects and individual scholars' initiatives. China's two leading research institutes under the aegis of the State Council – the Chinese Academy of Sciences (CAS) and the Chinese Academy of Social Sciences (CASS) – undertake relevant work, some of which is sponsored by China's largest public research fund for social sciences, the National Social Science Fund of China. One project led by the Institute of Automation at CAS explores issues like the relationship between humans and AI and challenges associated with determining liability. CAS-sponsored researchers also apply social science research to practical problems, such as social ethics issues caused by the introduction of robots into families.²¹

Several prominent scholars are particularly influential in driving forward ethics research. At CASS, Duan Weiwen (段伟文) – one of China's most prominent thinkers on philosophical, ethical and social issues surrounding AI and Big Data – leads a Science, Technology and Society Research Center. Duan frequently emphasizes that innovation runs faster than ethics, which requires targeted work to tackle ethical risks in specific technology application scenarios rather than abstract prescriptions. He also advocates for public participation and oversight in ethics matters.²²

Some researchers approach AI ethics from the perspective of traditional Chinese philosophy. CAS-affiliated Zeng Yi (曾毅) spearheaded the formulation of Harmonious Artificial Intelligence Principles, which are based on the concept of “harmony” in Chinese philosophy. These principles emphasize harmony between humans and machines, a concept that is also present in the Beijing AI Principles, and advocates for a positive symbiosis between the two. In addition to playing a leading role in drafting several seminal documents mentioned in this report, Zeng drives major applied ethics research efforts in areas like brain-inspired neural network architectures.²³

Renmin University's Guo Rui (郭锐), another prominent scholar and government advisor, focuses on translating ethical guidelines into an actionable governance system. Guo has advocated for companies to set up ethics committees, and in his latest book examines the ethical risks of specific AI applications, from precision marketing and content recommendation algorithms to sex robots and smart courts.²⁴

Chinese academia seems to be gaining influence in official government efforts to govern AI

Chinese academia actively engages in global exchanges on AI ethics. This aligns with the government's call to increase the country's “discourse power” (话语权) in the field. A prominent example of the interplay between scholarly exchanges and the state' soft

power ambitions is the Berggruen China Center, established by Peking University and the Berggruen Institute in 2018 with the stated goal of engaging Chinese thinkers to “examine, share and develop ideas to address global challenges.”²⁵ AI ethics is one of the center’s main research areas. Additionally, in 2020 Tsinghua University established the Institute of Artificial Intelligence Global Governance to “actively contribute Chinese wisdom” and shape the field.²⁶

While promoting official Chinese global governance concepts is an important goal behind these initiatives, it would be wrong to view all academic research and collaborations as being driven by the state’s aims. The diverse range of individual research initiatives reflects scholars’ genuine aspiration to make AI beneficial for mankind, as well as to overcome political tensions and cultural barriers between China and the West to advance cooperation. Xue Lan (薛澜), the Director of Tsinghua’s abovementioned institute, has warned that geopolitical tensions between China and the United States are having a chilling effect on industry and policy exchanges in the AI field, which may hinder cooperation on global AI governance.²⁷

The BAAI has emerged as China’s leading AI research institute and a hub for multi-stakeholder and international collaboration. The institute has a research center, led by Zeng, which is dedicated to investigating AI ethics, governance and solutions for sustainable development. To foster international dialogue, a recently published study by BAAI and researchers at Cambridge University urges academia to play a greater role in overcoming cultural barriers to collaboration on AI ethics and governance.²⁸

Chinese academia seems to be gaining influence in official government efforts to govern AI. Xue and Zeng, for instance, are also members of MOST’s AI Governance Committee. Yet it remains to be seen to what extent scholars will be able to directly influence government policy, corporate practices and regulation towards higher ethical standards.

3.3 Public pushback on AI risks has led to some regulatory changes

While, generally, the public is not seen as the decisive force in China’s AI development, Chinese citizens are pushing for ethical constraints on some use cases. Despite the common perception in the West that Chinese people are particularly trusting of new technologies, there is growing awareness, debate and occasionally pushback related to the risks of AI. In some cases, this has led to policy changes and corporate self-regulation.

Chinese consumers care about the protection of their personal information. When in 2018 Baidu’s CEO Robin Li said Chinese people were less sensitive about privacy and more willing to trade it for convenience, he faced intense opposition on social media. During the Covid-19 outbreak, the use of monitoring apps that collect health information and location data also provoked public criticism due to concerns over discrimination and the erosion of privacy.²⁹

In recent years, consumer backlash has played a key role in holding Chinese tech companies accountable for data privacy violations and spurring on regulators to create more stringent regulations. The resulting data regime, which notably includes a Personal Information Security Specification and the soon-to-be finalized Personal Information Protection Law, imposes wide-ranging restrictions on companies’ ability to handle personal information.³⁰

The prevalence of AI-powered surveillance technologies also worries citizens. In one 2019 survey, over 70 percent of respondents expressed concerns over privacy violations in the rollout of facial recognition systems.³¹

Consumer backlash holds tech companies accountable for data privacy violations and spurs on authorities to create more stringent regulations

A series of civil lawsuits have also drawn attention to citizens' privacy concerns with China's growing use of facial recognition in public spaces. The first was filed in 2019 against a wildlife park in Hangzhou for introducing a facial recognition access-control system. The individual won the landmark case, which sparked intense online discussions about the excessive collection of facial data. Several other high-profile cases followed. Meanwhile, prominent scholars have recommended the use of tech measures and processes such as ethics by design in addition to regulation to ensure the responsible use of biometric recognition.³²

It seems no coincidence that facial recognition regulation has received increasing attention from China's top lawmakers. Both the recently enacted Civil Code and the abovementioned personal information regulations tighten restrictions over biometric data collection. More recently, standard-setting authorities released a draft of a dedicated national data security standards for facial recognition data.³³ Several cities have also introduced or are considering regulations to restrict the use of facial recognition and are fining companies for data privacy infringements.³⁴

Another notable public backlash unfolded in response to the rise of deepfakes – false or altered images, videos and audio generated using AI. In August 2019, the release of the face-swapping app ZAO caused almost-immediate outrage among users over data privacy and copyright infringements.³⁵ Regulators responded swiftly, demanding that the app's parent company take corrective action. WeChat subsequently restricted some access to the app on its messaging platform, citing security risks.

Authorities soon thereafter took initial steps to regulate the use and spread of deepfakes. One policy document – released apparently in reaction to the ZAO incident – requires online information service providers to review and label any audiovisual content that is produced using new technologies such as deep learning.³⁶ The rules also prohibit the use of deep learning to create, publish or transmit fake news.

While these examples show that civil society influence government regulations and corporate actions related to AI ethics to some degree, their ability to do so is ultimately constrained by China's political system. China's data protection regime, for example, leaves the government with unrestrained power to harvest and use citizens' data for its intrusive public security and law enforcement activities.³⁷

4. AI ETHICS AND GOVERNANCE IN SPECIFIC SECTORS

Government agencies, tech companies, scholars and to some extent the public actively explore AI ethics and governance issues in China. Some of the resulting principles and research bear strong similarities to international efforts. This section examines whether and how these are put into practice by delving into specific ethics and governance issues confronted in relation to AI applications in healthcare, autonomous driving and public security. This shows that the government and companies have a long way to go in turning their normative discourse into practical governance mechanisms.

4.1 Healthcare

China's government has prioritized healthcare as an area for boosting AI applications to alleviate some of the pressure on its strained system. Chinese companies have made headway

Government and companies have a long way to go in turning their normative discourse into practical governance mechanisms

in introducing AI in healthcare, with many applying AI technologies to disease control and prevention as well as medication and vaccine development during the Covid-19 outbreak.³⁸

Nevertheless, there are numerous legal and ethical concerns surrounding medical AI, such as the data security of health IT systems. According to the Committee of Health Information Security and New Technology Application, 60 percent of the reported data breaches that happened in China in 2017–2018 occurred in the healthcare industry.³⁹ Scholars have pointed out that China’s approach to medical AI ethics puts more emphasis on public health than individual wellbeing, as exemplified by the mass collection of citizens’ health data at the cost of individual privacy.⁴⁰ It also remains to be seen whether AI will improve access to healthcare and social equity or exacerbate inequalities by improving the quality of care only for the rich.

Chinese research investigating the legal and ethical challenges brought by medical AI has skyrocketed since 2019.⁴¹ Researchers have examined a variety of issues such as the legal liability of AI systems when errors occur, the transformation of health workers’ role, data protection and privacy and ethical decision-making in using autonomous systems.

One predominant approach to regulating AI in healthcare is to apply existing risk assessment frameworks for medical equipment. The authorities’ focus has been mainly on setting up industry standards and equipment-assessment procedures to ensure data security and safety in each application. For example, Chinese regulators have established evaluation and certification measures for medical products and services that incorporate big data and issued assessment criteria for medical devices that incorporate deep learning.⁴²

One predominant approach to regulating AI in healthcare is to apply existing risk assessment frameworks for medical equipment

AI regulation appears to be more advanced in healthcare than in other sectors, with a dozen policies, standards and assessment criteria issued in recent years by government entities ranging from the State Council to healthcare regulators such as the National Health Commission and the National Medical Products Administration.

However, the government’s focus on data security and its approach to treating medical AI as medical equipment leave unaddressed many ethical issues, such as bias embedded in AI systems. Algorithmic bias refers to outcomes that are systematically less favorable to a particular social group due to unrepresentative, incomplete or flawed training data. Moreover, the ethical review process in the healthcare system, which allows the examination of individual AI use cases, also appears to be underdeveloped: since the developers of most AI systems used in the sector are not healthcare companies but tech firms, they are not under the purview of existing ethical review committees. Meanwhile, in several provinces and cities these committees lack technical AI expertise.⁴³

4.2 Autonomous driving

Autonomous driving is one field of AI application where China is widely believed to be on its way to taking the global lead. In 2020, pilots for Chinese self-driving vehicles advanced rapidly while the United States suspended testing on pandemic-related grounds. However, advances in autonomous driving – fueled by the government’s ambition for China to become a global leader in smart-car development by 2035 – have been accompanied and at times moderated by concerns about ensuring the safety of intelligent connected vehicles (ICV).

A small number of prominent fatal accidents caused by automated driving systems in China and abroad have caused widespread public discussions that highlighted to government

and industry players the need to pay closer attention to safety. Though the government is still pushing the commercialization of ICVs, safety concerns seem to have taken on a greater importance.

For example, the National Development and Reform Commission, China's industrial planner, downgraded previous ambitious market-share targets for autonomous vehicles in its finalized development strategy for intelligent vehicles released in February 2020.⁴⁴ Regulators aim to complete a basic standards system for autonomous vehicles by 2025, including a standards framework for driver-assistance functions and low-level autonomous driving that was scheduled for completion by the end of 2020. Since 2017, China has also rolled out regional standards for road-testing autonomous vehicles, alongside national road-testing rules for ICVs.⁴⁵

Advances in autonomous driving have been accompanied and at times moderated by safety concerns

Chinese auto and tech companies are actively involved in domestic and international efforts to improve the safety of autonomous vehicles. Baidu – the designated AI national champion for autonomous driving – plays a key part in leading industry discussions through its Apollo open autonomous driving platform and technical research into safety frameworks. Baidu was also one of 11 major international companies that published a white paper proposing a framework for implementing existing auto industry safety standards in autonomous driving.⁴⁶

Governance efforts have also tackled information and algorithmic security. These feature strong involvement from and cooperation between government-affiliated research institutes, industry players and academia. Industry players and research institutes jointly issued China's first technical specification for ICVs, which proposes evaluations for different information-security dimensions including data, network and hardware.⁴⁷ The effort was spearheaded by Baidu's Apollo Cyber Security Lab, which has been a key promoter of the implementation of information-security standards for ICVs in China.

Another research center affiliated with the MIIT has proposed a data-security system that classifies different types of autonomous-driving data and their associated risks.⁴⁸ Baidu has also tackled issues related to algorithmic security, with the company's Security X-Lab conducting extensive research on adversarial attacks that pose threats to deep-learning models in autonomous driving.⁴⁹

4.3 AI-enabled surveillance and the targeting of ethnic minorities

AI plays a pivotal role in the CCP's vision of data-driven governance and control. The AIDP calls for the "intelligentization of social governance" and the enhancement of public safety and security capabilities through "intelligent monitoring."⁵⁰ A 2018 White Paper by the China Academy of Information and Communications Technology (CAICT) recommends that AI be more vigorously integrated with public-security work to improve social governance by automating tasks like censorship, "smart" security and public-opinion monitoring.⁵¹ The use of AI in public security in China has grown explosively, amid an effort to apply it to mass surveillance.⁵²

In Xinjiang, where at least one million Muslims are being held in detention camps, technology complements conventional police methods in the repression of Uighurs and other ethnic minorities. While not all surveillance methods deployed there involve advanced technologies and decisions to designate people as terrorists are still made by humans, AI plays a big role.⁵³ AI companies profit from government demand for high-tech surveillance applications and tailor products and solutions to the needs of public security.⁵⁴



Megvii's contradictory actions on AI ethics

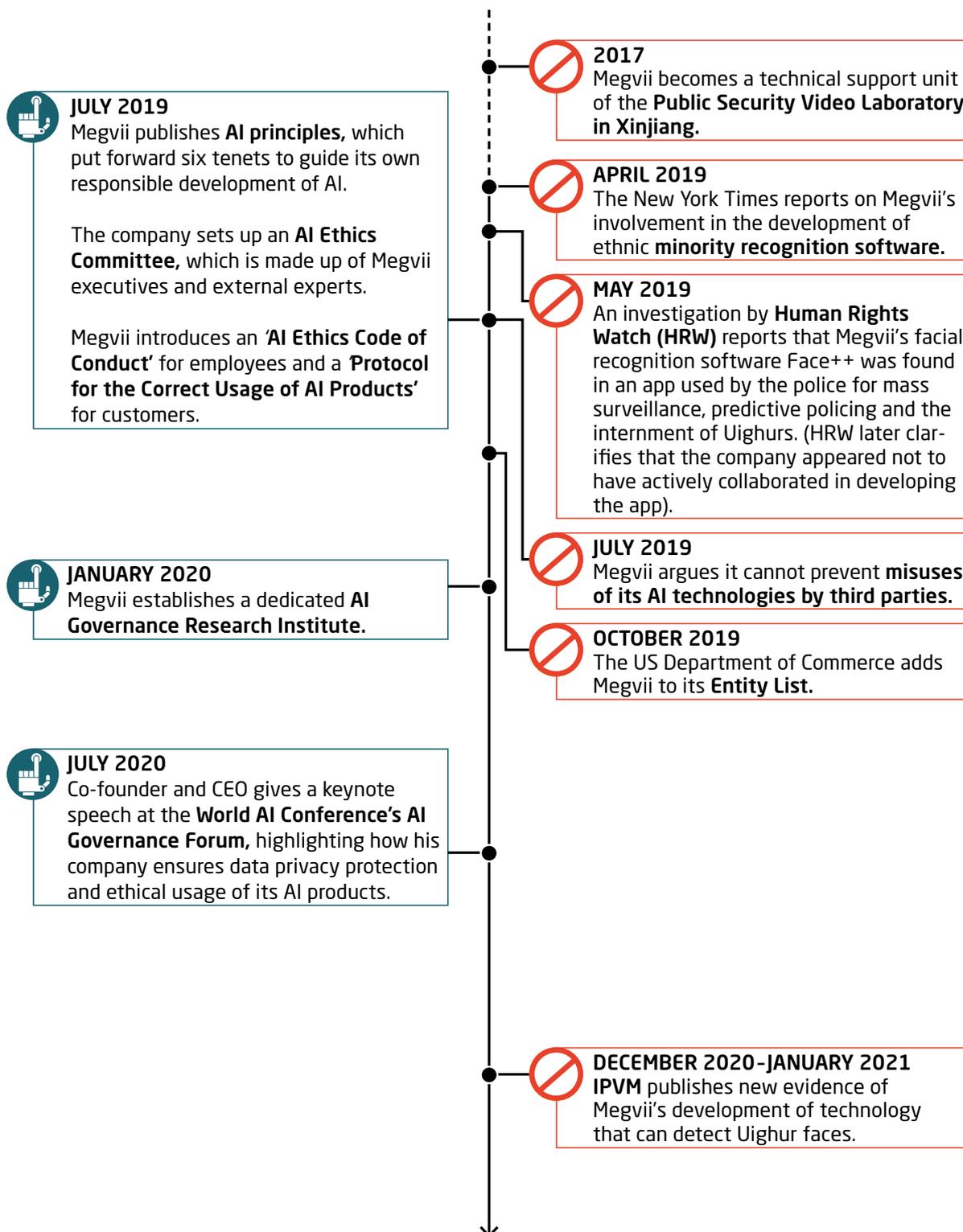
AI unicorn spearheads self-regulation efforts on ethical AI while developing minority recognition software



AI ethics champion



Human rights entanglements



Sources: MERICs; Human Rights Watch; IPVM; The New York Times

© MERICs

Algorithms designed to perform racial profiling are particularly troubling. The aim is to further the government’s goal of monitoring, tracking and controlling minorities, and to support predictive policing and extra-judicial confinement in detention camps. Yitu, Megvii, SenseTime and CloudWalk, among others, have been associated with the development of minority-tracking machine-learning software.⁵⁵ iFlytek has provided speech recognition technology to help the police monitor communications in minority languages.⁵⁶

The ethical issue goes much beyond algorithmic bias: the government weaponizes AI against specific ethnic groups. Across China, several video-surveillance projects have mandated Uighur-detection capabilities. Even Alibaba has developed facial recognition software that can specifically detect Uighurs and other ethnic minorities in images and videos.⁵⁷

The government weaponizes AI against specific ethnic groups

The government’s approach to AI ethics will continue to reflect priorities that, though conflicting from a liberal democratic perspective, are aligned in China’s political system. These are making AI beneficial for society, addressing citizens’ privacy concerns and positioning China as a constructive player in global AI governance while harnessing AI to safeguard political security. Companies are bound by these considerations, no matter whether they align with their own visions for how technology should be developed and used.

More importantly, the growing importance of AI in the surveillance state creates conflicting incentives for private companies. Not only is their ability to meaningfully push back against the party-state’s interests limited, but they also have political, financial and reputational incentives to develop and supply AI products tailored to the system’s needs (see Exhibit 3).

Many companies do not even seem to view the use of AI in the party-state’s public security work as troubling.⁵⁸ Ultimately, Beijing’s security-centric approach risks leading to the prioritization of political goals over ethical considerations in how AI research, development and adoption will take shape in China in the years to come.

5. IMPLICATIONS FOR EUROPE: AI RESEARCH COLLABORATIONS WITH CHINESE ACTORS SHOULD BE GUIDED BY ROBUST ETHICAL STANDARDS

China wants to play an driving role in global AI ethics and governance activities. MOST’s Governance Principles emphasize “open collaboration” across borders and call for “a broad consensus on an international AI governance framework, standards and norms.” The AIDP expresses China’s ambition to drive such a consensus and strengthen international research into common AI challenges.

China’s participation in intergovernmental AI governance efforts is still limited. Some of the highest-profile multilateral efforts lack Chinese participation, which is unsurprising given their emphasis on democratic values and human rights. These include the Global Partnership on AI and the OECD Principles on AI.⁵⁹

Nevertheless, Chinese actors’ participation in international initiatives is growing. As a member of the G20, China signed on to the group’s non-binding AI principles, drawn from the OECD principles.⁶⁰ At UNESCO, Chinese experts spearheaded a consensus on AI and education and contributed to an expert group drafting AI ethics recommendations.⁶¹

Chinese actors are also driving global AI standard-setting. In 2019, Beijing hosted the first meeting of an influential AI standardization committee and Chinese companies are

involved in ethics-related projects at leading international standards bodies.⁶² In the academic sphere, Chinese scholars conduct research with international counterparts on issues like the use of facial recognition during the Covid-19 pandemic.⁶³ Researchers at tech companies also engage in international research projects on technical challenges.⁶⁴

China's eagerness to engage in global AI ethics and governance initiatives could provide opportunities for the EU, where ethics is a cornerstone of the EU's "human-centric" approach to AI.⁶⁵ In April 2021, the European Commission proposed an AI regulation that introduces a regulatory structure centered on a risk-based classification of AI systems.⁶⁶ Working with other countries on standards and regulations will be crucial to ensuring a beneficial AI future. And given China's rapid advances in AI applications and the growing global reach of its companies, it will be critical for the EU to engage with Chinese actors.

Broad similarities between European and Chinese ethical interests and standards could pave the way for constructive collaboration. Some key tenets of MOST's Governance Principles closely resemble EU guidelines, indicating that Chinese researchers are already looking to European insights and recommendations on AI.⁶⁷ Stakeholders on both sides have shared interests in many areas, such as managing AI's impact on sustainability or developing aligned safety frameworks for self-driving cars. In pursuing cooperation, European actors should be receptive to the multitude of ethical concerns and solutions identified by Chinese lawmakers, academics and citizens.

However, clear differences exist at the governance level that create substantive challenges and barriers to cooperation. The Chinese party-state's prioritization of social stability and political security in its AI development is incompatible with Europe's approach, which is rooted in the values of democracy, rule of law and respect for universal human rights. Chinese concepts like harmony and the construction of a "community of common destiny" (cited in the Beijing AI principles) are part of a political project aimed at optimizing social governance to ensure regime stability while promoting the party-state's interests in international relations.

European policymakers must take the Chinese government's international outreach on AI ethics with great caution. In China, as in other countries, there are notable gaps between the lofty ethical principles adopted by the government, industry and academia, and their implementation. Europe should closely watch how those differences and contradictions develop as China becomes a key player in global AI governance.

Any engagement with Chinese government actors must be informed by an awareness that political interests outweigh ethical considerations. China's government and businesses are contributing to a global discourse that focuses on ethics without addressing crucial human rights questions.⁶⁸ China is certainly not alone in using (and exporting) AI for surveillance, including in ways that raise serious ethical concerns.⁶⁹ However, what sets China apart is the scale of the party-state's ambitions to harness AI to strengthen its authoritarian governance system.

Against this backdrop, European countries must work with like-minded democracies to advance standards rooted in liberal-democratic values. They must develop robust ethical standards and guidelines for AI research collaborations with Chinese actors and use the recently reformed EU export controls to prevent European hardware, software and research from enabling unethical applications in China and elsewhere. Lastly, the EU should continue pushing Beijing to respect non-binding principles it endorsed, like the G20 AI principles.

In pursuing cooperation, European actors should be receptive to the multitude of ethical concerns and solutions identified by Chinese lawmakers, academics and citizens

ENDNOTES

- 1 | Xinhua (2020). “习近平：推动我国新一代人工智能健康发展 (Xi Jinping: Promoting the Development of China's New Generation of Artificial Intelligence).” October 31. http://www.xinhuanet.com/politics/leaders/2018-10/31/c_1123643321.htm. Accessed: November 5, 2020.
- 2 | State Council (2017). “国务院关于印发新一代人工智能发展规划的通知 (State Council Notice on Issuing the New Generation Artificial Intelligence Development Plan).” http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm. Accessed: November 5, 2020.
- 3 | Chinese Electronics Standards Institute (2018). “人工智能标准化白皮书 (Artificial intelligence Standardization White Paper).” <http://www.cesi.cn/images/editor/20180124/20180124135528742.pdf>. Accessed: November 5, 2020.
- 4 | These principles are: harmony and friendliness; fairness and justice; inclusivity and sharing; respect for privacy; safety and controllability; shared responsibility; open collaboration; agile governance. See: MOST (2019). “发展负责任的人工智能：新一代人工智能治理原则发布 (Developing Responsible Artificial Intelligence: Governance Principles for a New Generation of Artificial Intelligence Are Published).” http://www.most.gov.cn/kjbgz/201906/t20190617_147107.htm. Accessed: November 15, 2020.
- 5 | These are 15 principles covering AI R&D, use, and governance. BAAI (2019). “Beijing AI Principles.” <https://www.baai.ac.cn/news/beijing-ai-principles-en.html>. Accessed: November 15, 2020.
- 6 | DigiChina (2019a). “Translation: Chinese AI Alliance Drafts Self-Discipline 'Joint Pledge'.” June 17. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-ai-alliance-drafts-self-discipline-joint-pledge/>. Accessed: November 10, 2020.
- 7 | Nesta (2020). “China's Approach to AI Ethics.” <https://www.nesta.org.uk/report/chinas-approach-to-ai-ethics/mapping-official-ai-ethics-discussions-china/>. Accessed: December 6, 2020.
- 8 | State Council (2017)
- 9 | Chinese Politics and Law (2019). “Analysis of the Chinese Communist Party's Political-Legal Work Directive - First Impressions.” January 20. <http://lilingsblog.blogspot.com/2019/01/highlights-of-chinese-communist-partys.html>. Accessed: December 10, 2020.
- 10 | Lewis, Dev (2017). “Interview with Dr. Rogier Creemers: AI + Social Credit + Algorithmic Governance + Cybersecurity + VPNs.” August 14. <https://www.digitlasiahub.org/2017/08/14/interview-with-dr-roger-creemers-ai-social-credit-algorithmic-governance-cybersecurity-vpns-cross-border-dataflows/>. Accessed: December 15, 2020.
- 11 | The updated version of China's new Personal Information Security Specification stipulates that people need to explicitly consent to the collection of their biometric data, but to what extent it will be enforced remains unclear. Even the draft Personal Information Protection Law, a key privacy legislation if enacted, is unlikely to limit government organs' surveillance powers even though it would apply to them too, as national security and the public interest can always be invoked. See: APP Governance Working Group (2020). “GB/T 35273-2020《信息安全技术 个人信息安全规范》正式发布 (GB/T 35273-2020 'Information Security Technology - Personal Information Security Specification' is formally published).” March 7. <https://mp.weixin.qq.com/s/A4jnrKozCQOJzaZhkQYxw>. Accessed: November 4, 2020; DigiChina (2021). “Personal Data, Global Effects: China's Draft Privacy Law in the International Context.” January 4. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/personal-data-global-effects-chinas-draft-privacy-law-in-the-international-context/>. Accessed: January 15, 2021.
- 12 | CAICT (2020). “人工智能治理白皮书 (Artificial Intelligence Governance White Paper).” <http://www.caict.ac.cn/kxyj/qwfb/bps/202009/P020200928368250504705.pdf>. Accessed: October 20, 2020.
- 13 | The BAAI's founding members include Baidu, ByteDance, Megvii, Meituan, and Xiaomi. Members of the Artificial Intelligence Industry Alliance (AIIA) and the Artificial Intelligence Industry Technology Innovation Strategic Alliance (AITISA) include Baidu, Alibaba, Tencent, JD.com, Huawei, Qihoo 360 and Yitu.
- 14 | China Daily (2019a). “Baidu CEO Wants Ethics Research in AI Strengthened.” March 10. <https://www.chinadaily.com.cn/a/201903/10/WS5c851998a3106c65c34edc5b.html>. Accessed: September 28, 2020; South China Morning Post (2019). “China's Tech Billionaires Back Ethical Rules to Guide Development of AI and Other Technologies.” March 3. <https://www.scmp.com/tech/enterprises/article/2188449/chinas-tech-billionaires-back-ethical-rules-guide-development-ai>. Accessed: September 28, 2020.
- 15 | For example, companies contribute to CESI's AI standardization work as well as a group focused on AI software standard-setting. See: Ding, Jeffrey (2018a). “ChinAI Newsletter #22: An Open Source AI Strategy - China's New White Paper on AI Open Source Software.” August 6. <https://chinai.substack.com/p/chinai-newsletter-22-an-open-source-ai-strategy-chinas-new-white-paper-on-ai-open-source-software>. Accessed: October 10, 2020.
- 16 | Tencent Research Institute (2020). “‘ARCC’: An Ethical Framework for Artificial Intelligence.” April 9. <https://www.tisi.org/13747>. Accessed: September 28, 2020.
- 17 | For example, Tencent's Research Institute, AI Lab and Open Platform co-authored a 2017 book with the CAICT that includes one of China's earliest in-depth discussions of AI safety issues. See: Ding, Jeffrey (2018b). “Deciphering China's AI Dream.” https://www.fni.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf. Accessed: September 28, 2020.
- 18 | China Daily (2019b). “China's Tencent advocates 'AI for Good' at AI Everything Summit in Dubai.” May 2. <http://www.chinadaily.com.cn/a/201905/02/WS5ccaca1ba3104842260b98ao.html>. Accessed: September 28, 2020.
- 19 | Stock Exchange of Hong Kong Limited (2019). “Application Proof of Megvii Technology Limited.” <https://ipvm-uploads.s3.amazonaws.com/uploads/7490/c099/megvii-ipo.pdf>. Accessed: October 10, 2020; IPVM (2020a). “Huawei / Megvii Uyghur Alarms.” December 8. <https://ipvm.com/reports/huawei-megvii-uyghur>. Accessed: December 9, 2020.
- 20 | Based on a review of relevant academic papers published in Chinese journals and accessible via China National Knowledge Infrastructure (CNKI), the largest repository of PRC academic journals.

- 21 | National Laboratory of Pattern Recognition (2018). “人工智能伦理研究项目启动会暨人工智能伦理、政策与法律论坛”在京举办 (Artificial Intelligence Ethics Research Project Launch Conference and Artificial Intelligence Ethics, Policy and Law Forum is Held in Beijing).” <http://www.nlpr.ia.ac.cn/cn/news/664.html>. Accessed: December 20, 2020; The Paper (2019). 中国社科院举办“人工智能的社会、伦理与未来研究”研讨会 (CASS Holds ‘Research on the Social [Impact], Ethics and Future of Artificial Intelligence’ Seminar).” May 1. https://www.thepaper.cn/newsDetail_forward_3381346. Accessed: November 30, 2020.
- 22 | Academic Resources (2019). “段伟文：人和机器要互相理解对方缺陷，协同进化 (Duan Weiwen: Humans and Machines Must Understand Each Other’s Drawbacks and Evolve in Coordination).” November 19. <https://homest.org.cn/article/detail?id=509997>. Accessed: November 10, 2020; Jia, Hepeng (2020). “Research Ethics: A Safeguard for Advanced Technologies.” National Science Review 7(11): 1787-1792.
- 23 | Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences. Harmonious Artificial Intelligence Principles. <http://harmonious-ai.org/>. Accessed: October 15, 2020; Research Center for Brain-inspired Intelligence, Institute of Automation, Chinese Academy of Sciences. Brain Cog: Brain-inspired Cognitive Engine for Brain Simulation and Brain-inspired Artificial Intelligence. <https://bii.ia.ac.cn/braincog/>. Accessed: January 10, 2021.
- 24 | Beijing News (2019). “人工智能企业要组建道德委员会，该怎么做? (What Should Artificial Intelligence Companies Do if They Want to Form an Ethics Committee?)” July 26. <http://www.bjnews.com.cn/feature/2019/07/26/608130.html>. Accessed: November 20, 2020; Guo, Rui (2020). 人工智能的伦理和治理 (The Ethics and Governance of Artificial Intelligence). Beijing: Falu Chubanshe.
- 25 | The Berggruen China Center plays an active role in promoting the Chinese party-state’s discourse and governance model for an international audience. For example, together with government-affiliated think tanks it has co-hosted several editions of the Understanding China Conference, which promotes the government’s vision for global governance. See: Berggruen Institute (2018). “The Third Understanding China Conference Co-Hosted by the Berggruen Institute 21st Century Council.” December 16. <https://www.berggruen.org/events/the-third-understanding-china-conference/>. Accessed: January 12, 2021.
- 26 | Tsinghua University (2020). “清华大学成立人工智能国际治理研究院 (Tsinghua University Establishes the Institute for International Governance of Artificial Intelligence).” June 25. <https://news.tsinghua.edu.cn/info/1003/80208.htm>. Accessed: December 12, 2020.
- 27 | WeChat account of China Institute for Science and Technology Policy at Tsinghua University (2020). “薛澜教授接受《网络传播》杂志专访谈人工智能发展与治理 (Professor Xue Lan Gives Special Interview to ‘Internet Communication’ Magazine on the Development and Governance of Artificial Intelligence).” August 24. <https://mp.weixin.qq.com/s/OoG9zLuKAIZH6bHLAXeKgw>. Accessed: November 25, 2020.
- 28 | China-UK Research Centre for AI Ethics and Governance (2020). “Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance.” <https://ai-ethics-and-governance.institute/2020/05/15/overcoming-barriers-to-cross-cultural-cooperation-in-ai-ethics-and-governance/>. Accessed: January 12, 2021.
- 29 | Backlash was particularly intense among citizens in Hangzhou, where officials contemplated expanding the use of those apps for long-term health assessments. See for example: Weibo (2020). Weibo user post. May 25. <https://weibo.com/1686709997/J3wKexkD1?type=comment>. Accessed: October 27, 2020.
- 30 | TC260 (2020). “信息安全技术 个人信息安全规范 (Information Security Technology - Personal Information Security Specification).” May 3. <https://www.tc260.org.cn/front/postDetail.html?id=20200918200432>. Accessed: May 7, 2021; DigiChina (2021a). “China’s Draft Privacy Law Adds Platform Self-Governance, Solidifies CAC’s Role.” <https://digichina.stanford.edu/news/chinas-draft-privacy-law-adds-platform-self-governance-solidifies-cacs-role>. Accessed: May 7, 2021.
- 31 | Southern Metropolis Daily (2019). “使用人脸识别 超七成受访者担心信息泄露 (Using Facial Recognition: Over 70% of Survey Respondents Are Worried about Data Leaks).” December 6. http://epaper.oeeee.com/epaper/A/html/2019-12/06/content_52097.htm. Accessed: October 22, 2020.
- 32 | Zeng, Yi et al. (2019). “Responsible Facial Recognition and Beyond.” <https://arxiv.org/abs/1909.12935>.
- 33 | TC260 (2021). “关于国家标准《信息安全技术 人脸识别数据安全要求》征求意见稿征求意见的通知 (Notice on the Release for Public Comments of the National ‘Information Security Technology - Facial Recognition Data Security Requirements).” https://www.tc260.org.cn/front/bzzqyjDetail.html?id=20210423182442&norm_id=20201104200034&recode_id=41855. Accessed: May 10, 2021.
- 34 | SCMP (2021). “Increasing Use of Facial Recognition Technology in China Faces Backlash from City Governments.” April 28. <https://www.scmp.com/tech/tech-trends/article/3131442/increasing-use-facial-recognition-technology-china-faces-backlash>. Accessed: April 29, 2021.
- 35 | The Verge (2019). “Another Convincing Deepfake App Goes Viral Prompting Immediate Privacy Backlash.” September 2. <https://www.theverge.com/2019/9/2/20844338/zao-deepfake-app-movie-tv-show-face-replace-privacy-policy-concerns>. Accessed: October 27, 2020.
- 36 | NRTA (2019). “网络音视频信息服务管理规定 (Provisions on the Management of Online A/V Information Services).” http://www.nrta.gov.cn/art/2019/11/29/art_113_48908.html. Accessed: October 27, 2020; CAC (2019a). “制度创新与技术创新同步——评《网络音视频信息服务管理规定》的特点与亮点 (When Institutional Innovation and Technological Innovation Go Hand in Hand: Comment on the Highlights of the ‘Provisions on the Management of Online A/V Information Services’).” http://www.cac.gov.cn/2019-12/10/c_1577513351255547.htm. Accessed: October 27, 2020.
- 37 | DigiChina (2021b). “How Will China’s Privacy Law Apply to the Chinese State?” January 26. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/how-will-chinas-privacy-law-apply-to-the-chinese-state/>. Accessed February 10, 2021.
- 38 | AIIA (2020). “人工智能助力新冠疫情防控 调研报告 (Artificial intelligence in the Prevention and Control of the Coronavirus Epidemic: Research Report).” <http://aiaa.org.cn/upload-file/2020/0324/20200324060228591.pdf>. Accessed: November 10, 2020.
- 39 | Sohu (2019). “2017-2018年，551起数据泄露事件中医疗行业占六成 (The Healthcare Industry Accounts for 60% of 551 Data Breaches in 2017-2018).” June 6. https://www.sohu.com/a/319005119_161795. Accessed: November 10, 2020.

- 40 | Roberts, Huw et al. (2010). “The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation.” In: *AI & Society* 36, 59-77.
- 41 | Based on a list of relevant academic papers on the topic of AI ethics accessible through the CNKI database.
- 42 | Sohu (2018). “人工智能在医疗领域的准入、管理与前景 (The Access, Management and Prospects of Artificial Intelligence in Healthcare).” December 19. https://www.sohu.com/a/282928551_456032. Accessed: November 22, 2020.
- 43 | Wang, Yue and Dai, Haiyang (2020). “医疗AI安全风险的伦理与法律保障机制研究 (Research on Ethical and Legal Safeguard Mechanisms for Medical AI Safety Risks).” August 25. <https://mp.weixin.qq.com/s/5gRTjG-YP8n4K38HuKGUiA>. Accessed: November 22, 2020.
- 44 | NDRC (2020). “智能汽车创新发展战略 (Intelligent Vehicle Innovation Development Strategy).” <https://www.ndrc.gov.cn/xxgk/zcfb/tz/202002/P020200224573058971435.pdf>. Accessed: October 5, 2020; Caixin (2020). “China Plans to Mass Produce Driverless Cars by 2025 Later Than Previous Forecast.” February 25. <https://www.caixinglobal.com/2020-02-25/china-plans-to-mass-produce-driverless-cars-by-2025-later-than-previous-forecast-101520085.html>. Accessed: October 5, 2020.
- 45 | China Daily (2020). “China Pushes Forward with Autonomous Driving Standards.” April 20. <https://www.chinadaily.com.cn/a/202004/20/WS5e9d025da3105d50a3d175co.html>. Accessed: October 5, 2020; CAC (2020). “三部门关于印发《国家车联网产业标准体系建设指南（车辆智能管理）》的通知 (Notice on Three Ministries Issuing the “Guide for the Construction of a National Connected Vehicle Industry Standards System).” http://www.cac.gov.cn/2020-04/24/c_1589274576115317.htm. Accessed: October 5, 2020; China Daily (2018). “Country Issues National Standards for Autonomous Vehicle Testing.” August 13. <http://www.chinadaily.com.cn/a/201808/13/WS5b70f60aa310add14f385697.html>. Accessed: October 5, 2020.
- 46 | Baidu Apollo (2019). “A Whitepaper on Automated Driving Safety.” <https://apollo.auto/platform/white-paper.html>. Accessed: October 5, 2020.
- 47 | People’s Daily Online (2019). “中汽协联合企业机构发布技术规范，提升智能网联信息安全 (CAAM Joins with Industry to Issue Technical Specifications to Improve the Information Security of Intelligent Networked Vehicles).” June 14. <http://auto.people.com.cn/n1/2019/0614/c1005-31136829.html>. Accessed: October 5, 2020.
- 48 | Smart City Ecosystem WeChat Account (2020). “自动驾驶数据安全白皮书 (White Paper on Automated Driving Information Security).” September 18. <https://mp.weixin.qq.com/s/6vCA8jv0eS-duFCeGtIVCA>. Accessed: October 5, 2020.
- 49 | Baidu Security X-Lab (2019). “Tackling AI Challenges in Safety-Critical Scenarios — A Review on Robustness of Deep Learning Models and the Release of Perceptron Robustness Benchmark Tools.” June 18. <https://medium.com/baidulab/tackling-ai-challenges-in-safety-critical-scenarios-a-review-on-robustness-of-deep-learning-8e0e30ff1018>. Accessed: October 5, 2020.
- 50 | State Council (2017). The term “intelligentization” (智能化) is a uniquely Chinese concept that refers to the integration of AI technologies into military operations and the consequent transformation and reshaping of the form and character of warfare through the integration of human and machine intelligence. For a detailed discussion of the term see Kania, Elsa (2019). “Chinese Military Innovation in Artificial Intelligence.” Testimony before the U.S.-China Economic and Security Review Commission Hearing on Trade, Technology, and Military-Civil Fusion. June 7. https://www.uscc.gov/sites/default/files/June%207%20Hearing_Panel%201_Elsa%20Kania_Chinese%20Military%20Innovation%20in%20Artificial%20Intelligence_o.pdf. Accessed: March 3, 2021. Outside the military innovation discourse, the term also frequently appears in official discussions about social governance and public security management, where the emphasis is on the integration of Big Data, AI and other technologies to improve predictive governance and policing capabilities. The term also frequently appears in official discussions about social governance and public security management, where the emphasis is on the integration of Big Data, AI and other technologies to improve predictive governance and policing capabilities. The idea is to “make everything smart” by collecting and analyzing vast amounts of data to preemptively detect and eliminate any risks – perceived or real – to social stability and regime security.
- 51 | DigiChina (2019b). “Translation: Key Chinese Think Tank’s “AI Security White Paper” (Excerpts).” February 21. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-key-chinese-think-tanks-ai-security-white-paper-excerpts/>. Accessed: November 15, 2020.
- 52 | The New York Times (2018). “Inside China’s Dystopian Dreams: A.I., Shame and Lots of Cameras.” July 8. <https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html>. Accessed: December 6, 2020; on the integration of AI into the surveillance infrastructure, see also: CSET (2020a). “Designing Alternatives to China’s Repressive Surveillance State.” <https://cset.georgetown.edu/research/designing-alternatives-to-chinas-repressive-surveillance-state/>. Accessed: November 17, 2020.
- 53 | Human Rights Watch (2019). “China’s Algorithms of Repression.” <https://www.hrw.org/report/2019/05/02/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass>. Accessed: December 5, 2020; The New York Times (2019a). “How China Uses High-Tech Surveillance to Subdue Minorities.” May 22. <https://www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang>. Accessed: December 5, 2020; Financial Times (2019). “The Role of AI in China’s Crackdown on Uighurs.” December 10. <https://www.ft.com/content/e47b33ce-1add-11ea-97df-cc63de1d73f4>. Accessed: December 5, 2020.
- 54 | CSET (2020b). “China’s System of Oppression in Xinjiang: How It Developed and How to Curb It.” <https://cset.georgetown.edu/research/chinas-system-of-oppression-in-xinjiang-how-it-developed-and-how-to-curb-it/>. Accessed: January 12, 2021; the successes of video-surveillance and facial recognition giants Hikvision and Dahua illustrate how lucrative the business of repression in Xinjiang is. They won bids worth over USD 1 billion for surveillance contracts with public security bureaus in Xinjiang between 2016 and 2018. See: IPVM (2018). “Dahua and Hikvision Win Over \$1 Billion In Government-Backed Projects in Xinjiang.” April 23. <https://ipvm.com/reports/xinjiang-dahua-hikvision>. Accessed: January 12, 2021.

- 55 | The New York Times (2019b). “One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority.” April 14. <https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html>. Accessed: December 15, 2020; IPVM (2021). “Patenting Uyghur Tracking - Huawei, Megvii, More.” January 12. <https://ipvm.com/reports/patents-uyghur>. Accessed: January 15, 2021.
- 56 | ASPI (2019). “Mapping more of China’s tech giants: AI and surveillance.” <https://www.aspi.org.au/report/mapping-more-chinas-tech-giants>. Accessed: December 10, 2020; Human Rights Watch (2017). “China: Voice Biometric Collection Threatens Privacy.” October 22 <https://www.hrw.org/news/2017/10/22/china-voice-biometric-collection-threatens-privacy>. Accessed: December 12, 2020.
- 57 | IPVM (2019). “China Government Spreads Uyghur Analytics Across China.” November 25. <https://ipvm.com/reports/ethnicity-analytics>. Accessed: January 15, 2021; IPVM (2020b). “Alibaba Uyghur Recognition as A Service.” December 16. <https://ipvm.com/reports/alibaba-uyghur>. Accessed: January 15, 2021.
- 58 | China’s third-largest video surveillance manufacturer, Uniview, insists that it pursues “tech for social good” through an ethics code and an ethics panel when interviewed about its software used for tracking Uighurs. See: IPVM (2020c). “Uniview Racist Uyghur Recognition Revealed.” November 16. <https://ipvm.com/reports/uniview-uyghur>. Accessed: January 15, 2021.
- 59 | The Global Partnership on AI is an international coalition initiated by G7 leaders that aims to ensure that AI is used responsibly. See: OECD (2020). “OECD to Host Secretariat of New Global Partnership on Artificial Intelligence.” June 15. <https://www.oecd.org/going-digital/ai/oecd-to-host-secretariat-of-new-global-partnership-on-artificial-intelligence.htm>. Accessed: December 10, 2020; OECD (2019). “Forty-Two Countries Adopt New OECD Principles on Artificial Intelligence.” May 22. <https://www.oecd.org/science/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence.htm>. Accessed: December 10, 2020.
- 60 | The non-binding resolution on human-centric AI was adopted by the G20 in June 2019, but it is unclear whether this will be followed by concrete actions. G20 (2019). “G20 Ministerial Statement on Trade and Digital Economy.” https://trade.ec.europa.eu/doclib/docs/2019/june/tradoc_157920.pdf. Accessed: December 10, 2020.
- 61 | United Nations (2018). “United Nations Secretary-General Appoints High-level Panel on Digital Cooperation.” July 12. <https://digitalcooperation.org/united-nations-secretary-general-appoints-high-level-panel-on-digital-cooperation/>. Accessed: December 8, 2020; UNESCO (2019). “Beijing Consensus on Artificial Intelligence and Education.” <https://unesdoc.unesco.org/ark:/48223/pfo000368303>. Accessed: December 8, 2020; UNESCO (2020). “UNESCO’s International Expert Group Begins Work on Drafting the First Global Recommendation on the Ethics of AI.” April 27. <https://en.unesco.org/news/unescos-international-expert-group-begins-work-drafting-first-global-recommendation-ethics-ai>. Accessed: December 8, 2020.
- 62 | DigiChina (2018). “Chinese Interests Take a Big Seat at the AI Governance Table.” June 20. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/chinese-interests-take-big-seat-ai-governance-table/>. Accessed: December 10, 2020.
- 63 | BAAI (2020). “Facial Recognition and Public Health.” May 17. <https://www.baai.ac.cn/news/20200517.html>. Accessed: December 10, 2020; China-UK Research Centre for AI Ethics and Governance (2020).
- 64 | See for example: Cao, Yulong et al. (2019). “Adversarial Objects Against LiDAR-Based Autonomous Driving Systems.” <https://arxiv.org/abs/1907.05418>. Accessed: November 1, 2020
- 65 | European Commission (2019). “Building Trust in Human-Centric Artificial Intelligence.” <https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence>. Accessed: December 10, 2020. See also: European Commission (2020). “White Paper on Artificial Intelligence: A European Approach to Excellence and Trust.” https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Accessed: December 10, 2020; High-Level Expert Group on Artificial Intelligence (2019). “Ethics Guidelines for Trustworthy AI.” <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Accessed: December 10, 2020.
- 66 | European Commission (2021). “Proposal for A Regulation Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).” <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence>. Accessed: May 5, 2021.
- 67 | While similarity between terms does not imply equivalence in meaning, China’s Governance Principles broadly resemble the EU’s Ethics Guidelines in that they also stress conformity to human values, respect for privacy, fairness and non-discrimination, environmental sustainability, technical safety and robustness, and explainability. See also how EU principles are being actively interpreted by Chinese scholars in the context of specific application scenarios: Shen, Yuan and Wang, Qiong (2020). “人工智能在教育中应用的伦理考量——从教育视角解读欧盟《可信赖的人工智能伦理准则》 (Ethics Arguments on AI in Education: An Analysis of the EU’s Ethics Guidelines for Trustworthy AI from an Educational Perspective).” In: Peking University Education Review (2019) 17 (4): 18-34; Xiang, Xiejun et al. (2020). “欧盟医疗人工智能相关战略形成路径及启示 (The Formation and Teaching of the EU’ Strategy for Medical AI).” April 29. https://www.sohu.com/a/392088292_694090. Accessed: December 10, 2020.
- 68 | Stiftung Neue Verantwortung (2019). “Artificial Intelligence Needs Human Rights.” https://www.stiftung-nv.de/sites/default/files/ai_needs_human_rights.pdf. Accessed: December 8, 2020.
- 69 | Carnegie Endowment for International Peace (2019). “The Global Expansion of AI Surveillance.” <https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847>. Accessed: December 8, 2020; The New York Times (2019c). “Facial Recognition’s Many Controversies, From Stadium Surveillance to Racist Soft.” May 15. <https://www.nytimes.com/2019/05/15/business/facial-recognition-software-controversy.html>. Accessed: December 10, 2020.

ACKNOWLEDGEMENTS

The author would like to thank the many MERICS colleagues who engaged with this report through substantive input, ideas, suggestions and feedback, in addition to reviewing previous drafts of the text.

CONTACT

Rebecca Arcesati
Analyst, MERICS
rebecca.arcesati@merics.de

EDITORIAL TEAM

Claudia Wessling
*Director Communications
and Publications, MERICS*
claudia.wessling@merics.de

Nick Bouchet
Freelance Editor

GRAPHICS

STOCKMAR+WALTER Kommunikationsdesign

LAYOUT

Alexandra Hinrichs
Graphic Designer, MERICS

PUBLISHER

MERICS | Mercator Institute for China Studies
Klosterstraße 64 | 10179 Berlin
Tel.: +49 30 3440 999 0
Mail: info@merics.de
www.merics.org