The AI Powered State:

China's approach to public sector innovation

nesta

Acknowledgements

This project has been a real team effort and there are lots of people who have dedicated their time and energy to make this essay collection possible.

A first set of thanks goes to each of the authors for sharing their expertise and for working with such enthusiasm and commitment in spite of the extraordinary and difficult circumstances of the past few months.

A huge thank you to Michelle Pauli for her excellent editorial advice and support, and to Joana Pereira for her imaginative work on the collection's design.

At Nesta, Tom Symons has worked tirelessly to bring this project to life and I am deeply grateful for his constant support and guidance. Special thanks also go to Nyangala Zolho for her fantastic input and advice throughout the process of curating this collection. Many thanks to Wayne Holmes for offering his expertise and editorial support, and to Anna Cleary, Wallis Grant, Ella White and Georgina Roberts for their help on production and communications. Thank you to all the other Nesta colleagues who offered advice and support along the way: Geoff Muglan, Jonathan Bone, Aleks Berditchevskaia, Madeleine Gabriel, Laurie Smith, Juan Mateos-Garcia and Jen Rae.

The essays in this collection represent the views of each author and do not necessarily represent the views of Nesta. Nesta's views are reflected in the introduction and conclusion to the collection.

Hessy Elliott

About Nesta

Nesta is an innovation foundation.

For us, innovation means turning bold ideas into reality and changing lives for the better. We use our expertise, skills and funding in areas where there are big challenges facing society.

Nesta is based in the UK and supported by a financial endowment. We work with partners around the globe to bring bold ideas to life to change the world for good.

www.nesta.org.uk

If you'd like this publication in an alternative format such as Braille, large print or audio, please contact us at: information@nesta.org.uk

May 2020

Contents

Introduction

What can we learn from China's use of artificial intelligence in public services?

China is striding ahead of the rest of the world in terms of its investment in artificial intelligence (AI), rate of experimentation and adoption, and breadth of applications.[1] In 2017, China announced its aim of becoming the world leader in AI technology by 2030.[2] AI innovation is now a key national priority, with central and local government spending on AI estimated to be in the tens of billions of dollars.[3]

While Europe and the US are also following AI strategies designed to transform the public sector,[4] there has been surprisingly little analysis of what practical lessons can be learnt from China's use of AI in public services. Given China's rapid progress in this area, it is important for the rest of the world to pay attention to developments in China if it wants to keep pace.

This essay collection finds that examining China's experience of public sector innovation offers valuable insights for policymakers. Not everything is applicable to a western context – there are social, political and ethical concerns that arise from China's use of new technologies in public services and governance – but there is still much that can be learnt from its experience while also acknowledging what should be criticised and avoided.

Taking a closer look at China's use of AI in public services

A central theme of China's 2017 AI development strategy is that AI can help improve public services and governance,[5] and this has created significant opportunities for public sector innovation. Local governments and authorities, research institutions and tech companies are making rapid progress in pursuing AI projects that address real-world problems, from medical imaging for disease diagnosis to smart city applications for the improvement of transport and urban infrastructure.[6] Such applications have tended to be overlooked by analysts outside China, who too often default to narratives about authoritarianism or the foreign policy implications of a global AI arms race.

This essay collection aims to offer policymakers new ideas that may inform their own practice, by tapping into China's experience of AI in public services. The essays – written by experts on AI innovation in China – offer detailed analyses of China's development and use of AI technologies, examining the impact on public service delivery and society, and considering what can be learnt.

While each essay represents the views of the author and does not necessarily reflect the position of Nesta, we hope that the collection as a whole will offer a nuanced understanding of public sector AI innovation in China – challenging misconceptions, highlighting promising approaches and applications, and cautioning against problematic uses.

Learning from China's experience

Through the essays, we discover what's good, bad and unexpected about China's public sector AI innovation. There are six key lessons that emerge:

01. Rapid AI development is driven by local innovation
China is experiencing rapid innovation in many AI fields, thanks to its strategy of actively cultivating local innovation ecosystems. Local governments in China are empowered to drive forward with experimentation and implementation, and are establishing cross-sector partnerships with industry and academia. In light of the UK government's regional 'levelling-up' ambitions to address geographic economic imbalances by, among other things, boosting research and innovation throughout the UK,[7] China's innovation ecosystem is instructive. China's increased R&D spending, combined with devolved local governments that have the power and authority to drive collaboration with technology and research partners, highlights how flourishing innovation ecosystems can be fostered outside of capital cities. There are echoes here of previous calls from Nesta and others for more of the UK's R&D budget to be devolved to cities and regions to spread the impact and benefits of innovation across the UK.[8] China's experience would suggest there is merit in pursuing this approach.

02. 'Experiment first and regulate later' leads to quick progress
China has made swift progress in practical AI applications because it experiments first and regulates later. China's agile and pragmatic approach to innovation is resulting in the fast-paced development of AI applications, especially in the field of healthcare, with promising results in areas such as medical image processing and pharmaceutical research. The urgency of health challenges such as an ageing population are universally applicable and there are major gains to be had for governments that can move quickly and find new and effective approaches to tackle such challenges. The regulatory environment in Europe may not allow us to replicate China's fast-paced approach to experimentation and implementation (and there are reasons why we wouldn't want it to),[9] but it is worth considering how experimentation methods like regulatory sandboxes and innovation testbeds could be used to adapt this approach for a western context. This resonates with Nesta's work on 'anticipatory regulation' which has explored the role these methods can play in allowing new products and services to be safely tested and brought to market more quickly,[10] and is an approach that has also been called for by NHS Digital in the UK.[11]

03. Innovation isn't always the result of cutting-edge tech
China's AI innovation successes are not always due to cutting-edge technologies. Instead, success in China is often down to rapid deployment and scaling of existing AI technologies. China's innovation model, which prioritises 'fusion and speed' over breakthrough technologies,[12] demonstrates that an AI strategy doesn't have to focus solely on the most advanced technologies in order to be successful. While conversations about innovation in the UK often focus on R&D for new, advanced technologies, equally important is ensuring better adoption of existing technologies[13] – yet the uptake of available AI technologies among UK councils, for example, remains strikingly low.[14] There is clear scope for dedicated efforts to be made in the UK and other European countries to consider how existing AI technologies can be adopted and scaled to improve public services.

04. Good-quality services shouldn't come at the expense of widening inequalities

While AI has the potential to widen access to good-quality services, it could also end up reinforcing inequalities. China provides a useful lesson of what should be avoided. Developments driven by private sector tech companies and then not universalised by the public sector risk widening inequalities and the marginalisation of certain groups. This is the case with China's adoption of AI educational tools. With development led by private sector ed tech companies, these tools are often only accessible to wealthier families. Nesta has highlighted the importance of ensuring that innovation tackles inequality rather than increases it,[15] and has called for improved understanding among policymakers about how innovation – and innovation policies – impact different social groups, and for interventions that serve the needs of those who are particularly marginalised.[16]

05. The issue of AI ethics in China is complicated

Contrary to dominant western media narratives, China is talking about AI ethics. Various multi-stakeholder expert committees have been established and have released documents outlining AI ethics principles, many of which align with existing global standards. The international community could do more to engage with China on these issues, and foster greater global co-operation on key conversations about AI ethics and safety. Europe has the opportunity to draw on its strengths and be a co-operation partner on matters such as data protection and AI ethics.[17]

Yet China's discussions around AI ethics don't detract from the fact that, as an authoritarian state, there are elements of the Chinese government's use of AI which are profoundly alarming and violate civil liberties. Privacy laws that protect consumers from tech companies do not limit the government's access to and use of private data. There are many instances where citizens' privacy is being infringed, or where the operation of AI systems lacks transparency.

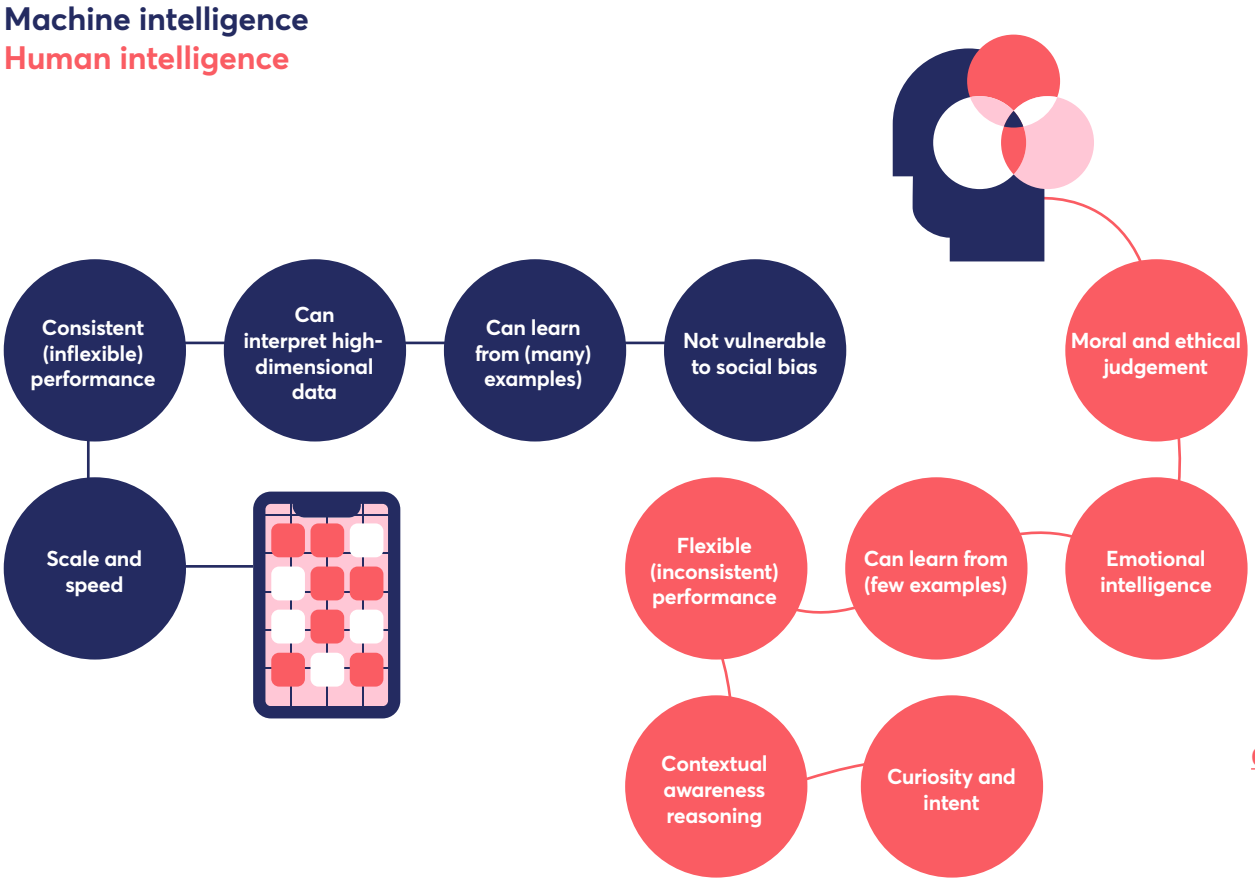
Most disturbingly, reports and leaked documents have revealed the government's use of facial recognition technologies to enable the surveillance and detention of Muslim ethnic minorities in Xinjiang province.[18]

China is a global leader in AI surveillance and facial recognition technologies. But companies in the US, France, Germany and elsewhere are also active in developing and supplying surveillance technologies worldwide,[19] and governments in Europe and the US have been criticised by rights groups for their use of facial recognition.[20]

The relationship between AI and surveillance deserves deep consideration by policymakers and technologists the world over. The ethical tensions highlighted in this essay collection may focus on China but should give us cause to reflect on fundamental questions about the use of AI wherever it is being deployed.

06. AI is not a panacea

China's experience also demonstrates that AI is not a silver bullet – challenging the techno-optimism of Silicon Valley. In areas with clearly defined and generally accepted outcomes, such as smart city traffic optimisation, AI is a promising tool to improve the efficiency and reach of services. However, in areas of social life that are complex, unpredictable or subject to debate – such as education or the judiciary – AI may not be a well-placed or desirable tool. In these areas, fundamentally human qualities such as flexibility, emotional intelligence, contextual awareness and moral judgement remain vital.[21] Here AI may be able to play a useful role as a decision support tool but should not replace humans.



Infographic from *The Future of Minds and Machines*. Visual asset created by Margherita Cardoso of Soapbox.

Conclusion

We hope this collection gives people a deeper understanding of how China is innovating with AI in public services, challenging current misconceptions and illuminating nuances in the debate. Taken together, the essays demonstrate the broad spectrum of activity and innovation in China. The collection offers both new ideas and insights to policymakers and innovators about how AI can be used to benefit public services, as well as highlighting the social, political and ethical issues that arise from China's applications.

We are sure that readers will learn something new and encounter ideas that challenge their thinking. Understanding and learning from China's experience offers us the opportunity to reflect on some of the fundamental questions that the use of AI in public services raises. This collection provides a fresh perspective and adds a new dimension to this debate.

Hessy Elliott
Government Innovation Research, Nesta

More about the essays in this collection

The collection begins with Jeffrey Ding's essay on China's AI innovation ecosystem. Contrary to western perceptions that China's AI development is driven by a centrally planned industrial strategy, it is at the local level that real innovation is happening. Cities are creating successful 'hybridised industrial ecosystems' by connecting the public sector, leading tech companies, research institutes and networks of SMEs to drive technological advances and implementation. Increased R&D spending is only part of the equation: for such ecosystems to develop successfully, cities also need a strong local technology base, an elite university, a co-operative local government, a good skill supply and inter-regional or international linkages.

Andy Chun is optimistic about AI's potentially revolutionary impact on healthcare in China, which faces the dual challenges of an ageing population and a scarcity of medical services in rural areas. AI is being developed and deployed in several key areas: improving diagnostic accuracy; assisting with treatment decisions; predicting potential health problems through analysis of patient data generated by wearables; automating certain tasks to reduce physician workload and increase remote access; and reducing R&D timeframes for new pharmaceuticals using AI modelling and simulations. Following the outbreak of COVID-19 in Wuhan, these key AI applications are being deployed in China to tackle the pandemic and contribute to vaccine research efforts. China's 'experiment first, regulate later' approach is aiding its rapid advances in healthcare AI – in a European context, regulatory innovations such as sandboxes could help boost the pace of progress.

Yi-Ling Liu examines the emergence of AI educational tools and adaptive learning systems, and the potential benefits that these innovations are said to offer. However, with development led in large part by private sector ed tech startups, it remains to be seen whether these tools can translate to the public education system and help balance the rural-urban educational divide, as touted, or whether in fact they are likely to entrench existing socioeconomic inequalities. Moreover, there are fundamental questions around whether AI educational tools are beneficial in a pedagogical sense, capable of teaching students beyond mere rote learning, as well as serious concerns about the ethical implications of collecting student data. Learning from China's experiences of AI in education becomes of vital importance as the world looks to online learning as a potential solution to the global disruption to schools due to the coronavirus pandemic.[22]

Eugeniu Han examines the use of AI in smart cities in China, exploring two illuminating case studies: 'City Brain', an intelligent urban management system designed in Hangzhou, which has now been rolled out to 23 cities across Asia, and Shanghai's Smart Court system. While the judicial system is not often analysed alongside urban management and transport services in typical smart city case studies, this original perspective reflects two key drivers behind the development of China's smart cities: the growing demand for improved services in the context of urbanisation, and the desire to strengthen China's main governing institutions – this latter theme is also explored in the essays by Dev Lewis and Rogier Creemers. The topic of smart cities in China is complex: some applications have great potential to benefit cities and citizens and make urban services more efficient, yet other elements are problematic, especially regarding issues of privacy, bias and political interference.

Dev Lewis examines the development of China's Social Credit System (SCS). The SCS has gained notoriety outside China and, while critics raise legitimate concerns, there is much misunderstanding about the system's functioning, aims and reach. The SCS forms part of the Chinese government's attempt to strengthen its weak institutions and boost trust between governments, firms and individuals by improving and harnessing its public data platforms. Rather than being a single overarching system, the SCS is a cluster of different emerging initiatives. These include the national Blacklist-Redlist Joint Sanctions and Rewards regimes, that attempt to remedy the judicial system's failure to enforce laws, and municipal-level Citizen Score experiments being developed in some cities, that offer benefits to individuals with 'good credit'. Western narratives tend to conflate SCS with other types of surveillance in China, obscuring what is distinctive about the system and the motivations behind it. Of course, there are causes for concern too – particularly over fairness, accountability and the lack of transparency around what is deemed to constitute 'trustworthy' behaviour.

Danit Gal explores China's approach to AI ethics, dispelling a common assumption that these issues are not being considered in China. China has established various expert committees and published several documents outlining AI ethics principles. In some areas, especially around privacy, China's approach aligns with universal ethical principles adopted globally, and the international community could do more to include China in global AI governance discussions. But China also has its own unique ethical and cultural foundations, and questions around how to apply these distinct philosophical concepts to AI are still being debated in China. Applying AI in public services, as explored in the collection's other essays, carries with it a host of ethical considerations and implications, including privacy, bias, accessibility, accountability and transparency. While China has implemented robust consumer privacy laws, the government's use of private data for surveillance is a cause for serious concern. Yet it is worth remembering that governments in Europe and the US are also using facial recognition in problematic and concerning ways. The scope and extent of China's deployment of AI in public services may be beyond anything occurring elsewhere, but the ethical considerations it raises are universally relevant, and examining China's experience helps us reflect on shared ethical issues that must be tackled collectively.

Finally, Rogier Creemers offers a critical reflection on why China is seeking to deploy AI in public services. China's drive to deploy technology in social governance stems from the Communist Party's ideological view that social order is governed by an objective and intelligible set of 'laws', and that big data and AI can both help to understand these laws and 'engineer' society to solve social problems. The development of this ideological position is particular to Chinese intellectual history, yet there is an interesting parallel to be drawn between this ideological position and Silicon Valley 'solutionism', where technology and data are seen as a panacea for social problems. Strikingly, China's State Council explicitly identifies AI as a tool to assist with both sides of 'social governance': providing better public services as well as preventing and controlling social unrest. Understanding this helps us see that the smart courts in Eugeniu Han's essay and the Social Credit System in Dev Lewis' essay are part of the same story as the other public service AI applications that are intended to improve the lives of citizens.

Endnotes

[1] Daniel Castro, Michael McLaughlin, and Eline Chivot, 'Who is Winning the AI Race: China, the EU or the United States?', *Center for Data Innovation*, 19 August 2019, https://www.datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/.

[2] Graham Webster et al., 'Full Translation: China's "New Generation Artificial Intelligence Development Plan"', *DigiChina*, 1 August 2017, New America, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

[3] Gregory C. Allen, 'Understanding China's AI Strategy', *Center for a New American Security*, 6 February 2019, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

[4] 'AI for the Public Sector', *European Commission*, https://ec.europa.eu/knowledge4policy/ai-watch/topic/ai-public-sector_en.

[5] Paul Triolo and Graham Webster, 'Visions of Leadership for China in AI, and for Tech in Chinese Politics', in 'China's Plan to "Lead" in AI: Purpose, Prospects, and Problems', *New America*, 1 August 2017, https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/.

[6] Paul Triolo and Jimmy Goodrich, 'From Riding a Wave to Full Steam Ahead', *New America*, 28 February 2018, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/riding-wave-full-steam-ahead/.

[7] Chris Skidmore, 'Levelling up Research and Innovation Right Across the United Kingdom', speech delivered on 24 January 2020, https://www.gov.uk/government/speeches/levelling-up-research-and-innovation-right-across-the-united-kingdom.

[8] Nesta and BritainThinks, *Is the UK Getting Innovation Right?* (London: Nesta, February 2020), https://media.nesta.org.uk/documents/Is_the_UK_Getting_Innovation_Right.pdf.

[9] 'Why is Regulating Artificial Intelligence Important in Europe?', *European Parliament*, 20 February 2020, https://www.europarl.europa.eu/news/en/headlines/society/20200213STO72575/why-is-regulating-artificial-intelligence-important-in-europe.

[10] Harry Armstrong, Chris Gorst, and Jen Rae, *Renewing Regulation: 'Anticipatory Regulation' in an Age of Disruption*, (London: Nesta, March 2019), https://media.nesta.org.uk/documents/Renewing_regulation_v3.pdf.

[11] Andrea Downey, 'Regulatory Sandbox for AI Needed to Test and Build Systems, NHSX Says', *Digital Health*, 12 February 2020, https://www.digitalhealth.net/2020/02/regulatory-sandbox-for-ai-needed-to-test-and-build-systems-nhsx-says/.

[12] Kai-Fu Lee, 'Dr Kai Fu Lee: AI Superpowers: China, Silicon Valley, and the New World Order', speech filmed 28 September 2018 at Asia Society Northern California Center, introduction plus video, https://asiasociety.org/northern-california/dr-kai-fu-lee-ai-superpowers.

[13] 'Driving the Adoption of Digital Technology', CBI, https://www.cbi.org.uk/our-campaigns/driving-the-adoption-of-digital-technology/.

[14] William Eichler, '"Shockingly Small" Number of Councils Embrace Automation, Study Reveals', *LocalGov*, https://www.localgov.co.uk/Shockingly-small-number-of-councils-embrace-automation-study-reveals/47387.

[15] Nesta & BritainThinks, Is the UK getting innovation right?, 5.

[16] Isaac Stanley, Alex Glennie and Madeleine Gabriel, *How Inclusive is Innovation Policy? Insights from an International Comparison* (London: Nesta, November 2018), 21, https://media.nesta.org.uk/documents/How_inclusive_is_innovation_policy__Insights_from_an_international_comparison_v3.pdf.

[17] Kai von Carnap, Kristin Shi-Kupfer, 'Values as a Quality Seal: Germany Should be More Active in Shaping Cooperation with China', *MERICS blog: European Voices on China*, 4 September 2019, Mercator Institute for China Studies (MERICS), https://www.merics.org/en/blog/values-quality-seal-germany-should-be-more-active-shaping-cooperation-china.

[18] Human Rights Watch, *China's Algorithms of Repression: Reverse Engineering a Xinjiang Police Mass Surveillance App* (New York: Human Rights Watch, 1 May 2019), https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance.

Paul Mozur, 'One Month, 500,000 Face Scans: How China Is Using AI to Profile a Minority', *The New York Times*, 14 April 2019, https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html.

Catherine Shu, 'Leaked Chinese Government Documents Detail How Tech is Used to Escalate the Persecution of Uighurs', *TechCrunch*, 25 November 2019, https://techcrunch.com/2019/11/24/leaked-chinese-government-documents-detail-how-tech-is-used-to-escalate-the-persecution-of-uighurs/.

[19] Steven Feldstein, *The Global Expansion of AI Surveillance* (Washington DC: Carnegie Endowment for International Peace, 17 September 2019), https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847.

[20] Janosch Delcker and Cristiano Lima, 'Fight Against Facial Recognition Hits Wall Across the West', *Politico*, 27 December 2019, https://www.politico.eu/article/fight-against-facial-recognition-hits-wall-across-the-west/.

[21] Aleks Berditchevskaia and Peter Baeck, *The Future of minds and machines*, (London: Nesta, March 2020), 16, https://media.nesta.org.uk/documents/FINAL_The_future_of_minds_and_machines.pdf.

[22] UNESCO, '1.37 Billion Students Now Home as COVID-19 School Closures Expand, Ministers Scale up Multimedia *Approaches to Ensure Learning Continuity*', UNESCO, 24 March 2020, https://en.unesco.org/news/137-billion-students-now-home-covid-19-school-closures-expand-ministers-scale-multimedia.

The dynamics of China's AI strategy

01. Promoting Nationally, Acting Locally: China's next-generation AI approach

By Jeffrey Ding
DPhil researcher and China lead at the Centre for the Governance of AI, University of Oxford's Future of Humanity Institute

When China's State Council published its Next Generation Artificial Intelligence Development Plan (AIDP) in July 2017, many interpreted the move as China planning its way to AI supremacy from the top down. 'There's a tendency to place this AI mobilisation within China's longstanding tradition of centrally planned engineering achievements that have wowed the world,' writes Matt Sheehan.[1] However, that China's approach to AI is defined by its top-down and monolithic nature was one of the myths I sought to debunk in an earlier report on China's AI landscape.[2]

In this essay, I underscore the importance of provincial and local governments in implementing AI policy in China. Specifically, to spur advances in AI, subnational governments are connecting leading firms, research institutes and networks of small and medium enterprises in order to build 'hybridised industrial ecosystems'.[3] The AI ecosystems with the most potential are those that are adapted to local conditions. But, while provinces and cities are taking bold action on AI, it is important not to overstate the scale and success of local activity. Additionally, some aspects of China's locally driven approach to AI development may be counterproductive. Finally, I suggest some lessons that European governments can take from the Chinese experience, including shifting focus to potential clusters beyond major capital cities and the opportunities that lie in specialising in particular AI subdomains.

The importance of local AI policy

Provincial and local governments play an outsized role – one only increasing in significance – in implementing innovation policy. Their proportion of China's overall fiscal expenditures on science and technology rose from 48 per cent during 2007–2011 to 59 per cent in 2015–2016.[4] With respect to implementing China's national science and technology initiatives, provinces and localities have significant autonomy in experimenting with different approaches.[5]

China's AI approach does not deviate from this trend. In one sense, the central government's release of the AIDP served as a 'lead geese effect' (头雁效应) of sorts, spurring more than 20 provinces to issue at least 30 specific policies to support AI development by the end of 2018.[6] In another sense, many forward-thinking local governments implemented AI-related policies that preceded national government action, such as Hangzhou and Hefei – the two cases covered in detail in this essay.

Irrespective of the order, when the targeted benchmarks from the subnational plans that outline specific figures are aggregated, the combined target – a core AI industry worth 429 billion RMB by 2020 – far exceeds the AIDP's target of 150 billion RMB.[7] On their own, these figures are difficult to interpret as the line between core AI and AI-related industries is fuzzy. However, the ratio reveals the scale of local ambition.

Hybridised industrial ecosystems

The first-tier cities, especially Beijing, Shanghai and Shenzhen, have drawn most of the attention in analyses of local government incentives for China's AI development. According to one estimate, these three cities are home to more than 70 per cent of Chinese AI firms, with Beijing at 43 per cent, Shanghai at 15 per cent and Shenzhen at 12 per cent.[8] Given such an unbalanced distribution in their favour, it is no surprise that these cities have also established strong and robust hybridised industrial ecosystems for AI.

For instance, Beijing's National New-Generation AI Innovation and Development Pilot Zone, announced in February 2019, has provided a supportive environment for corporate and academic joint AI labs. These include research centres, such as the Center for Open Data Research, which works on establishing open data standards, as well institutes such as the Beijing Academy of Artificial Intelligence, which explores issues related to AI ethics and safety. Described by one observer as a 'first of its kind' pilot project, Beijing's AI pilot zone model is being replicated in other municipalities, with Shanghai opening its AI pilot zone in May 2019 and 18 other pilot zones expected by 2023.[9]

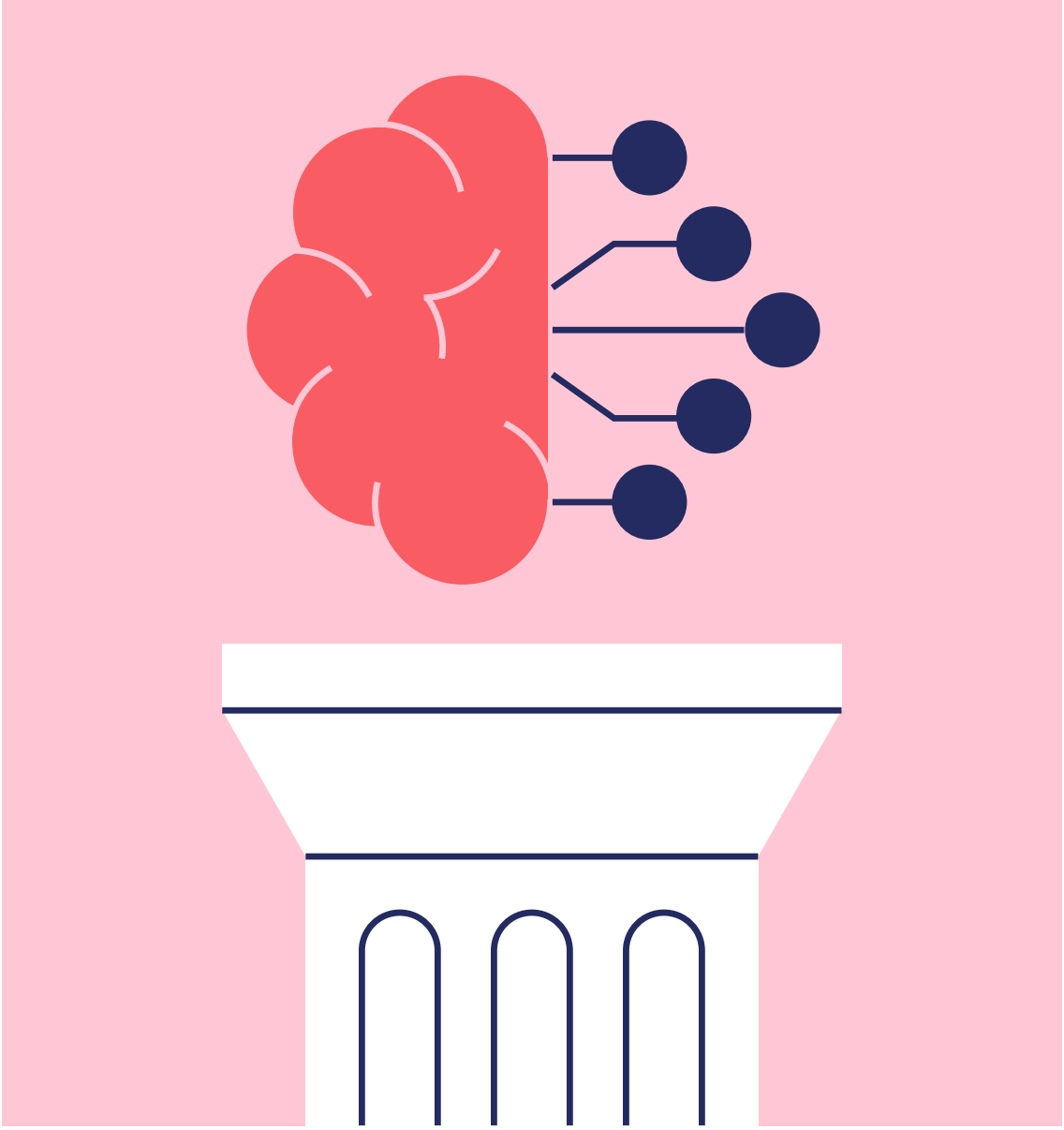
Whether other cities and regions can also support competitive ecosystems for AI development will be a crucial factor in the widespread diffusion of AI advances across China. This essay examines local efforts to build robust AI industrial ecosystems in Hangzhou, in Zhejiang province, as well as Hefei, in Anhui province. Both have the potential to achieve this goal. One ranking of Chinese cities by their development of computing power placed both within the first tier, with Beijing, Shenzhen and Shanghai sandwiched between Hangzhou (ranked first) and Hefei (ranked fifth).[10]

The Hangzhou AI Town

Eleven days before the AIDP was released, on 9 July 2017, the Hangzhou AI Town (杭州人工智能小镇) opened for business, with the mission to link together e-commerce company Alibaba, Zhejiang University, graduates returning from overseas and local businesses in an AI cluster. This AI park does not exist in a vacuum. It is housed within the Hangzhou Future Sci-Tech City (HFSTC) which itself is connected to a larger infrastructure of science and technology (S&T) parks in Hangzhou. Initiated after visits by Zhejiang provincial leaders to Silicon Valley, the driving force behind HFSTC is to cultivate knowledge spillovers and agglomeration benefits similar to those that flow to Silicon Valley.[11]

The Hangzhou AI Town's 2019 audit report, which provides a breakdown of funding allocations, offers some preliminary indications of the town's progress.[12] It disbursed 43 million RMB in funding in 2019, separated into research and development (R&D) funds, subsidies for office fees and cloud services funds. Across the three categories, the 123 accepted project proposals spanned an extensive range of subdomains (from sign language translation using computer vision to predictive analytics of smart city data) and parts of the AI stack (from open source software to end devices).[13] Moreover, the very existence of transparent, budget-line disclosures of how much funding was allocated to which companies is refreshing in that it demonstrates that there is something substantive happening.

One concern is whether Hangzhou AI Town can sustain development beyond Alibaba, whose influence dominates much of the funding breakdown. For instance, a project to industrialise the use of big data to discern risks to public safety received the most funding in both the R&D and office fees category. The company, Chengying Shuju (杭州橙鹰数据), is a subsidiary of Alibaba. Alibaba Cloud was the sole supplier or co-supplier for 25 of the 27 projects in the cloud services category.



While the benefits of S&T parks are often viewed through the lens of local spillover effects, extra-regional and international linkages may be even more significant. For HFSTC, one connection that is particularly meaningful is to Silicon Valley, which inspired the creation of the S&T park. The Bay Area Council has an office in HFSTC that helps Californian companies register new enterprises. In addition, HFSTC issued guidelines to attract domestic and international talents to accelerate the development of the AI industry in July 2017, offering 3 million RMB in settling expenses and 15 million RMB in subsidies for office space costs.[14] Overall, Hangzhou boasted the highest growth in attracting returning graduates from international universities from 2017 to 2018, according to a report by recruitment site Boss Zhipin.[15]

Hefei's China Speech Valley

The history of China Speech Valley (中国声谷), in Hefei, the capital of Anhui province, dates back to 2012 when the Ministry of Industry and Information Technology and Anhui province signed an agreement to establish China's first national-level AI industrial area.[16] In the wake of the AIDP, the Anhui provincial government and Hefei's municipal government released a flurry of guidelines to support AI development in China Speech Valley, aiming to spend 3.2 billion RMB on intelligent speech technology by 2020.[17] Initial indicators are promising. From 2017-2018 more patent applications came from Hefei than Shanghai or Guangzhou. As of January 2019, China Speech Valley hosted more than 300 companies, including iFlytek, Huami and Tencent's Youtu lab.[18]

'Both Hangzhou and Hefei have a key 'anchor tenant' tech company and an elite university that glue the ecosystem together...Global and extra-regional linkages are also necessary to build competitive clusters in a world of globalised innovation'

As with Hangzhou AI Town, Hefei's China Speech Valley contains many of the ingredients needed for successful hybridised industrial ecosystems. Designated as the 'national team' member (see Eugeniu Han's essay on smart cities for more on the 'national team' concept) on intelligent speech, iFlytek could play a role similar to Alibaba in Hangzhou, circulating research from local universities such as the University of Science and Technology of China (USTC) and knowledge spillovers to local companies. To that end, iFlytek has built an open platform (iFLYOS) to make it easier for developers to build speech recognition applications.[19] USTC, like Zhejiang University, is a member of the elite C9 group of Chinese universities. In fact, of the top 10 Chinese universities in terms of publication output from 2003–2012, the University of Science and Technology of China has the highest average citation impact.[20]

There are reasons to doubt the fitness of Hefei's AI ecosystem, at least relative to that of Hangzhou and the first-tier cities. Compared to Hangzhou, China Speech Valley's international linkages are not as strong and there is less appeal for graduate returnees to work in Hefei. The entrepreneurial spirit of Hefei's AI ecosystem still drastically trails that of the leading cities. From 2013 to November 2018, only 20 AI startups emerged from Hefei; this comprised 2 per cent of Beijing's total.[21] While Hefei's specialisation in speech recognition is a rational adaptation to local conditions, a more diverse base of AI capabilities may be required to better sustain the ecosystem.

How effective is local action?

While the initiatives in Hangzhou and Hefei are an extensive front of subnational action in China's AI development, it is important not to overstate the scale and success of local activity. Consider the over-hyping of Tianjin's $15 billion AI fund, highlighted in a 2020 report by the Institute for Defense Analyses. Upon deeper investigation, the unfavourable comparisons to US federal government expenditures on AI R&D were misleading because 'the announced expenditure from the Tianjin government does not appear to be annualised, focused on R&D, come from the central government, or consist of an actual outlay of money.'[22]

Sometimes local government targets for AI development are simply infeasible. For instance, Sichuan province's target for the scale of its core AI industry is 1000, double that of Beijing's target even though Beijing is home to more than 20 times the number of AI firms as Sichuan.[23] Similar tendencies have informed other Chinese S&T policies. As one analysis of financing schemes to implement 'Made in China 2025' concludes, 'local authorities clearly tend to overstate the size of collected funds in order to signal compliance with central government policies, and funds pledged are often much higher than those eventually deployed.'[24] Another danger is that overeager intervention by local governments distorts the market, as was the case with the overinvestment in the solar photovoltaics industry.[25]

As the brief studies above of Hangzhou's AI Town and Hefei's China Speech Valley demonstrate, a better understanding of local industrial ecosystems is a first step to getting empirical evidence on the effect of local implementation of China's AI strategy.

Conclusion and lessons for Europe

Are there any lessons that Europe and the UK could adapt from China's approach to developing industrial AI ecosystems? After all, Chinese provinces are governed like large socioeconomic entities with populations and economic outputs that surpass some European countries, and 'it is often said that a province is to China what a country is to Europe.'[26]

One starting point is to shift the focus toward clusters outside of Berlin, London and Paris, in order to find the dark horse AI ecosystems that could facilitate Europe-wide diffusion of AI technologies. These ecosystems cannot be conjured from thin air. One PhD dissertation on S&T parks found that parks built in areas with a weak local technological base resulted in a glorified real estate project that was disconnected from the local economy.[27] The ecosystems in Hangzhou and Hefei provide some additional guidance. In both cases, there is a key 'anchor tenant' (iFlytek for Hefei and Alibaba for Hangzhou) and elite university (USTC for Hefei and Zhejiang University for Alibaba) that glue the ecosystem together. Empirical analysis has shown that the presence of a large technology firm can enhance the productivity of local innovation systems by enabling local university research.[28]

With an eye towards attracting the global and extra-regional linkages that are necessary to build competitive clusters in a world of globalised innovation, ambitious policymakers could provide increased access to capital and other policy supports to AI ecosystems in the making in cities such as Amsterdam, Barcelona and Stockholm.[29] In the United Kingdom, outside of London, other important clusters are forming in Cambridge, Manchester, Oxford and elsewhere. Already a leading hub for advanced materials, Manchester could have the ingredients to build a dark horse AI ecosystem around the University of Manchester and the establishment of a government communications headquarters as anchors.[30]

Finally, in line with Hefei's approach, European governments should better assess opportunities to specialise in specific AI subdomains or parts of the value chain. There is no need to reinvent the wheel. Manchester, for instance, has adapted the national Artificial Intelligence and Data Grand Challenge, which is focused on health care applications, toward local advantages in smart manufacturing and cybersecurity. The European Commission's Key Enabling Technologies strategy has examined European comparative advantages in six technology groups: micro and nanoelectronics, nanotechnology, industrial biotechnology, advanced materials, photonics, and advanced manufacturing technologies.[31] Given its potential to enable all six of these enabling technologies, incorporating AI into the Key Enabling Technologies framework could help sustain the competitiveness of Europe's industrial base.

Endnotes

[1] Matt Sheehan, 'How China's Massive AI Plan Actually Works', Macro Polo, 12 February 2018, https://macropolo.org/analysis/how-chinas-massive-ai-plan-actually-works/.

[2] Jeffrey Ding, 'Deciphering China's AI Dream', *Future of Humanity Institute Technical Report*, March 2018, 3-4.

[3] The concept of 'hybridised industrial ecosystems' comes from Kim 2019. Sung-Young Kim, 'Hybridized Industrial Ecosystems and the Makings of a New Developmental Infrastructure in East Asia's Green Energy Sector', *Review of International Political Economy* 26, no. 1 (2019): 158-182.

[4] Loren Brandt and Thomas G. Rawski, 'Policy, Regulation, and Innovation in China's Electricity and Telecom Industries', in *Policy, Regulation and Innovation in China's Electricity and Telecom Industries* (Cambridge: Cambridge University Press, 2019), 1-51, 13.

[5] Dan Breznitz and Michael Murphree, *Run of the red queen: Government, innovation, globalization, and economic growth in China.* (New Haven: Yale University Press, 2011).

[6] Economic Reference News, 'From "Looking Up at the Stars" to "Landing at Scale": More Than 20 Chinese Provinces Issue AI Policies', [从"仰望星空"到"规模化落地" 中国逾20省发布人工智能政策], *Economic Reference News* [经济参考报], 27 February 2019, http://www.sinovision.net/finance/201902/00459048.htm.

[7] Jeffrey Ding, 'ChinAI Newsletter no. 4: 2018 is the Year of AI Policy at the Local Level', *ChinAI Newsletter*, 2 April 2018, https://chinai.substack.com/p/chinai-newsletter-4-2018-is-the-year-of-ai-policy-at-the-local-level.

[8] Yue Hu et al., *Research on the Impact of AI and Rig Data on the Trade Between China and other Countries along the Belt: Take the Application of Medical Economy as an Example* (Changchun: SSER 2019, 2019), https://webofproceedings.org/proceedings_series/ESSP/SSER%202019/SSER30212.pdf.

[9] Thomas Lehmann, 'AI Politics is Local', *DigiChina Special Report: AI Policy and China*, 23 January 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/ai-politics-local/.

[10] One proxy was demand for AI servers. Jeffrey Ding, 'ChinAI No. 28: Chips and Compute: Alternative Perspectives and City Rankings', *ChinAI Newsletter*, 17 September 2018, https://chinai.substack.com/p/chinai-newsletter-28-chips-and-compute-alternative-perspectives-and-city-rankings.

[11] Julie T. Miao et al., 'The Trials of China's Technoburbia: The Case of the Future Sci-tech City Corridor in Hangzhou', *Urban Geography* 40, no. 10 (2019): 1443-1466.

[12] 'Announcement on the First Batch of Subsidy Funds Allocated for the Government Subsidy Project of Artificial Intelligence Towns in 2019', [关于拨付人工智能小镇政府补助项目2019年第一批补助资金的公示], Hangzhou Future Sci-Tech City, 2 December 2019, http://www.yuhang.gov.cn/art/2019/12/2/art_1601762_40715127.html.

[13] Some companies showed up multiple times.

[14] 'Policy Analysis: Summary of AI Policies', *Deqing Makerspace* [政策解读 | 人工智能政策汇总], 31 January 2019, https://www.sohu.com/a/292560106_100179411.

[15] ECNS, 'Hangzhou Leads Nation in Attracting Overseas Returnees',*ECNS Wire*, 28 June 2018, http://www.ecns.cn/news/cns-wire/2018-06-28/detail-ifyvrptq6363759.shtml.

[16] Fan Nana, 'Known as "China's Speech Valley", How Strong is Hefei's AI Technology Foundation?', *The Beijing News Smart City Institute* [新京智库], 11 January 2019, https://mp.weixin.qq.com/s/hLKc1gBk63N5yRKl-sPQGw?fbclid=IwAR1mcf0zNZHkhDWxMqADldK4_3PabgvHspX8lFEQeGNf4nxN2II5eAPFnNA.

[17] Jaqueline Ives and Anna Holzmann, 'Local Governments Power Up to Advance China's National AI Agenda', *MERICS blog: European Voices on China*, 24 April 2018, Mercator Institute for China Studies (MERICS), https://www.merics.org/en/blog/local-governments-power-advance-chinas-national-ai-agenda.

[18] Nana, 'China's Speech Valley'.

[19] Jeffrey Ding, 'ChinAI no. 51: China's AI "National Team"', *ChinAI Newsletter*, 20 May 2019, https://chinai.substack.com/p/chinai-51-chinas-ai-national-team.

[20] Kristen Bound et al., China's *Absorptive State*, (London: Nesta, October 2013), 45.

[21] Nana, 'China's Speech Valley'.

[22] Thomas J. Colvin et al., *A Brief Examination of Chinese Government Expenditures on Artificial Intelligence R&D*, (Washington DC: The Institute for Defense Analyses, February 2020), https://www.ida.org/-/media/feature/publications/a/ab/a-brief-examination-of-chinese-government-expenditures-on-artificial-intelligence-r-and-d/d-12068.ashx.

[23] 'An Article That Takes You Through the Latest Policies for the AI Industry Throughout the Country in 2018', *Forward Looking Industry Research Institute* [前瞻产业], 30 March 2018, http://www.qianjia.com/html/2018-03/30_288481.html.

[24] Max J. Zenglein and Anna Holzmann, 'Evolving Made in China 2025', *Mercator Institute for China Studies*, 2019, 12.

[25] Xiaolan Fu, Wing Thye Woo, and Jun Hou, 'Technological Innovation Policy in China: The Lessons, and the Necessary Changes Ahead', *Economic Change and Restructuring* 49, no. 2-3 (2016): 139-157.

[26] Cheng Li, 'Political Localism Versus Institutional Restraints: Elite Recruitment in the Jiang Era', in *Holding China Together*, ed. Dali L. Yang and Barry J. Naughton (Cambridge: Cambridge University Press, 2004), 29-69, 34.

[27] Achilleas Tsamis, 'Science and Technology Parks in the Less Favoured Regions of Europe: An Evaluation of Their Performance and the Parameters of Success' (PhD diss., London School of Economics, 2009), 389. etheses.lse.ac.uk/2543/1/U615492.pdf.

[28] Ajay Agrawal and Iain M. Cockburn, *University Research, Industrial R&D, and the Anchor Tenant Hypothesis* (Cambridge, MA: National Bureau of Economic Research, 2002).

[29] James Manyika, *10 imperatives for Europe in the age of AI and automation* (McKinsey Global Institute, October 2017), https://www.mckinsey.com/featured-insights/europe/ten-imperatives-for-europe-in-the-age-of-ai-and-automation.

[30] There are few large technology giants based in Manchester, though Microsoft has established an office in the city. 'Greater Manchester Local Industrial Strategy', *HM Government*, June 2019, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/808579/greater-manchester-local-industrial-strategy-single-page.pdf.

[31] 'Key enabling technologies policy', *European Commission*, https://ec.europa.eu/growth/industry/policy/advanced-technologies_en.

AI in action: China's use of AI in public services


abcl

02.
Breaking the Iron Triangle: AI in China's healthcare system

By Dr Andy Chun
Adjunct professor at City University of Hong Kong
Council member and convenor of the AI Specialist Group, Hong Kong Computer Society

Over the past decade China has shifted rapidly from a manufacturing-based economy to a knowledge-based economy, with technology and innovation a major policy focus. It is becoming a leading country in the application of artificial intelligence (AI) across key industries. AI relies heavily on data for machine learning and predictive analytics, and China has no shortage of data with its 1.4 billion population constantly generating massive amounts of it in real time. The Chinese people have also shown themselves to be eager users of technology and early adopters of AI.

In healthcare, China, like many other countries, is facing the formidable challenge of providing easy access to affordable and high-quality services for its ageing population. With its use of AI combined with readily available medical and biometric data, China is on its way to providing personalised healthcare to more people at a lower cost, while keeping them healthier in the first place through continuous monitoring and alerts. It is also using AI to assist medical professionals in decision-making, track patient outcomes, reduce physician workload and speed up drug development. China's investment in AI and healthcare over the past few years is now helping it deal with the coronavirus crisis, explored further in a case study in this essay.

While privacy around medical data is a clear issue that requires appropriate regulation, there are lessons for European policymakers in China's 'experiment and innovate first' approach to AI and its appreciation that time to market is crucial if medical innovation is not to be stifled.

AI for healthcare

Countries around the world are turning to digital technology and AI to provide more cost-effective and highly scalable means to deliver services. This is particularly important for China with its ageing population and shortage of medical staff. Medical services can be scarce in China's rural areas, while in urban areas services are highly strained due to large patient volumes. According to the Organisation for Economic Cooperation and Development (OECD), China has two practicing doctors per 1,000 persons, compared with 2.9 for UK and 4.3 for Germany.[1] It also has only 2.7 nurses per 1,000 persons, compared with 7.8 for UK and 12.9 for Germany.[2] In addition, China is ageing more rapidly than almost any country in the world, due mainly to its previous one-child policy. By 2050, China will have around 330 million people over 65 with nearly a third of its population over 60.[3]

AI offers the potential to break out of the 'iron triangle' dependencies of healthcare priorities: access, cost and quality. These inherent trade-offs traditionally mean we cannot improve one priority without sacrificing another. However, AI could democratise health care and boost access for underserved communities while lowering costs and maintaining, if not also improving, quality. According to the consultancy firm Frost and Sullivan, AI has the potential to improve the outcomes of medical treatment by 30–40 per cent and reduce costs by as much as 50 per cent.[4]

In 2017, when China's State Council laid out its AI strategic plan, it called for the development of a whole gamut of AI and AI-related healthcare technologies, such as intelligent diagnosis, wearables, AI health monitoring, robot-assisted surgeries, intelligent medical image recognition and medical genomics, with a strong emphasis on elderly care. In terms of AI research, a study by the Allen Institute for Artificial Intelligence predicts that China will overtake the US in total number of top 10 per cent AI research papers by 2020, and top 1 per cent papers by 2025.[5]

Ways AI can help healthcare

AI captures medical knowledge

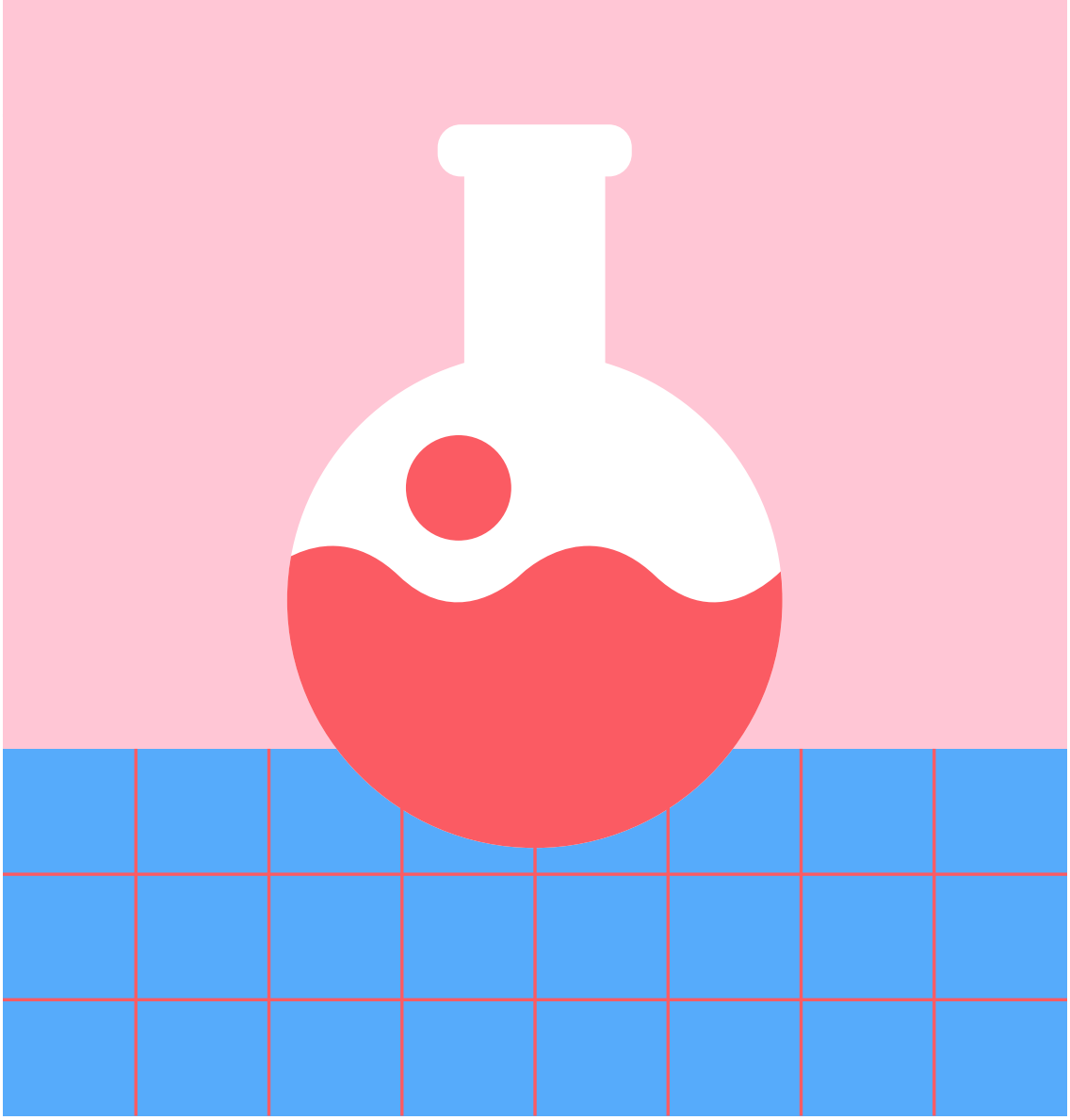
Medical professionals must learn, recall and apply vast amounts of medical textbook knowledge while keeping up to date with the newest medical research/journals and pharmaceutical products. This is exactly what AI and cognitive computing are best at. AI can use natural language processing (NLP) and machine learning to read and understand millions of online documents and data points to help diagnose and recommend treatments. An example of AI cognitive computing is IBM Watson for Oncology, which has been shown to offer cancer treatment recommendations that are on par with human doctors. It has been reported that IBM Watson is already in use at 65 large hospitals in 39 cities in China,[6] the ethical implications of which Danit Gal explores.

Researchers in China are also using AI to capture general medical knowledge. For example, in 2017, iFlytek and Tsinghua University successfully created an AI system that not only passed the Chinese medical licensing exam but scored better than over 96 per cent of people taking the exam.[7] This is an important AI accomplishment as the exam not only tested breadth of knowledge but also the ability to understand intricate connections between facts and use them to reason and make decisions.

In 2018, the Guangzhou Women and Children's Medical Center created an AI prototype that can diagnose common childhood diseases with accuracy comparable to experienced paediatricians. It uses NLP and deep learning to extract clinically relevant information from 1.4 million electronic health records (EHR) in paediatric patients.[8] Applications like this can help streamline frontline patient care, such as triaging patients.

Deep learning for medical image processing

The use of AI deep learning for medical image processing has been one of the most promising applications of AI in the medical field. Artificial neural networks (ANN) and deep learning have been used to process various medical images such as MRI scans, CT scans and X-ray images. With a large enough database of medical images, AI deep learning can learn to recognise visual symptoms of potential diseases and illnesses. For example, using deep learning on retinal images can provide early detection of diabetic retinopathy with accuracy similar to that of experts. The same retina scanning technique has been used to predict heart disease. In China, researchers have been using AI on eye scans to diagnose congenital cataracts as accurately as human doctors; an estimated 200,000 children are bilaterally blind from cataracts annually.[9] Besides retina scans, radiology departments at top Chinese hospitals routinely handle tens of thousands of scanned images per day. AI deep learning is already used to relieve workload by automatically analysing and highlighting abnormalities. Big tech players, such as Alibaba and Tencent, are investing heavily in this area. Alibaba uses AI to help interpret CT scans, while Tencent's Miying uses AI to help detect early signs of cancer.



In a competition designed by the Stanford Machine Learning group to compare AI to human experts in interpreting chest X-rays, a Chinese startup, JF Healthcare, became the first to beat Stanford University radiologists in July 2019.[10] JF Healthcare provides remote diagnostic services, focusing on chest X-rays, for rural township hospitals in China where certified radiologists are often not available. It also provides mobile screening trucks to screen for tuberculosis in more than 10 provinces in China. By the end of 2019, it expects to have offered half a million people the possibility of early detection while preventing the spread of disease to broader populations.

AI helps drug design

AI is also changing how drugs are discovered, making research for new pharmaceuticals quicker, cheaper and more effective. Traditional drug development may take over a decade to complete, costing hundreds of millions to billions of USD. AI greatly reduces the time and money needed by using machine learning to generate candidates automatically from existing drugs.[11] [12]

AI helps people to stay healthy

More affordable healthcare wearables that track activities and heart rate mean that consumers are actively taking responsibility in monitoring their own health. According to a Mintel 2017 report, 52 per cent of Chinese consumers own a smart wristband and 42 per cent own a smartwatch,[13] making China the largest market for wearables globally.[14] Big data and AI predictive analytics can continuously monitor and alert users if there are any abnormalities and before the outset of more major medical problems. Insurance companies, such as China's Ping An Health, are starting to integrate wearables into their offerings, providing customers who live healthier lifestyles with discounts and rewards.

AI doctors fill a gap
China, with its shortage of doctors, has turned to AI chatbot technology and cognitive computing to make getting medical advice more accessible. For example, China's Ping A Good Doctor has an AI Doctor chatbot that uses AI to gather medical history and generate a diagnosis plan before passing it to a specialist consultant via telemedicine.[15][16]

The AI Doctor is also used in unstaffed booths, called One-Minute Clinics, helping people in rural areas to get 24/7 medical consultations from doctors in other cities. About 1,000 of these booths were to be installed across eight provinces and cities in China in 2019 to service more than three million patients. The booths also fill prescriptions and dispense up to 100 common medications.[17] Tencent has a similar service called WeDoctor, which village doctors in rural areas use to support their work. Doctors simply input a patient's symptoms and the AI system suggests diagnoses and treatments, based on knowledge of more than 5,000 symptoms and 2,000 diseases. WeDoctor claims an accuracy rate of 90 per cent and serves more than 90,000 patients daily.[18]

AI makes health protection affordable
Creating affordable medical protection is another way to reduce strain on public sector health services. One example is Ant Financial's Xiang Hu Bao, an online mutual aid platform targeting China's poor who cannot afford traditional commercial medical insurance. More than 100 million people signed up in the first year, mostly migrant workers or people living in rural areas. It is free to sign up with no premiums or upfront payments and the monthly fee is based on actual payouts, with risks and expenses shared equally across all members. For 2019, this fee averaged a few US dollars, while covering 100 critical illnesses. The monthly contribution is low because AI automates processing without needing humans.

AI also automatically processes claims and reads medical records. Fraud is prevented by AI verifying authenticity of hospital documents through image recognition of official seals from more than 10,000 hospitals in China. Blockchain technology further ensures trust in data stored in the system.[19]

Case study: Using AI to control COVID-19

There is no greater challenge to the public health system than a pandemic, such as the novel coronavirus outbreak that emerged in Wuhan, China. For a fast-spreading virus like COVID-19, time is of the essence and China's experience in quickly converting AI technology into practical applications has proved useful. AI-powered applications are being used to combat the virus by speeding up screening, diagnosis and new drug development.

Screening
AI is used in infrared cameras to scan and screen people en masse for high temperature at airports and railway stations across China. Together with facial recognition cameras, the cameras can identify the person with a high temperature, even when wearing a surgical mask.[20]

Sequencing
AI was used soon after the outbreak was discovered to help sequence the new coronavirus genome, so that research on a vaccine could get started. In the previous SARS outbreak in late 2002, scientists had to wait months before they knew what the germ looked like. With COVID-19, it only took Chinese scientists using AI a few weeks to sequence the virus's genome before sharing it online with researchers around the world.[21]

'China's experience in quickly converting AI technology into practical applications has proved useful in tackling COVID-19'

Testing
Knowing the genome sequence allows researchers to design lab tests to identify the presence of the virus. Unfortunately, the new COVID-19, like SARS and MERS, is also a single-stranded RNA virus, which means it is susceptible to mutation and harder to test. Using AI, Alibaba was able to counter this problem and reduce lab test time to diagnosis coronavirus infection from hours to only 30 minutes, with a new 'whole genome detection' approach.[22]

Scanning
Besides lab tests, a CT scan of the lungs is also an effective way of detecting signs of coronavirus infection. During an epidemic, radiologists are overwhelmed with thousands of scans to inspect each day. China turned to AI to automate this process. Alibaba trained an AI deep learning system using thousands of CT scans from confirmed cases. The resulting AI model was then able to analyse a CT scan within 20 seconds with 96 per cent accuracy of detecting COVID-19.[23] [24]

R&D
AI companies in China, such as Alibaba and Baidu, have been offering free and open source AI technology, datasets and AI compute power to public research institutions around the world to help shorten R&D time in combating the disease. Baidu also open-sourced its AI LinearFold algorithm that predicts RNA virus 3D structures and behaviours and allows scientists to understand better how the virus invades our cells so that matching vaccines can be created. Using LinearFold, prediction time can be reduced from 55 minutes to 27 seconds.[25] Alibaba also made its AI drug discovery platform available to researchers. Historical data on drug R&D efforts for coronaviruses, such as SARS and MERS, allows AI prediction models to provide insights on how existing drugs might be repurposed for COVID-19 and thus improve the efficiency of new drug screening.[26]

Advice
AI chatbot technology is used to help reduce pressure on hospital and government frontline personnel by automatically answering inquiries from the public. AI chatbots also provide advice to individuals on whether they need to go to a hospital for screening or stay at home for the 14-day quarantine. In Shanghai and other cities, AI chatbots automatically call high-risk patients to assess their condition.[27] Physical AI robots are also used in segregated wards in Shanghai hospitals to disinfect those wards, as well as deliver food and medications and check body temperatures.

Quarantine
Facial recognition is used to ensure affected citizens are following the 14-day self-quarantine requirement. According to reports[28], people who violated the quarantine and did not stay at home were automatically detected by facial recognition cameras and contacted by the police as well as employers. A Chinese company created an AI system that allows users to check if they have recently travelled with someone who has contracted the new coronavirus. The system uses public data and other information sources to correlate with an individual's travel itinerary. It was reported[29] that more than 21 million people used the service within two days of its launch.

Potential issues with using AI

As in other countries, the use of AI has raised some concerns about ethics and data privacy, particularly AI applications that involve facial recognition and surveillance. However, individuals' privacy need not be sacrificed if technology is implemented properly[30] and related regulations are put into place. Towards this end, last year China introduced its Governance Principles for the New Generation Artificial Intelligence[31] to provide guidelines for legal and ethical use of AI. These Principles are considered further in Danit Gal's essay. When used for public health and social good, such as analysing CT scans, AI algorithms may not need to consider personal information. In other cases, where facial recognition and surveillance of quarantined individuals is involved, balancing privacy with the pressing needs of the health of a vast population might be challenging. This is particularly so in the dire situation of a pandemic, where potentially hundreds of thousands, if not millions, of lives are at stake. Achieving a greater good for humanity might be more urgent. For example, countries including China, South Korea, Israel and Iran all track citizens' smartphones to enforce quarantine and to check the movements of people tested positive for the virus.

Conclusion

With its 2017 AI Strategic Plan, it is evident that China sees AI as a crucial component for future economic growth. However, advances in AI not only benefit businesses and industries but can also solve many of China's deep-rooted social challenges, such as the uneven distribution of resources in health care. The use cases described in this essay are examples of how AI, with its ability to learn from massive data, can reduce inefficiencies as well as stepping in when human professionals are scarce. AI, together with other digital technologies, can transform vital public services to be highly scalable and readily available at a very low cost, even to rural areas of China.

It is less than three years since China announced its AI strategy yet it has achieved widespread use of AI across all industry sectors, benefiting all aspects of daily life in China, and at a pace unseen elsewhere. For Europe and North America, there are some lessons to be learned. China's approach to AI and other new technologies is to experiment and innovate first, while putting in place supporting policies and regulations as it goes. Time to market is crucial. This is in stark contrast to western economies, where lengthy policies and regulations usually precede innovation use. There may be a place for a middle ground where new technologies can be tried out within regulatory sandbox environments so as not to stifle innovation growth. In addition, China has a unique edge on AI development – the vast amounts of data constantly produced by its 1.4 billion population are the fuel for AI's machine learning. Chinese citizens are also more willing to share data in exchange for convenience and low-cost services. To mitigate the data disadvantage, western countries may need to find ways to share their datasets while maintaining data privacy, such as using new federated approaches to machine learning. Finally, the stark reality of a global pandemic has highlighted where China's investment in AI is yielding results in speeding up some of the processes involved in combating the coronavirus. While there may be social and democratic implications in some elements of this approach, it offers a path that other countries might learn from.

Endnotes

[1] 'Doctors', OECD Data, accessed March 12, 2020, https://data.oecd.org/healthres/doctors.htm.

[2] 'Nurses', OECD Data, accessed March 12, 2020, https://data.oecd.org/healthres/nurses.htm.

[3] Yanzhong Huang, 'Population Aging in China: A Mixed Blessing', *Council on Foreign Relations*, 4 November 2013, https://www.cfr.org/blog/population-aging-china-mixed-blessing.

[4] 'From $600 M to $6 Billion, Artificial Intelligence Systems Poised for Dramatic Market Expansion in Healthcare', *Frost & Sullivan*, 5 January 2016, https://ww2.frost.com/news/press-releases/600-m-6-billion-artificial-intelligence-systems-poised-dramatic-market-expansion-healthcare/.

[5] Field Cady and Oren Etzioni, 'China May Overtake US in AI Research', *AI2 Blog, 13 March 2019*, Medium, https://medium.comai2-blog/china-to-overtake-us-in-ai research-8b6b1fe30595.

[6] Wang Xiaodong, 'Merger of IT, Medical Care to be Fostered', *China Daily*, updated 7 May 2018, http://www.chinadaily.com.cn/a/201805/07/WS5aef9beea3105cdcf651c45a.html.

[7] Ma Si and Cheng Yu, 'Chinese Robot Becomes World's First Machine to Pass Medical Exam', *China Daily*, updated 10 November 2017, https://www.chinadaily.com.cn/bizchina/tech/2017-11/10/content_34362656.htm.

[8] Huiying Liang et al., 'Evaluation and Accurate Diagnoses of Pediatric Diseases Using Artificial Intelligence', *Nat Med* 25, 433–438 (2019), https://doi.org/10.1038/s41591-018-0335-9.

[9] Charles Q. Choi, 'Computer Diagnoses Cataracts As Well As Eye Doctors Can', *Live Science*, 31 January 2017, https://www.livescience.com/57696-artificial-intelligence-cataract-diagnosis.html.

[10] Jasmine Pennic, 'AI vs Humans: AI Solution Beats Stanford Radiologists in Chest X-ray Diagnostics Competition', *HIT Consultant*, 22 August 2019, https://hitconsultant.net/2019/08/22/ai-tech-beats-radiologists-in-stanford-chest-x-ray-diagnostic-competition/#.XmrDrqhKjb1.

[11] Zachary Hendrickson, 'New research Shows an AI-Powered System Can Develop Potential New Drugs in Just Three Weeks', Business Insider, 4 September 2019, https://www.businessinsider.com/insilico-medicine-ai-system-develops-drugs-in-three-weeks-2019-9.

[12] 'Faster, Cheaper, Better: How China and AI are Helping Pharmaceutical Development', technode, 24 April 2018, https://technode.com/2018/04/24/better-drugs-through-chinese-ai/.

[13] '43% of Chinese Consumers Would Buy Wearable Devices for Themselves', *Mintel*, 13 February 2017, https://www.mintel.com/press-centre/technology-press-centre/43-percent-chinese-consumers-would-buy-wearable-devices.

[14] IDC, 'New Product Launches Drive Double-Digit Growth in the Wearables Market, Says IDC', *IDC*, 3 December 2018, https://www.idc.com/getdoc.jsp?containerId=prUS44500418.

[15] Telemedicine allows health care professionals to evaluate, diagnose and treat patients in remote locations using telecommunications technology, such as smartphones and video chat.

[16] Tim Hornyak, 'What America can learn from China's use of robots and telemedicine to combat the coronavirus', *CNBC*, 18 March 2020, https://www.cnbc.com/2020/03/18/how-china-is-using-robots-and-telemedicine-to-combat-the-coronavirus.html.

[17] Dean Koh, 'Ping An Good Doctor Launches Commercial Operation of One-Minute Clinics in China', *MobiHealthNews*, 7 January 2019, https://www.mobihealthnews.com/news/asia-pacific/ping-good-doctor-launches-commercial-operation-one-minute-clinics-china.

[18] Amy Hawkins, 'How Elderly, Sickly Farmers are Quenching China's Thirst for Data', *Wired*, 12 April 2019, https://www.wired.co.uk/article/china-ai-healthcare.

[19] Alipay, 'Alipay's Xiang Hu Bao Online Mutual Aid Platform Attracts 100 Million Participants in One Year', *Business Wire*, 27 November 2019, https://www.businesswire.com/news/home/20191126005952/en/Alipay%E2%80%99s-Xiang-Hu-Bao-Online-Mutual-Aid.

[20] Isaac Scher, 'Chinese Subways are Using Artificial Intelligence Facial Recognition Scanners to Help Detect Whether People Have Coronavirus', *Business Insider*, 11 February 2020, https://www.insider.com/chinese-tech-companies-artificial-intelligence-surveillance-coronavirus-2020-2.

[21] 'Novel Coronavirus – China', *World Health Organization*, 12 January 2020, https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/.

[22] Alibaba Clouder, 'How Is Alibaba DAMO Academy Helping to Fight the Outbreak of the Novel Coronavirus?', *Alibaba Cloud*, 6 March 2020, https://www.alibabacloud.com/blog/how-is-alibabadamo-academy-helping-to-fight-the-outbreak-of-the-novelcoronavirus_595954.

[23] Sun Henan, 'Alibaba Says AI Can Identify Coronavirus Infections With 96% Accuracy', *Nikkei Asian Review*, 19 February 2020, https://asia.nikkei.com/Spotlight/Coronavirus/Alibaba-says-AI-can-identify-coronavirus-infections-with-96-accuracy.

[24] Tom Simonite, 'Chinese Hospitals Deploy AI to Help Diagnose Covid-19', *Wired*, 26 February 2020, https://www.wired.com/story/chinese-hospitals-deploy-ai-help-diagnose-covid-19/.

[25] David Aikman and Alan Chan, 'Five Ways Chinese Companies are Responding to Coronavirus', *World Economic Forum*, 20 February 2020, https://www.weforum.org/agenda/2020/02/coronavirus-chinese-companies-response/.

[26] Sun Henan, 'Chinese Tech Giants Open Cloud Computing Technology to Researchers Amid Coronavirus Emergency', *KrAsia*, 31 January 2020, https://kr-asia.com/chinese-tech-giants-open-cloud-computing-technology-to-researchers-to-help-fight-coronavirus-outbreak.

[27] Sarah Dai, 'AI Applications Surge as China Battles to Contain New Coronavirus Epidemic', *Tech in Asia*, 4 February 2020, https://www.techinasia.com/ai-surge-china-coronavirus.

[28] Shawn Yuan, 'How China is Using AI and Big Data to Fight the Coronavirus', *Aljazeera*, 1 March 2020, https://www.aljazeera.com/news/2020/03/china-ai-big-data-combat-coronavirus-outbreak-200301063901951.html.

[29] Zen Soo, 'Worried You Were on a Train With Someone Infected With the Wuhan Coronavirus? This AI-Backed Platform Lets You Find Out', *South China Morning Post*, 30 January 2020, https://www.scmp.com/tech/apps-social/article/3048232/worried-you-were-train-someone-infected-wuhan-coronavirus-ai.

[30] One example of technology implemented properly is 'privacy-preserving' approaches, such as: https://www.apple.com/covid19/contacttracing.

[31] One example of technology implemented properly is 'privacy-preserving' approaches, such as: https://www.apple.com/covid19/contacttracing.

'Governance Principles for the New Generation Artificial Intelligence: Developing Responsible Artificial Intelligence', China Daily, updated 17 June 2020, https://www.chinadaily.com.cn/a/201906/17/WS5d07486ba3103dbf14328ab7.html.


abcl

03. The Future of the Classroom? China's experience of AI in education

By Yi-Ling Liu
Visiting scholar at the NYU Arthur L Carter Journalism Institute

In the wake of the COVID-19 pandemic, China's education system has been forced to adapt. Public schools have suspended all spring term classes and 260 million students have taken their studies online. In this climate of uncertainty, students, educators and policymakers have increasingly realised that they must turn to technological solutions – in particular, the use of artificial intelligence (AI) technologies – to overhaul the existing education system.

But the push for AI education in China began earlier, before the crisis hit. Three years ago, in July 2017, the Chinese government published an ambitious master plan[1] to build the nation into a powerhouse of AI innovation, particularly in education. The government's mandate, fuelled by the massive quantities of student data available to drive new technologies and funded by the significant spending power of Chinese parents in a highly competitive academic culture, laid the foundations for the country's AI education boom. Education technology startups in the private sector, with their access to large datasets, ample venture capital funding and flexibility to experiment with new technologies, have been spearheading the development of AI education.

On one hand, purveyors of intelligent technologies – such as smart schools and adaptive learning systems – believe that the development of AI education holds great potential to address teacher scarcity, offer alternative models of schooling and reshape the traditional learning paradigm. On the other hand, these developments face significant limitations and risks. Can new innovations be implemented and made widely accessible in existing public school systems or will they deepen existing educational inequities? Are AI technologies capable of meaningfully teaching students the complex skills of creativity, collaboration and critical thinking, or will they simply magnify outmoded pedagogical practices? How will schools address and regulate the ethical and privacy concerns around the mass collection and use of student data? These are all questions that apply not only to the Chinese education system but should also be of concern to educators and policymakers in Europe as AI begins to make rapid inroads into education to fill a gap in these uncertain times.

Factors driving China's AI education boom

Three key factors have laid the unique conditions for China's current AI education boom: the government's national push to develop AI technologies, Chinese parents' willingness to pay for their children's education and the vast quantities of student data available to fuel new algorithms. The government's AI innovation plan[2] served as a symbolic mandate, sending a signal to educators, schools and enterprises that both implementing AI education technologies and teaching the skills[3] necessary to develop AI in the nation's schools and universities would be a top priority.

Government rhetoric, paired with tax breaks and incentives[4] for AI ventures improving student learning, have encouraged investors to pour money[5] into education technology and for schools and parents to be open to new models of online learning.

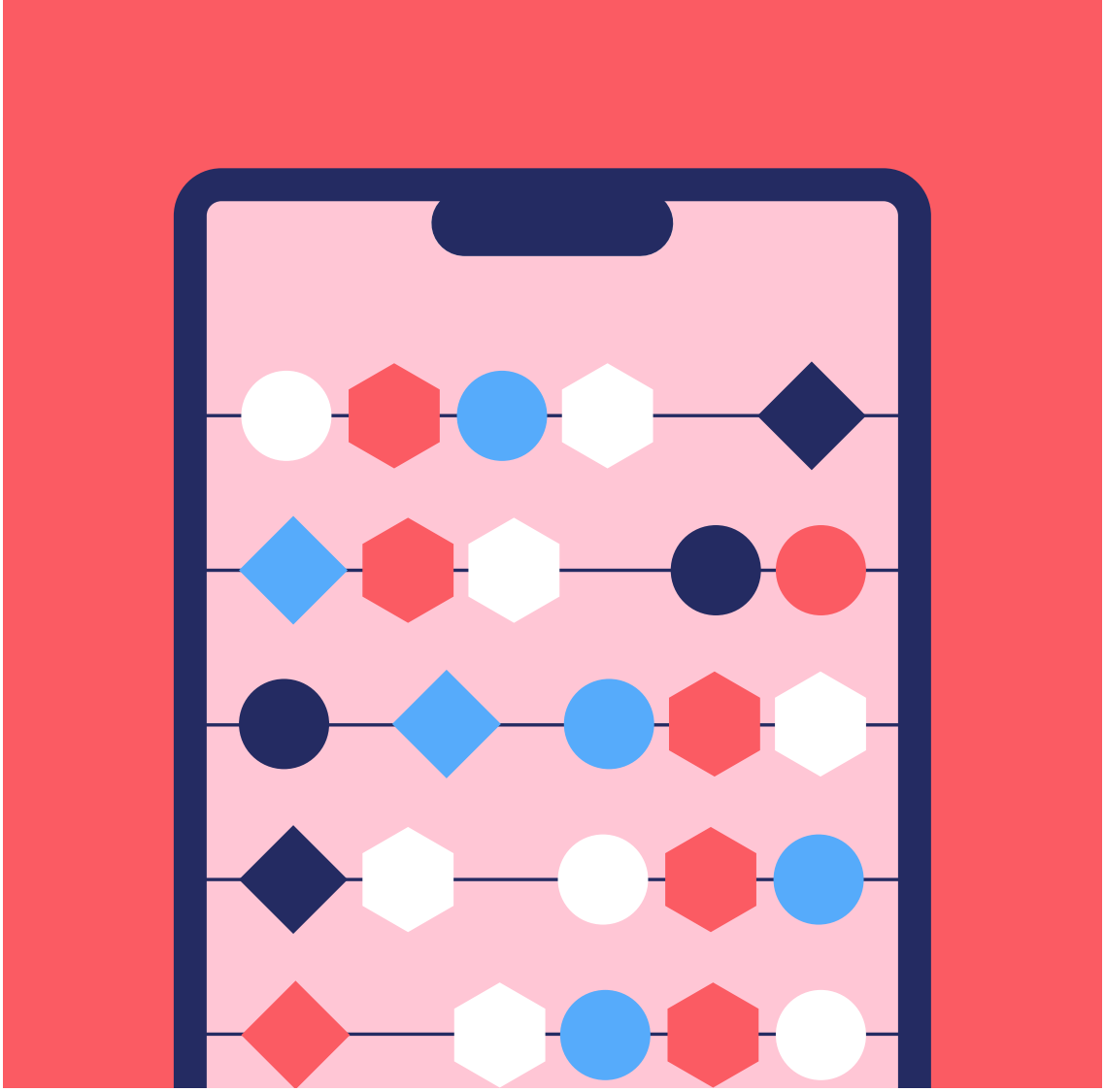
Indeed, the openness of the average Chinese parent to pay for their child's education has been crucial in funding the development of new technologies. In China's highly competitive academic environment, where the national college entrance examination – the gaokao – is a key determinant of a student's success in life, Chinese parents are willing to spend up to half their monthly salary on their children's education, according to venture capitalist and AI expert Kai-Fu Lee. This is in stark contrast to the United States and Europe where parents are much less willing to pay for afterschool education. Total annual spending by households on preschool and primary education reached 1.9 trillion RMB (296 billion USD) in the 2016-2017 school year, comprising some 2.5 per cent of China's GDP in 2016, according to a survey[6] conducted by the China Institute for Education.

Funded by parents and driven by tech entrepreneurs, the Chinese education industry has invited a greater culture of experimentation.[7] Underpinning the growth of the sector is China's large population, which translates into a vast quantity of student data. According to Zhang Yingqian, a professor of computer science teaching a course on AI applications at Xiamen University, 'with more data, new technologies are able to extract more important information and sort students, teachers and learning content much with greater precision, at a more granular scale.'[8] With a larger database to draw from, companies can grow more quickly and refine their machine learning models. For example, a Chinese online tutoring company such as VIPKid,[9] which began as a video conferencing app connecting American teachers with Chinese students, can accumulate more user interactions, make more precise student-teacher pairings and, in turn, make more complex adjustments to its technology.

AI applications in Chinese education

AI technologies are being implemented in China's education system in two key ways: firstly, through the creation of educational tools that require basic automation technology, such as test grading and homework correction, and secondly, through the introduction of adaptive learning systems – a form of AI-driven educational technology that responds to a student's interactions in real time and automatically tailors support to their individual needs. VIPKid, for example, will automatically pair students with teachers according to their preferred learning styles and needs; the online platform 17zuoye[10] uses AI algorithms to streamline the homework distribution and grading process; and Liulishuo,[11] an English-language learning app similar to the American Duolingo, offers an intelligent tutor equipped with real-time personalised feedback systems and automatic grammar correction.

Squirrel AI[12] combines physical and digital schooling, where students can learn remotely online or go to one of the company's bricks-and-mortar learning centres. Through a combination of a human coach and virtual, AI-driven tutor, they receive a personalised lesson plan tailored to learning needs and gaps in their knowledge. Each course is subdivided into the smallest possible conceptual pieces, such that the algorithm can diagnose student gaps in understanding as precisely as possible, adjusting learning pathways in real time. Middle school maths, for example, is broken into 10,000 'knowledge points' such as rational numbers, the properties of a triangle and the Pythagorean theorem. By tailoring educational experiences to the individual student and optimising for their abilities, adaptive learning technologies can reshape the paradigm of learning, says Richard Tong, Squirrel AI's chief technology officer.[13] He compares Squirrel's transformation of the public education system to the way ride-sharing companies such as Uber and Didi changed public transportation systems – providing the same service through a more convenient and personalised pathway.



But this claim is based on the assumption that the most efficient form of education is also the most effective. Like an Uber cab or a GPS navigation device, adaptive learning systems can 'get a user from point A to point B the most efficient way possible,' says Professor Yong Zhao, who specialises in global and online education at the University of Kansas. 'But they do not show the user how to choose their own path, navigate the roads and familiarise themselves with the fabric of the city.'[14] In other words, students taught through intelligent learning systems are not necessarily active decision-makers in their own learning experience. If designed with the goal of efficiency and scalability alone, intelligent learning systems not only overlook but could also strip away one of the most fundamental skills that students need to survive in the 21st century – self-actualisation.

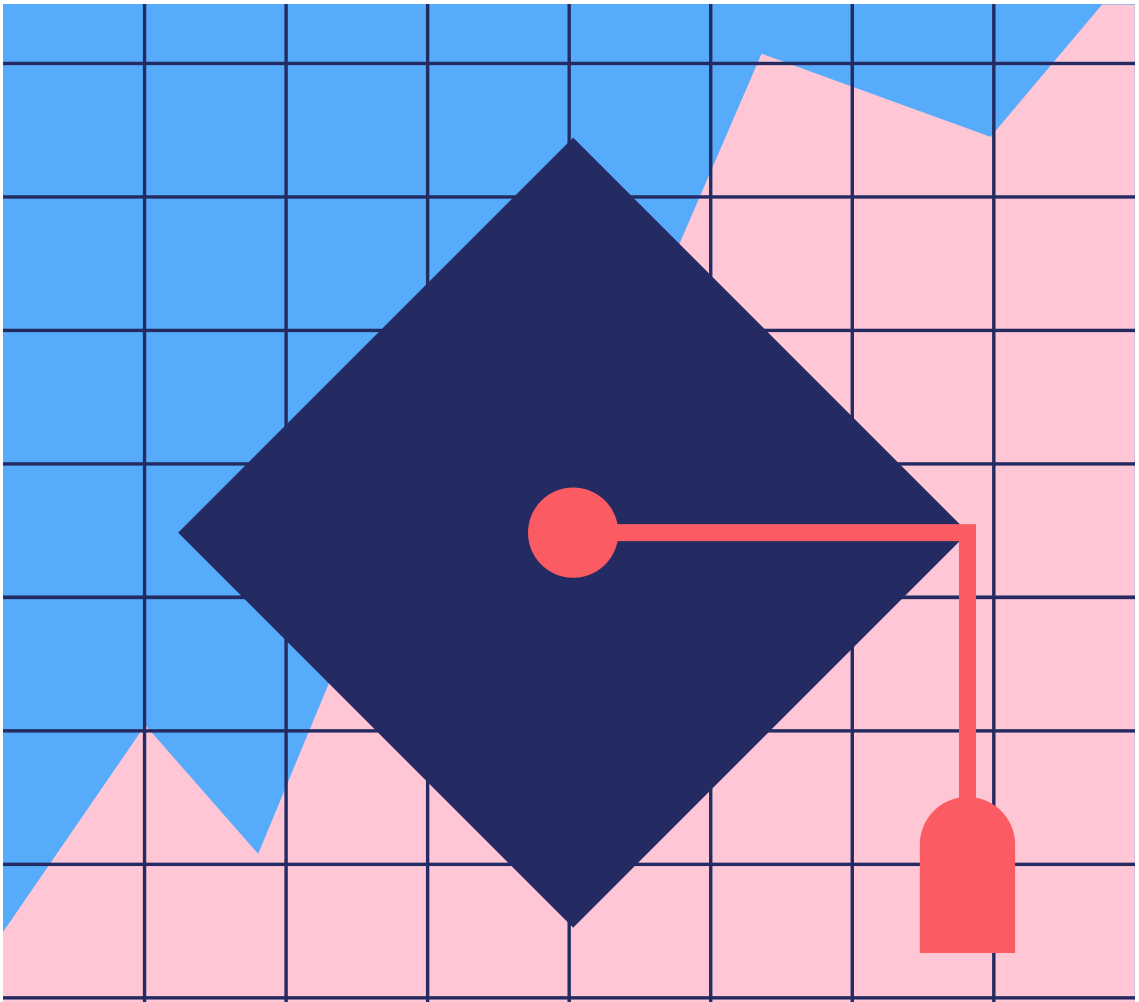
An example of a simpler AI technology that allows for greater student agency is the Smart Learning Partner, a collaboration between Beijing Normal University's Advanced Innovation Center for Future Education and Tongzhou district of Beijing. An AI-driven platform that enables students to connect with a human tutor via their mobile phones, the smart learning partner is used like a dating app, but one that matches students and tutors according to student queries, tutor expertise, availability and ratings. In contrast to the previous examples, the system may 'put students more in control of their own learning'.[15]

Another argument advanced for adaptive learning systems is that they have the potential to reshape the role of the traditional teacher. According to Tong, Squirrel AI's technologies can, for example, replace tasks such as face-to-face instruction with video tutorials, and automate the delivery of student diagnosis and feedback with adaptive learning systems. Freed from the rote tasks of homework correction and test-grading, teachers are able to dedicate more of their time and effort to more complex tasks such as motivating students and facilitating interpersonal collaboration. Their role might shift from that of the instructor, standing in front of the classroom giving lectures, to that of the motivator, focused on the emotional and interpersonal lives of their students.

Despite its promise, it is important to note that teaching is an integrative process that cannot be mechanically divided into the discrete categories of 'automatable' and 'non-automatable'. A good lecture (what an ed tech company would describe as 'content delivery') is a dynamic process that cannot simply be replaced by standardised video tutorials. Grading a paper (what might be described as 'diagnosis') is an activity that familiarises a teacher with a student's personality and thought process, and to automate is to strip away a crucial way in which a teacher might relate to their student. The rhetoric of 'scalability' overlooks the fact that good teaching is highly context dependent – a method of teaching a wealthy student in urban Shanghai may be totally inappropriate for a child raised in poverty in rural Shanxi.

Promises and limitations of AI applications in Chinese education

Given that developments in AI education have been spearheaded by the private sector, one of the most significant challenges is whether or not AI technologies can be widely implemented in the existing Chinese public school system. To continue the ride-sharing analogy, although Uber was able to integrate itself into existing public transportation systems in some countries, elsewhere the company's presence disrupted the local taxi cab industry. Public school systems do not operate at the speed of startups – they are embedded in long-standing bureaucracies, local authorities driven by their own interests and teachers unfamiliar with novel technologies. In order for new innovations to achieve widespread impact, public schools will need to collaborate with private companies and ensure that these collaborations are fundamentally driven by student interests, not maximising the company's profits.



Moreover, for these innovations to have impact, they must be made widely accessible. China's education system remains deeply divided, often between wealthy schools in the nation's first-tier cities and poorer schools in its rural regions. Schools that have the ability to introduce innovative new pedagogies are most likely the nation's elite. A 2019 report by the Mercator Institute for China Studies on creativity in Chinese schools points to a widening gap between China's poorer and more prosperous regions. Wealthy schools 'offer a variety of novel teaching approaches, like maker spaces, escape rooms and advanced computer classes … [Their poorer counterparts] are struggling with limited resources and are struggling to establish creativity-fostering environments.'[16] Some companies, such Squirrel AI, have diverted their attention away from China's urban centres, focusing their efforts on opening over 1700 schools in the nation's second and third-tier cities.[17] Nonetheless, relying on the private sector alone may perpetuate rather than narrow existing inequities within the education system.

Furthermore, while intelligent technologies may be capable of teaching skills that are easily quantifiable, such as middle-school mathematics, we do not know whether they can teach more complex, soft skills such as creativity, critical thinking and collaboration. 'Intelligent technologies are useful for the rapid acquisition of knowledge,' says Professor Yong Zhao, University of Kansas. 'But does that amount to a higher quality of education?'[18] China's current high-school education system – rigidly standardised, outcome driven and metrics based – revolves entirely around preparing students for the gaokao, its national university entrance exam, and is criticised for its excessive emphasis on test scores and rote memorisation. In this system, students are still prepared for the workplace of the industrial age, treated as passive recipients of knowledge transfer, like minds on a conveyor belt. If tailored for the current system, intelligent tutoring systems may simply bolster the existing gaokao education system and train students to become better test-takers.

However, as Rogier Creemers highlights in his essay in this collection, there are limitations to the 'engineering' approach when it comes to those areas of life that are more complex and unpredictable. If China were to define a high-quality education as one that cultivates creative and collaborative critical thinkers, capable of contributing meaningfully to society, what role can intelligent technologies play in teaching the next generation of students? Can an adaptive learning system teach students how to cultivate supportive, interpersonal relationships – crucial to effective collaboration? Can a robot tutor teach emotional resilience – a skill critical to navigating the uncertainty of the 21st century? Can a video tutorial show students how to make meaning from their own experiences and learn from their mistakes? Even if there existed technology that could successfully teach these skills (for example, Squirrel AI is collaborating with Stanford University to research AI applications in the teaching of leadership and creativity),[19] would it be in the interests of China's one-party government to implement such programmes in the nation's public schools?

Finally, any implementation of AI technology raises the question of the ethical use and collection of data. Students, parents and educators should be informed of the data being collected and how it is being used, companies should articulate transparent best practices and governments should establish clear regulations. What data is helpful and what is harmful to a student's educational experience? On one hand, data is a crucial asset that allows for intelligent technologies to deliver feedback and address student needs with much greater precision and efficacy. An education company might collect information on the time it took a student to answer a question to better evaluate their abilities. On the other hand, data could be used as a means of maximising a company's profit.

'Whereas the human teacher assumes change, AI assumes continuation. These are fundamentally different approaches to an individual's capacity for change and growth'

Personal identifiable information – say, a student's sexual orientation or biometric data – could be used to discriminate or target a student on the basis of their sexuality or physical experience. Reports even show some schools deploying AI-powered gates and facial recognition cameras to monitor everything from student concentration levels to their emotional state.

Although the Chinese government has taken some steps to 'curb and regulate'[20] the use of facial recognition and new technologies in schools, without stringent regulations these technologies have serious implications for student privacy, as Danit Gal discusses in her essay. If schools are capable of tracking every keystroke, knowledge point and facial twitch, they are effectively furnishing either a technology company or the Chinese state with an eternal ledger of every step of a child's development. This is potentially problematic because, whereas the human teacher assumes change, AI assumes continuation. Today, when a kindergarten student makes a mistake, the human teacher will try to help the student overcome it and the mistake will eventually be forgotten. In contrast, an intelligent tutoring system could not only store that information and tailor a personalised pathway for the student in the first grade, it may extrapolate that information many years later, when the student is in high school. These are fundamentally different approaches to an individual's capacity for change and growth. Should schools use your behaviour in high school to judge or predict your trustworthiness as an adult?

Should colleges have access to a student's genetic predisposition for mental illness, when sorting them into dormitories? If algorithms can identify pre-existing signs of prodigy and talent, can schools pre-emptively recruit them into special programmes? These are all questions that must be grappled with when the task of evaluating student potential has been handed over to an algorithm.

Conclusion

Weeks after China shut down its schools due to COVID-19, 22 other countries announced school closures. The coronavirus pandemic has revealed to us that no nation is immune to what the United Nations has called an 'unparalleled'[21] disruption to the world's education systems. In order to continue preparing our next generation of students for the future, our schools must adapt to the uncertainties and rapidly changing realities of the 21st century. In the last few years, China has increasingly turned to technology as a solution – pumping money into AI education and fostering the growth of education startups, which have introduced new technologies that claim to revolutionise the future of learning. But, despite its potential, as China's development of AI applications in education accelerates, it is imperative that students, parents, educators and policymakers around the world observe China's push for AI in education with a critical eye, think through important questions of educational equity, effectiveness and ethics and learn from its missteps.

Endnotes

[1] Arjun Kharpal, 'China Wants to Be a $150 Billion World Leader in AI in Less than 15 Years', CNBC, 21 July 2017, https://www.cnbc.com/2017/07/21/china-ai-world-leader-by-2030.html.

[2] Kharpal, 'China Wants to Be World Leader in AI'.

[3] Lin Shujuan et al., 'Students Set Sights on AI Majors', *China Daily*, updated 16 July 2018, accessed 9 April 2020, http://www.chinadaily.com.cn/a/201807/16/WS5b4bdd4ba310796df4df68af_2.html.

[4] Karen Hao, 'China Has Started a Grand Experiment in AI Education. It Could Reshape How the World Learns', *MIT Technology Review*, 2 August 2019, accessed 2 April 2020, https://www.technologyreview.com/2019/08/02/131198/china-squirrel-has-started-a-grand-experiment-in-ai-education-it-could-reshape-how-the/.

[5] Michelle R. Davis, 'Global Artificial Intelligence Boom Predicted in Education, Particularly in China', *EdWeek Market Brief*, 4 June 2019, https://marketbrief.edweek.org/marketplace-k-12/global-artificial-intelligence-boom-predicted-education-particularly-china/.

[6] Yan Dai, 'Household Education Spending Totals $296 Billion', *China Daily*, 16 January 2018, accessed 9 April 2020, http://www.chinadaily.com.cn/a/201801/16/WS5a5d5e8aa3102c394518f794.html.

[7] Richard Tong, CTO, Squirrel AI, said in interview with author, 18 February 2020, New York City.

[8] Said in interview with author, 11 February 2020, Xiamen, China.

[9] 'Teach English Online', VIPKid, accessed 9 April 2020, https://www.vipkid.com/teach.

[10] '17zuoye', 一起教育科技_让学习成为美好体验, accessed 9 April 2020, https://ucenter.17zuoye.com/login.vpage?returnURL.

[11] Liulishuo, accessed 9 April 2020. https://m.liulishuo.com/en/aboutus.html.

[12] '1st AI-Powered Adaptive Education Provider in China', Squirrel AI Learning, accessed 9 April 2020. http://squirrelai.com/.

[13] Said in interview with author, 18 February 2020, New York City.

[14] Said in interview with author, 14 February 2020, Lawrence, Kansas.

[15] Wayne Holmes and Charles Fadel, 'The Promises and Implications for AI Teaching and Learning' in *Artificial Intelligence in Education*, ed. Maya Bialik (Boston, MA: Center for Curriculum Design, 2019), 149–150.

[16] Didi Kirsten Tatlow, 'Manufacturing Creativity and Maintaining Control: China's Schools Struggle to Balance Innovation and Safeguard Conformity', *China Monitor Perspectives*, 14 February 2019, Mercator Institute for China Studies (MERICS), https://www.merics.org/en/china-monitor/manufacturing-creativity-and%20maintaining-control.

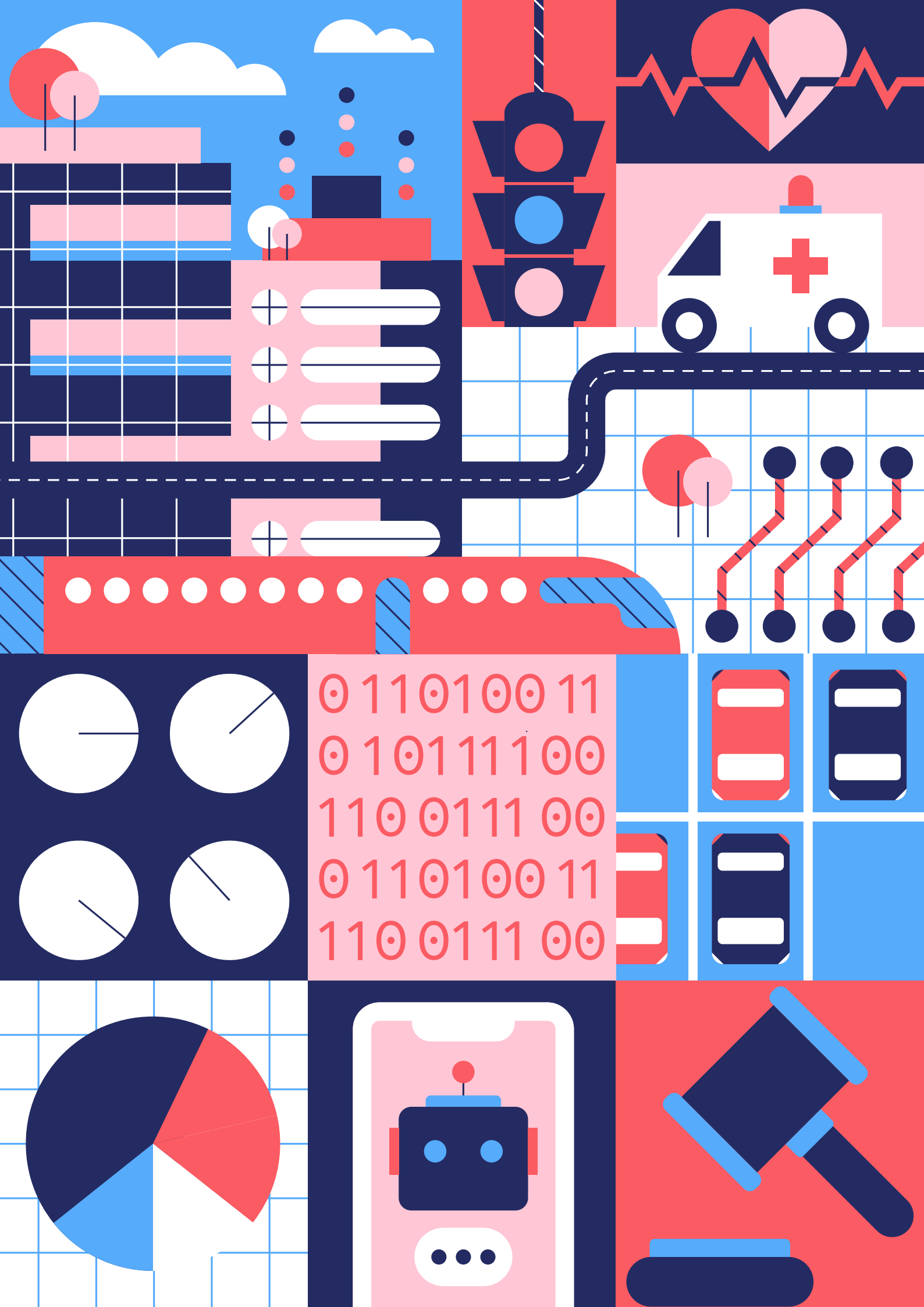
[17] Said in interview with author, 18 February 2020, New York City.

[18] Said in interview with author, 14 February 2020, Lawrence, Kansas.

[19] Said in interview with author, 18 February 2020, New York City.

[20] 'China to Curb Facial Recognition and Apps in Schools', *BBC News*, 6 September 2019, https://www.bbc.com/news/world-asia-49608459.

[21] 'Coronavirus Update: 290 Million Students Now Stuck at Home', *UN News*, 5 March 2020, accessed 9 April 2020, United Nations, https://news.un.org/en/story/2020/03/1058791.


0 1 1 0 1 0 0 1 1
0 1 0 1 1 1 1 0 0
1 1 0 0 1 1 1 0 0
0 1 1 0 1 0 0 1 1
1 1 0 0 1 1 1 0 0

04.
From Traffic Management to Smart Courts: China's approach to smart cities

By Eugeniu Han
PhD candidate in Policy Analysis at the Pardee RAND Graduate School
Assistant policy researcher at the RAND Corporation[i]

A 'smart city' generally refers to a city that has enhanced the quality of its services by adopting advanced information and communications technologies. The need for better services is usually motivated by two fundamental drivers: increasing urbanisation and growing demands for better quality of life. Added to these factors are economic considerations, such as the desire to increase gross domestic product, develop the services sector and attract innovative people and firms. In China's case, the push to develop smart cities seems to have an additional critical motive – the desire to strengthen the country's main governing institutions.[1] Most overarching policy frameworks, such as the five-year plans, emphasise upholding leadership by the Communist Party of China (CPC) as one of the basic principles of development, while China's New Generation of Artificial Intelligence Development Plan recognises that artificial intelligence (AI) can improve 'social governance, and is indispensable for the effective maintenance of social stability'.[2] Rogier Creemers explores this motivation in more depth in his essay in this collection on the ideology behind China's AI strategy. Finally, developing and exporting smart city technologies is part of China's international ambitions outlined in the Digital Silk Road component of its flagship Belt and Road Initiative.[3] For all of these reasons, the development of smart cities has become a policy priority in China.

Of China's more than 500 smart city pilots, this essay focuses on two flagship projects – the City Brain system of urban traffic management and the System 206 smart court case-handling system. The first is what we might consider to be a 'traditional' smart city initiative. Smarts courts are rather more unusual – although they are usually listed in Chinese planning documents as integral parts of smart cities – and demonstrate the complexity of smart projects and how different AI applications are bundled together. They also highlight some of the ethical questions and concerns that emerge when AI algorithms are used in justice systems. Both these smart city projects present possible lessons for the UK and the EU, including a focus on policy consistency and competition between different localities in the Chinese system, skilful deployment of proven commercial technologies (rather than necessarily cutting-edge technologies), and different models of interaction between government and the private sector.

[i] Any opinions expressed in this paper belong to the author and do not necessarily represent the views of the RAND Corporation or its research clients or sponsors.

'The evolution of City Brain resembles the growth of a startup, where the focus is on rapid development with mistakes and bugs fixed along the way'

Institutional and policy framework

The development of smart cities is governed by a complex mix of institutions and policies. At the highest policy level, the development of smart cities has been coded in the 13th five-year plan, which calls for the development of 'a number of exemplary smart cities, focusing on developing smart infrastructure, convenient public services, and refined social governance'.[4] This call confirms the pilot approach where national-level ministries are involved in the development of technologies and standards, while local governments generally retain decision-making on the specific implementation aspects of smart-city programmes.[5] The choice of pilots does not follow a particular pattern, with some projects resulting from a rigid top-down approach while others appear to be primarily driven by local initiative. This pattern of public services innovation diffusion in China has been documented in earlier studies.[6]

To accelerate the development of the core AI technologies, the government developed a 'national team' approach – where different tech giants are tasked with developing open innovation platforms in their respective fields,[7] as Jeff Ding also describes in his essay in this collection on local AI policy and hybridised industrial ecosystems. This approach is different from the classic 'national champions' approach since the members of the national teams are not shielded from competition, can partner with foreign firms and appear to be less dependent on government financial support.[8]

AI applications in smart cities

Since China has more than 500 smart city pilots, we will not attempt a comprehensive overview of the field. Instead, we will briefly discuss two flagship projects – traffic management and smart courts – that might provide important lessons. The former example is indicative of more traditional applications of AI in smart city technologies, while the latter may offer a glimpse into how the Chinese government is planning to use technology to improve the core functions of the government within the parameters of its political system. At the same time, we must caution that these cases describe relatively successful and well-known flagship projects and are not necessarily a reliable indicator of the overall performance of China's smart city programme.[9]

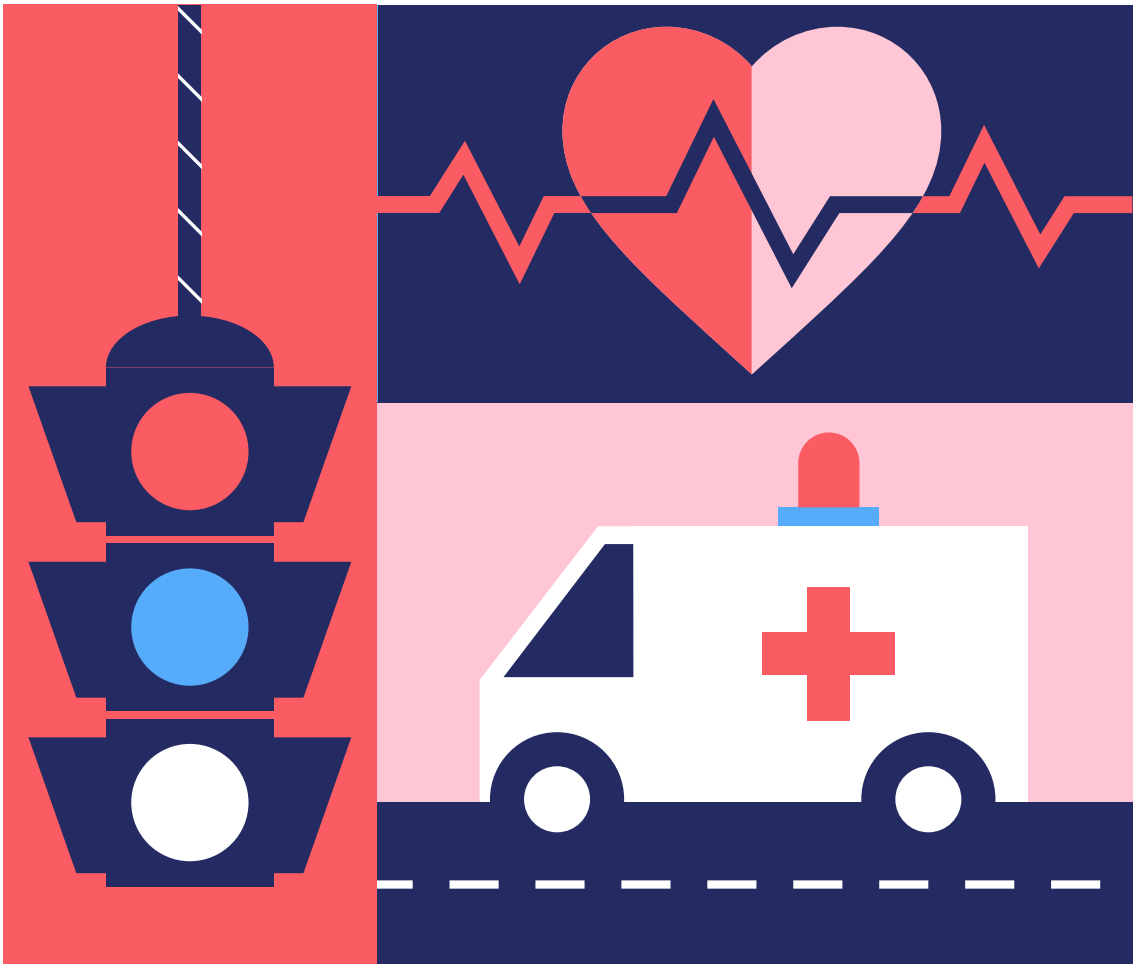
Urban management – City Brain

City Brain is a system of urban traffic management first launched in Hangzhou in 2016 with the primary goal of reducing traffic congestion. It was jointly developed by 13 companies in collaboration with the city government and is operated from Alibaba's cloud service. While City Brain was initially built around image recognition algorithms to detect traffic from traffic camera video feeds, it gradually developed into a data co-ordination centre that consolidates data feeds from more than 700 IT systems from different government agencies. It attempts to identify and solve an array of problems ranging from traffic management (including traffic light optimisation and facilitation of emergency response for ambulances and firefighters) to waste collection, parking lot management and health monitoring of the city's aging population. City Brain is available in several different modifications and has been implemented in 23 cities across Asia,[10] sometimes within the Digital Silk Road component of the Belt and Road Initiative.[11]

The evolution of the City Brain project has several notable features that are common to many smart city projects in China. The first is rapid growth. City Brain started as a small pilot for traffic management and grew into a comprehensive large-scale system that is governing many aspects of urban management. Its rapid evolution resembles the growth of a startup, where the focus is on rapid development with mistakes and bugs fixed along the way. Since the project was developed in a short time and quickly exceeded its initial scope, we can conclude that it was unhindered by lengthy deliberations and the participation of civil society groups and other non-government stakeholders.[12] The second notable feature is the emphasis on infrastructure and data collection. City Brain depends on data from thousands of sensors and dozens of government agencies. Data is viewed as a source of knowledge and a valuable resource. Some of this data can be shared with a pool of private firms to help develop solutions to specific urban problems.[13]

Third, the project is led by a big company (Alibaba) that provides the core technology (cloud service platform) and manages the co-ordination with various government departments. Notably, the project requires continuous endorsement from the highest level of the local government to allow data integration between different agencies and departments.

Finally, there is a symbiotic relationship between the leading firm and the government, where the government is helping the firm develop a business product that can be adopted by other Chinese cities and, potentially, exported to other countries. The essence of this process can be described as 'special deals' where local governments provide opportunities to selected firms on a non-competitive basis and help them expand to other domestic and international markets. There is some evidence that the benefits of these arrangements outweigh the costs in the Chinese context.[14]



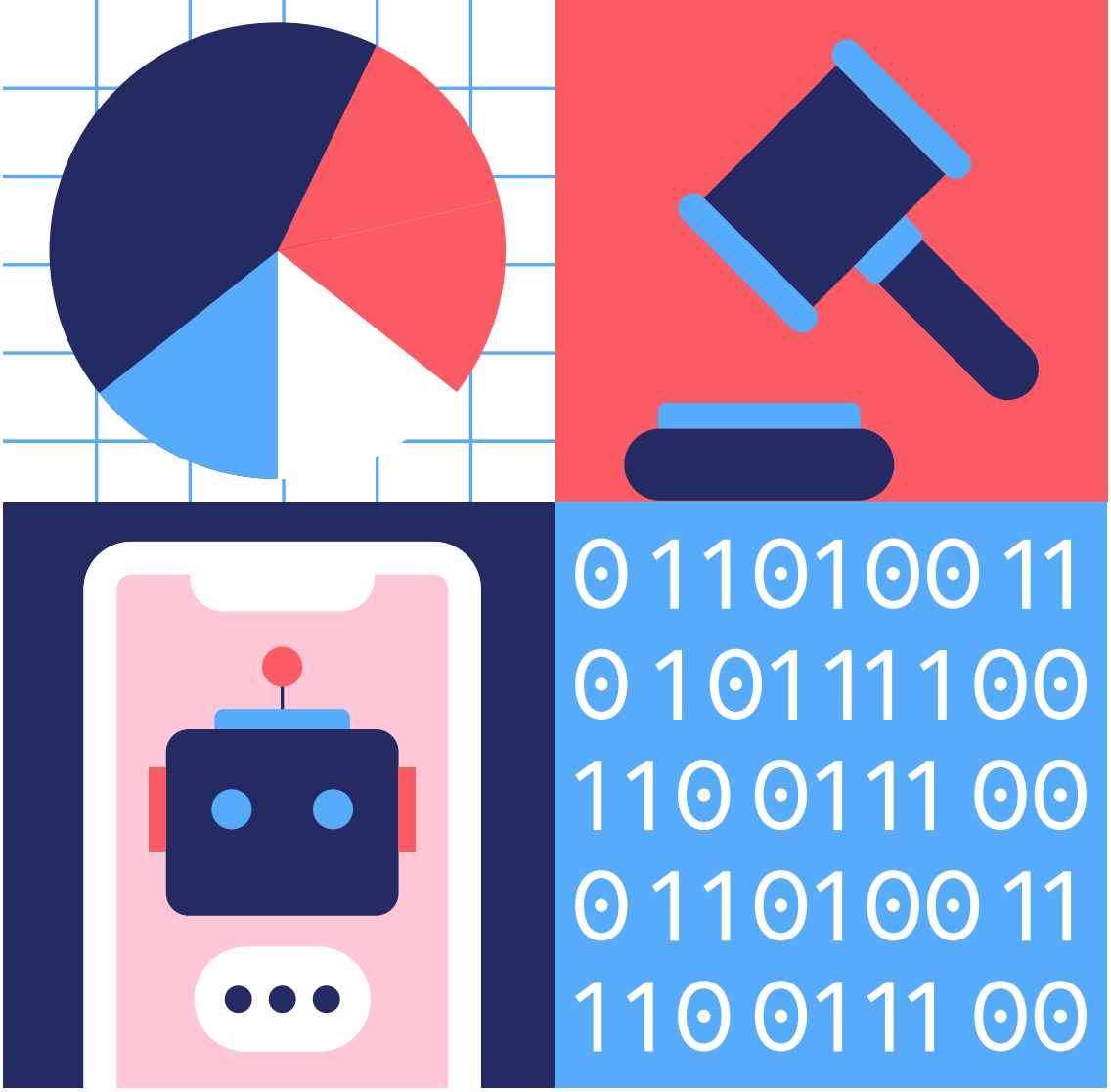
Smart courts – System 206

System 206 is the codename for the Shanghai High People's Court intelligent assistive case-handling system for criminal cases, which aims to improve quality and reduce 'false, unjust, or wrong' charges and sentences.[15] The system has two main components. The first is a cross-referencing system that uses speech recognition and natural language processing to compare different types of evidence and alert the judge to any contradictions in the fact patterns.[16] The second is a sentencing reference tool based on machine learning algorithms that use the defendant's basic information (criminal record, age, 'social harm' incurred, etc) to search a large database of past court records for similar cases and make sentencing recommendations. The output can take the form of a scatterplot graph, where each dot contains the information about the facts of the case and the corresponding sentence. Notably, the system can also be used to judge the judges and prosecutors by spotting the outliers – judges and prosecutors whose sentencing pattern is too far away from the mainstream.[17] In addition, the Shanghai Smart Court features an AI-powered chatbot for information enquiries, a blockchain data transmission system that secures links with other government agencies, a natural language processing system that records and transcribes the deliberations of the collegiate panels, a smart search system and a quick-access function that uses facial recognition.[18]

An analysis of different applications used in the Shanghai Smart Court reveals a complex picture. Applications such as chatbots are used extensively in the private sector and have been piloted in other countries to provide basic information and advice, and may even file a standard contestation of a parking ticket.[19] In the same vein, speech recognition applications could reduce the workload of the clerks, legal transcribers and court reporters and increase the efficiency of the system, while integrated mobile payments systems and remote access to court documents could save time and reduce congestion.

Other applications, such as 'assessment of social harm' and sentencing references, even in an advisory capacity, could skew the decision of the prosecutor or judge to the detriment of the defendant.[20] In addition, it could be argued that an AI-generated assessment may appear as a de facto witness against the defendant, but without the opportunity to test the quality of the data and verify the soundness of the algorithm. Some defendants (and their defenders) may lack the technical knowledge, resources and access to challenge AI processes for generating a sentencing reference and assess its potential biases.[21] The ability to 'judge the judges' by focusing on outliers is also ambivalent since it could be used to discipline biased judges, but also target those judges who may appear to be more lenient in political cases.

System 206 shares a number of remarkable similarities with City Brain. Both were built in partnership with tech giants on the 'national team' – iFlytec in the case of System 206 and Alibaba in the case of City Brain. As with City Brain, System 206 was developed at startup speed with the primary goal of creating a viable prototype that could be perfected during operation. In the same vein, the deliberations were mainly limited to different government agencies where interagency co-ordination was facilitated by high-level political support.[22] In another similarity with City Brain, System 206 rapidly scaled up to handle all criminal cases in Shanghai and is currently expanding to other jurisdictions.[23]



Conclusions and further considerations

Given the novelty of these systems and the lack of first-hand evidence, more research is needed to understand how different AI applications affect the provision of public services in China. A cursory overview of the field allows us to make several observations:

- The Chinese government places a heavy emphasis on the development of AI in general, and for the provision of public services in particular. AI is meant to improve the functioning of the existing institutions, not replace or reform them.
- National-level policies tend to focus on technology development, while local governments (for the most part) can choose their own pilots.
- Policies play an important role in agenda setting at national, local and even company levels. Most policies focus on speeding up technology development, data collection and implementing pilots. Issues like risk management, data privacy and accountability are coded in policy, but appear to be secondary to the development imperative.
- Individual projects take the form of public-private partnerships and are led by big technology companies rather than SMEs or startups, at least in the case of flagship initiatives.
- A flagship project may involve many different applications of AI. Some of these applications are straightforward and are in line with global best practices; others raise significant risks to privacy and may have built-in biases.
- Some of the issues related to the development and deployment of AI pertain to values and moral choices and could be less amenable to simple cost-benefit analysis.

Given these observations, we can briefly discuss the implications and possible lessons for the UK and the EU. From a policy perspective, the Chinese smart city policy model has traits of both traditional industrial policy and open methods of co-ordination. The industrial policy component is primarily focused on technology development, while localities are mostly free to choose their own pilot projects in line with a softer approach that resembles the open method of co-ordination,[24] as Jeff Ding explores in his essay. One important lesson of the Chinese policy model is its attempt to reduce policy fragmentation and to nudge (or sometimes force) local governments to develop smart city plans that can be later used to assess progress and induce competition between different regions and localities. While most Chinese policies are rather context-specific and cannot (and should not) be directly transferred to other countries, the overall focus on policy consistency and competition between different localities is an important element that might render further examination.

The second consideration refers to specific technologies and applications. While China has undoubtedly made very significant strides in AI and machine learning technologies over the past few years, most AI technologies that appear in its flagship smart city projects probably couldn't be described as cutting edge.[25] At the same time, the development of integrated systems with multiple applications of well-known technologies and the ability to scale up rapidly (and cheaply) show that successful development and deployment of machine learning technologies need not be conditional on being at the technological frontier. In fact, a skilful deployment of proven commercial technologies could significantly improve the provision of public service under the right circumstances.

The third consideration refers to the model of interaction between the government and the private sector. The Chinese model features partnerships between large players and local governments, usually with political approvals from the higher level to facilitate interagency co-ordination. When the technologies are sufficiently developed in the domestic market, the government may help domestic firms export its technologies in foreign markets through the Digital Silk Road and other similar initiatives. This approach generally favours large firms and leaves relatively little room for SMEs and startups. The benefits of this approach may include speed of deployment and economies of scale, possibly at the cost of increasing the influence of big tech and suppressing smaller players. This approach is markedly different from the more decentralised practices in Europe, and these differences are likely to persist.

Finally, the issues of data privacy and security and the real (and perceived) differences between Europe and China in this respect have been extensively discussed elsewhere,[26] and are addressed by Danit Gal in this collection. We will only note that learning from actual data (for example, in the case of traffic management or, more recently, epidemic control) may provide unique insights and capabilities that cannot be acquired from synthetic data or human observation. The final decisions about the extent of data collection will be context specific and should involve the relevant stakeholders from the government, industry and civil society.

Endnotes

[1] The primacy of the CCP leadership has been coded in most planning documents. Most initiatives that deal with the provision of public services and governance issues are evaluated not only on their efficiency, but also on their potential impact on the Party leadership.

[2] 国务院关于印发新一代人工智能发展规划的通知, 国发〔2017〕35号.

[3] Andrew Kitson and Kenny Liew, 'China Doubles Down on Its Digital Silk Road', *Reconnecting Asia*, 14 November 2019, accessed 20 February 2020, Center for Strategic and International Studies, https://reconnectingasia.csis.org/analysis/entries/china-doubles-down-its-digital-silk-road/.

[4] 'The 13th Five-Year Plan for Economic and Social Development of the People's Republic of China (2016–2020)', State Council, 2016.

[5] The role of local governments was outlined in the State Council's 'Several Opinions of Promoting Consumer Spending on Information Technology and Expanding Domestic Demand' in 2013. This document calls for pilot smart cities projects financed by local governments through local treasury bonds and with participation of private firms.

[6] For a description of innovation diffusion mechanisms, see Xufeng Zhu, 'Mandate Versus Championship: Vertical Government Intervention and Diffusion of Innovation in Public Services in Authoritarian China', *Public Management Review* 16, no. 1 (2014): 117-139.

[7] In 2017, the Ministry of Science and Technology assigned Baidu to develop autonomous driving; Alibaba, smart cities; Tencent, medical imaging; and iFlytek, intelligent voice technologies. Later on, the team was expanded to 15 companies. Sheldon Chin, 'What could we learn from China's National AI Team [国家队] Strategy', *Medium*, 12 August 2019, accessed 8 April 2020, https://medium.com/@smuaiclub/what-could-we-learn-from-chinas-national-ai-team-国家队-strategy-3cb4732b3d25.

[8] Jeffrey Ding, 'China's Current Capabilities, Policies, and Industrial Ecosystem in AI', testimony before the US-China Economic and Security Review Commission's *Hearing on Technology, Trade, and Military-Civil Fusion: China's Pursuit of Artificial Intelligence, New Materials, and New Energy*, 7 June 2019: 16-25.

[9] In fact, a recent study of 'ordinary' smart city projects in Wuhan notes that 'it is difficult to find visible or tangible evidence of the "smart"' upon arrival in town. This observation mirrors authors' own experience from visiting several Chinese cities in 2016. Robert Cowley et al., 'Ordinary Chinese Smart Cities: The Case of Wuhan' in *Inside Smart Cities: Place, Politics and Urban Innovation*, eds. Andrew Karvonen, Federico Cugurullo, and Federico Caprotti (Routledge, September 2018), 45-64.

[10] Alibaba Clouder, 'City Brain Now in 23 Cities in Asia', *Alibaba Cloud*, 28 October 2019, accessed 8 April 2020, https://www.alibabacloud.com/blog/city-brain-now-in-23-cities-in-asia_595479.

[11] Barry Naughton, 'Chinese Industrial Policy and the Digital Silk Road: The Case of Alibaba in Malaysia', *Asia Policy* 27, no. 1 (2020): 23-39.

[12] We couldn't find evidence of extensive engagement with non-government stakeholders prior to the commencement of the project in either the Chinese or English press.

[13] Abigail Beall, 'In China, Alibaba's Data-Hungry AI is Controlling (and Watching) Cities', *Wired*, 30 May 2018, accessed 10 April 2020, https://www.wired.co.uk/article/alibaba-city-brain-artificial-intelligence-china-kuala-lumpur.

[14] This model is described in detail in Chong-En Bai, Chang-Tai Hsieh, Zheng Michael Song, *Special deals with Chinese characteristics* (Cambridge, MA: National Bureau of Economic Research, May 2019).

[15] The factual information about the functions of the 'System 206' is derived from a comprehensive study by Yadong Cui, the former president of Shanghai High People Court, who oversaw the development and deployment of the system. Yadong Cui, *Artificial Intelligence and Judicial Modernization* (Singapore: Springer, 2020), 22.

[16] A fact pattern is a summary of the key facts of a particular legal case, presented without any associated discussion of their legal consequences.

[17] Kai-fu Lee, 'Judging the Judges', *AI Superpowers*, 5 February 2019, https://aisuperpowers.com/blog/judging-the-judges.

[18] Cui, *Artificial Intelligence and Judicial Modernization*, 72-76.

[19] Samuel Gibbs, 'Chatbot Lawyer Overturns 160,000 Parking Tickets in London and New York', *Guardian*, 26 June 2016, accessed 7 April 2020, https://www.theguardian.com/technology/2016/jun/28/chatbot-ai-lawyer-donotpay-parking-tickets-london-new-york.

[20] These issues are also debated in the US, where some courts use criminal-sentencing AI algorithms to estimate the defendant's likelihood of committing a future crime. Derek Thompson, 'Should We Be Afraid of AI in the Criminal-Justice System?', *The Atlantic*, 20 June 2019, accessed 9 April 2020, https://www.theatlantic.com/ideas/archive/2019/06/should-we-be-afraid-of-ai-in-the-criminal-justice-system/592084/.

[21] For a summary of risks of AI use for sentencing reference, see Noel L. Hillman, 'The Use of Artificial Intelligence in Gauging the Risk of Recidivism', *The Judges' Journal* 58, no. 1 (2019): 36-39.

[22] Cui, *Artificial Intelligence and Judicial Modernization*, notes support from Meng Jianzhu and Han Zheng, both members of CPCCC's powerful Political Bureau at the time.

[23] Cui, *Artificial Intelligence and Judicial Modernization*, xvii, states that 'System 206' was widely applied to the whole city since March 2018 and 'public security organs, procuratorates, and courts used the 206 System to handle 100% cases'.

[24] 24 With the notable exception of public security applications of the smart city technologies, which appear to be mandatory rather than optional.

[25] See Cui, *Artificial Intelligence and Judicial Modernization*, 72-76, for a list of technologies used in the Shanghai Smart Court.

[26] Samm Sacks, 'China's Emerging Data Privacy System and GDPR', *Commentary*, 9 March 2019, accessed 6 April 2020, Center for Strategic and International Studies, https://www.csis.org/analysis/chinas-emerging-data-privacy-system-and-gdpr. See also https://medium.com/the-balance-of-privacy/chinas-dataprivacy-law-vs-gdpr-566fde8c213c and https://epernot.com/data-privacy-law-china-comparison-europe-usa/.

05.
Separating Myth from Reality: How China's Social Credit System uses public data for social governance

By Dev Lewis
Fellow and program lead at Digital Asia Hub
Yenching scholar at Peking University

China's growing technological prowess is capturing the attention of the world as its companies and government rapidly build and deploy new technologies, with an eye on global AI leadership.[1] While Chinese bureaucracy has traditionally been a technology laggard, it too is waking up to the promise of previously untapped data within government. The Social Credit System (SCS) is the government's attempt to modernise its governance capacity through data and technology. In western media it is often caught up in a web of misconceptions centred around a perceived 'Orwellian' system that involves using big data to score citizens' morality.[2] In reality, the initiative is a cluster of experiments harnessing public data to improve governance by boosting trust between government, firms and individuals. This includes larger national efforts, such as Blacklist-Redlist Joint Sanctions and Rewards regimes, as well as smaller efforts taking place in a handful of cities, where efforts at scoring citizens are being experimented with – albeit in fairly mundane fashion compared to reporting in mainstream western media.

The SCS is best understood as an overarching policy initiative consisting of multiple sub-systems, some with different policy goals and rules, rather than one distinct system. Ambitiously, it takes aim at nearly all of China's development ills – from environmental protection to IP and financial fraud to academic plagiarism – which the government believes stems from firms and individuals not following laws and regulations.[3] This is seen to be largely due to the judicial system's failure to enforce laws, but is also put down to a general weakness in the country's institutions, which the government believes is holding China back in its path to becoming a developed country with a 'decisive market economy'.[4] The SCS is an attempt to 'encourage individuals, business, legal institutions and government itself to be more "trustworthy" (守信, shouxin)'.[5]

While, as this essay shows, Chinese government efforts to aggregate public data and attempt regulatory experiments to boost trust and improve law enforcement are much more limited in scope, and possibly effectiveness, than they are frequently portrayed outside China, the system does raise questions of fairness, transparency and privacy. However, it is a mistake to conflate all surveillance in China into the SCS. It may be that the issue is less the SCS itself than the laws it is seeking to enforce more effectively.

The data layer: Public data platforms

Included within the SCS policy umbrella is building a modern financial credit reporting system, as exists in developed economies. This is the kernel for the SCS and can be traced back to the late 1990s, when it was led exclusively by the People's Bank of China (PBOC). Today, financial credit reporting remains a key priority but the scope of the SCS as a whole goes far beyond it, with the government's chief policy planning agency, the National Development and Reform Commission (NDRC), now jointly leading the policy initiative.

The SCS aims to channel public data across departments so that it can be used for decision-making by other government or private sector actors. In government and legal documents this data is referred to as 'public credit information', defined to include broadly any information and data 'collected or generated by public agencies that can be used in evaluating how well a subject observes its legal and contractual obligations'.[6]

To enable the collection of this data, governments across the country have built Public Credit Information Platforms to 'aggregate data generated from public management functions by various departments and units'.[7] The data collection typically begins at the municipal (city) level where respective city governments aggregate credit information from government ministries, departments and public enterprises in the city. For instance, in Fuzhou, the capital city of Fujian province, the data platforms receive credit information from as many as 630 different entities.[8] Organised in a hierarchical fashion, the city level platforms connect up to provincial platforms, which finally connect to the national level platform in Beijing. This reportedly receives data from 44 other national government ministries and 32 provincial-level platforms.

'While the social credit system is much more limited in scope, and possibly effectiveness, than it is frequently portrayed outside China, the system does raise questions of fairness, transparency and privacy'

The platforms are also designed to have public-facing portals through which government agencies, the private sector and the wider public can query information and obtain other credit services, which is discussed in the next section.

In the absence of a national law, the data collection processes and details of the systems in general are usually governed by the credit regulations of the respective city or provincial government. Thus the build out is often not uniform across the country, as local governments co-operate and compete with each other, working to the same objectives but with some freedom to experiment, as is the norm in modern governance in China[9], and as Jeff Ding explores in his essay. These credit regulations also control the collection of certain types of data, forbidding the collection of religious faith, genetics, fingerprints, blood type and medical history. Some are more progressive than others, for instance Shanghai's social credit regulations include provisions for a 'right to be forgotten'.[10]

While the core scope for data collection within the SCS is credit information, the initiative has also spurred on existing efforts to upgrade China's open government data practices. In the past couple of years, more than 46 open government data portals have been set up by governments, intended to include a variety of datasets such as administrative penalties, administrative licenses, land ownership, tender notices, credit rating, corporate credit, foreign business, revocation, credit services and rights protection.[11]

Pooling the credit information in these platforms creates the crucial data layer. On top of this the government has been able to operationalise some regulatory initiatives aimed at enforcement of the law.

The application layer: Blacklists, redlists, sanctions and rewards

The use of blacklists and redlists, combined with what is known as the Joint Sanctions and Rewards (JSR) regime, can be seen as the key initiative of the emerging SCS so far.

Essentially, agencies and government departments blacklist individuals or entities found to break laws, regulations or legally binding decisions, and 'redlist' those with exemplary records within their jurisdiction.[12] Names and related information are aggregated in the public credit information platforms and those on blacklists face a number of sanctions, such as being disqualified from participating in government tenders or obtaining credit from banks. Their names can be publicly searchable on the credit platform portals for the benefit of other government or private sector actors. In some cases there is even an element of public shaming, displaying names in public spaces, such as cinemas or cellphone caller tunes.[13]

The first target of this initiative was debt default. This was addressed through what is known as the Nationwide Judgement Defaulter Blacklist, maintained by the Supreme People's Court, which contains a list of individuals and firms with outstanding court judgements against them, in most cases willful debt defaulters.

'While the blacklist-redlist and JSR regime does not involve computing scores of any sort, some city governments are experimenting with using credit information to generate scores that assess a citizen's level of promise keeping'

Now there are a growing number of industry blacklists ranging from tax evasion and unpaid salaries to customs and even academic plagiarism.

The blacklists gain their bite when combined with the JSR regime. Departments that sign up to this regime agree to recognise each other's blacklists and take enforcement actions within their power. For instance, the names from the Judgement Defaulter Blacklist are shared with about 40 other central government agencies who then impose sanctions within their respective authority.[14] These punishments can range from being disqualified from obtaining bank credits and government subsidies to, more famously, bans on purchasing flights or high-speed rail tickets. Increasingly, private companies are also entering into agreements with the government, committing to limit services or access to their platforms. For instance, ride hailing service Didi may prevent blacklisted drivers or riders from using its platforms.[15] Ant Financial's Sesame Credit score, used for a number of services, is negatively affected by blacklist status.

The key policy motive behind the JSR is to strengthen the enforcement power of regulations by coercing individuals or firms to fulfil their legal obligations. In the case of the judgement defaulter list, the aim is clearly to strengthen the Supreme People's Court and coerce firms and individuals to pay back loans. According to data published on the China Credit website and collected by this author, from June 2018 to May 2019 1.3 million firms and 2.2 million people were added to various government blacklists. During the same time 813,000 firms and 1.56 million people were removed.

Citizen scoring

While the blacklist-redlist and JSR regime does not involve computing scores of any sort, some city governments are experimenting with using credit information to generate scores that assess a citizen's level of promise keeping. In 2018, Suqian and Suzhou in Jiangsu province, Hangzhou in Zhejiang province, and Xiamen and Fuzhou in Fujian province, all relatively prosperous cities on the east coast, rolled out city-level personal credit scores (个人信用评分).[16]

The scoring system is highly experimental at this stage but the in-depth case studies[17] I have previously undertaken of Xiamen's Bailu score and Fuzhou's Moli score reveal the general design.

Both cities assign a score from 0-1000 to each citizen as a reflection of their 'adherence to laws, promise keeping, and credit in daily life'. The scores, accessed by citizens via smartphone apps, are computed using models not dissimilar to existing credit scoring models that have been in use globally in the finance industry, such as the FICO model, but remixed with a different set of variables and data points towards the goals of the system.

Figure 1. Summary of Xiamen and Fuzhou's scoring systems

	Bailu Score	Moli Score
Date of introduction	August 2018	April 2018
Registered users	210,000	6.9 million
Registered users as % of city population	5%	21%
Active users*	85,000	1.19 million
Score range	0-1000	0-1000
Average score	666 'good credit'	650-700 'good credit'
User-interface (UI)	WeChat Miniprogram	e-Fuzhou app
Algorithm model	FICO model	Multi-dimensional model
Operated and managed	Xiamen Big Data Company	Fuzhou NDRC Big Data Office

*Users that have used their score to avail a service at least once.
Source of data: interviews with stakeholders and various media reports[18]

Bailu Score	Moli Score
Basic information: refers to the behaviour of those living in Xiamen, reflecting the relationship between the individual and the city and the individual's participation and stability in the city.	Public credit: whether or not citizens are abiding by public morals as well as actively contributing to public interest activities.
Keeping promises: public welfare behaviours that show care for others and a positive return to society is a manifestation of personal social responsibility and trustworthiness.	Credit ability: a citizen's basic credit circumstances and initial credit condition.
Untrustworthy breach of contract: violating regulations and owing fees reflects a negative manifestation of individual's willingness and ability to obey the law.	Working ability: whether in daily work-life a citizen is industrious, hard-working/conscientious and meticulous (爱岗敬业), and strictly observes ethics and morals.
Credit repair: refers to the individual's efforts to repair credit and regain integrity through active participation in public interest events, social services and public welfare activities.	Economic ability: includes delayed payment to common reserve fund, defrauding/cheating the healthcare fund, common reserve fund or wages, or embezzling (侵占) people's legitimate interests and rights.
	Administrative law ability: whether citizens are abiding by the system of regulations and rules and staying away from illegal activities and unlawful practices.
	Legal ability: whether or not citizens are abiding by laws and legal duties.

As with the blacklists, the focus remains on a citizen's standing with the law, which plays the most decisive component in determining whether a score is high or low. The general message is that citizens should fulfil the basic expectations in a timely fashion: paying utility bills, income tax, healthcare, social security and other mandated government payments on time, and thus having no 'notices' against one's name. Individuals with legal charges against them stand to lose points, in some cases as much as 150 points per legal offence.

Unlike in the JSR system, which mainly focuses on disincentives in the form of sanctions, these systems focus exclusively on incentives for those with high scores in an attempt to encourage the types of behaviour the government considers desirable. In interviews with this author, officials revealed that activities such as volunteering, donating blood, using public transport, separating waste or working in areas of public interest, such as teachers or doctors, could be interpreted as 'good credit' and boost one's score.

Citizens in the 'good credit' region (600 and above), can unlock a range of benefits that can be placed in three categories:

- Deposit-free access such as public cycles.
- Discounted access such as subway rides.
- Priority access to services such as government administrative services.

In its present iteration the scores seem more like a government version of a loyalty scheme — all citizens get access to the basic service but some can opt-in for fringe benefits for convenience and comfort. Those with a low score face no direct punishments but the low score is the punishment itself, as it denies access to the benefits afforded to high scores.

The scores are voluntary at the moment and, so far, there has been fairly low engagement, with less than one-third of either city's population registered on the app and the average score lying between 600 and 650. Government officials in charge of implementing the system suggest that they are hoping the score can be adopted by more government agencies and reduce the friction citizens may face when using government services.[19]

Conclusion

The SCS has catalysed the Chinese government's efforts to digitise and pool public data, particularly within the realm of administrative regulations and laws, towards its use as a form of reputation that is now institutionalised in government decision-making around allocation of resources and services, evident in both the blacklists and citizen scores.[20]

The creation of public credit platforms can be seen as a significant landmark in Chinese government efforts at collecting public data, chipping away at the 'data islands' between ministries and departments,[21] an aspect of government data that Rogier Creemers also discusses in his essay in this collection. Studies of open government data platforms reveal they meet industry benchmarks, are fitted with APIs and offer data in a variety of machine readable formats. However, there are question marks around the consistency of data gathering, especially between platforms in the less developed cities away from the east coast.[22] Conversations with officials also suggest there is still work to do in building a culture of digital record keeping, with some departments still lagging behind and relying on physical documents.

The blacklist-JSR regime has strengthened considerably with real implications for individuals and businesses. Data published on the blacklist platform indicates that, for every two names added to the list each month, one is removed. This suggests that the system is encouraging individuals and firms to take desired actions. The expansion from the judgement defaulter list to other areas reflects the government's view on its effectiveness, although it is difficult to answer conclusively if the policy is truly achieving the goal of individuals and firms behaving more in a more 'trustworthy' manner.

Examining the systems, as this essay has done, largely shows that western media narratives of social credit as a morality score, which is devised by capturing all citizen behaviour, from social media posts to supermarket purchases, and which regulates citizens' movements and access to basic services,[23] is unfounded. As demonstrated above, the type of data collected is far more limited in scope and national schemes do not involve scoring at all. Even in the pilot cities the scoring is far more benign in nature.

That is not to suggest there are no reasons to be concerned. The system in its present form raises questions of fairness, especially given the growing span of punishments over and above those mandated by existing laws. While important terms such as 'trustworthiness' remain undefined in law, the blacklist JSR system continues to expand and impact millions of people.[24] The system itself relies on transparency to achieve some of its goals but individuals have very little knowledge about the type of data collected, which leaves them vulnerable to incorrect data and the implications it can bring, as Danit Gal explores in her essay on ethics. These are some of the problems that deserve further research rather than focusing on potential implications of what the system could one day evolve into. Analysing the SCS from the lens of China as a surveillance state tends to miss the actual story and conflate all surveillance in China, such as law enforcement, into the SCS. As Jeremy Daum, editor of China Law Translate, has emphasised, the problem may be less the SCS itself and rather the 'bad laws' that it is trying to enforce more effectively,[25] many of which run counter to core democratic values, such as freedom of speech.

Going into 2020, the final year of the 2014 State Council Plan, discussions around assessing the system are taking place inside and outside of government. These range from assessing the merits of the system and the potential harms it may pose, as well as engaging with questions around fairness, to thorny issues such as privacy and ethics. Parallel to the SCS, legal regimes around data protection are expected to expand and mature over the next couple of years, which will place new boundaries on what type of data can be collected, as well as address a critical question of how data from the private sector can be 'imported' into SCS's initiatives. A 'credit law' is currently being drafted, although most experts expect it remains several years away as critical questions around ethics, privacy and fairness are addressed. Yet the biggest question that the Chinese government will have to answer is whether the system is meaningfully improving trust in institutions and catalysing the country's development trajectory in the way the government believes it can.

Endnotes

[1] Graham Webster et al., 'Full Translation: China's "New Generation Artificial Intelligence Development Plan"', *DigiChina*, 1 August 2017, accessed 23 February 2020, New America, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

[2] Anna Mitchell and Larry Diamond, 'China's Surveillance State Should Scare Everyone', *The Atlantic*, 2 February 2018, https://www.theatlantic.com/international/archive/2018/02/china-surveillance/552203.

Steven W. Mosher, 'China's New 'Social Credit System' is a Dystopian Nightmare', *New York Post*, 18 May 2018, https://nypost.com/2019/05/18/chinas-new-social-credit-system-turns-orwells-1984-into-reality.

[3] State Council, 'Planning Outline for the Construction of a Social Credit System (2014-2020)', *China Copyright and Media*, 14 June 2014, updated 25 April 2015. https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/.

[4] State Council, 'Planning outline'.

[5] Shazeda Ahmed, 'The Messy Truth About Social Credit', *Logic Magazine* issue 7 (1 May 2019), retrieved 16 January 2020, https://logicmag.io/china/the-messy-truth-about-social-credit/.

[6] China Law Translate, '上海市社会信用条例', *China Law Translate*, 30 June 2017, https://www.chinalawtranslate.com/shanghai-municipal-social-credit-regulations/.

[7] State Council, 'Planning outline'.

[8] '《福州市公共信用信息管理暂行办法》_行政法规、规章和规范性文件_福州市人民政府', accessed 4 February 2020, http://www.fuzhou.gov.cn/zfxxgkzl/szfbmjxsqxxgk/szfbmxxgk/fzsrmzf/zfxxgkml/xzfggzhgfxwj/201711/t20171128_1894930.htm.

[9] Lizhi Liu and Barry R Weingast, 'Taobao, Federalism, and the Emergence of Law, Chinese Style', Minnesota Law Review, n.d., 28.

[10] Translate, China Law. "上海市社会信用条例." *China Law Translate* (blog), June 30, 2017. https://www.chinalawtranslate.com/shanghai-municipal-social-credit-regulations/

[11] 'China Open Data Report 2018', Fudan University, National Center on Information Research, n.d.

[12] Jeremy Daum, 'The Redlists Are Coming! The Blacklists Are Coming!' *China Law Translate*, 30 March 2018, https://www.chinalawtranslate.com/the-redlists-are-coming-the-blacklists-are-coming/.

[13] Yin Yijun, 'Chinese App Gives Blacklisted People "Dishonest' Caller ID"', *Sixth Tone*, 30 August 2017. http://www.sixthtone.com/news/1000762/chinese-app-gives-blacklisted-people-dishonest-caller-id.

[14] Daum, Jeremy. "The Redlists Are Coming! The Blacklists Are Coming!" *China Law Translate* (blog), March 30, 2018. https://www.chinalawtranslate.com/the-redlists-are-coming-the-blacklists-are-coming/

[15] "国家发改委与滴滴签署信用信息共享协议-中新网." Accessed February 21, 2020. http://www.chinanews.com/business/2018/03-07/8461965.shtml

[16] 厦门市大数据. (n.d.). 厦门市民个人信用"白鹭分"正式发布. WeChat Official Accounts Platform. Retrieved February 4, 2020, from https://mp.weixin.qq.com/s?__biz=MzU1NjY4MjA2MA==&mid=2247483658&idx=1&sn=44bff6b198c6262504adf28c025cea1a&chksm=fbc01f29ccb7963f8e4ce5d40fac59eff8fb59a7ffc22ddd2209b7e2b585fe89c037631ef338#rd

[17] Dev Lewis "All Carrots and No Sticks: A Case Study on Social Credit Scores in Xiamen and Fuzhou", 4 Oct 2019 https://medium.com/berkman-klein-center/social-credit-case-study-city-citizen-scores-in-xiamen-and-fuzhou-2a65feb2bbb3

[18] 《征信》. (n.d.). 国内城市个人信用分，这一篇讲透了！. *WeChat Official Accounts Platform*. Retrieved 4 February 2020, from http://mp.weixin.qq.com/s?__biz=MzI3ODM1MzI3NA==&mid=2247493124&idx=1&sn=5d56f7418c97895d12a1e990a64958d2&chksm=eb5ae211dc2d6b0785664c07d364376a6f8bdb633820a28d9472ba7894ce8345a91693bf8769#rd

[19] Information shared with the author in interviews with NDRC officials implementing these scores.

[20] Xin Dai, 'Toward a Reputation State: The Social Credit System Project of China', *SSRN Electronic Journal* (10 June 2018), https://doi.org/10.2139/ssrn.3193577.

[21] Mareike Ohlberg, Shazeda Ahmed, and Bertram Lang, 'Central Planning, Local Experiments', *MERICS China Monitor*, 12 December 2017, Mercator Institute for China Studies (MERICS), https://www.merics.org/sites/default/files/2017-12/171212_China_Monitor_43_Social_Credit_System_Implementation.pdf.

[22] Xiao Diyu [肖荻昱]. 'An analyses of open government data and credit service platforms in China', *Journal of Library and Information Science Volume* 7 (2019) [图书 情报导刊] , Yuandian Credit, https://mp.weixin.qq.com/s/xugzyprTafQYRaoP8Ve9bQ.

[23] Mitchell and Diamond, 'China's Surveillance State'.

[24] Luo Peixin-罗培新-社会信用法 Social Credit Law：原理，规则，案例 principles, rules, cases (2018), 15, Peking University Press.

[25] Elizabeth M. Lynch , 'What's the T on China's Social Credit System - Jeremy Daum Explains - Part 1 of 2', *China Law & Policy*, 12 November 2018, accessed 10 March 2020, https://chinalawandpolicy.com/2018/11/12/whats-the-t-on-chinas-social-credit-system-jeremy-daum-explains-part-1-of-2/.

The ethics and ideology of China's approach to AI

06.
China's Approach to AI Ethics

By Danit Gal
Technology advisor to the UN Secretary General's High-level Panel on Digital Cooperation
Associate fellow at the Leverhulme Centre for the Future of Intelligence at the University of Cambridge

In its 2017 New Generation AI Development Plan, the Chinese State Council paved the path for a government-endorsed process of debating and adopting artificial intelligence (AI) ethics principles.[1] Since then, ministries, businesses, academic institutions and expert committees have stepped up to formalise their existing, and institutionalise new, discussions about AI ethics. Nearly three years down the line, these discussions are becoming more common, taking place in government consultations, corporate meetings, academic fora and online conversations.

It is therefore surprising that the western misconception that China lacks debate around AI ethics seems to prevail, when in fact: i) existing Chinese AI principles largely align with global ones; ii) Chinese discussions enjoy unprecedented government support; and iii) the country is already investigating the technical and social implementation of these principles by exploring how they interact with its distinct cultural heritage.

In this essay I explore AI ethics in China through the lens of some of the key topics discussed in depth in the other essays in this collection – education, healthcare, smart cities and social credit systems – to consider which, if any, ethical issues are unique to China and which should be seen as global concerns.

Mapping official AI ethics discussions in China

Notable examples of bodies discussing AI ethics in a formal, government-endorsed capacity are found in three sets of documents that contain AI ethics principles.

The first one is the Beijing AI Principles (BAIP), released in May 2019, backed by the Ministry of Science and Technology and drafted by the Beijing Academy of AI in a multi-stakeholder consultative process.[2] While the contents of the BAIP closely align with existing international principles, it has some distinctive features. The BAIP takes an applied approach to AI ethics by suggesting detailed and action-oriented goals, norms and best practices in research and development, AI use and AI governance. In addition, the document references the 'philosophy of "optimising symbiosis"'.[3] One of the drafters, Yi Zeng, explains that optimising symbiosis refers to the principle of 'harmony' in Chinese philosophy, where all life forms evolve in harmony rather than competition and is especially relevant in the context of human-machine interactions.[4] This particular application of Chinese ethics and philosophy to AI is a distinctive approach that I will return to in this essay.

In line with its applied focus, theBAIP also touches on long-term planning for the expected development trajectory of AI. The BAIP's forward-looking, hands-on approach and long-term outlook are unique among other global principles, as is the way it links existing principles to Chinese practices and philosophical traditions.

The second document is a draft Joint Pledge on AI Industry Self-Discipline, also released in May 2019. It was backed by the Ministry of Industry and Information Technology and drafted by the China Academy of Information and Communication Technology in another multi-stakeholder consultative process.[5] Again, this document largely aligns with other global AI ethics principles but has one notable point of distinction. Article 10: Self-discipline and self-governance, is unique in the global landscape as well as the Chinese one.

This article calls to 'strengthen awareness of corporate social responsibility, integrate ethical principles into all aspects of AI-related activities and implement ethical reviews. Promote industry self-governance, formulate norms of behaviour for practitioners, and progressively build and strengthen industry supervision mechanisms.'[6] The suggestion of ethical reviews, self-governance and industry supervision mechanisms introduces a significantly stronger degree of implementation, enforcement and oversight on AI ethics compliance, a field still largely considered soft law. This lends a more applied nature to AI ethics principles in the Chinese AI industry, in the event that this joint pledge materialises.

Finally, the third document is the Governance Principles for a New Generation of AI, released in June 2019, also backed by the Ministry of Science and Technology and drafted by the National New Generation AI Governance Expert Committee.[7] As well as being aligned with other international AI ethics principles, these principles also share considerable similarity with the two previous Chinese documents. Predominantly applied in nature, this document again references harmony, offering a more direct translation as 'human-machine harmony'. In a global landscape of human-centred AI ethics, this vision is unique in its holistic approach to the interactions between humans and AI, and the way they shape and are being shaped by each other.[8]

An additional official committee still in discussion is the Professional Committee on AI Ethics under the Chinese Association for Artificial Intelligence, a state-level AI organisation under the Chinese Ministry of Civil Affairs. In a media interview, the committee's leader, Professor Chen Xiaoping, shared his goal of developing 'an entire system of AI ethics, with ethical guidelines and a set of operating mechanisms to realise them'.[9]

These official committees are joined by ongoing corporate efforts to align with the developing AI ethics discussions. These efforts received considerable attention during the Chinese Government's Two Sessions in March 2019, when the CEOs of technology companies Baidu and Tencent submitted proposals calling for AI ethics principles and rules. Baidu's proposal called for the sharing of distinct Chinese wisdom with the international AI ethics community,[10] and Tencent's proposal focused on the societal applications and implications of AI.[11]

AI ethics in Chinese applications of AI

What does the applied approach to Chinese AI ethics detailed above look like in real-life AI applications? Generally speaking, Chinese discourse revolves around two axes, the application of Chinese ethics and philosophy to AI and the ethical implications of AI use in public services.

The Chinese cultural and philosophical concept of harmony (和) occupies a special place in the official discussions of AI ethics. While many of the other principles are, at least in theory, universally shared, the concept of harmony as applied to ethics is unique to China.[12] Attempts to apply this concept, among other Chinese ethical and philosophical principles, to AI are actively debated in China.[13] A number of experts have explored how harmonisation between different ethical and philosophical principles within China and abroad can be achieved, so that they evolve in a complementary manner rather than a mutually exclusive one.[14] However, it has not yet moved beyond discussions to play a more practical role. Privacy is among the most visible ethical issues when it comes to AI in public services. Chinese authorities have been particularly receptive to public discourse and have been strict in regulating personal data collection and using sharp-toothed enforcement mechanisms. Public services that fail to comply with the Cybersecurity Law of the People's Republic of China[15] and Personal Information Security Specification[16] are suspended until proven otherwise or entirely banned. Authorities periodically review such services and respond to the public's complaints on privacy breaches and data leaks.[17]

AI applications in healthcare: privacy and bias

As illustrated by Andy Chun's essay on AI in public healthcare services, China is rushing ahead in deploying AI technology to deliver consistent, higher-quality medical care to more than 1.4 billion citizens. Such rapid and wide-scale use of AI requires near-constant collection, storage and analysis of large personal data troves. These pervasive and data-intensive operations have raised enduring public concerns about the privacy of patients during and after treatment and the security of their data. The sensitivity of public health information has long been a focal point for government action,[18] but has yet to be fully addressed. With mounting public pressure over frequent medical data leaks affecting tens of millions of patients,[19] we can expect this issue to continue capturing a significant portion of the applied AI ethics discussion in China.

While this ethical issue is not unique to China, the country's strong push to apply AI to the world's largest healthcare system (by patient numbers) does carry distinct implications by virtue of its size. There are no clear differences in the ethical discourse and approaches taken by western and Chinese entities in this regard, with one exception. China is more proactive in curbing privacy breaches under its aforementioned regulatory frameworks by developing and enacting national laws and actively monitoring and punishing violators.[20]

An additional, albeit less central, ethical concern relates to bias. While bias within Chinese AI use in healthcare is still in early discussion stages and mostly applies to the urban-rural divide, a study has highlighted the bias that exists in foreign systems used to service Chinese patients. An example of that is the trialling of IBM's Watson for Oncology, mentioned in Andy Chun's essay, which was trained on western datasets.

'The use of facial recognition and biometric identification to commute, pay for transactions and use public infrastructure...establishes a culture of pervasive monitoring of citizens'

The discrepancies between the western and Chinese datasets led to some biased system recommendations that did not apply to Chinese patients.[21] This joins other western-trained technologies that present such biases, such as iPhone's facial recognition software not distinguishing between different Chinese faces.[22] This is a more general issue, with people of Asian descent being among the underrepresented groups in western training datasets.

The ethical question surrounding bias in AI is often discussed globally but is relatively new to internal Chinese ethical discourse. With a largely homogenous population, Chinese applications of AI in public services have made significant attempts to alleviate biases in other aspects of diversity, such as dialects,[23] but this is typically as far as it extends before venturing into broader urban-rural divides, which fall under ongoing discussions on equity and equality.

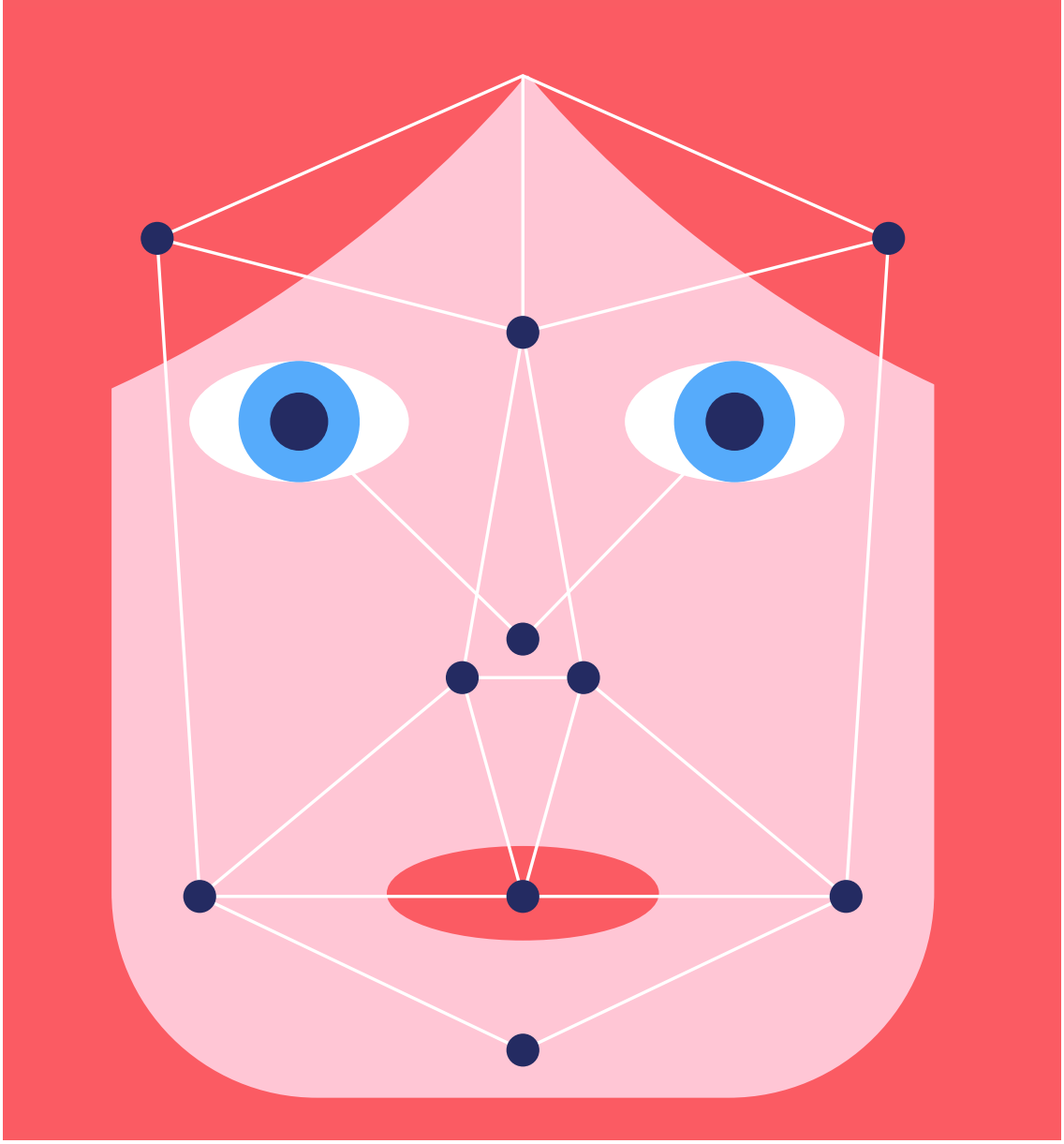
AI applications in education: Privacy and accessibility

Yi-Ling Liu's essay on China's use of AI in its education system highlights two ethical issues: privacy and accessibility. On the privacy side, AI-enabled education systems collect, store and analyse students' facial expressions, although this is pending further regulation,[24] to identify how students are responding to the content being taught by the system. In addition to the pervasive privacy implications this has for individuals, there are the legal implications of these individuals being minors. This is expected to be addressed, at least partially, by the Law of the People's Republic of China on the Protection of Minors, which is being amended to include cyberspace.[25]

Many other countries are facing similar predicaments, which UNICEF is working to address by advancing children's rights in the age of AI.[26] The approach that the Chinese government is taking to addressing the privacy vulnerabilities of children in the application of AI in public services is also in line with other international efforts. If anything, China's rush to deploy AI-enabled education throughout the country is magnifying such privacy efforts and underscoring the importance of their implementation.

The ethical issue of accessibility to China's AI-enabled education applications is shaped by its rural-urban divides. With advanced technology being less available in rural areas, the efficacy of these applications for rural education efforts is put into question. This is further exacerbated by the lack of teaching capacity to supplement areas of educational enrichment not provided by AI. While improved access to some unified education resources is granted, this still falls short of the high level of education and resources accessible in urban schools.[27]

Urban-rural education divides are being tackled through AI in other countries, too, of course. What does make China stand out is the scale of its ambition as it leads the way with extensive, government-backed efforts to introduce AI in public education services. As such, the solutions that China is testing now will diffuse into other developing countries later,[28] making them a pressing global concern.

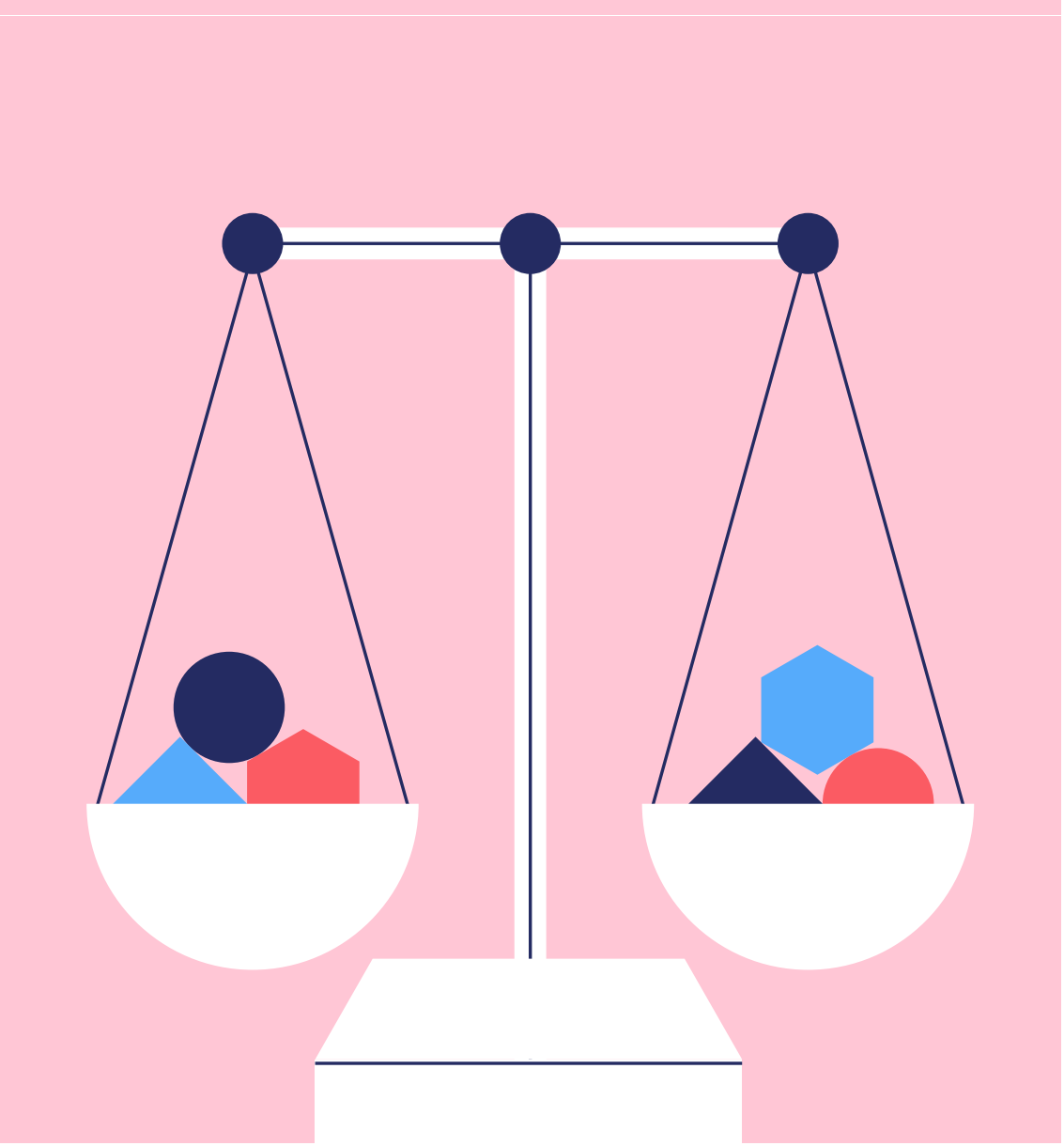


AI applications in smart cities: Privacy and accountability

As Eugeniu Han demonstrates in his essay, AI-enabled smart cities are abundant in China, which has about 50 per cent of the world's smart cities in-the-making.[29] However, considering that residents' facial information is collected to board subways,[30] sort trash and enforce recycling,[31] and obtain toilet paper in public toilets,[32] this raises serious privacy concerns. The use of facial recognition and biometric identification to commute, pay for transactions and use public infrastructure is powered by concentrated and vulnerable public data treasure troves. It also establishes a culture of pervasive monitoring of citizens. Recently, China also came under fire for using facial recognition in Xinjiang for racial profiling-based monitoring.[33]

While this elevates the ethical issue of privacy breaches via smart infrastructure monitoring in China to new heights, it is, again, a concern that extends beyond China. Many other countries are using facial recognition in their smart infrastructure for racial profiling-based monitoring.[34] However, what once again makes China different is the magnitude and scope of operations. While privacy laws in China protect citizens' data from abuse by non-governmental actors, they do not limit the government's access to and use of private data. In this regard, China is lagging behind other countries seeking enhanced user privacy while balancing security concerns.

China is also unusual in its attempt to co-create smart courts that automate case handling, as Eugeniu Han describes in Shanghai. It raises the ethical concern of accountability given the role fulfilled by the AI application in holding others accountable. With smart courts and AI judges making opaque decisions that directly affect humans, we are also left with the question of who is accountable for the AI systems' actions?



In addition, with these courts handling cases where Alibaba, the defendant in most cases, is also the co-creator of the smart court, questions of conflict of interest arise, further exacerbating legal accountability for decisions made by using these systems.

The use of AI to support court systems is not unique to China but the instance of a smart court and AI judge handling claims against a corporate actor, while also being developed by that very same corporate actor, is. With China's move to advance the rule of law throughout the country, the partial privatisation of smart courts and AI judges highlights the ethical issue of questionable accountability for such legal actions carried out by opaque systems.

AI applications in social credit systems: Privacy and transparency

As Dev Lewis shows in his essay on social credit systems, pervasive data collection and linking pose significant privacy concerns. By breaking data silos and linking public data collected by different government departments and corporate actors, these systems enhance access to this data and, at the same time, the risk of privacy invasion. The essay also shows that addressing these pervasive data collection concerns are regulations limiting the types of sensitive information that can be collected (such as religious faith). However, private data is still collected and linked in order to achieve social credit scoring, putting citizen's privacy at risk if abused.

'While privacy laws in China protect citizens' data from abuse by non-governmental actors, they do not limit the government's access to and use of private data'

While the mechanisms of credit score systems are not unique to China (and use similar methods to systems already used in the USA,[35] for example) the governmental-level development and deployment of such an initiative is, again, unique in scale. With many competing initiatives, personal data is collected and linked by multiple actors, creating multiple vulnerability points. Given China's observed approach to privacy violations and data leaks, we may expect the forthcoming Credit Law to address privacy protection (although Dev Lewis notes that local experts believe this is still far from enough).

Another ethical issue highlighted by this case study is transparency. In particular, transparency around how these social credit scores are calculated and the repercussions of errors in calculations. Without knowledge of how these mechanisms work, citizens will have a hard time appealing their decisions. The models and variables used to calculate a citizen's score remain largely opaque, making auditing these systems nearly impossible. Without such basic transparency, inequality between citizens with high and low scores may further exacerbate, unchecked.

Social credit scores are still largely voluntary at the moment and improving transparency is essential before making them mandatory. The lack of transparency in credit scoring schemes is a globally shared ethical issue we must collectively tackle before using AI in public services that have the ability to grant or deny access to other critical public services.

Conclusions

While AI use in Chinese public services is unique in scope, the majority of ethical issues such use highlights, and the approaches taken by the government to address them where relevant, are not. The exception is the concept of 'harmony'. As AI ethics discussions in China develop, we can expect more explorations of how Chinese philosophical and cultural practices apply to AI ethics principles. This, alongside the implementation of other principles, will mark further progress from principles to action, where we can better observe the similarities and differences that emerge between countries.

As this essay demonstrates, there are more similarities than differences at this point, which is not surprising given the unified nature of AI and commonalities among available public services worldwide. It does, however, create a great opportunity for fostering global co-operation on these shared concerns, especially considering China's robust AI ethics discussions and the country's key role as a developer and exporter of AI applications. Both the above-mentioned Chinese AI ethics and governance bodies and the international community should do more to bring a diverse group of relevant stakeholders into key conversations on AI ethics. We're all in the same boat, using very similar oars.

Endnotes

[1] Graham Webster et al., 'Full Translation: China's "New Generation Artificial Intelligence Development Plan"', *DigiChina*, 1 August 2017, New America, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

[2] BAAI, 'Beijing AI Principles', *Beijing Academy of Artificial Intelligence*, 28 May 2019, https://baip.baai.ac.cn/en.

[3] BAAI, 'Beijing AI Principles'.

[4] Yi Zeng, remote interview by Danit Gal, 14 March 2019.

[5] Graham Webster, 'Translation: Chinese AI Alliance Drafts Self-Discipline "Joint Pledge"', *DigiChina for New America*, 17 June 2019, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-ai-alliance-drafts-self-discipline-joint-pledge/.

[6] Webster, 'Self-Discipline Joint Pledge'.

[7] Lorand Laskai and Graham Webster, 'Translation: Chinese Expert Group Offers "Governance Principles" for "Responsible AI"', *DigiChina for New America*, 17 June 2019, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/.

[8] Danit Gal, 'Perspectives and Approaches in AI Ethics: East Asia' in *Oxford Handbook of Ethics of Artificial Intelligence*, eds. M. Dubber, F. Pasquale, & S. Das (Oxford: Oxford University Press, 2020).

[9] Xinhua, 'China Hosts Conference to Promote AI Ethical Standards', *XinhuaNet*, 12 June 2019.

[10] 新浪财经综合, '李彦宏三大提案:完善电子病历 加强人工智能伦理研究', 新浪财经, 4 March 2019, http://finance.sina.com.cn/review/jcgc/2019-03-04/docihsxncvf9583301.shtml.

[11] 腾讯研究院, '2019年两会，马化腾提了这7份建议案 | 2万 字全文版', 6 March 2019, https://mp.weixin.qq.com/s/yb6OyulvHQjlmcNMJsJYSw.

[12] Chenyang Li, 'The Ideal of Harmony in Ancient Chinese and Greek Philosophy', *Dao* 7 (2008): 81-98.

[13] See for example 'About Linking Artificial Intelligence Principles (LAIP)', LAIP, http://www.linking-ai-principles.org/about.

[14] John Thornhill, 'Formulating Values for AI is Hard When Humans Do Not Agree', *Financial Times*, 29 July 2019, https://www.ft.com/content/6c8854de-ac59-11e9-8030-530adfa879c2.

[15] Rogier Creemers, Paul Triolo, and Graham Webster, 'Translation: Cybersecurity Law of the People's Republic of China [Effective June 1, 2017]', *DigiChina for New America*, updated 29 June 2018, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-cybersecurity-law-peoples-republic-china/.

[16] Mingli Shi et al., 'Translation: China's Personal Information Security Specification', *DigiChina for New America*, 8 February 2019, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinas-personal-information-security-specification/.

[17] Minghe Hu, 'China Issues Rules to Stop Apps from Abusing User's Personal Information in Latest Data Privacy Effort', *South China Morning Post*, 31 December 2019, https://www.scmp.com/tech/apps-social/article/3044051/china-issues-rules-stop-apps-abusing-users-personal-information.

[18] Shan Juan, 'New Rules Will Keep Patients' Info Secret', *China Daily*, updated 3 May 2017, https://www.chinadaily.com.cn/china/2017-05/03/content_29174728.htm.

[19] Eliza Gkritzi, 'Chinese Tech Companies Still Can't Stop Medical Data Leaks', *TechNode*, 9 January 2020, https://technode.com/2020/01/09/chinese-tech-companies-still-cant-stop-medical-data-leaks/.

[20] Liu Shuangshuang and Denise Jia, 'China Starts Crackdown on Apps' Violations of User Privacy', *CX Tech*, 5 November 2019, https://www.caixinglobal.com/2019-11-05/china-starts-crackdown-on-apps-violations-of-user-privacy-101478807.html.

[21] Chaoyuan Liu et al., 'Using Artificial Intelligence (Watson for Oncology) for Treatment Recommendations Amongst Chinese Patients with Lung Cancer: Feasibility Study', *Journal of Medical Internet Research* 20, issue 9 (2018): 1 – 10.

[22] Zhuang Pinghui, 'Chinese Women Offered Refund After Facial Recognition Allows Colleague to Unlock iPhone X', *South China Morning Post*, 14 December 2017, https://www.scmp.com/news/china/society/article/2124313/chinese-woman-offered-refund-after-facial-recognition-allows.

[23] Qian, Zhecheng, 'Saving China's Dialects? There's an App for That', *Sixth Tone*, 25 June 2019, https://www.sixthtone.com/news/1004175/saving-chinas-dialects%3F-theres-an-app-for-that.

[24] BBC News, 'China to Curb Facial Recognition and Apps in Schools', *BBC News*, 6 September 2019, https://www.bbc.com/news/world-asia-49608459.

[25] Wei Sheng, 'China Amends Minor Protection Law to Include Cyberspace', *TechNode*, 25 October 2019, https://technode.com/2019/10/25/chinas-draft-amendment-to-minors-protection-law-highlights-cyber-regulations/.

[26] Erica Kochi, Jennie Bernstein, and Tushar Ghei, 'Generation AI', *UNICEF*, 10 December 2018, https://www.unicef.org/innovation/stories/generation-ai.

[27] China Power Team. 'How Does Education in China Compare With Other Countries?', *China Power*, 15 November 2016, https://chinapower.csis.org/education-in-china/.

[28] Karen Hao, 'China Has Started a Grand Experiment in AI Education. It Could Reshape How the World Learns', *MIT Technology Review*, 2 August 2019, https://www.technologyreview.com/s/614057/china-squirrel-has-started-a-grand-experiment-in-ai-education-it-could-reshape-how-the/.

[29] Jamil Anderlini, 'How China's Smart-City Tech Focuses on its Own Citizens', *Financial Times*, 5 June 2019. https://www.ft.com/content/46bc137a-5d27-11e9-840c-530737425559.

[30] Sarah Dai, 'China's Subways embrace Facial Recognition Payment Systems Despite Rising Privacy Concerns', *South China Morning Post*, 4 December 2019, https://www.scmp.com/tech/apps-social/article/3040398/chinas-subways-embrace-facial-recognition-payment-systems-despite.

[31] Celia Chen, 'China's War on Trash Goes Hi-tech With AI-driven Apps for Sorting and Facial Recognition to Enforce Recycling', *South China Morning Post*, 18 July 2019. https://www.scmp.com/tech/policy/article/3018992/chinas-war-trash-goes-hi-tech-ai-driven-apps-sorting-and-facial.

[32] Nick Statt, 'China is Fighting Toilet Paper Thieves with Facial Recognition Software', *The Verge*, 20 March 2017. https://www.theverge.com/2017/3/20/14986640/china-toilet-paper-theft-facial-recognition-machine.

[33] Paul Mozur, 'One Month, 500,000 Face Scans: How China Is Using AI to Profile a Minority', *New York Times*, 14 April 2019, https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html.

[34] See for example: Niraj Chokshi, 'Facial Recognition's Many Controversies, From Stadium Surveillance to Racist Software', *New York Times*, 15 May 2019, https://www.nytimes.com/2019/05/15/business/facial-recognition-software-controversy.html; Geoff White, 'Use of Facial Recognition Tech "Dangerously Irresponsible"', *BBC News*, 13 May 2019, https://www.bbc.com/news/technology-48222017; and Sahil Chinoy, 'The Racist History Behind Facial Recognition', *New York Times*, 10 July 2019, https://www.nytimes.com/2019/07/10/opinion/facial-recognition-race.html.

[35] Mike Elgan. 'Uh-oh: Silicon Valley is Building a Chinese-Style Social Credit System', *Fast Company*, 26 August 2019, https://www.fastcompany.com/90394048/uh-oh-silicon-valley-is-building-a-chinese-style-social-credit-system.

07.
The Ideology Behind China's AI Strategy

By Dr Rogier Creemers
Assistant professor in the Law and Governance of China at Leiden University
Associate fellow of the Hague Program for Cyber Norms, Institute of Security and Global Affairs

As China devotes significant efforts to realising its goal of becoming a leading global player in artificial intelligence (AI) technology, the question of which values will be reflected in and furthered by these new capabilities has rapidly gained prominence. In western countries, these questions relate to how China's government might use AI to enhance a digital authoritarian state or to expand surveillance of its citizens. Yet within China, the ethics and governance of AI have also received considerable attention. In the spring and summer of 2019, several companies, business associations and expert groups released principles and policies for the ethical use of AI, which usually highly resemble similar documents in western countries. The New Generation AI Governance Expert Committee, convened by the Ministry of Science and Technology, declared AI should 'conform to human values, ethics and morality [...] should be based on the premise of safeguarding social security and respecting human rights, avoid misuse, and prohibit abuse and malicious application'.[1] The Beijing AI Principles, formulated by the Beijing Academy of Artificial Intelligence, state that 'human privacy, dignity, freedom and rights should be sufficiently respected'.[2] These principles are explored in more detail in Danit Gal's essay in this collection.

To foreign observers, such language appears counterintuitive. China is not usually seen as a staunch defender of human rights or civil liberties, and its use of AI in repressive surveillance programmes is widely reported. How, then, does one make sense of these initiatives? Are they merely a hypocritical charade intended to cover up the unsavoury exercise of power by an autocratic regime and ensure Chinese businesses remain palatable abroad, or do they reflect a way of seeing the world – and the role of digital technology in it – in a different manner? As a set of high-priority technologies with considerable political, economic, social – and international – impact, debates surrounding AI governance are a microcosm of the questions about China's future continually addressed by its leaders. What goals should the technology achieve? Where will it be encouraged, tolerated or rejected? Who will have access to its levers, or the vast amount of data, and who will be exempt from it? This essay argues that Communist Party ideology is the best framing device to understand how China's leadership intends to steer the development of AI. With specific regard to the deployment of AI in governance, the key ideological assumption is that social order is governed by an objective, external and intelligible set of 'laws' (guilü). Big data and AI technologies not only assist in better understanding these laws, they can also help in 'engineering' society to solve development problems. This, in turn, will help the Party achieve its utopian goal: the Chinese dream of the great rejuvenation of the Chinese nation.[3]

'The key ideological assumption is that social order is governed by an objective and intelligible set of laws. Big data and AI technologies not only assist in better understanding these laws, they can also help in 'engineering' society to solve development problems'

This essay will elaborate this point in three sections. First, it will review the different components of China's AI development plans and their driving motivations. Second, it will briefly sketch how Party ideology developed to the point that AI would become a top priority for the leadership. Third, it will discuss the inherent tensions within Party ideology that AI may well lay bare, including the tension between the desire for autonomous decision-making systems and the necessity to maintain the primacy of Party authority.

What are China's AI objectives?

The Chinese leadership finds AI technologies attractive for three major reasons. First and foremost, the 2017 AI Development Plan claims that AI is the 'core driving force for a new round of industrial transformation', which will trigger 'significant changes in economic structure, profound changes in human modes of production, lifestyle, and thinking; and a whole leap of achieving social productivity'.[4] As China's previous growth model of export-led, low-added value manufacturing has run out of steam and its economy is entering a 'new normal' of lower annual GDP growth, AI promises new sources of productivity growth, as well as potential new markets for China's increasingly sophisticated digital goods and services. Second, as a 'strategic technology', AI is also important in national security, most significantly competition with the United States. The People's Liberation Army (PLA) has identified AI as a revolutionary force in military affairs and it is a central element of the 'military-civil fusion' initiative that aims to leverage the strengths of the public sector for China's military build-up.[5] More broadly, as tensions with the United States are growing, and observers increasingly see this as a 'Sino-US technology Cold War',[6] China strives for self-reliance on core technologies, of which AI is one.

The third component is the use of AI technologies in all areas of social governance. The 2017 Plan calls for the widespread use of AI in 'education, medical care, pensions, environmental protection, urban operations, [and] judicial services', in order to enhance citizens' quality of life. But it also proposes coercive applications: 'AI technologies can accurately sense, forecast and provide early warning of major situations for infrastructure facilities and social security operations; grasp group cognition and psychological changes in a timely manner; and take the initiative in decision-making and reactions – which will significantly elevate the capability and level of social governance, playing an irreplaceable role in effectively maintaining social stability.' In other words, China's leadership – in the same paragraph – identifies AI both as a promising conduit for better provision of public services and as a channel for more effective prevention and control of social unrest. The tension between these two sides of the coin has been touched on throughout the essays in this collection, through the lens of different areas of social policy, from education and smart cities to social credit systems.

However, what is striking about the 2017 Plan is not merely the scale of the objectives or the ambition to become an AI world leader by 2030. It is also the level of detail and intricacy with which the path to those objectives is laid out. In terse phrases, it covers interlocking

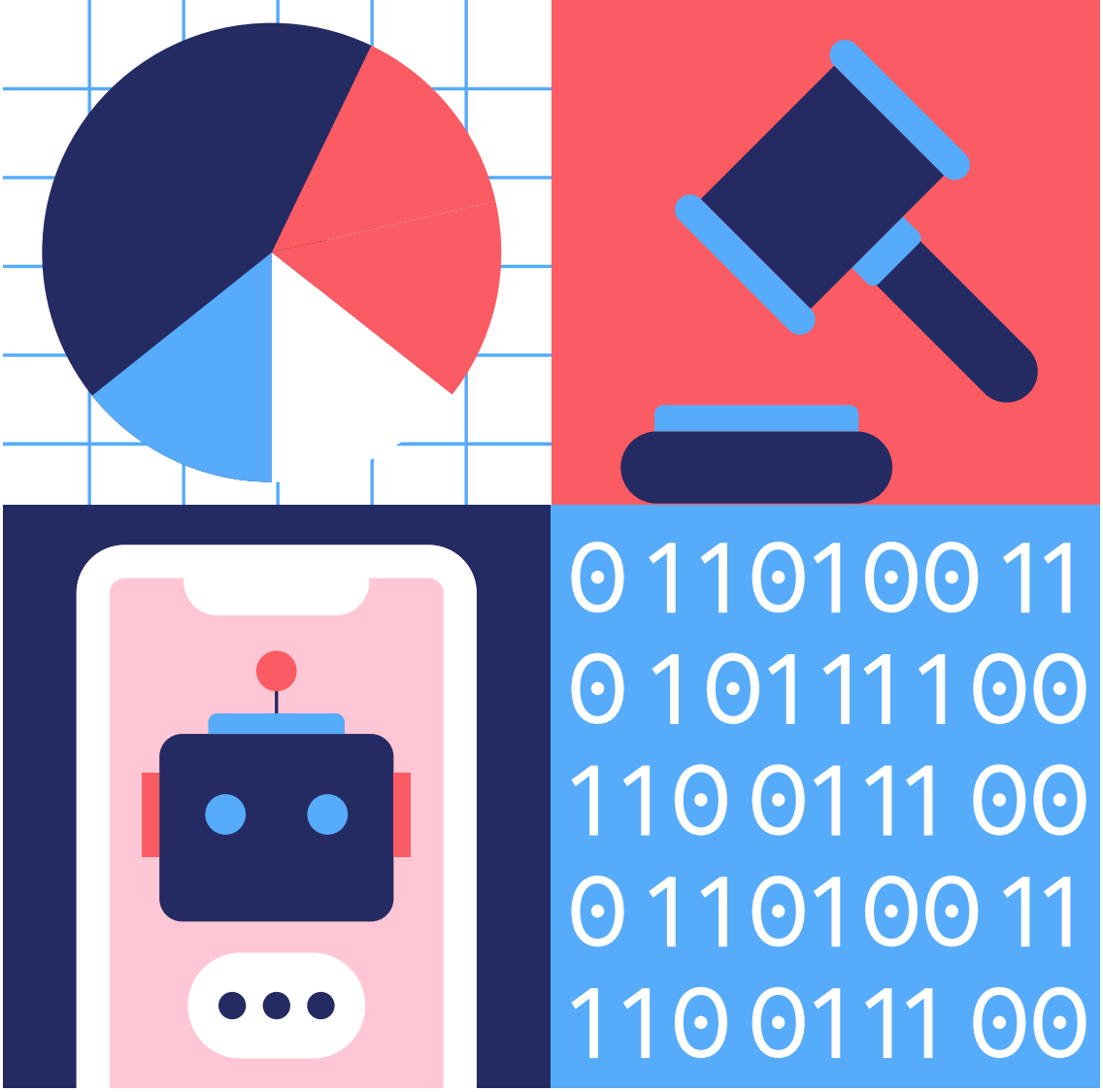
and mutually supportive elements, ranging from supporting scientific research and development to specific AI development focus projects to human resources and financial market reform. This whole-of-society, systems-based view of AI development is only one of many salient examples of the extent to which systems engineering theory has come to underpin Party ideology. It not only reflects how society as a whole is seen as a complex system amenable to intervention, or how the structure of policymaking echoes that system, it also clearly indicates the engineering-type solutions that AI technologies are envisaged to generate.

Why does Party ideology need digital technology?

It is nearly impossible to overestimate the importance of Party ideology in the politics of today's China. It is not only the intellectual glue that holds 90 million Party members together, it is also an intellectual framework that frames the tasks and challenges the Party faces, and shapes and constrains solutions. This does not mean that all Party members hold similar, monolithic opinions, but there is a carefully monitored and manicured corpus of beliefs to which they must demonstrate allegiance.

Like a seabed, Party ideology is a complex accretion of sedimentary ideas, forming over centuries. Some go back to imperial practice, where dominant state doctrine held social order and harmony as the highest good. This required that, in China's highly hierarchical society, everyone learned to know their place and carry out the concomitant duties. After the fall of the empire, Marxism and Leninism were introduced to China in response to deep political crisis, as reformers sought to restore the nation to 'wealth and strength' (fuqiang). Marxism was attractive as it provided a scientific theory for historical progress, aimed at achieving a state of material abundance and, thereby, social harmony. In other words, the essence of what became Communist Party authority is to implement an objective and external scientific theory. Leninism provides the organisational component, in which a disciplined, professional party forms a vanguard leading society towards the promised destination. Yet safeguarding the CCP's supreme authority also became a goal in itself. In the same way that airplane passengers should don their own oxygen mask before helping others, the Party believes enabling its own survival is the chief necessary condition for it to achieve its goal of the 'great rejuvenation of the Chinese nation'.[7] In the past decade, this has led to the construction of an increasingly capable security state,[8] in which AI technology is now starting to play an important role.

Apart from maintaining its dominance, the Party has sought a specific approach to ensure national wealth and strength throughout its tenure in power. The first three decades of Party rule saw a pendulum between Maoist voluntarism and careful Soviet-style economic and social planners. Mao triumphed twice, initiating the disastrous Great Leap Forward and Cultural Revolution. After Mao's death, the new paramount leader, Deng Xiaoping, looked for another way forward. As one of the only Chinese leaders to have travelled abroad, Deng recognised how far China lagged behind developed countries and envisioned a path forward based on science and technology. In this view, science not only provided material progress, it also played an important role in social governance and management. Luminaries such as Qian Xuesen, the progenitor of China's nuclear programme, explored how complex systems theory could both explain the functioning of the social order and provide tools to predict and control it.[9] In this view, social reality is a system conditioned by objective 'laws' (guilü) in the same way as the natural world. These laws are intelligible and can be used to engineer society in pursuit of historical progress. It may be trite (and, indeed, incorrect) to refer to China's leadership as consisting largely of engineers but the metaphor of social management as engineering is strong. Artists, for instance, are still referred to as 'engineers of human souls', while the complex development initiatives such as the 2017 AI Plan display clear influence of systems engineering thinking.



In this ideological view, the keyword is unity. There is only one accepted form of truth, which is both scientifically correct and morally good. There is only one political organisation with a sufficient grasp of that truth to ensure progress. There is only one integral nation and in the words of Xi Jinping, 'only when the country does well, can the individual do well'.[10] This has major implications for notions of social governance. Ideology assumes the world can be perfected and made harmonious. Conflicts and contradictions, therefore, are indications that someone has, consciously or unconsciously, made a mistake. Government should create systems to educate and govern the well-meaning, vast majority of the population so they do not make mistakes, and to protect the Party and the nation from the small number of 'enemies' who seek to undermine its integrity. It should not be encumbered through legal niceties such as powerful individual rights. On the contrary, it should be empowered to make the required changes effectively. This means rights are only possible if there is a powerful state and are always conditional upon them not hindering the execution of the political programme.

However, in practice, the social 'machine' is highly complex. Consequently, realising the vision of engineering social change requires considerable information gathering and processing capabilities. This is where AI and big data technologies come in. For the first time in history, the leadership believes, digital technologies provide the data management abilities and the computational power to permit it to realise its project and propel China into the future. In this new, data-driven future, a number of perennial problems plaguing the Party will be eliminated. Smart devices, for instance, will gather data more efficiently and communicate them more sincerely than notoriously unreliable local governments. Nationwide integrated databases will liberate government-held information from the paper forms and incompatible formats in which they have been imprisoned – the 'data islands' as Dev Lewis describes them in his essay on Social Credit Systems – and thus enable more efficient and automated forms of governance. Curiously, Chinese political leaders seem to share the typical Silicon Valley belief that Evgeny Morozov calls 'solutionism': that the right code, algorithms,

data sources and applications can solve all of mankind's problems and make life 'frictionless' and problem-free.[11] In both China and California there is a drive to use technology to eliminate imperfections in social life and make everything more efficient, through 'programming' the behaviour of individuals.

How does this ideological view influence the use of AI in governance?

As indicated above, both the AI strategy and the Party's overall programme have three main pillars: economic growth, national security and domestic governance. The latter has, over the past decades, been mobilised in order to realise the two former. Yet China is now more wealthy and powerful than it has been in recent history and is on the verge of achieving the first of two markers Xi Jinping put forward: the Party's 100th anniversary in 2021. By this time, China should have become a moderately prosperous society, growing into a fully-fledged developed nation by the second marker: the centenary of the People's Republic in 2049. If China manages to continue along this path, and leaves the crises of its past further behind, the question what sort of society it wishes to be is now becoming more acute. Digital technologies and AI will lie at the centre of these debates.

At this point in the argument, it would be customary to reflect on the likelihood that AI might contribute to political liberalisation in China. However, such expectations are wide of the mark. At least within the foreseeable future, the position of the Party seems to be secure and if fundamental political change were to come to China, chaos and disruption would be far more likely than successful democratisation. Instead, the interesting questions are how the Party could use AI to adapt itself to ever-changing circumstances, and how the development of AI might be constrained by its political context.

First and foremost, it is clear that the use of AI in the domestic security environment will be further expanded in the everlasting pursuit of social and political stability. Yet even if one takes the cynical view that everything the Chinese leadership does is aimed at consolidating Party rule, it must equally be recognised the leadership believes an efficient way of doing so is ensuring the population is, by and large, satisfied. This is often couched in economic terms and GDP growth rates but perhaps equally important is the provision of social services to an ever more demanding populace. Under Xi Jinping's 'new normal', the Party seeks to combine a gradually reducing growth rate with greater attention to public welfare and social development. AI will thus also be deployed to remedy problems ranging from an ageing population and shrinking workforce to the shortage of doctors and teachers in remote and rural areas. China's giant digital corporations will be supportive partners, able to score political points and derive considerable income from helping the state to achieve its goals.

At the same time, two major issues will need to be addressed

The first is the question of data quality and data protection. China does not have, nor is it likely to develop, an all-encompassing privacy right for individuals. It is in the process of developing data frameworks detailing which specific actors can collect, access and process which data. Yet this process has advanced slowly, as behind-the-scenes negotiations between businesses and related government departments continue to unfold. Businesses are, unsurprisingly, not keen to share full access to data with other businesses or government offices. They might lose their competitive advantage, see the risk of data leaks increase, and lose user confidence, and have thus often managed to stave off requirements for doing so.

One reason business data is attractive is that it tends to be more reliable than government data, where incentives for corruption and abuse have proliferated for decades. As the success of any AI application depends on the quality of the underlying data, the success of projects in the public sector may remain below par for quite some time.

The second is the nature of decision-making itself. In many ways, China seeks to use AI and other digital technologies to automate governance, creating self-correcting social systems requiring far less government intervention and supervision than the present model. However, if this requires autonomous decision-making technologies, involving self-improving machine learning algorithms, it threatens the monopoly on political decisions the Party has reserved for itself since 1949. While this is of less importance in areas with clearly defined and generally accepted outcomes, such as healthcare, it may well become a source of tension in politically sensitive areas such as the judiciary. Eugeniu Han touches on this when he considers System 206 smart courts in his essay on smart cities in this collection, and raises the question of the difficulty of challenging AI processes and AI-generated assessments. In contrast, keeping a human in the loop may mean the existing pathologies in the Party-state architecture are sustained. This issue also raises a bigger point about the relationship between AI and means and ends in governance. Based as it is on the idea of perfectibility, the engineering approach may work very well in those areas of social life where the ends are clear and broadly shared, and there is a direct, well-understood connection between inputs and outcomes. Here, AI could play a very important role in raising efficiency and improving social outcomes.

However, many areas of social life are complex and unpredictable, while others are beset by conflicts between values that are all seen as desirable, yet may be incommensurable. In these areas, the engineering approach is highly likely to backfire. This tension can be clearly seen in the use of AI in education, as Yi-Ling Liu explores in her essay, where the most 'efficient' form of education is not necessarily the most effective pedagogically. Technocratic decision-making enhanced by technology is simply not a substitute for the agility and flexibility required to navigate the ultimate difficulty of governance: operating under uncertain circumstances, with unpredictable outcomes, on the basis of incomplete information. This, incidentally, is the same criticism made against the Silicon Valley approach discussed earlier: if the technocratic logic is all, where is the room for judgment and accountability?

Endnotes

[1] 1 Lorand Laskai and Graham Webster, 'Translation: Chinese Expert Group Offers "Governance Principles" for "Responsible AI"', *DigiChina for New America*, 17 June 2019, accessed 5 March 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-expert-group-offers-governance-principles-responsible-ai/.

[2] BAAI, 'Beijing AI Principles', Beijing Academy of Artificial Intelligence, 28 May 2019, https://baip.baai.ac.cn/en.

[3] Xi Jinping, 'Speech at "The Road to Rejuvenation"', *China Copyright and Media*, 19 December 2012, updated 26 February 2014, accessed 5 March 2020, https://chinacopyrightandmedia.wordpress.com/2012/11/29/speech-at-the-road-to-rejuvenation/.

[4] Graham Webster et al., 'A New Generation Artificial Intelligence Development Plan', *DigiChina for New America*, July 20, 2017, accessed 5 March 2020, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

[5] Elsa B. Kania, 'In Military-Civil Fusion, China is Learning Lessons From the United States and Starting to Innovate', *Center for a New American Security*, 27 August 2019, accessed 5 March 2020, https://www.cnas.org/publications/commentary/in-military-civil-fusion-china-is-learning-lessons-from-the-united-states-and-starting-to-innovate.

[6] Adam Segal, 'Year in Review 2019: The U.S.-China Tech Cold War Deepens and Expands', *Council on Foreign Relations*, 18 December 2019, accessed 5 March 2020, https://www.cfr.org/blog/year-review-2019-us-china-tech-cold-war-deepens-and-expands.

[7] Rogier Creemers, 'Party Ideology and Chinese Law', in *Law and the Party in China: Ideology and Organisation*, eds. Sue Trevaskes and Rogier Creemers (Cambridge: Cambridge University Press, forthcoming).

[8] Yuhua Wang and Carl Minzner, 'The rise of the Chinese security state', *The China Quarterly* 222 (2015): 339-359.

[9] Børge Bakken, *The Exemplary Society: Human Improvement, Social Control, and the Dangers of Modernity in China* (Oxford: Clarendon Press, 2000); Mara Hvistendahl, 'A Revered Rocket Scientist Set in Motion China's Mass Surveillance of its Citizens', Science, 14 March 2018, accessed 5 March 2020, https://www.sciencemag.org/news/2018/03/revered-rocket-scientist-set-motion-china-s-mass-surveillance-its-citizens.

[10] Xi, 'Road to Rejuvenation'.

[11] Evgeny Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism* (New York: Public Affairs, 2013)

Conclusion

China is innovating with AI in public services at breathtaking speed. It is striding ahead in terms of its levels of investment, rate of experimentation and adoption, and breadth of applications. Policymakers around the world need to pay attention to developments in China and learn from what's happening if they want to keep pace.

The essays in this collection show the benefit of a nuanced understanding of AI in China and offer some insights that might be unexpected. Challenging common misconceptions helps us understand what is really happening in China – in terms of what technologies are being used, who is involved and what the underlying motivations are.

We learn, for example, that China's innovation hasn't only been driven by cutting-edge technologies but also the skilful deployment and scaling of existing AI technologies. Rather than only looking to the most advanced technologies as solutions, policymakers in Europe could learn from China's approach and consider how harnessing and scaling already-existing technologies might be an equally innovative way to improve services and solve problems.

We also learn that China is engaged in discussions about the ethics of AI. This provides an opportunity to bring China into global AI governance conversations, and for Europe to draw on its strengths and be a co-operation partner in areas such as data protection and AI ethics.[1]

While China's authoritarian politics mean that there are significant elements of the government's approach that should be condemned, it is important not to lose sight of all of China's advances. Even if we disagree with China's political system, there are still elements of its AI innovation that we could usefully learn from, and interesting ideas that could be adapted for a European context to benefit citizens.

There are insights to be gained from China's local innovation ecosystems and how they drive AI development and implementation – policymakers in Europe could assess which cities have the right conditions to build this kind of AI innovation ecosystem and support these to develop, through capital and other policy supports. In light of the UK's regional levelling-up ambitions, this approach might be a particularly promising way to boost research and innovation throughout the UK.

China's agile and experimental approach to finding new AI applications for real-world problems is also worth attention. While it would not be replicable in a European context due to the regulatory environment, it is worth considering how experimentation methods such as regulatory sandboxes and innovation testbeds could be used to foster rapid AI innovation in important application areas.

Observing China's use of AI in public services with a critical eye helps us pinpoint what's concerning about China's approach and identify where AI systems are being designed or implemented in ways that perpetuate inequalities, diminish citizens' rights or give too much power to tech companies or governments. These insights offer a cautionary tale for what we should strive to avoid, and remind us of the need for open and equitable AI systems.

There is a need for transparency about where, how and why AI is used in public services, and for accountability structures to be put in place, so that algorithmic decision-making processes remain open and auditable and citizens are able to challenge decisions they deem unfair.[2]

Citizens and civil society stakeholders should have a voice in decisions around data collection and the development of AI, to ensure that systems are designed and implemented in ways that are inclusive and protect citizens' digital rights.[3]

It is critical that policymakers around the world pay attention to how China is innovating with AI in public services if they want to learn from and keep pace with China's advances. We hope these essays give readers a deeper understanding of China's use of AI and challenge existing thinking. As well as offering new and useful ideas and insights, this collection should help readers to reflect on fundamental questions around the broader ethical, social and political implications of applying AI in public services.

Hessy Elliott,
Government Innovation Research, Nesta

Endnotes

[1] Kai von Carnap and Kristin Shi-Kupfer, 'Values as a Quality Seal: Germany Should be More Active in Shaping Cooperation with China', *MERICS Blog* - European Voices on China, 4 September 2019, Mercator Institute for China Studies (MERICS), https://www.merics.org/en/blog/values-quality-seal-germany-should-be-more-active-shaping-cooperation-china.

[2] Eddie Copeland, '10 principles for public sector use of algorithmic decision making', *Nesta (blog)*, 20 February 2018, https://www.nesta.org.uk/blog/10-principles-for-public-sector-use-of-algorithmic-decision-making/.

[3] 'Digital Rights Policy-Making', Digital City Toolkit, https://sokotech_team.gitlab.io/unhabitat/digital_city_toolkit.html; Tom Saunders, 'How to involve the public in the development of artificial intelligence', *Nesta (blog)*, 18 December 2017, https://www.nesta.org.uk/blog/how-to-involve-the-public-in-the-development-of-artificial-intelligence/.

Contributors

Andy Chun

Dr Andy Chun is a council member and convenor of the AI Specialist Group of the Hong Kong Computer Society and an adjunct professor at the City University of Hong Kong. Andy is a seasoned technology executive and an AI pioneer with over 30 years of experience. Currently, he is with the global financial services group Prudential plc as the regional director of technology innovation for Asia, providing leadership in the use of AI for digital transformation. Previously, Andy was the chief information officer at City University of Hong Kong and an associate professor at its Computer Science Department. Andy holds a PhD in Electrical Engineering from the University of Illinois at Urbana-Champaign and attended Executive MBA training at Boston University.

Jeffrey Ding

Jeffrey Ding is a DPhil researcher at the Centre for the Governance of Artificial Intelligence, within the University of Oxford's Future of Humanity Institute, where he leads the centre's research on China's AI development. Interested in the relationship between technological change and the rise and fall of great powers, Jeff is reading for a DPhil (PhD) in International Relations as a Rhodes Scholar at the University of Oxford. His writing has appeared in forums such as *Foreign Affairs* and *MIT Technology Review*, and his research has been cited in the *Washington Post*, *Financial Times* and other outlets. Jeff is the creator of a weekly *ChinAI* newsletter featuring translations of Chinese media and scholarship on AI-related topics.

Rogier Creemers

Dr Rogier Creemers is an assistant professor in the law and governance of China at Leiden University. He holds degrees in Chinese Studies and International Relations, and a doctorate in Law. His research focuses on two major projects: China's efforts to build itself into a digitised 'Smart State', and the impact of China's technological rise on global cyber affairs. Rogier's research has been published in journals including the *China Journal* and the *Journal of Contemporary China*, and he is a frequent commentator in news media including the *New York Times*, the *Economist* and the BBC. He is a co-founder of the DigiChina programme at New America, and a fellow of the Hague Project for Cyber Norms.

Danit Gal

Danit Gal is technology advisor to the UN Secretary General's High-level Panel on Digital Cooperation. Previously, she was project assistant professor at the Cyber Civilizations Research Centre at the Keio University Global Research Institute in Tokyo, Japan and a Yenching fellow at Peking University. Danit chairs the IEEE P7009 standard on the Fail-Safe Design of Autonomous and Semi-Autonomous Systems and is an associate fellow at the Leverhulme Centre for the Future of Intelligence at the University of Cambridge, and an affiliate at the Center for Information Technology Policy at Princeton University.

Eugeniu Han

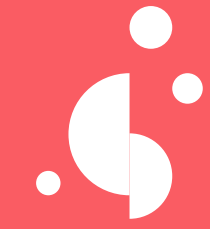
Eugeniu Han is a PhD candidate in Policy Analysis at the Pardee RAND Graduate School and an assistant policy researcher at the RAND Corporation. His research interests lie at the intersection of technological innovation, international economics and national security. He has also co-authored several studies in education and health policy. Eugeniu received a master's degree in Public Policy from the University of Chicago, specialising in economics and data analysis, and a Master's degree in Management from Tsinghua University. He completed his undergraduate studies in Chinese and Russian languages at the University of Bucharest.

Yi-Ling Liu

Yi-Ling Liu is a nonfiction writer covering technology, culture and society in China. She is visiting scholar at the NYU Arthur L Carter Journalism Institute, a recipient of the Matthew Power Literary Reporting Award and the Overseas Press Club Foundation Fellowship and her work has been published in the *New York Times Magazine*, the *New Yorker*, the *Economist* and *Foreign Policy*. As an educator, she has designed workshops and directed summer camps for high-schoolers around the world. She is interested in storytelling as pedagogy, cultivating self-authorship and preparing the next generation of students for the age of automation.

Dev Lewis

Dev Lewis is a fellow and program lead at Digital Asia Hub and a Yenching scholar at Peking University. His research interests lie at the intersection of technology, politics and policy, and he is currently conducting research on the Social Credit System in China. Dev has an International Relations degree from Roger Williams University and studied Mandarin in Shanghai and Zhengzhou. He frequently writes for several regional publications and think tanks, and previously spent time at Gateway House, a foreign policy think tank in Mumbai, and Infosys China, an IT services Multinational. He also writes a bi-weekly newsletter *ChinaIndia Networked*.



58 Victoria Embankment
London EC4Y 0DS

+44 (0)20 7438 2500

information@nesta.org.uk
facebook.com/nesta.uk
nesta.org.uk
@nesta_uk