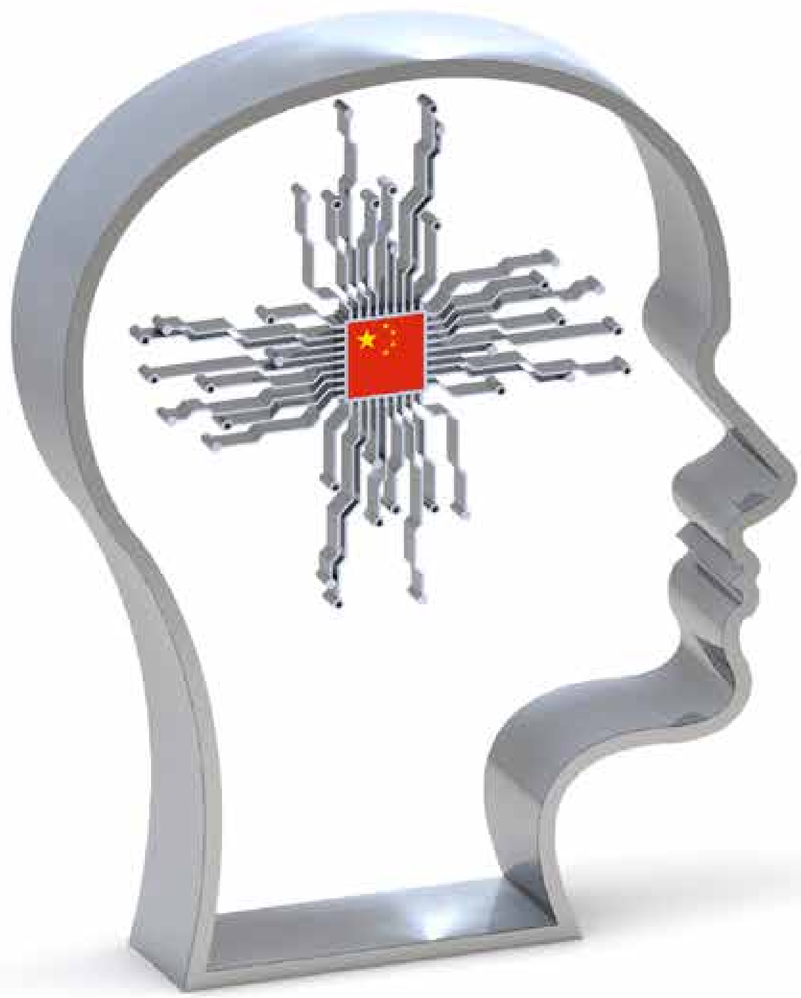
Centre for International Governance Innovation

SPECIAL REPORT

Competing in Artificial Intelligence Chips

China's Challenge amid Technology War

Dieter Ernst



Centre for International
Governance Innovation

SPECIAL REPORT

Competing in Artificial Intelligence Chips

China's Challenge amid Technology War

Dieter Ernst

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

À propos du CIGI

Le Centre pour l'innovation dans la gouvernance internationale (CIGI) est un groupe de réflexion indépendant et non partisan dont les recherches homologuées par des pairs et les analyses fiables incitent les décideurs à innover. Grâce à son réseau mondial de chercheurs pluridisciplinaires et de partenariats stratégiques, le CIGI offre des solutions politiques adaptées à l'ère numérique dans le seul but d'améliorer la vie des gens du monde entier. Le CIGI, dont le siège central se trouve à Waterloo, au Canada, bénéficie du soutien du gouvernement du Canada, du gouvernement de l'Ontario et de son fondateur, Jim Balsillie.

Credits

Director, Global Economy Robert Fay
Program Manager Heather McNorgan
Senior Publications Editor Jennifer Goyder
Publications Editor Susan Bubak
Graphic Designer Brooklynn Schwartz

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org

Table of Contents

About the Author

Senior Fellow Dieter Ernst joined CIGI in May 2016. At CIGI, Dieter's research explores unresolved challenges for the global governance of trade, intellectual property (IP) and innovation, addressing three issues in particular: finding out what adjustments are needed in the development and use of IP, especially patents and trade secrets, to cope with the requirements of increasingly complex and diverse global corporate networks of production and innovation; dealing with the effects of the proliferation of strategic patenting behaviour on the organization and governance of these global networks; and assessing US-China technology competition in information technology.

Based in Hawaii, Dieter is an adjunct senior fellow at the East-West Center. He has served as a member of the US National Academies' Committee on Global Approaches to Advanced Computing, as a senior adviser to the Organisation for Economic Co-operation and Development in Paris and as a research director of the Berkeley Roundtable on the International Economy at the University of California, Berkeley.

Previously, Dieter was a professor of international business at the Copenhagen Business School and served as a scientific adviser to governments, private companies and international institutions, including the World Bank, the UN Conference on Trade and Development and the UN Industrial Development Organization.

He holds a Ph.D. in economics from the University of Bremen.

Acronyms and Abbreviations

3GPP	3rd Generation Partnership Project
5G	fifth-generation
AGI	artificial general intelligence
AGI-19	12th Conference on Artificial General Intelligence (2019)
AI	artificial intelligence
AIDP	Next Generation Artificial Intelligence Development Plan
ASICs	application-specific integrated circuits
BIS	Bureau of Industry and Security (US Department of Commerce)
CFIUS	Committee on Foreign Investment in the United States
CISTP	China Institute for Science and Technology Policy
CPUs	central processing units
DARPA	Defense Advanced Research Projects Agency
DoD	Department of Defense
DRAM	dynamic random-access memory
ECRA	Export Control Reform Act
fabs	semiconductor fabrication plants
FDD	frequency division duplex
FinFET	fin field-effect transistor
FIRRMA	Foreign Investment Risk Review Modernization Act
FPGAs	field-programmable gate arrays
GPU	graphics processing unit
GVCs	global value chains
HR	Horizon Robotics
IaaS	infrastructure-as-a-service
IC	integrated circuit
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IP	intellectual property
IT	information technology
LOB	line of business
MIC 2025	Made in China 2025
nm	nanometre
OECD	Organisation for Economic Co-operation and Development
R&D	research and development
RMB	renminbi
SASAC	State-owned Assets Supervision and Administration Commission
SEPs	standard-essential patents
SMIC	Semiconductor Manufacturing International Corporation
SOEs	state-owned enterprises
TDD	time division duplex
TSMC	Taiwan Semiconductor Manufacturing Company
WIPO	World Intellectual Property Organization

Executive Summary

This special report assesses the challenges that China is facing in developing its artificial intelligence (AI) industry due to unprecedented US technology export restrictions. A central proposition is that China's achievements in AI lack a robust foundation in leading-edge AI chips, and thus the country is vulnerable to externally imposed supply disruptions. Success in AI requires mastery of data, algorithms and computing power, which, in turn, is determined by the performance of AI chips. Increasing computing power that is cost-effective and energy-saving is the indispensable third component of this magic AI triangle.

Research on China's AI strategy has emphasized China's huge data sets as a primary advantage. It was assumed that China could always purchase the necessary AI chips from global semiconductor industry leaders. Until recently, AI applications run by leading-edge major Chinese technology firms were powered by foreign chips, mostly designed by a small group of top US semiconductor firms. The outbreak of the technology war, however, is disrupting China's access to advanced AI chips from the United States.[1]

Drawing on field research conducted in 2019, this report contributes to the literature by addressing China's arguably most immediate and difficult AI challenges:

→ How to ensure that the country's AI developers, implementers and users have secure access to specialized semiconductors that are needed for training an algorithm and for conducting inference with an already trained algorithm?

→ In the face of US technology restrictions, what realistic options does China have to substitute AI chip imports from the United States through local design and fabrication or through imports from other non-US sources?

The report highlights China's challenge of competing in AI, and contrasts America's and China's different AI development trajectories. Starting much later than the United States, Chinese universities and public research institutes have conducted a significant amount of AI research (some of it at the frontier), but knowledge exchange with industry remains limited. Drawing on deep integration with America's AI innovation system, Chinese AI firms, in turn, have focused primarily on capturing the booming domestic mass markets for AI applications, investing too little in AI research.

To find out what is happening today in China's AI chip design, capabilities and challenges are assessed, both for the large players (Huawei, Alibaba and Baidu) and for a small group of AI chip "unicorns." The report concludes with implications for China's future AI chip development.

1 As this report goes to press, the rapid spread of the coronavirus (COVID-19), first identified in Wuhan, China, in late 2019, is further decoupling China from international trade and technology flows.

Introduction

China's Next Generation Artificial Intelligence Development Plan (AIDP), released by the State Council in July 2017, provides a detailed road map for developing an increasingly integrated AI ecosystem.[2] An essential part of the Made in China 2025 (MIC 2025) plan, the AIDP seeks to use AI as a catalyst for upgrading China's manufacturing and service industries and for catching up with the United States and other advanced economies.[3] With full backing from President Xi Jinping, support institutions were established within a few months to coordinate interministerial implementation of the AIDP and to involve China's leading AI industry players.

In the United States, a broad consensus exists among defence and international relations elites and economic policy makers that the AIDP poses a serious threat to America's leadership in science and advanced technology (National Security Commission on Artificial Intelligence 2019, 18–20). For Michael Kratsios (2019), US President Donald Trump's chief technology officer, China has become a strategic competitor, an "adversarial nation and bad actor" that should no longer be allowed "to steal our ideas, copy our technology and cheat their way to leadership, in a field central to our national security."

The Outbreak of Technology War

These words have been followed by action. As part of the trade war that the United States is prosecuting against China, the Trump administration has sharply increased the range of restrictions on China's access to advanced US technology. The Commerce Department has placed Huawei and China's leading AI start-up companies on its so-called Entity List, which is basically a trade blacklist that bars anyone on it from buying advanced semiconductors and software from US companies without the government's approval first.[4] The US Treasury's Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA) has drastically expanded the mandate of the Committee on Foreign Investment in the United States (CFIUS),[5] signalling a fundamental paradigm shift in US policy toward foreign investment. Washington now has in place "the most comprehensive — and activist — regime of national security regulation of inbound foreign direct investment (FDI) among the advanced economies" (Broadman 2019). As CFIUS has tightened the screws, Chinese investment in US firms with "sensitive" technology faces almost insurmountable hurdles.[6]

The Export Control Reform Act (ECRA) of 2018 seeks "to enhance protection of U.S. technology resources by imposing greater restrictions on the transfer to foreign persons — particularly through exports to China — of certain key emerging and foundational technologies and cybersecurity considered critical to U.S. national security, including technical capabilities, specifications and related knowledge, including through joint ventures" (Braverman 2018). It is important to emphasize that the United States defines an export in extremely broad terms. It is not just the

2 For an English translation of the AIDP, see www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/.

3 Three reports provide a detailed analysis of China's AI strategy. The *China AI Development Report 2018* (China Institute for Science and Technology Policy [CISTP] (2018) examines data on China's AI talent, AI research papers and AI patenting, and its critical importance for implementing China's MIC 2025 plan. The China AI Index 2018 report, jointly prepared by the Center for AI and Institutions (CAII) of the Cheung Kong Graduate School of Business, and the Big Data and Cloud Computing Lab of Wuhan University, www.ckgsb.edu.cn/uploads/《中国人工智能指数2018》.pdf (in Chinese); and the report *China New Generation Artificial Intelligence Development Report 2019*, jointly prepared by the Ministry of Science and Technology and the China Academy of Science, which examines implementation and progress of the AIDP, but only a summary is available thus far (see www.xinhuanet.com/tech/2019-05/24/c_1124539084.htm). See also the study prepared for the European Commission, China's *"1+N" funding strategy for Artificial Intelligence* (Development Solutions for Europe Ltd. 2018).

4 According to a Federal Register posting, "the US government will review license applications under a policy presumption of denial" (see www.federalregister.gov/documents/2019/05/21/2019-10616/addition-of-entities-to-the-entity-list). "In this rule, the Bureau of Industry and Security (BIS) amends the Export Administration Regulations (EAR) by adding Huawei Technologies Co., Ltd. (Huawei) to the Entity List. The U.S. Government has determined that there is reasonable cause to believe that Huawei has been involved in activities contrary to the national security or foreign policy interests of the United States" (ibid.).

5 For details, see US Department of the Treasury (2019) and Congressional Research Service (2019).

6 Note, however, that these hurdles as yet are not watertight. An example is the CFIUS decision in April 2019 to clear the acquisition of Beijing OmniVision Technologies Co., Ltd. by Shanghai Will Semiconductor Co., a China-listed company. California-based OmniVision designs advanced digital imaging technologies that could be used for security and surveillance purposes. It is unclear what exactly motivated CFIUS's decision. At the very least, this indicates that even in the midst of the US-China technology war, "the two largest economies in the world continue to find ways to do business together" (Zou 2019).

sale of a final good across borders. Basically, any transfer of information to a foreign national can be deemed an export — it is a "deemed export," which requires an export licence. A primary focus of US technology warfare thus is on blocking knowledge exchange, which is the lifeblood of the global AI research. According to David Shambaugh (2019, 3), a leading China scholar, "academic exchanges have…suffered and decoupled to a significant extent. Joint academic research has always been difficult to forge outside of the sciences, but now it is nearly impossible. Further, but related, both governments have increasingly restricted visas for scholars and non-governmental organizations. This is a 'race to the bottom' that harms both sides. Thus, there has been a partial decoupling of the intellectual communities of both countries already."

For instance, if a US citizen speaks to a foreign national or just sends them an email, that can be declared an export by the Department of Commerce and can be prohibited. US firms or research and development (R&D) labs may have to sequester and lay off portions of their foreign staff, otherwise they may be in violation of US export control laws.

A primary focus of US technology warfare is on blocking knowledge exchange, which is the lifeblood of the global AI research.

Chinese citizens who work at US firms or at research labs in American universities have long accounted for most of these licences. Since 2017, however, the United States has sharply reduced export licences of Chinese nationals (see figure in Appendix 1). US visa restrictions have increased, obstructing knowledge exchange across borders. As a result, US university labs and US companies will find it more difficult to recruit and retain Chinese talent, which is likely to constrain the diffusion of AI technologies.[7]

The ECRA not only controls the export of items produced in the United States. The law also controls the export of items produced in foreign states incorporating technology previously exported from the United States, if the share of US technology exceeds 25 percent, the so-called 25 percent US content rule.[8] In January 2020, the US government agencies responsible for export controls — Commerce, State, Energy and Defense — have sent to the Office of Management and Budget new regulations that would largely eliminate a loophole that allowed US companies to sell to Huawei from their overseas facilities. "The new proposal is that, for sales to Huawei only, the de minimis threshold would be lowered to 10% or even zero" (Kroeber and Wang 2020). These regulations will broaden the list of restricted semiconductors (Leonard and King 2019). In addition, they would "enable Commerce to impose export controls on foreign-made products even if they had no direct inputs of controlled US technology, but were manufactured based on US designs and blueprints. One of the biggest potential targets here is Taiwanese chip fabricator TSMC, which manufactures a large share of the semiconductors used by Huawei" (Kroeber and Wang 2020).[9]

Such extreme export restrictions describe a unique feature of US trade policy — the exterritorial reach of US trade law reflects a long-established US policy of pursuing certain foreign policy goals through the extraterritorial application of its export control laws.[10] In addition, ECRA requires the Commerce Department to create lists of "emerging" and "foundational" technologies that are essential to US national security.[11] Far-reaching new restrictions are expected soon on the export of "emerging technologies" to China, including biotechnology, AI and machine-learning technology, quantum information and sensing technology, additive manufacturing (for example, 3D printing) and robotics. Appendix 2 provides the list of AI-related emerging technologies considered to be essential to US national security, proposed by the US Commerce Department's Bureau of Industry and Security (BIS).[12]

7 In response to the coronavirus crisis, the US government has increased visa and travel restrictions for Chinese citizens, further constraining knowledge exchange that is critical for the AI industry.

8 See www.bis.doc.gov/index.php/documents/pdfs/1382-de-minimis-guidance/file.

9 For details, see the section "What's Happening in China's AI Chip Industry?" (pages 27–44).

10 See Editors (1984).

11 See www.congress.gov/bill/115th-congress/house-bill/5040/text?format=txt.

12 With regard to AI, the Commerce Department noted that its review falls under three categories: analyzing whether new processors continue to be captured appropriately under existing controls for AI technologies; identifying AI technologies that are not currently controlled; and identifying very specialized applications of AI that should be controlled.

From China's perspective, it is crystal clear that US policy seeks to prevent China from catching up. No country in China's position could reasonably allow itself to be put under such existential economic pressure. In response, the Chinese government is pushing toward increasing self-reliance. China has long factored in the risk of future technology restrictions and has searched for ways to develop alternative sources of supply. In fact, the possibility of the United States mounting technology restrictions on China has long been the subject of intense study in both countries — various scenarios have been war-gamed by both parties. The Trump administration's policies are thus hardly unstudied — and China has been preparing contingency plans for this eventuality for quite some time.

To all appearances, both parties were caught somewhat flatfooted. The United States appears to have been surprised by the speed with which Chinese technology has developed and found itself actually trailing China in fifth-generation (5G) capabilities (the wireless technology for digital cellular networks). Moreover, the US restrictions have had unanticipated negative consequences for US information technology (IT) firms, resulting in significant pushback from US interests. In turn, China's leadership appears to have been unprepared for the sudden escalation of the trade conflict into a technology war and finds itself with its contingency plans inadequately advanced.

Today, the die is effectively cast for growing conflict. As established rules of trade are broken, mutual distrust and rising uncertainty are beginning to result in a "decoupling" of trade, investment and knowledge networks between the United States and China. Even if the initial US-China trade deal, signed on January 15, 2020, would lead to some kind of temporary truce, technology competition in AI thus is likely to intensify between both countries. An extended period of technology warfare might have quite distorting implications on both the United States and China, as well as for the global AI industry at large. To counter the political pressure for decoupling, it is necessary to revisit the assumptions underpinning America's claim that China is about to overtake it as the AI technology leader. The real issue missing in this narrative is that China, despite impressive achievements in this industry, still has a long way to go to reduce the huge technology gap that separates it from the United States.

Key Question

This special report assesses the challenges that China is facing in developing its AI industry in the face of the unprecedented US technology export restrictions. A central proposition is that China's achievements in AI lack a robust foundation in leading-edge AI chips, and thus the country is vulnerable to externally imposed supply disruptions.[13] Success in AI requires mastery of data, algorithms and computing power, which, in turn, is determined by the performance of AI chips. Increasing computing power that is cost-effective and energy-saving is the indispensable third component of this magic AI triangle.[14]

The literature on China's AI strategy has shown that data arguably constitutes China's primary AI advantage. With fewer obstacles to data collection and use, China has amassed huge data sets that do not exist in other countries (Knight 2017, 5). Of critical importance is China's huge cost advantage in big data management — a huge population of low-cost college students work long hours doing the repetitive work of categorizing huge troves of data needed to train algorithms (Lee, 2018, 14). In addition, progress in algorithms has been impressive. While open-source platforms TensorFlow and Caffe, developed by US academics and companies, are widely used to design and train AI algorithms, China's leading AI firms have succeeded in using and upgrading existing algorithms (often available on an "open-source" basis) to develop new and low-cost mass applications of AI.

Research on China's AI strategy has largely neglected the role of computing power, as it was assumed that China could always purchase the necessary AI chips from global semiconductor industry leaders. In fact, until recently, AI applications run by leading-edge major Chinese technology firms were powered by foreign chips, mostly designed by a small group of leading US semiconductor firms. The outbreak of the technology war, however, threatens to disrupt China's access to advanced AI chips from the United States.

This report contributes to the literature by addressing China's arguably most immediate and difficult AI challenges:

→ How to ensure that the country's AI developers, implementers and users have secure access to specialized semiconductors that are needed for training an algorithm and for conducting inference with an already trained algorithm?

→ In the face of US technology restrictions, what realistic options does China have to substitute AI chip imports from the United States through local design and fabrication or through imports from other non-US sources?

The analysis introduces new findings based on field research in China's AI industry, conducted in April 2019 with a group of researchers from the CISTP at Tsinghua University.[15] Compared to my earlier field research in China (which goes back many years), this time companies and government agencies were much more reluctant to accept interviews. Initially, there was resistance against a foreigner participating.[16] We were, however, able to conduct official interviews with a carefully selected sample of major companies and academic advisors.[17] I also gave invited talks on the topic of this research to leading universities in Beijing, Shanghai and Hangzhou, in April 2019 and, most recently, in November 2019. These talks were well attended and provided some unique insights in the extended Q&A sessions. In all of my discussions during that visit, Chinese participants were

13 This is in line with the assessment by Gao Wen of the Chinese Academy of Engineering, who highlights AI chips as one of China's key AI weaknesses (presentation at the China Big Data Industry Expo in May 2019, quoted in Nelson [2019]). In a similar vein, Li Deyi, the president of the Chinese Association for Artificial Intelligence, highlights China's lack of basic infrastructure, especially AI chip design, as a major weakness (presentation at the Wu Wenjun AI Science and Technology Award ceremony in Suzhou, November 30, 2019, quoted in Laskai and Toner [2019]).

14 See the path-breaking study by Tim Hwang (2018), a former Google researcher, who now directs the Harvard-MIT Ethics and Governance of AI Initiative, entitled "Computational Power and the Social Impact of Artificial Intelligence."

15 This field research was exclusively funded by CIGI. I am grateful to Liang Zheng, director of CISTP, for organizing this field research, and for joining me in the interviews, together with Yu Zhen, Li Dan and Dai Tian. In Hangzhou, our research team was also joined by Yu Hanzhi, from Zhejiang University.

16 This shows that the intensifying US-China trade and technology war has made objective field research much more difficult.

17 Between April 16 and 28, 2019, official interviews were conducted with Microsoft Asian Research Institute; Baidu; SenseTime/商汤; the Department of Microelectronics, Tsinghua University; Megvii; Internet Development Research Institute at Peking University/北京大学互联网发展研究院; Beijing Academy of Artificial Intelligence/北京智源人工智能研究院; Unisound/云知声; Huawei; School of Microelectronics, Zhejiang University; Netease/网易; Alibaba; Transwarp/星环科技; and Horizon Robotics (HR). Given the sensitivity of this type of field research, the names and identities of interviewees are kept in confidence.

very concerned whether US technology export restrictions might lead to a progressive bifurcation of the global AI value chain, strictly separating a US-centred AI system from one centred on China.

The report is organized as follows. To highlight China's challenge of competing in AI, the first part provides background information, highlighting patterns of competition in this emerging technology. The analysis focuses on fundamental cognitive limitations of current systems of "narrow AI"; the struggle between competing technologies whose design features remain fluid; and the resultant increasing complexity of global value chains (GVCs). To cope with these challenges, close interaction is required between AI research and the development of AI applications. At the same time, knowledge diffusion across countries should be facilitated through an open rules-based international trading system, and should not be systematically obstructed through trade and technology warfare.

The second part contrasts America's and China's different AI development trajectories. In the United States, breakthrough AI research has played a prominent role from the very beginning, closely interacting with the development of AI applications. A defining characteristic of AI development in the United States has been the constant back and forth of new ideas between academic research, vibrant start-up clusters in California and Massachusetts, and the vast network of projects funded by the Defense Department / Defense Advanced Research Projects Agency (DARPA).

China's approach is very different. Starting much later, Chinese universities and public research institutes have conducted a significant amount of AI research (some of it at the frontier), but knowledge exchange with industry remains limited. Drawing on deep integration with America's AI innovation system, Chinese AI firms, in turn, have focused primarily on capturing the booming domestic mass markets for AI applications, investing too little in AI research. China's resulting heavy reliance on foreign sources of AI technology (especially for AI chips) has become a major vulnerability with the outbreak of the US-China technology war.

To deepen this analysis, the third part will take a closer look at China's efforts to ensure secure access to AI chips in the face of intensifying technology warfare. What are China's realistic options to develop its domestic AI chip industry? What new opportunities might open up for China to leapfrog into at least some niches of the AI chip market, given the big changes in chip architectures that are disrupting the global semiconductor industry? A review of China's position in the design and fabrication of semiconductors will show that, despite rapid catching up, major weaknesses remain. To find out what is happening in China's AI chip design, capabilities and challenges will be assessed, both for the large players (Huawei, Alibaba and Baidu) and for a small group of promising AI chip "unicorns."

The conclusions present implications for China's future AI chip development.

券交易所行情显示屏
涨跌%
成交量
今开盘
最高价
最低价
昨收盘
深圳
股票名称
代 码
怡 亚 通
海得控制
华天科技
全 聚 德

Competing in AI: Major Challenges

Defined as machines that mimic human cognitive functions, AI seeks to enhance pattern recognition and statistical analysis, and provides basic learning and problem-solving capabilities. It has taken more than 60 years for AI to advance to the point where it is now beginning to affect our daily lives, how we work and how we conduct manufacturing and services, sciences and the government. A recent study summarizes the expected wide-ranging impact: "Artificial intelligence promises to improve existing goods and services, and, by enabling automation of many tasks, to greatly increase the efficiency with which they are produced. But it may have an even larger impact on the economy by serving as a new general-purpose 'new method of invention' that can reshape the nature of the innovation process and the organization of R&D" (Cockburn, Hendersen and Stern 2017).

Fundamental Limitations

Fundamental limitations remain, however. Despite its long gestation, something even remotely akin to human intelligence is not within reach. A particular AI program can only address a unique application — Google's AlphaGo program does not play chess; nor can AI-aided interpretation of cancer images do any other task. In addition, it is very costly to develop the software for each of these specific applications. Take LawGeex AI, a widely used program to automate the review and approval of everyday business contracts before signing. A recent study shows that leading US-trained legal academics and experts are no match for the LawGeex AI algorithm. However, there is a substantial hidden cost —- the algorithm's development required $21.5 million in development funding.[18] Most importantly, if an attempt was made to expand this software to a different legal application such as writing an affidavit, the algorithm would then need to be trained all over again just to reach a comparable level of competence to that of a human lawyer.[19]

There is, of course, "blue-sky" fundamental research that seeks to advance, however gradually, toward artificial general intelligence (AGI) that could successfully perform any intellectual task that a human being can master.[20] This research agenda is outlined in the proceedings of the annual AGI conferences, which have been organized since 2006 by the AGI Society, in cooperation with the Association for the Advancement of Artificial Intelligence (AAAI). In fact, the most recent AGI conference on August 6–9, 2019, took place in Shenzhen, China.[21] Despite the current US-China technology war, broad international participation has continued to shape the conference committees.[22]

Yet, most of the research dollars and corporate investments are spent on so-called "narrow AI," which remains focused on the use of software and statistics to study or accomplish specific problem-solving or reasoning tasks (Nilsson 2010; Norvig 2003). This has given rise to a boom in AI applications covering a wide array of activities, such as visual perception, speech recognition, translation, as well as smart and interactive robots, cars, drones, medical equipment and weapon systems.[23]

Enablers of AI Applications: Vertical Specialization Facilitates Access to AI Tools

What explains the rapid growth of AI applications despite the persistent limitations of narrow AI? An important enabling factor has been the rapid development of the tools and techniques by specialized engineers that enable a computing system to learn with data, without the need to be explicitly programmed. The development of an integrated AI ecosystem (an "AI stack" in industry parlance) has enabled AI application developers to explore "make-versus-buy"

18 All dollar figures are in US dollars.

19 This is the essence of the steep economies of scale in AI applications — large upfront costs that then permit monopolization of the market. Hence, access to global markets becomes important for highly refined and specialized AI. In areas such as law, where national systems (sometimes even regional systems) are idiosyncratic, the scale might not be there to make AI development commercially viable (Dan Ciuriak, email to the author, September 29, 2019).

20 See https://en.wikipedia.org/wiki/Artificial_general_intelligence#cite_ref-K_1-2.

21 See http://agi-conf.org/2019/.

22 See http://agi-conf.org/2019/committees/.

23 While much of the debate focuses on commercial applications for civilian purposes, defence and security applications of AI are rapidly growing. See Horowitz et al. (2018); Lamm (2019).

options for critical AI tools. Typically, an AI stack consists of three components — infrastructure, developer environment and line of business (LOB) applications and services.

Infrastructure refers to the tools, platforms and techniques used to run and store data, build and train AI algorithms, and the algorithms themselves. This includes data, the computing power required to run AI algorithms, the AI algorithms themselves and the critically important machine learning platforms (such as TensorFlow or Caffe2). There is a wide choice of data platforms available — structured and non-structured databases, big data platforms, managed databases and cloud-based databases. Sources of computing power include a wide variety of specialized AI chips and servers.[24]

The developer environment refers to the tools that assist in developing code to bring out AI capabilities. This includes a variety of libraries, either for advanced mathematical operations or for adding cognitive capability, say in computer vision. LOB applications and services are technically not part of the AI stack. They derive value from the AI stack. For the AI application developer, this provides a convenient short-cut — they can develop an application without being forced to develop all the different components needed. Vertical specialization through outsourcing is the organizational principle. As in manufacturing, outsourcing of the diverse building blocks of AI provides the application developer with a low-cost solution, and enables them to move ahead fast. Paraphrasing Richard Baldwin (2014, 39), the modern integrated AI stack provides the "fast lane to AI development."

Vertical specialization through selective outsourcing of AI stack components has facilitated the rapid development of AI applications. However, despite this impressive organizational innovation, data scientists and engineers keep emphasizing that the current boom in AI applications is based on a relatively narrow knowledge base. A lot more science is needed to overcome these fundamental limitations. No doubt, as the pool of data keeps exploding, AI applications are developing at an exponential pace. But this proliferation of AI applications must be matched by continuous progress in AI research.

This is in line with the findings of our recent interviews in China's AI industry. Most interviewees agreed with the following two propositions:

- While AI applications can race ahead of AI research for a limited time, it is now time to reduce that gap to enable further progress.
- The resurgence in AI research needs to cover both basic and applied research in computing hardware, as well as in data analytics, algorithms, digital platforms, software libraries and development frameworks.

AI's Development Paradigm: Competing Research Approaches

Throughout its history, AI has moved through multiple cycles of progress and optimism (periods of "AI fever") followed by setbacks and pessimism (the so-called "AI winters"). There is a paradigm of something new emerging, being hyped (inevitably overhyped), disappointment setting in, a continued steady underlying build-up of capacity, and, finally, the realization of the hype some decades later.

AI's development paradigm primarily reflects the intense competition between very different science and research methodologies that are lumped together under the heading of AI. Three approaches are distinguished.

Symbolic Systems

First came symbolic systems, which are an attempt to replicate the logical flow of human decision making through processing symbols, i.e., algorithms.[25] Since the 1960s, one influential subfield was knowledge engineering through expert systems. The goal was to package the expertise of a scientist, an engineer, or a manager and apply it to the data of an enterprise or a weapon system.

Early on, robotics provided an important application. Together with powerful sensors, symbolic AI systems made it possible to develop more responsive robots that are endowed with precisely programmed response algorithms as they encounter certain types of stimuli. This approach has focused on providing feedback mechanisms that would allow for practical and effective robots

24 For details, see the section "What's Happening in China's AI Chip Industry?"

25 These concepts were pioneered by Newell, Shaw and Simon (1958) and Newell and Simon (1976).

for specified applications, such as adaptable industrial robots that could interact with humans.

Outside of robotics, symbolic systems have not been central to a broad-based commercial application of AI. However, research on symbolic systems now seems to be experiencing a revival, driven by Stanford's Symbolic Systems program, which has become one of the top five undergraduate majors at Stanford.[26]

Machine Learning

AI's real breakthrough came with machine learning, which feeds a huge volume of data into an algorithm that is essentially a generalized strategy for learning. It then "trains" the machine to derive a rule or procedure for interpreting data or making predictions.[27] Most computer programs codify human knowledge, step-by-step, mapping inputs to outputs as prescribed. In contrast, "machine learning systems figure out the relevant mapping on their own, typically by being fed very large data sets of examples. Using these methods, machines have made impressive gains in perception and cognition" (Brynjolfsson, Rock and Syverson 2017, 2).

Machine learning represents a fundamental rethinking of the relationship between algorithm and data. Until around 2009, the prevailing AI doctrine was: "A better algorithm makes better decisions, regardless of the data." Machine learning instead studies algorithms that improve themselves through data. This shift to a data-centred approach gives rise to new powerful AI applications, in particular in computer vision. To implement this paradigm shift, Li Fei-Fei,[28] during her time at Princeton University, created the ImageNet project (later transferred to Stanford University) — a large database designed for use in visual object recognition software research. The project culminated in an open competition, run annually from 2010 to 2017, attracting participation from more than 50 institutions worldwide.[29] The 2012 ImageNet competition brought the decisive breakthrough — Geoffrey Hinton and colleagues from the University of Toronto submitted a deep convolutional neural network architecture called AlexNet, still used in research to this day.[30] AlexNet gave rise to the current dominant AI technology — deep learning[31] through neural networks. These networks are computing systems inspired by the biological neural networks. Such AI systems "learn," i.e., they progressively improve performance on tasks by considering examples, generally without task-specific programming. "Neural networks involve repeatedly interconnecting thousands or millions of simple transformations into a larger statistical machine that can learn sophisticated relationships between inputs and outputs. In other words, neural networks modify their own code to find and optimise links between inputs and outputs" (Organisation for Economic Co-operation and Development [OECD] 2019a, 28.)

Take image recognition. A neural network might learn to identify images that contain cats by analyzing example images that have been manually labelled as "cat" or "no cat," and by using the results to identify cats in other images. They do this without any a priori knowledge about cats, for example, that they have fur, tails, whiskers and cat-like faces. Instead, they evolve their own set of relevant characteristics from the images (the "learning material") that they process. Using deep learning algorithms thus gives rise to an incredibly tedious approach to learning, even when the goal is to distinguish something as simple as a cat.

It is important to emphasize that neural networks took off only once large data sets became available and thanks to growing computing power at affordable cost. For computer vision, the growth of the internet generated a huge library of images for machine-learning systems to train on ImageNet. An annotated data set of 14 million images in 20,000 categories was assembled by low-paid workers on the Amazon Mechanical

26 See https://symsys.stanford.edu/ssp_description.

27 As defined in Barton et al. (2017, 2).

28 See https://profiles.stanford.edu/fei-fei-li.

29 In a classic paper, Li, Olga Russakovsky and a large team of collaborators describe the creation of this benchmark data set and the advances in object recognition that have been possible as a result (Russakovsky et al. 2015).

30 For a brief history, see Gershgorn (2017).

31 The term "deep learning" is a misnomer. "The word 'deep' in deep learning refers to a technical, architectural property (the large number of hidden layers used in a modern neural network, where their predecessors used only one) rather than a conceptual one (the representations acquired by such networks don't, for example, naturally apply to abstract concepts like 'justice', 'democracy' or 'meddling')" (Marcus 2018, 7).

Turk platform,[32] and provided a common data set for researchers to work with (Markoff 2012).

Equally important, advanced specialized AI chips can now provide increasing computing power at much, much lower cost (both in processing and storage).[33] Marvin Minsky,[34] a well-known cognitive scientist at MIT, was already working with AI in 1957. "Computers at that time were a billion times slower than they are now....Computers then did an OK job and cost a couple million dollars. Now, what used to be thought of as supercomputers are inside smartphones. They cost a million times less, are a million times faster and have a million times as much memory" (Soley, quoted in *Forbes* 2018).

The rise of the neural network approach to AI has given rise to new forms of information asymmetries that greatly expand the scope for market failure in the data-driven economy (Ciuriak 2019b). Extremely large neural networks (with more than one billion parameters) have accelerated the rise of the big data industry, which now has the capacity to monitor and collect data on virtually every aspect of human behaviour, interaction and thought.[35] Neural networks have created "a dramatic new tool for companies like Google, Microsoft and Apple that were anxious to deploy Internet services based on vision, speed and pattern recognition" (Markoff 2015, 148). This emerging data-driven economy "features large economies of scale and scope, often accompanied by network externalities, and pervasive information asymmetry. These characteristics tend to result in 'winner take most' economics, with the prize to the winner being the capture of international rents; these rents promise to be very large and thus serve as an inducement for strategic trade and investment policy" (Ciuriak 2019a).

The Search for New Hybrid Approaches

The seemingly inexorable rise of neural networks is now facing limitations, with the result that leading-edge AI research is searching for new approaches.

32 Amazon Mechanical Turk, the hidden low-tech underbelly of deep learning neural networks, is a service established by Amazon where large numbers of humans sitting at computers around the world would complete small online tasks for pennies. For details, see Crockett (2019).

33 For details, see the section "What's Happening in China's AI Chip Industry?"

34 See Knight (2016).

35 See Zuboff (2019).

In a widely quoted paper, Gary Marcus (a leading neural scientist at New York University) argues that deep learning is now generating decreasing returns and that it needs to be complemented with new approaches (Marcus 2018). Marcus quotes Geoff Hinton, the grandfather of deep learning neural networks: "'Science progresses one funeral at a time.' The future depends on some graduate student who is deeply suspicious of everything I have said" (Hinton, quoted in ibid., 2).

And Stanford University's Anthony Ng, another neural network pioneer, adds: "Today's supervised learning software has an Achilles' heel: It requires a huge amount of data. You need to show the system a lot of examples of both A and B. For instance, building a photo tagger requires anywhere from tens to hundreds of thousands of pictures (A) as well as labels or tags telling you if there are people in them (B). Building a speech recognition system requires tens of thousands of hours of audio (A) together with the transcripts (B)" (Ng 2016).[36]

In short, deep learning neural networks do not actually understand what they are seeing. This is mirrored in speech recognition, and even in much of natural language processing. While AI today knows what things are, understanding these objects in the context of the world is next. How AI researchers will get there is still unclear. All of this has given rise to an intense search within the global AI community for hybrid models that combine the strengths of symbolic systems and deep learning. According to Marcus, "the right move today may be to integrate deep learning, which excels at perceptual classification, with symbolic systems, which excel at inference and abstraction" (Marcus 2018, 20).

The speed of new discoveries in AI research is mindboggling. "As many as 50 technical papers on AI are published daily, and it's going up — it couldn't be a more exciting field," according to David Patterson, former professor of computer science at UC Berkeley, who became a distinguished engineer at Google in 2016, working on Google's Tensor Processing Unit (quoted in Merritt 2018).

Most recently, the 12th Conference on Artificial General Intelligence (AGI-19) reflects this return to the roots of AI research. According to the AGI-19 website, "more and more researchers have recognized the necessity — and feasibility — of returning to the original goals of the field by treating intelligence as a whole. Increasingly, there is a call for a transition back to confronting the more difficult issues of 'human-level intelligence' and more broadly artificial general intelligence.... [The focus of research is now shifting to]...the versatility and wholeness of intelligence....The AGI conference series...encourages interdisciplinary research based on different understandings of intelligence and exploring different approaches."[37]

The International Dimension: AI Increases GVC Complexity

The development of AI has an important international dimension — AI is increasing the complexity of GVCs, with the result that disruptions through trade and technology conflicts are likely to be more distortive. Such impacts may well persist far into the future.[38] This will require adjustments in national policies on AI as well as in its global governance.

As long as the PC dominated GVCs through the Wintel standard, the number of GVC stakeholders remained limited to the flagship (for example, HP as the notebook brand name vendor), the core component suppliers (such as Intel) and a few layers of lower-tier suppliers. The network organization is hierarchical — the flagship defines and outsources innovation, mostly to offshore suppliers in Asia (Ernst 2009, chapter 3).

GVCs have become longer and deeper, involving a greater diversity of stakeholders on multiple GVC layers.

As mobile communication technology has taken over as the main driver of information technology, through the smartphone and the promise of 5G communication networks, the GVCs have become longer and deeper, involving a greater diversity of stakeholders on multiple GVC layers. In this second stage of GVC development, issues of intellectual property (IP) protection and standardization became crucial for achieving a rapid and broad-based diffusion of innovation, giving rise to

36 Andrew Ng co-founded and led Google Brain and was a former vice president and chief scientist at Baidu, building the company's Artificial Intelligence Group.

37 See http://agi-conf.org/2019/.

38 Since January 2020, the proliferation of COVID-19 adds further uncertainty to the future of GVCs.

aggressive patent monetization strategies (Bekkers, Verspagen and Smits 2002). Standard-essential patents (SEPs) became a central battleground for competition (Pohlman and Blind 2016).[39] Over the years, an institutional framework has been developed for the governance of those mobile technology GVCs through standard consortia, such as the 3rd Generation Partnership Project (3GPP)[40] and the recourse to court decisions.[41]

With the rise of AI, GVCs have increased in complexity to such a degree that disruption through trade and technology warfare is now likely to be much more serious. Control of big data is the core asset. Data covers anything "that is captured in digits; facial recognition, commercial transactions, web surfing history, the location of cell phones, and the immense variety of machine data generated by IoT [Internet of Things]....Today as much as 90% of the value of the S&P500 (currently $24 trillion in total) is comprised in intangible assets and a very large share of that is accounted for by the value of data" (Ciuriak 2019c, 4).

The purpose of AI-related GVCs is to capture the very large rents that exclusive data control provides. Somewhat simplified, the typical AI value chain (the AI stack) consists of multiple levels that combine data analytics, algorithms, AI chips and machine-learning platforms (such as, for instance, Tensor Flow or Cafffee2, or Huawei's Mindspore Framework). The case study in the third part of this report ("What's Happening in China's AI Chip Industry?") of China's efforts to develop its AI chip industry will provide a sense of the heightened complexity. We will see that developing a national AI chip industry requires more than just a new AI-optimized chip. That chip can only function if it is integrated into a multi-layered ecosystem and if developers are willing to develop AI applications around this specific chip. No country, not even the United States or China, can bring together all the different layers of that ecosystem. Hence, access is needed to highly specialized GVCs that transcend national borders.

It will not be easy for GVCs to comply with these complex requirements. In fact, the architecture and the governance of these new AI-related GVCs are still emerging, little is fixed and there is still a lot of experimentation. All that can be said at this stage is that AI is transforming the architecture of GVCs. It is in this fluid state of GVC development that the US-China tech war has broken out. One could argue, of course, that in the current formative stage of AI, there might still be ample opportunities to prepare strategies to cope with bifurcated GVCs caused by technology warfare. On the other hand, in light of the increasing complexity of these new AI-related GVCs, it may well be plausible to assume that the disruption through trade and technology warfare will be more severe than in the established GVCs for PCs or mobile communications. But to find out what is happening exactly, empirical research is required to identify all the different building blocks that are needed to make a specialized AI chip work and to find out where these building blocks can be sourced.[42]

To summarize, AI's development paradigm reflects the persistent struggle between competing AI technologies, and hence necessitates permanent progress in AI research. Today's excitement around AI — and more specifically the recent breakthroughs in the subfield of machine learning — represent only the latest upswing in this historical pattern.

The breakthrough in AI applications was based on the extraordinary expansion in the availability of vast new data sets for training algorithms. Data sets thus are as important as algorithms. This gave rise to deep learning through neural networks, which are, however, extremely data hungry. In fact, these big data sets are costly to collect, assemble and vet.

In order to reduce the cost of data management, AI research today has to address three inter-related challenges:[43]

→ Computing power needs to be increased, in order to improve the training of algorithms and for conducting inference with an already-trained algorithm. This explains why AI chips are of critical importance for the development of AI (as discussed in the third part of the report).

39 For an analysis of China's approach to SEPs, see Ernst (2017).

40 The 3GPP is a standards organization that develops protocols for mobile telephony (see https://en.wikipedia.org/wiki/3GPP).

41 For a good overview, see Contreras (2015).

42 In other words, an accounting exercise is required such as that conducted with the iPhone to show the source of its components. See, for instance, Kraemer, Linden and Dedrick (2011).

43 Interviews with AI experts who have requested anonymity, February and March 2019.

- At the same time, data quality needs to be improved through standardization and enhanced data governance.[44]
- To improve the speed of data transmission for the training of algorithms, a successful transition to 5G is necessary. As one interviewee puts it: “If data can’t move where it is needed, it’s useless.”[45]

In short, continuous AI research is needed to enable the development of new AI applications. As we will see below, implementing China’s AIDP will thus face important challenges, in particular in AI chips.

44 These issues will be explored in a future paper (Ernst, forthcoming).

45 Phone interview with AI expert who requested anonymity, March 19, 2019.


YORK STOCK EXCHANGE

Contrasting America's and China's Different AI Development Trajectories

America's Approach: AI Research and Technology Diffusion Are Inseparable

In the United States, serious work on AI started more than 60 years ago with the 1956 Dartmouth Summer Research Workshop on Artificial Intelligence (Kline 2011). The focus right from the beginning was on fundamental breakthrough research. As stated in the proposal to the Rockefeller Foundation: "The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves" (McCarthy et al. 1955).

Boundless optimism was the defining feature — the proposal suggested "that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer" (ibid.). This overly ambitious agenda is still very much "work in process" today.

Soon after the Dartmouth workshop, the walls between academia and industry were coming down. First propelled by massive Department of Defense funding but soon giving rise to a start-up fever, this gave birth to today's global digital platform leaders, such as Google, whose founders happened to be students of Stanford University's Terry Winograd, one of the towering figures of the US AI community (Markoff 2015).[46] Since 1962, DARPA's Information Processing Techniques Office has invested in breakthrough technologies and seminal research projects that led to major advances in computer graphics, networking, advanced microprocessor design, parallel processing and other fundamental building blocks of AI.[47] By inventing the digital protocols that gave birth to the internet, DARPA's defence-related research provided many of the essential advances that support speech recognition, machine intelligence and semi-autonomous systems.

However, the initial AI fever did not last. Limited data available to feed the algorithms as well as insufficient computing power gave rise, since the 1970s, to an extended "AI winter," with DARPA clearly taking a negative view on further AI funding (Correa 2017, 4-5). Nevertheless, despite the AI winter, America's leading research universities, and especially Stanford University, remained thriving hubs for academic, commercial and military AI (Markoff 2015, 299).

With the rise of neural networks, DARPA has again become a key promoter of America's AI development. A defining strength of the DARPA approach is to "promote the follow-on development and implementation of technologies they support in their mission areas....They are therefore much more activist than more standard American R&D agencies, which do not pursue conscious technology strategies oriented to specific mission technology challenges" (Bonvillian 2018).[48] DARPA's focus of technology diffusion has enabled the growth of venture capital-funded start-up companies, which rapidly adopted and expanded upon many of DARPA's inventions. These companies were able to attract experienced top engineers and recruit the best graduates from America's leading research universities, many of them from China and other foreign countries. The main attraction of these companies was that they could offer exciting projects.[49]

This gave rise to a lively, sometimes quite acrimonious debate about what AI should do, and which technologies might be best suited

46 More than anyone else, Terry Winograd's research has defined the boundaries of AI. See, for instance, Winograd and Flores (1986). Debates within the AI community are shaped by Winograd's insight that intelligence was not simply a matter of pattern recognition and processing data; it involved being and existing. As a result, computers alone could not possess true intelligence. The proper role of AI was therefore to help humans live more fully human lives, not to replace them. See Fisher (2017).

47 See www.darpa.mil/about-us/timeline/modern-internet.

48 See also the earlier influential paper by Fuchs (2009).

49 Interviews with AI experts in the United States who emphasize that "attractive projects that push the technology frontier" were much more important for their job decisions than remuneration.

to achieve those goals. At every turn of the United States' AI development, critical voices were questioning fundamental features of its design and of its purpose. Radically different approaches were pursued simultaneously, with intense competition among them.

To understand this characteristic of the US AI development trajectory, it is useful to go back to late 1945, when a young refugee from Hungary, Janos Neumann, or, as he is better known, John von Neumann, brought together a small group of engineers at the Institute of Advanced Study in Princeton, New Jersey. The goal was "to begin designing, building and programming an electronic digital computer...the physical realization of Alan Turing's Universal Machine, a theoretical construct invented in 1936" (Dyson 2012, ix). In contrast to the typical "big science" approach introduced by the Manhattan Project that produced the first atomic bombs, von Neumann's path-breaking computer was developed largely by a dozen engineers in their twenties and thirties "outside the bounds of industry, breaking the rules of academia and relying largely on the US government for support" (ibid.). That this was achieved despite the Kafkaesque hurdles imposed by the powerful Defense Department bureaucracy, is one of the great miracles of the US approach to AI development.

This informal, flexible and undogmatic approach to innovation is, arguably, the root cause for the resilience of the United States' AI development trajectory. Its defining features are technology diffusion through knowledge networks, combined with intense contests among competing ideas. A primary driver has been this constant back and forth of new ideas between academic research, private industry and the vast network of Department of Defense-funded DARPA projects.[50] The DARPA approach facilitated rapid scaling of investment, organizations and experience-based specialized AI skills.

China's Approach Is Different

A Latecomer to AI

As a latecomer to AI, China has pursued a very different approach than the United States. China's AI research began during the 1980s, much later than in the United States. China's industry has only started to get involved over the last few years (Zhu et al. 2018).

The National Natural Science Foundation of China (which supports basic research) and the 863 Program (which supports applied research) began funding, since the mid-1980s, a broad range of AI-related research topics. This included hardware and software for intelligence, human-computer interaction, intelligent application systems, neural networks, genetic algorithms, machine learning, natural language processing, computer vision and robotics. After 2000, both the Ministry of Science and Technology and local governments provided funding that enabled Chinese researchers to attend leading international AI conferences (such as the AAAI Conference on Artificial Intelligence), and to publish in high-impact international journals.

These policies were spectacularly successful. They did increase the role of Chinese AI researchers in leading AI conferences and journals. For instance, the AAAI Society had to postpone its 2017 annual meeting by a week when it found out the planned date coincided with the Chinese New Year. A nearly equal number of accepted papers came from researchers based in China and the United States, so the change was necessary to ensure that Chinese participants were able to attend (Zhang 2017). And Chinese researchers' contributions to the best 100 AI journals/ conferences rose as a percent of total papers from 23.2 percent in 2006 to 42.8 percent in 2015 and as a percent of cited papers from 25.5 percent to 55.8 percent (Press 2017). This is an impressive achievement, establishing China's AI research community as a serious and respected player.

A Limited Nexus between AI Research and Industry

What really sets China's AI development trajectory apart from the United States, however, is that the nexus between academic research and industry has remained quite limited. This may be perplexing. For some observers in the United States, China's innovation policy often seems to present a homogenous picture of a top-down, unified "model of neo-mercantilist state developmental capitalism" (Wolff 2011, 3). That picture represents the current "China bashing" narrative in Washington. However, the reality of China's policy making is very different.

Harvard's Mark Wu (2016, 270) has recently laid to rest the "myth of the Chinese economy...

50 See, for instance, Potember (2017).

that the state's presence is everywhere." He writes, "What makes China complicated is that, while the Party-state holds vast control levers, it allows market forces to play out in huge swaths of the economy" (ibid., 282). For Wu, the intertwined nature of private enterprises and the party-state "involves a complex web of overlapping networks and relationships — some formal and others informal — between the state, Party, SOEs [state-owned enterprises], private enterprises, financial institutions, investment vehicles, trade associations, and so on. This economic structure is not static: Chinese leaders have been remarkably adaptive and pragmatic in their economic stewardship" (ibid., 284).

What needs to be added to Wu's framework is an analysis of the surprisingly fragmented Chinese innovation system. Like most latecomers, China's innovation system is constrained by multiple disconnects between research institutes and universities, on the one hand, and industry, on the other; between "civilian" and "defence" industries; between central government and regional governments; and between different models of innovation strategy (Ernst 2011). In addition, China had to cope with the institutional heritage of the Soviet planning system. As a result, R&D remained locked into different layers of the public administration (both central and regional), while enterprises were assumed to be pure "production units" without adequate research and engineering capabilities and no role to play in marketing and strategic planning. Despite many efforts of "market reform and organizational change," as analyzed by Gu Shulin (the great Chinese innovation economist) (1999),[51] China's institutional framework and support programs still search for ways to enhance knowledge exchange between diverse stakeholders with conflicting interests through an increasing reliance on market forces.

Centralized control has recently regained in importance in China — this is likely to slow down a transition to a more market-oriented approach to innovation policy. According to Loren Brandt and Thomas G. Rawski (2019), "Recent policy initiatives, with Made in China 2025 in the forefront...[emphasize]...top-down technological choice, relying on state-run firms, and insulating priority sectors from potential rivals. Current policy trends magnify plan-era weaknesses that four decades of reform have never squarely confronted. Worse yet, China's leaders seem intent on reviving Mao Zedong's economically counterproductive veneration of self-reliance and suppression of criticism. Beijing's mercantilism, amplified by exclusionary echoes among China's provinces and localities, threatens to undermine product quality, a central component of success in the advanced industries that dominate China's ambitious innovation agenda."

Unfortunately, as US-China technology warfare is heating up, it is even less likely that China's innovation policy will return to a more open market-oriented approach. There is no doubt that China will continue to generate a significant flow of AI-related innovations. Yet institutional structures surrounding Chinese innovative efforts are likely to create large-scale misallocation, waste and resource leakage. The resultant disconnect between AI research and industry is real, immediate and unlikely to fade away any time soon.

China's Three Separate AI Development Trajectories

Somewhat simplifying, it seems that in China three separate AI development trajectories have uneasily coexisted, with only limited exchange of new ideas and sharing of projects among them.

The first trajectory is centred on state-related institutions, i.e., public research organizations and universities, but also some powerful SOEs that conduct AI research in line with the objectives outlined in the AIDP and its implementation plans. Apart from the Chinese Academy of Science and the Chinese Academy of Engineering, this includes specialized research institutes, such as the Internet Development Research Institute at Peking University, or the Department of Microelectronics at Tsinghua University. The first trajectory also includes large SOEs with substantial AI research activities, such as the State Grid Corporation of China, the state-owned electric utility monopoly of China. Noteworthy in particular is the important role played by institutions in the defence sector, as analyzed by Samm Sacks (2019) and Gregory C. Allen (2019).

51 Reform efforts since the 1990s, starting with the Torch Program, have not yet fundamentally strengthened the nexus between academic research and industry.

China has many AI-related plans, so it is not always easy to identify what is really important.[52] Overall, it seems that a principal objective is the industrialization of AI applications in the service of the MIC 2025 plan. MIC 2025 actually "is at the center of the China AI policy citation network and has served as a programmatic document for local governments' AI policymaking as they respond to the national AI development strategy" (CISTP 2018, 5). Local governments continue the old pattern of "following the steps of the central government" and "chasing after hot areas" (ibid., 7).

Success is measured primarily by the number of Science Citation Index publications and by the number of patent applications. However, there is little systematic attempt to ensure that industry is exposed to the research findings included in those publications and the patents. China's share in global academic AI research papers has substantially increased, from 4.26 percent in 1997 to 27.28 percent in 2017 (ibid., 21). However, interaction with domestic industry has remained limited. In fact, most responses in our interview sample indicated that knowledge exchange with public research institutes has only very recently gained in importance. Two consumer-oriented AI companies noted that they use experts from Chinese public research organizations for specific consulting projects. This may also involve moonlighting arrangements through personal contacts.

Both for AI publications and for AI patents, interactions between public research institutes and industry remain limited — "a lot of AI knowledge is lying idle at universities and research institutions."

As for AI patents, China filed more than 30,000 public patents in 2018, an impressive tenfold jump in five years and about 2.5 times more than the United States, which it surpassed for the lead in 2015 (Okoshi 2019).[53] It is necessary, however, to take these findings with a fairly large grain of salt. As AI is a relatively new technology, it is still easier to obtain patent awards than in other more established fields such as mobile communications or biotechnology.[54] This has given rise to a patent filing race for AI. For a latecomer like China, it makes perfect sense to try to use this window of opportunity to catch up in patent filing. However, this rush to file patents has given rise to quality issues (Schulte 2019). In fact, China's patent policies are still primarily focused on pushing up the number of patent applications, while little attention seems to be paid to what happens with these patents once they are registered (Rotenberg 2016). Most importantly, no significant efforts are made to identify and foster patents that could achieve high citations.

In short, both for AI publications and for AI patents, interactions between public research institutes and industry remain limited — "a lot of AI knowledge is lying idle at universities and research institutions" (CISTP 2018, 7).

The second trajectory is shaped by China's digital platform leaders, such as Huawei, Alibaba, Tencent, Baidu and Lenovo. These large companies have the scale and the resources that allow them to conduct their own AI research. In fact, some official data claims that China's top internet firms, led by Alibaba, have invested 153.9 billion renminbi (RMB) ($21.85 billion) in 2018 in R&D (Yu 2019).[55] The problem with this data, however, is that it uses an excessively broad concept of R&D, where most of these investments are directed to product development for AI mass applications.

In our interviews, we found that, for these national AI leaders, investment in AI applications was the top priority. By contrast, investments in the development of algorithms and AI chips were considered to be "too little" and "insufficient." Importantly, we were also told that close interaction with public AI research has only gained in importance since the outbreak of

52 According to the *China AI Development Report 2018*, "Since 2013, China has released a series of AI and related policy documents, including State Council Guidelines on Promoting the Healthy and Orderly Development of the Internet of Things, State Council Notice on Issuing 'Made in China 2025,' State Council Guidelines on Promoting the 'Internet+' Action, State Council Notice on Issuing the Action Outline for Promoting the Development of Big Data, Thirteenth Five-year Plan on National Economic and Social Development, and State Council Notice on Issuing the 'Next Generation Artificial Intelligence Development Plan' released in 2017… which identifies the development directions and priority areas of China's AI development" (CISTP 2018, 62-63).

53 Similar findings are reported in the *China AI Development Report 2018*, drawing on the Derwent World Patent Family Index (see https://en.wikipedia.org/wiki/Derwent_World_Patents_Index).

54 According to *Nikkei Asian Review*, "Patents are awarded for about 70% of applications across all fields, but that figure raises to more than 90% for a subset of AI known as machine learning" (Okoshi 2019, 2).

55 Quoting figures released by the Internet Society of China, which is affiliated with the Ministry of Industry and Information Technology.

the technology war. That Chinese firms until recently innovated in their areas of comparative advantage is in line with trade theory. They are now being forced by the technology war to innovate in areas of US comparative advantage.

Since 2018, the government has actively pursued a policy that seeks to develop an A-list of dominant AI platforms that will rely heavily on China's three digital platform leaders: Baidu for autonomous driving, Tencent for AI in health care and Alibaba for smart cities. Of particular interest is the active role of local governments. For instance, since 2016, the Alibaba Group and Foxconn have partnered with the city of Hangzhou for the "City Brain" project, which uses AI to analyze data from surveillance cameras and social feeds. However, a defining characteristic of this project seems to be a clear division of labour between the local government, which bears responsibility for broader aspects of traffic management, and the much narrower business objectives pursued by the Alibaba group, which, through its affiliate Ant Financial, uses facial recognition for payments at Alibaba-owned retail stores.[56]

More recently, the government seems to direct China's digital platform leaders to invest in underperforming SOEs (Yeung 2019). The powerful State-owned Assets Supervision and Administration Commission (SASAC) has announced a so-called mixed-ownership reform to improve the services and earnings of state firms. SASAC emphasizes in particular Tencent's participation in China Unicom's reform efforts. It is unclear what exactly might be the expected. Are Tencent, Alibaba and others supposed to fix up the SOEs and transform them? Or is the main purpose to access underutilized R&D and to translate their efforts into commercially successful innovations?

It remains an open question whether such top-down knowledge network construction by fiat will succeed in breaking down the barriers to knowledge exchange that separates China's first and second AI development trajectories.

China's third AI development trajectory covers a great variety of companies that develop, implement or use AI technologies and applications. Much research has focused on the growing population of so-called AI unicorns, i.e., start-up companies valued at more than $1 billion (CB Insights 2019). Typically, these companies use existing machine-learning algorithms, primarily neural networks, to sell AI application software. A second group of AI unicorns, involved in the design of AI chips, are receiving substantial support from the government. Some of these companies have made good progress in specific areas, but they still need to build up scale and capabilities. There is no doubt, however, that, despite current weaknesses, some of these AI chip start-ups will play an important role in the future.[57]

In our interviews, we found that many of the first group of AI unicorns can hardly keep up with the rapid demand growth for AI applications in the Chinese market. To move ahead quickly in application markets, these companies are trying to recruit as many young engineering graduates as they can from across China. They also are fiercely competing for experienced top talent from overseas. Investment in AI research (both applied and basic) has low priority, and interaction with public AI research organizations remains limited.

Of particular concern is that most Chinese AI start-ups do not invest nearly enough in cutting-edge technology to compete. This reflects a peculiar feature of China's stock exchanges. In fact, "excessive R&D spending can hamper Chinese businesses' ability to go public. Unlike in the U.S., China's stock exchanges require companies to be profitable for at least three years before making an initial public offering. R&D spending shows up on income statements as an operating expense and thus keeps young companies in the red for longer" (Ren 2019).[58]

But what about linkages to China's digital platform leaders? In fact, all of these national AI champions have been actively acquiring AI start-ups.[59] Many of China's tech giants have established investment arms, which invest in a broad range of companies in the country's internet and AI sectors. Primary motivations

56 Interview with Alibaba, Hangzhou, April 24, 2019.

57 For details, see the section "What's Happening in China's AI Chip Industry?"

58 The article adds that this may change: "Starting in late July, technologically ambitious companies can list on a new Nasdaq-style platform: the Shanghai Stock Exchange's Science and Technology Innovation Board, colloquially known as STAR, where profitability doesn't matter....STAR stocks gained 140% on the board's first day, *but the enthusiasm may not last, especially after waves of corporate scandals this year* [emphasis added]."

59 This is in line with the US model where Google and other digital platform leaders have acquired AI start-ups.

for these investments seem to have been access to promising AI applications and to teams of experienced engineers. In addition to Alibaba and Tencent, Huawei established its own investment arm Hubble Technologies on April 23, 2019, in Shenzhen (Jill Shen 2019). It is a wholly owned subsidiary of Huawei with registered capital of RMB 700 million ($104 million), and Bai Yi, president of Huawei's global financial risks control centre, was named legal representative and chairman of the company. Xiaomi founder Lei Jun established a venture capital firm, Shunwei, in 2012. With a focus on intelligent devices and internet services, Shunwei co-invested with Xiaomi for nearly 100 of the total of 300 start-ups it holds shares in, folding them into the Xiaomi IoT ecosystem.

Prominent examples of such AI-related acquisitions include Tencent's leadership role in an RMB 340 million ($50 million) funding of AI chip start-up Suiyuan Technology, together with Chinese private equity funds Zhen Fund, Delta Capital, Yunhe Partners and Summitview Capital (Tian 2018). In contrast to Tencent's primarily financial investments, Alibaba seems to have pursued a more strategic investment approach, with a focus on expanding the company's capability set into chip design. For instance, in 2018, Alibaba acquired C-Sky Microsystems, one of China's more promising AI chip design companies (Demler 2018).[60] So far, the results of these acquisitions remain unclear. It will take some time to separate the wheat from the chaff.

It would seem fair to conclude that China's approach to AI development remains fragmented. Public research institutions conduct AI research (some of it at the frontier). Interactions with industry, however, remain limited, as industry's primary concern is to forge ahead in China's rapidly growing mass market for AI applications.

Until well after 2010, very little of China's academic AI research has found its way into the business plans and strategies of China's emerging digital platform leaders. Our interviews found that the main concern of these companies has been to keep up with China's "AI applications fever." Hence, business investment in AI research (both applied and basic) has remained limited, focused on a select group of "star" AI unicorn start-up companies.

Nor is there much evidence that China's dominant AI industry players have developed anything similar to the vast and thriving networks that have constituted a major strength of America's AI industry. Hectic wheeling and dealing there are aplenty. This is especially the case when bouts of venture capital fever occur in the midst of a top-down government-centred approach to innovation policy. But whether this will give rise to a sustainable AI ecosystem, remains an open question.

What Explains China's Fragmented AI Development?

It is useful to remember that, before the outbreak of the US-China technology war, it was widely believed that China can win the global AI race because it has a larger data set than any other country.

Take Lee Kai-Fu, a former Apple, Microsoft and Google executive who is now a venture capitalist. His widely quoted book, published in 2018, captures the boundless optimism that prevailed before the technology war. Lee (2018, 12) argues that China can outcompete the United States because AI has moved "from the age of discovery to the age of implementation, and from the age of expertise to the age of data." For China, what matters now is "the power of data" (ibid., 14).

The main challenge for China's ascending AI industry is thus speedy exploitation of China's large data sets. The primary concern is to exploit China's huge cost advantage in big data management. Andrew Chen, a partner at Andreessen Horowitz,[61] provides the following example: "Let's say you invest $10 million into a small AI company. In the US, the company would spend $2 million of that money on labeling data, whereas in China that effort might take a quarter or a tenth of the cost.... [In China], you get much more throughput for your $10 million investment" (Chen quoted in World Intellectual Property Organization [WIPO] 2019, 89).

According to Lee, the key to success is China's huge population of relatively low-cost college graduates who will toil away for long hours to do the repetitive work of categorizing huge troves of data needed to "train" AI algorithms. He suggests that China should use its big data treasure trove

60 For more on Alibaba's link with C-Sky Microsystems, see the section "What's Happening in China's AI Chip Industry?"

61 See https://a16z.com/author/andrew-chen/.

to forge ahead in mass markets for lower-cost AI applications. While some applied AI R&D might be of use, China should concentrate most of its efforts on capturing markets for AI applications. China could achieve this goal by simply using existing algorithms and by purchasing leading-edge AI chips from global semiconductor industry leaders.

Much of China's AI industry in fact largely followed Lee's strategic advice. Lee's explanation of why this happened, however, is problematic. His central proposition is that the big intellectual breakthroughs in AI have already occurred, and that, as a result, barriers to late entry into AI are reduced, as long as the focus is on AI applications. The reality is more complex. As documented in the first part of this report, "Competing in AI: Major Challenges," continuous AI research is needed to enable the development of new AI applications. Lee's proposition is based on a fundamental misreading of AI's permanent research revolution. As one AI expert put it, "One twist on his…[Lee's]…analysis might be that AI is in both phases at once. Conventional deep learning is in the exploitation phase, because we know pretty well how to implement it, even if we don't know how to document it or predict its behavior. But much more interesting things than deep learning are coming, and they are still very much in the research stage."[62]

The real explanation for China's focus on AI applications can be found in an important organizational innovation in the emerging AI value chain. As described in the first part of this report, it is the development of an integrated AI ecosystem (an AI stack in industry parlance) that has enabled AI application developers to execute fast and low-cost applications through selective outsourcing of critical AI tools.[63] This explains why, in contrast to the United States, China's AI industry started with the wheeling and dealing and the start-up fever, well before it engaged in significant AI research. This is a necessary consequence of China's late entry into AI in the 2000s rather than in the early formative stages of the 1950s. The conduct of AI research, however, was left to the academic "ivory tower," either in China or at America's leading research universities. For Chinese internet companies, "their ultimate goal is to make money, and they are willing to create any product, adopt any model, or go into a business that will accomplish that objective….The core motivation for China's market-driven entrepreneurs is not fame, glory, or changing the world. Those things are all nice side benefits, but the grand prize is getting rich, and it doesn't matter how you get there….Rigorous copying of perfection was seen as the route to true mastery" (Lee 2018, 27).

As Lee puts it, Chinese students are "taking advantage of AI's open research culture to absorb knowledge straight from the source and in real time. That means dissecting the latest online academic publications, debating the approaches of top AI scientists in WeChat groups and streaming their lectures on smartphones" (ibid., 83). This is in line with China's late entry into AI. The logical starting place for any newcomer is to tap into existing state-of-the-art technology and to find a way forward from there. Standing on the shoulders of industry leaders made perfect sense, rather than trying to go back to the 1950s to reinvent the wheel.

However, as will be seen below, there is a flipside to China's heavy reliance on foreign sources of AI technology. This global knowledge sourcing was not supported by a robust body of domestic research. With the outbreak of the US-China technology war, this lack of research has become a major vulnerability for China's AI industry.

This highlights a fundamental conundrum of China's approach to technology development. Before the outbreak of the US-China technology war, Chinese AI firms innovated in areas that reflected their comparative advantage. Exploiting their huge database through their vast pool of low-cost university graduates, they focused on competing in China's rapidly growing mass markets for AI applications.

This strategy was made possible by China's deep integration into international trade and global production networks, which has provided ample opportunities for global knowledge sourcing (Ernst 2018a). To the degree that Chinese companies were able to rely on foreign technology, they could grow and prosper without investing in in-house basic and applied research.

It is important to recall that China's reliance on global knowledge sourcing was actually in line with the "gains from trade" globalization doctrine that was widely shared before the

62 International AI expert who requested anonymity, email to the author, September 28, 2018.

63 See the section "Enablers of AI Applications: Vertical Specialization Facilitates Access to AI Tools."

outbreak of trade and technology wars. Paul Romer (1994, 21), for instance, argued that the most important question for a developing country is: "What are the best institutional arrangements for gaining access to knowledge that already exists in the rest of the world?" Richard Baldwin (2014, 39) has popularized Romer's insights into widely disseminated policy prescriptions that global network integration provides "the 21st century fast lane to industrial development."

How Will US Technology Warfare Affect China's Innovation Strategy?

This "gains from trade through globalization" doctrine was based on the assumption that globalization will continue unabated.[64] With rising US technology export restrictions, it has become much harder to reap such gains.

Our research finds that US technology restrictions are forcing China to strengthen basic and applied AI research to catch up in core foundational technologies. Under pressure from US export restrictions, the government, through SASAC and other agencies, is now searching for ways to reduce the fragmentation of China's AI innovation system. The comments from Chinese firms during our interviews indicate that interactions between China's three AI development trajectories may be beginning to change. Ironically, US technology export restrictions are thus forcing a reform of China's technological investment and innovation policy, which may help China to correct one of the fundamental weaknesses of its innovation system in AI.

Most of the companies interviewed mentioned multiple funding schemes and support institutions to implement China's AIDP. The companies are well aware of the government's renewed efforts to strengthen domestic innovations through an increasing array of support policies and incentives. There are strong reasons for these companies to participate in such schemes, if only to qualify for additional funds. But it is unclear how active and effective such participation might become, as long as these companies are overwhelmingly focused on the development of AI applications.

The outbreak of US technology warfare has therefore exposed how vulnerable China's AI achievements are to external constraints. The development of AI applications provides engineering and management knowledge and experience that constitute a valuable resource for China. It needs to be supported, however, by an increasing body of domestic AI research. In short, the US-China technology war may act as a catalyst for upgrading China's technological capabilities in AI. Chinese AI firms are now being forced to innovate in areas of US comparative advantage.

As a short-term response, the rhetoric of self-reliance is clearly heating up in China, with references to the heroic "Long March" (Nakazawa 2019). But substantial hurdles remain. No one really seems to know how long the trade and technology wars will continue. For Chinese AI companies, a substantial increase in AI research may be considered very risky, given the current uncertainty. This raises the question whether China's newly minted support policies and incentives are strong enough to overcome the negative effects of US technology warfare.

To understand what is happening on the ground, the report will take a closer look at China's approach to the development of AI chips.

64 For a detailed critique of the gains from trade through globalization doctrine, see Ernst (2018b).

What's Happening in China's AI Chip Industry?

Defining AI Chips

An AI chip[65] is a class of microprocessor or computer system designed as co-processors to support AI applications, in particular artificial neural networks, machine vision and machine learning. AI chips thus display distinguishing features and architectures "that promise even greater computational and machine learning capabilities" (Reuther et al. 2019). As explained in a recent *White Paper on AI Chip Technologies*, there is no widely accepted definition of AI chips (Beijing Innovation Center for Future Chips and Tsinghua University 2018, 5).[66]

The development and production of AI chips constitutes an essential building block in the development of AI technologies and applications. It is a rapidly growing industry that attracts massive amounts of investments, both from industry and governments. According to one authoritative estimate, the global market for AI chips in 2021 is estimated to exceed $20 billion (Zekun 2019; quoting data from Morningstar, the international fund ranking institution). While this is still a small share of the overall global semiconductor market, estimated to be around $510 billion in 2021, demand for AI chips is projected to contribute significantly to the industry's overall growth (PwC 2019).

Taxonomy of AI Chips

To understand the role of AI chips, it is useful to start with a taxonomy of hardware components used to provide the computing power needed to support AI. Three classes of such AI-related components can be distinguished:

Memory

- → Electronic data repository for short-term storage during processing
- → Memory, typically consists of dynamic random-access memory (DRAM)[67]

Storage

- → Electronic data repository for long-term storage of large data sets
- → Storage typically consists of NAND flash memory[68]

Logic

- → Processor optimized to calculate neural network operations
- → Logic devices are typically CPUs (central processing units[69]), GPU (graphics processing unit[70]) cards, FPGAs (field-programmable gate arrays[71]) and various custom ASICs (application-specific integrated circuits).[72]

Source: Batra et al. (2018, 3).

65 I am grateful for guidance from AI chip experts during interviews in the spring of 2019 who have requested anonymity.

66 Within industry, the most broadly used term is "AI accelerator." Others use the terms "deep-learning accelerator," "neural engine" or "neural processing unit" (email interview, February 28, 2019, with Linley Gwennap, CEO of the Linley Group, which publishes the influential *Microprocessor Report*).

67 See https://en.wikipedia.org/wiki/Dynamic_random-access_memory.

68 NAND flash memory is a type of non-volatile storage technology that does not require power to retain data.

69 See https://en.wikipedia.org/wiki/Central_processing_unit.

70 See https://en.wikipedia.org/wiki/Graphics_processing_unit.

71 See https://en.wikipedia.org/wiki/Field-programmable_gate_array.

72 See https://en.wikipedia.org/wiki/Application-specific_integrated_circuit.

Memory, storage and logic are important building blocks of an AI hardware ecosystem.

For the purpose of our research, we focus on logic devices. An AI chip is defined as a processor that is used to train AI algorithms and for inference, i.e., to draw conclusions. This definition encompasses all of the products commonly used for AI today, such as Intel CPUs, Nvidia GPUs, Intel and Xilinx FPGAs and various custom ASICs.

GPUs are specialized processors for the manipulation of images and calculation of local image properties. They are today's workhorse. GPUs were originally designed to support computer graphics and image processing applications. To accomplish this, GPUs feature an architecture that distributes computational tasks across a large number of integrated circuits (called "cores") to be processed in parallel. This parallel architecture makes GPUs useful for machine-learning applications. By contrast, CPUs feature a smaller number of more powerful cores that are optimized for handling just a few tasks simultaneously.

FPGAs are distinct from CPUs and GPUs in that they do not run programs in stored memory. Instead, they are a collection of standardized "logic blocks" whose relationships can be configured by a programmer once the chip is received from a manufacturer. ASICs are purpose-built chip boards that cannot be easily reconfigured after they are manufactured.

Both FPGAs and ASICs consume less energy than CPUs and GPUs, and their specialization allows for greater speed. They are, therefore, particularly attractive in the context of machine-learning inference where speed and energy consumption are critical. However, these gains come at a price: the loss of flexibility and an increased cost. FPGAs and ASICs cannot be as easily and quickly configured to run a wide range of tasks. Both are more expensive than CPUs and GPUs. ASICs are cost-effective only in significant quantities, for instance, when they are used for smart cars.

There has been some partial convergence between FPGAs and ASICs. In fact, speed and energy consumption gains of both FPGAs and ASICs have encouraged Intel and Google to develop software that maximizes data reuse and minimizes external memory bandwidth to boost training performance on FPGAs (Hwang 2018, 13). Hence, borderlines between established semiconductor product categories are constantly blurred, giving rise to new hybrid forms of AI chip design.

From the perspective of China's strategy as a latecomer catching up, such changes in AI chip design parameters raise an important issue: as long as technology road maps are largely predictable, China can reap the benefits of fast-following, lower-cost catch-up strategies through replication. This is in line with the "gains from trade" theory discussed above. However, latecomers are likely to face greater challenges with rapid changes in AI chip design and especially when new hybrid forms of chip design are emerging. The breakneck speed of change in AI chips thus increases the challenge for China's AI chip development.

Is Off-the-Shelf Purchase of AI Chips Still a Realistic Option?

Industry experts typically argue that China's best bet would be to continue using off-the-shelf AI chips from global semiconductor industry leaders, and then focus more on the application of AI. This has been the prevailing pattern — China's AI developers, implementers and users have been able to grow rapidly because they sourced these specialized semiconductors from a handful of leading US vendors. China was thus able to avoid the huge costs and risks of developing an integrated domestic AI chip value chain.

Note that the proposed "outsourcing" strategy assumes largely free and open markets for leading-edge AI chips. The real world, however, is different. Since the mid-1990s, a handful of leading US semiconductor firms control around 50 percent of the global semiconductor market and more than 50 percent of the Chinese market (Semiconductor Industry Association 2019). More importantly, the markets for both GPUs and FPGAs are tight oligopolies, controlled by US companies. In 2019, Nvidia dominated GPUs with 74 percent of the global market, with 26 percent controlled by AMD (Ma 2019). And FPGAs, the second-largest group of AI chips, again are a tight oligopoly, with Xilinx controlling 59 percent of the global market, while Intel/Altera controls the other 41 percent (Erickson 2019). In light of such an extremely unequal distribution of market power, a handful of US market leaders can shape technology trajectories, standards and pricing strategies for AI chips.

China's leadership believes that such a heavy dependence on a few global oligopolists could cripple its AI ambitions. If China wants to sustain its achievements in AI applications, it is argued, a robust domestic AI chip industry is urgently needed. With the outbreak of the US-China technology war, America's unprecedented arsenal of technology export restrictions is now seriously threatening to constrain China's secure access to leading-edge AI chips.

Benchmarking the Challenge: Capability Requirements

However, developing an integrated domestic AI chip value chain poses a major challenge. For China to develop a national AI chip industry requires more than just a new AI-optimized chip. That chip can only function if it is integrated into a multi-layered ecosystem and if developers are willing to develop AI applications around this specific chip. According to Paul Triolo, "Success is not just about developing a new AI-optimized chip, there needs to be an ecosystem around it and developers who are willing to develop around that a hardware architecture."[73] Triolo adds: "That... [i.e., the broad and multi-layered ecosystem] is the critical missing piece in China, as Nvidia, Intel, and AMD are still very dominant and moving to a new chipset/development environment is not something done easily, or out of patriotism."

A defining characteristic of AI chips is the broad range of capabilities needed. A great variety of technological capabilities and skills (and their combinations) is required to improve performance and energy consumption of these AI chips. Performance requirements for AI chips differ, depending on:

→ whether the chip is needed for training an algorithm or for conducting inference with an already trained algorithm;

→ whether the chip is used by end-users (in industry parlance "on the edge"), for instance, a mobile device/smartphone, or whether it is used in a data centre — if the chip is used on the edge, it is necessary to reduce energy consumption and latency, i.e., the time delay between input and output of a system; and

→ the type of AI models for which they are applicable, such as conventional statistical models, deep learning, recurrent networks, long- and short-term memory networks and neuromorphic models.

In light of these complex characteristics of AI chips, it is important to emphasize that research agendas and technology trajectories increasingly overlap. This again highlights the rapid pace of AI research and the still-largely unpredictable technology road map. If China wants to become a serious player in AI, it needs to participate in, and possibly even co-shape, this global AI chip R&D race.

New Opportunities: AI-induced Changes in Chip Design

A quick look at the global techno-economic dynamics explains what is at stake. Big changes are under way in mainstream chip architectures. A defining characteristic of chip design has been the observation that the number of transistors in a dense integrated circuit doubles about every two years, the so-called Moore's law.[74] However, Moore's law has now reached its limits. Researchers at the MIT Lincoln Laboratory Supercomputing Center argue: "Many of the technologies, tricks and techniques of processor chip designers that extended...[Moore's law]... have been exhausted. However, all is not lost, yet; advancements and innovations are still progressing. In fact, there has been a Cambrian explosion of computing technologies and architectures in recent years. One area in which we are seeing an explosion... is ML processors and accelerators [i.e., AI chips, in our terminology]" (Reuther et al. 2019, 1-2).

For instance, a major challenge is power consumption, which is limiting machine-learning capabilities.[75] The total amount of power consumed for machine-learning tasks is staggering, and most of the power consumed is waste. To overcome this power bottleneck, major changes are needed in algorithms and in computing architecture, i.e., in chip design. But all of this is still early in the development cycle, and so far, it is hard to tell how all of this will play out. For countries playing catch up, like China, this may offer an opportunity to leapfrog. Will China

73 Paul Triolo, practice head, geo-technology, at the Eurasia Group, email to the author, August 29, 2019.

74 The observation is named after Gordon Moore, the co-founder of Fairchild Semiconductor and former CEO of Intel. See https://en.wikipedia.org/wiki/Moore%27s_law.

75 As analyzed in Bailey (2019).

find a commercially viable solution first, ahead of the United States, and then surge ahead?

The proliferation of AI is in fact disrupting established technology paradigms in the semiconductor industry. AI has provoked a global "architectural race" in semiconductors that is driven by the need to process big data reliably, fast and at reasonable cost and energy consumption. According to a recent report, "Chipmakers are working on new architectures that significantly increase the amount of data that can be processed per watt and per clock cycle, setting the stage for one of the biggest shifts in chip architectures in decades. All of the major chipmakers and systems vendors are changing direction, setting off an architectural race that includes everything from how data is read and written in memories to how it is processed and managed — and ultimately how various elements that used to be on a single chip are packaged together" (Sperling 2018).

As a result of this paradigm shift, the focus of semiconductor innovation shifts from process technology and fabrication to architecture and design at the front end, and post-fabrication packaging[76] at the back end. We saw that process technology and fabrication are China's primary weaknesses in semiconductors, while its integrated circuit (IC) design and packaging operations are closer to the technology frontier. This raises the question whether the AI-induced changes in chip design could, in principle, open up new possibilities for Chinese semiconductor firms to catch up in chip design and packaging.

At the same time, the barriers to catching up are likely to increase, as both new processor architectures and new memory architectures need to be developed simultaneously. New processor architectures search for better ways to process larger blocks of data per cycle, while new memory architectures seek to alter the way data is stored, read, written and accessed.

This has given rise to two radically different approaches to AI chip design; both, however, are quite complex (Waters 2019). Companies such as Intel, but also start-ups such as Habana Labs (San Jose, CA), Graphcore (Bristol, UK) and Cambricon (Beijing, China) are seeking to create smaller, modular elements, known as "chiplets," out of which today's most advanced AI chips are assembled. An alternative approach is taken by Cerebras, a Californian start-up that has raised more than $200 million in venture capital funding. Cerebras has developed a processor, cut from the largest available 300 mm silicon wafer, whose surface area slightly exceeds a standard iPad and is 56 times bigger than its closest competitor. Note, however, that both approaches at present are facing much greater challenges than expected.

AI chip design faces the additional challenge that machine-learning algorithms keep changing,

76 "Packaging" in the semiconductor industry is defined as wrapping of the silicon wafer in a case, consisting of ceramics or other complex materials. It is an essential part of semiconductor manufacturing and design, affecting power, performance and cost on a macro level, and the basic functionality of all chips on a micro level. See https://semiengineering.com/knowledge_centers/packaging/.

almost daily, so that basic design features remain fluid. According to another report in *Semiconductor Engineering*: "The first generation of solutions is not very efficient. Both training and inferencing are done in a very brute-force fashion. This is bound to change" (Bailey 2018). Apart from the dominant GPUs, "there are many techniques that people have not had a chance to try out yet because the field is moving so quickly. Some people expect that neural network accelerators will gain in importance, significantly reducing the unit price of ML chips....Once the dust settles and the way that people do training becomes more uniform, and the algorithms do not change on a daily basis, you will see people pushing down the power consumption curve" (ibid.). Hence, "it is critical to have flexible, scalable and energy-efficient ML chips that cover a great variety of performance points" (ibid.).

As a result, the diversity of AI chips is likely to increase, covering a wide range of cost and performance points. "There will be AI chips...that cost less than a dollar. Big standalone chips may cost over a thousand dollars but will outperform a box full of GPUs costing far more. Some argue that most of the world's AI processing will shift from legacy platforms to optimized solutions as quickly as the new silicon can be manufactured" (ibid.).

The AI-induced revolution in chip design raises a fundamental question: given the level of ferment and uncertainty, will China be the first to find a commercially viable solution, and then use this head start to surge ahead? It is important, however, to emphasize that all of these changes in IC design will take time. Hence, catching up and forging ahead in AI chip design will be a question of years, not months. China would need to mobilize substantial resources, both financial and in terms of highly specialized and experienced engineering talent, to reach a goal that, at present, still seems to be a moving target with uncertain promise of success.

China's Position in Semiconductors: Rapid Catching Up but Incomplete Value Chain

To get a realistic picture of China's capacity to deal with these formidable challenges in AI chips, let us look at where the country stands today in semiconductors. Over the last 60 or so years, China's semiconductor industry has come a long way, from being a completely government-owned part of the defence technology production system, with SOEs as the only players, toward a gradually more market-led development model. The role of SOEs has dramatically declined, and deep integration into international trade and global networks of production and innovation has transformed decisions on pricing and investment allocation, with private firms as the main drivers.

China's earliest semiconductor was built in 1956, not long after the technology was invented in the United States. But China's progress was stifled early on by the Cultural Revolution, when engineers, scientists and students were caught up in turmoil, disrupting their education and research and their exposure to international science and technology. When China reopened for business under Deng Xiaoping, semiconductors soon became a poster child of China's economic planners. Yet, the legacy of a Soviet-style central planning economy continued to stifle innovation. Barriers to innovation remained substantial well into the first decade of the twenty-first century, ranging from severe quality problems in education to plagiarism in science.

Despite decades of efforts to develop a robust domestic semiconductor industry, China remains weak in the design and fabrication of leading-edge memory and processors (Ernst 2015; 2016). This weakness is particularly grave for fabrication. Semiconductor Manufacturing International Corporation (SMIC), China's largest semiconductor foundry,[77] continues to lag two generations (more than three years) behind in leading-edge process

77 According to Wikipedia, "Semiconductor Manufacturing International Corporation (SMIC) is a publicly held semiconductor foundry company and the largest in China....It is headquartered in Shanghai and... provides...IC manufacturing services on 350 nm [nanometre] to 14 nm process technologies" (see https://en.wikipedia.org/wiki/Semiconductor_Manufacturing_International_Corporation).

nodes.[78] While SMIC keeps investing heavily in upgrading its production, industry experts expect that it would take a decade for SMIC to close the gap with Taiwan's TSMC, the industry's global leader (Hruska 2019). According to Jay Huang, a former Intel managing director in China, "China should be prepared for a marathon of at least a decade, which will also be loss-making [along the way]" (quoted in Deng 2019). This weakness in semiconductor fabrication is China's Achilles heel that is likely to prevent it from capturing the market with a new blockbuster chip.

China's prospects in chip design are better. China's chip design industry is growing rapidly. Led by HiSilicon (Huawei's affiliate), Tsinghua Group subsidiary Unisoc Communications and Beijing OmniVision Technologies, the industry's total revenue reached RMB 251.5 billion ($35.3 billion) in 2018, up nearly 23 percent from a year earlier (Qu 2019). These companies, and especially HiSilicon, are at the technology frontier. Given the substantial financial support provided by the Chinese government, these companies have a realistic chance of becoming serious global competitors.[79]

Of particular concern is the persistent gap between semiconductor consumption and production. China has been the largest market for semiconductors since 2005. Yet only slightly more than 15 percent of China's total semiconductor consumption was supplied by China-based production in 2018.[80] And foreign companies with semiconductor fabrication plants (fabs) in China may account for almost half of that domestic fab capacity. While the US semiconductor industry has consistently retained nearly half of the global market, China-based production is only around five percent (Semiconductor Industry Association 2019).

Reflecting these weaknesses, China's semiconductor trade deficit has more than doubled since 2005, surpassing crude oil to become China's biggest import item. This massive import dependence explains why the Chinese leadership at the highest levels has made it a priority to catch up and forge ahead in this industry. Like in the United States, national security needs play an important role. However, all these motivations are dwarfed by economic considerations. Semiconductors are critical for sustaining Chinese exports — their share in China's exports exceeds 50 percent. Secure access to leading-edge semiconductors is thus of critical importance from a Chinese perspective.

Key policies to strengthen China's semiconductor industry include the National IC Industry Development Guidelines and the MIC 2025 plan, published by China's State Council in June 2014 and May 2015, respectively. Both plans are backed by huge investments — a total of approximately $45 billion for the two phases of the National IC Industry Investment Fund,[81] and $300 billion for MIC 2025. In addition to direct budgetary support, below-market equity investment has played an important role in subsidizing China's semiconductor industry (OECD 2019b).[82] For two of China's leading semiconductor firms (SMIC and Tsinghua Unigroup), total government support exceeded 30 percent of their annual consolidated revenue.

A range of support policies cover IP, cyber security, procurement, standards, rules of competition (through the Anti-Monopoly Law), and the negotiation of trade agreements, such as the Information Technology Agreement (Ernst 2018a). The objective is to strengthen simultaneously advanced manufacturing, product development and innovation capabilities in China's semiconductor industry as well as in strategic industries that are heavy consumers of semiconductors.

A unifying feature of these plans is to secure timely and cost-effective access to advanced semiconductors that are needed to upgrade China's manufacturing and service industries and for modernizing its defence and security

78 SMIC entered initial production of 28 nm technology in the fourth quarter of 2015, more than three years after Taiwan Semiconductor Manufacturing Company (TSMC) began fabricating wafers with its 28 nm process. SMIC has started volume production of its new 14 nm fin field-effect transistor (FinFET) technology in the third quarter of 2019 and plans to introduce 12 nm FinFET technology in 2020 — once again about three years behind TSMC's introduction of similar processes (see Shen 2020; Shilov 2019).

79 See analysis below of AI chip design in HiSilicon.

80 See Appendix 3: China IC Market versus China IC Production Trends.

81 In October 2019, China's Big Integrated Circuit Fund was rolling out its second phase of funding through a just-incorporated company called the National Integrated Circuit Industry Investment Fund Phase II Co., Ltd. (National Big Fund Phase II). Compared to Phase I, which began in 2014, Big Fund Phase II is twice as big at RMB 204.15 billion ($28.9 billion), slightly exceeding market expectations (Liu 2019).

82 But as the OECD study shows, China is not the only government subsidizing the semiconductor industry. The international subsidies tournament in semiconductors goes back to the 1980s when the United States was fighting against Japan's rise in that industry (see Howell et al. 1988). At that time, China was not even on the map.

sector. It is too early to assess how long it might take for these policies to reduce the still quite substantial technology gap that separates China from the United States in semiconductors. What is clear, however, is that these policies are having an important mobilization effect. They have also strengthened the market power of Chinese IT companies (such as Huawei and Lenovo) as buyers of semiconductors. And they signal a concerted effort to broaden China's semiconductor technology portfolio.

Mobilization Effect

China's semiconductor support policies convey a clear message to all domestic and foreign stakeholders within this industry as well as along its value chain:

- → No effort will be spared to implement a massive increase of the production/consumption ratio of semiconductors.
- → Markets will play a "decisive" role (i.e., a greater role than before, to put it plainly) in determining the range of products, markets and value chain stages. While "the state strikes back" in the broader economy — to paraphrase Nick Lardy (2019) — the rules of the game may be slightly more flexible in critical industries such as semiconductors.
- → Firms that participate in this contest will profit and grow, while firms that stay on the sidelines will lose out.
- → Benefits include preferential tax treatment, land and monetary subsidies, R&D and labour incentives, and access to RMB equity funds.
- → According to SEMI, the global industry association serving the global IT industry,[83] these benefits apply to both domestic and foreign players — at least for now (SEMI 2019).
- → Our interviews with industry experts show that this message has captured the attention of both domestic and foreign firms.

Increasing the Market Power of China's Leading IT Companies

A second important effect may be on the demand side.[84] As a result of the incentives provided by China's semiconductor policies, four leading Chinese IT companies (Huawei, Lenovo, BBK Electronics and Xiaomi) have now joined the top 10 chip buyers worldwide, significantly increasing their bargaining power relative to US semiconductor firms. In 2018, Huawei in fact increased its chip purchases by 45 percent, moving ahead of Dell into the third spot (Kharpal 2019).

As a result, the pull of the Chinese market for US chip vendors has been increasing even further in importance.

Broadening China's Semiconductor Technology Portfolio

Until recently, China has focused primarily on logic semiconductors and mixed-signal integrated circuits for mobile communication equipment (including smartphones), and on the assembly, testing and packaging of chips. Since the start of the 13th Five-Year Plan, China's semiconductor industry strategy now seeks to cover a much broader range of products and value chain stages, with a focus on memory semiconductors and AI chips.

Memory Semiconductors

China's massive push to develop from nothing a domestic memory industry has received investments of more than $40 billion in flash memory production, and another $10–$15 billion in DRAM memory production.[85] Why would China want to spend such vast resources to venture into the memory business, the "bleeding-edge" of the semiconductor industry with its enormous entry barriers and technological hurdles?

Looking at the demand side, there is reason to be cautiously optimistic about China's chances of success. China's massive efforts to upgrade its manufacturing and service industries will create a huge demand for sophisticated servers, where memory, especially 3D NAND flash memory will be critical. Main application markets will include automotive, AI, data centres and mobile platforms (IoT). This huge market has already

83 See www.semi.org/en/about/organization.

84 Interviews with industry experts who have requested anonymity.

85 Interviews with industry experts who have requested anonymity.

attracted massive foreign investment, with the result that China is emerging as a global memory cluster: Samsung's main 3D NAND fab is in Xi'an; SK Hynix has a DRAM fab in Wuxi; and Intel is expanding its Dalian fab to build 3D NAND and has packaging and test facilities in Chengdu. China's leadership has decided that it wants to participate in this new "memory rush."

Massive barriers, however, are looming on the supply side. The memory market is highly concentrated — with Samsung as the predominant leader. Late entry into such a market will come at a very high cost. China's scale disadvantages are mindboggling. According to a June 2019 report, China's largest memory supplier, Changxin Memory Technologies, has a few thousand employees and a capital spending budget of about $1.5 billion per year. In contrast, Samsung's memory division is estimated to have more than 40,000 employees, while the combined 2018 capital spending from the three industry leaders (Samsung, SK Hynix and Micron) exceeds $46 billion (Nenni 2019).

Most importantly, access to advanced memory technologies has been drastically reduced by the US government's tightening control of core technologies. The outbreak of the technology war thus raises China's cost of catching up in memory. Over time, however, there is little that can stop China from becoming a serious contender in this industry.

To sum up, China's semiconductor industry continues to lag well behind its American counterpart in terms of the depth of its domestic value chain and its international reach. Islands of emerging technological excellence continue to coexist with deeply entrenched structural weaknesses.

Chinese Players in AI Chip Design: Capabilities and Challenges

Most of our interviewees agreed that China is still way behind in the rapidly evolving markets for AI chips where CPUs, GPUs and FPGAs intensely compete with new special-purpose AI chips for algorithms used in deep learning neural networks. US firms are ahead in all of these fields. As explained before, entry barriers are very high. China faces substantial challenges in terms of access to experienced talent and intangible knowledge both for AI chip design and fabrication. As a result, AI applications run by major Chinese technology firms are still mostly powered by foreign chips. Now that China faces increasingly tight US technology restrictions, entry barriers into AI chips have become much more severe.

To find out what is happening in China's AI chip design, capabilities and challenges will be assessed, both for the large players (Huawei, Alibaba and Baidu), and for a small group of AI chip unicorns, which are emerging as credible new players in important market niches.

The Large Players

Huawei/HiSilicon

There is a broad consensus among AI chip experts that the only Chinese company with sufficient design and engineering talent to compete head-on with global US industry leaders is Huawei, due to its subsidiary HiSilicon.

Three factors have enabled HiSilicon to become a serious player in AI chips.[86] Size comes first, which enables large sales volumes, and hence economies of scale and scope. HiSilicon is China's national champion in IC design; rapid growth has made the company the largest chip designer in Asia and number four globally. But the size advantage goes further. As an affiliate of the Huawei group, HiSilicon also benefits from the mother company's huge market and its well-established relationships. Huawei is the world's largest supplier of telecommunications network equipment and the second-biggest maker of smartphones. With an annual revenue of RMB 721.2 billion ($107.13 billion) in 2018 (Kharpal 2019),[87] Huawei is in the same league as major US multinationals, such as Boeing's 2018 revenue of $101.1 billion (Boeing 2019).

In addition, Huawei has established a dense international network with key telecom operators, and with suppliers across the semiconductor value chains for telecom equipment, smartphones and servers. The company has pursued a two-pronged strategy: it is building a variety of linkages and alliances with leading global industry players and universities, while concurrently

86 Based on interviews with semiconductor industry experts, who have requested anonymity.

87 Note, however, that Huawei still ranks substantially behind Google's 2018 annual revenue of $136.22 billion, which is largely made up by advertising revenue.

establishing its own global innovation network.[88] Huawei's Global Innovation Network includes, in addition to at least eight major R&D centres in China, 10 overseas R&D centres in Europe and five R&D centres in the United States.

Huawei is the world's largest supplier of telecommunications network equipment and the second-biggest maker of smartphones.

There are no significant US companies that could fill the vacuum left by Huawei's absence. It would take quite a while for Nokia and Ericsson (the two main competitors from Europe) to take over. Telecom operators would not be happy, because they rely on Huawei's lower-cost offer of comparable, if not superior, technology. Economies of scale mean that Huawei can make equipment for 5G base stations for 20 percent to 30 percent less than its competitors.

Second, a fairly long history of persistently high investments in R&D has generated a large pool of top chip designers, whose experience and knowledge networks were essential for moving into AI chip design. Nearly 90,000 employees of Huawei are active in R&D, around 45 percent of Huawei's total workforce. In 2018, Huawei was the only Chinese company to make the top 10 global R&D investors, ranking fifth after spending $13.1 billion, up from being number six in 2017 (Hwang 2019, 31).[89]

Over the last three decades, high R&D investments in key enabling technologies have made HiSilicon a leader in the adoption of leading-edge process technology and in the design of specialized chips for 5G communications equipment. In fact, HiSilicon's leading AI chips (Ascent 310 and Ascent 910) are both fabricated at TSMC, the global foundry industry leader from Taiwan, using leading-edge 7 nm process technology.[90]

88 For an analysis of Huawei's global innovation network, see Ernst (2009).

89 With $15.1 billion, Samsung is in the number one spot. However, half of the top 10 R&D investors were from the United States.

90 See further discussion of Huawei's critical link with TSMC below.

Based on its long and persistent investments in R&D, Huawei has accumulated extensive experience in standard setting, especially for 5G mobile communications standards. China has focused on developing time-division duplex (TDD) technology as a matter of national policy since the 3G era, while US and European players have generally opted for frequency division duplex (FDD).[91] The two perform much the same when it comes to 4G, but TDD is expected to be the main choice for 5G, as FDD cannot manage the necessary transfer speeds.[92]

In addition, significant investments in the critically important SEPs for 5G technologies have enabled Huawei to sign cross-licensing agreements with all major IP rights holders in the wireless industry, including Ericsson, Nokia-Siemens, Alcatel-Lucent, Qualcomm, Nokia, Sony-Ericsson, Sisvel and other leading players (Ernst 2017). Given its huge portfolio of 5G-relevant SEPs, it will not be easy to push Huawei to the sidelines, not even for the powerful US government. Huawei thus has a fair chance to retain its leading market position in the telecom equipment market.[93] HiSilicon's AI chip design efforts will benefit from Huawei's strong position in relevant standards and SEPs.

Industry experts consider Huawei's AI chip development strategy to be "technologically brilliant."

A third factor was an early entry into AI chip technologies. Huawei was the first to bring AI to smartphones with HiSilicon's Kirin 970, designed in 2017, which draws on the Cambricon-1A design, which in turn used basic design features developed at the Chinese Academy of Sciences' Institute of Computing Technology.[94] It was followed by Kirin 980, which has delivered decent AI performance in an independent third-party benchmark. In October 2018, Huawei announced the Ascend series of its AI chips, for end-to-end applications ranging from smartphones (the Ascend 310) to data centre server applications (the Ascend 910), specifically addressing different computation needs and power consumption budgets.

An important indicator of Huawei's progress in technological capabilities is the introduction in September 2019 of the Kirin 990 5G chip, designed by HiSilicon, which is now used by the Huawei Mate 30 Pro smartphone. Fabricated on TSMC's enhanced 7 nm+ EUV (extreme ultraviolet) process, the 990 5G has substantially improved power efficiency, and incorporates an in-house developed DaVinci Neural Processor Unit (NPU) designed for AI acceleration.[95] Our interviews show that this new design is well received in the international AI chip design community.

All of these HiSilicon designs thus far are not sold in the open market; they are intended only for in-house use. As long as this does not change, this is bound to limit economies of scale and scope. Note, however, that in January 2020, Chinese AI companies such as SenseTime are now rumoured to be switching to HiSilicon AI chips in response to US technology restrictions. If confirmed, this would indicate that US technology restrictions are, ironically, helping HiSilicon to gain new markets outside the Huawei family.[96]

Huawei has heavily invested in HiSilicon's development of leading-edge AI chips, as part of an integrated AI technology ecosystem built around Huawei's Mindspore framework (Huawei 2019).[97] Industry experts consider Huawei's AI chip development strategy to be "technologically brilliant."[98] Huawei's "full-stack, all-scenario AI portfolio" was at the centre of discussion during the most important annual AI chip conference, HotChips 2019, sponsored by the Institute of Electrical and Electronics Engineers (IEEE) Technical Committee on Microprocessors and

91 TDD and FDD refer to techniques for spectrum usage.

92 See www.electronicdesign.com/datasheet/what-s-difference-between-fdd-and-tdd-pdf-download. To simplify, the main difference boils down to the usage of scarce frequency. FDD uses lots of frequency spectrum, generally at least twice the spectrum needed by TDD. In addition, there must be adequate spectrum separation between the transmit and receive channels. These so-called guard bands are not usable, so they are wasteful. Given the scarcity and expense of spectrum, these are real disadvantages of FDD relative to TDD.

93 See the recent decisions by the UK and German governments to retain Huawei as a supplier of non-core 5G telecom equipment (Kynge and Fildes 2020; Germano 2019).

94 See discussion below about China's AI chip start-up Cambricon.

95 See https://en.wikichip.org/wiki/Kirin_990#Overview.

96 Interviews in January 2020 with industry experts who have requested anonymity.

97 AI computing frameworks such as Mindspore are critical to making AI application development easier, expanding AI applications, making them more accessible and potentially enhancing privacy protection.

98 April 2019 interviews with industry experts who have requested anonymity.

Microcomputers, and held at Stanford University. Huawei's presentation on August 19, 2019, drew one of the largest crowds. Entitled "A Scalable Unified Architecture for Neural Network Computing from Nano-level to High Performance Computing," Huawei's approach signalled a broad-based AI chip design strategy so that one family of AI chips can be scaled both for inference at mobile devices and to large algorithm training jobs in data centres.

According to email interviews with AI chip experts after the HotChips 2019 conference, Huawei's push to develop its own AI chip approach clearly indicates that Huawei is prepared to deal with the worst-case scenario of a far-reaching technology export blockade. At the same time, Huawei also displays a strong sense of pragmatism. While racing ahead with its own integrated Mindspore framework, Huawei continues to support the leading AI frameworks, such as the open-source machine-learning framework Pytorch[99] and Google's TensorFlow, a free and open-source software library for data flow and programming used for machine-learning applications, including neural networks.[100]

It is noteworthy that the Huawei presenter at HotChips 2019, Heng Liao, had to give his talk by video, as visa restrictions imposed by the US government prevented him from attending the Stanford conference. During the Q&A session, an Intel engineer said, "This has been an international conference for years, and I'm glad you got to present here," generating spontaneous applause of approval from attendees (quoted in Merritt 2019). Such an emotional outburst is quite untypical for IEEE technical conferences. This shows the deep sense of frustration that the US technology warfare against Huawei has generated in this small but well-connected global network of leading AI chip designers.

Huawei's Persistent Dependence on Advanced Chips

Nevertheless, it is unclear whether and how soon Huawei could use its technological strengths to successfully compete against the leading US AI chip vendors. In fact, Huawei remains heavily dependent on advanced semiconductors from US and other foreign industry leaders, for advanced AI chips, as well as for high-end memory devices, and a long list of analogue semiconductors, including digital signal processors, radio frequency devices, filters, specialized antennas, power amplifiers and advanced power management integrated circuits.[101]

The outbreak of the technology war has further increased Huawei's dependence. The US technology ban imposed numerous supply chain constraints on Huawei, "including RF [radio frequency] components, EDA [electronic design automation] software, FPGAs, high-end analog chips, software and operating systems, which will make life difficult for Huawei in export markets. Huawei smartphones will lose access to Google services and may be forced to use open source Android…. Even Huawei smartphones sold in China will be less competitive if the ban remains during this year and next if the company lacks access to high-end RF chips, EDA software and IP, as well as Qualcomm chipsets" (Patterson 2019, 3).

As a recent international expert panel concluded: "It is very unlikely that Huawei can survive for very long without access to US technologies…. If a phone or a mobile operating system lacks even a single chip the whole thing may not work. And the US is often the main or the only provider of that chip technology" (Gavekal 2019, 7).

Of particular concern is that on August 21, 2019, the US Commerce Department blacklisted more than 20 percent of Huawei's global R&D centres, striking at the heart of the company's ability to innovate. This targeted blacklisting is likely to slow Huawei's impressive R&D capabilities, including HiSilicon's research on AI chips (Li and Cheng 2019).

In principle, Huawei could try to use alternative suppliers, primarily from Japan, Taiwan and Korea. This is exactly what Huawei is doing. Some of these attempts will eventually bear fruit. However, this will require often quite extensive redesign efforts, which will take time, and thus carry high costs and risks.

Access to Leading-edge Fabrication as a Major Challenge

Arguably the most serious challenge is that Huawei needs secure access to TSMC, the top

99 See https://pytorch.org/.

100 TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache License 2.0 on November 9, 2015. See https://en.wikipedia.org/wiki/TensorFlow.

101 Phone interviews with US semiconductor industry experts in December 2018 who requested anonymity.

provider of leading-edge semiconductor fabrication services (the so-called “foundry services”).

Huawei’s advanced AI chips critically depend on secure access to fabrication at leading-edge process nodes (at present 7 nm) provided by TSMC. In May 2019, TSMC announced “that after careful consideration, it will maintain its shipments to Huawei’s chip arm HiSilicon throughout this year.”[102] TSMC believes that “it doesn’t need to stop exports to Huawei as it does not have over 25% U.S.-origin technology in its manufacturing process.” About 90 percent of TSMC’s labour and overhead are in Taiwan and a substantial portion of its blank silicon wafers are from Japan, Europe and Taiwan. The only sizeable US inputs to TSMC are EDA software, IP and equipment.

Until late 2019, it seemed that TSMC had developed a strong legal position to counter the US government’s efforts to enforce the 25 percent US content rule. However, the US government continues to exert pressure on Taiwan’s government and on TSMC. As a well-connected Taiwanese source told me: “I think US government has absolute power to stop TSMC from serving Chinese firms like Huawei if it decides to do so. Then legally TSMC has to stop offering services to Huawei.”[103] As discussed at the beginning of the report, in January 2020, the US Commerce Department had submitted new regulations that would effectively close Huawei’s access to TSMC. In the meantime, rivalries within the US government seem to have complicated these efforts. In fact, the US Department of Defense (DoD) has blocked the proposed increase in technology restrictions against Huawei. Ironically, “the DoD is worried about the national security threats posed by Huawei building large parts of the world’s 5G networks, but it is even more worried about what would happen to US military hardware supply chains if key American semiconductor and component makers are deprived of the revenues they derive from selling to Chinese clients” (Wang 2020).

On purely economic terms, it would seem unlikely that the Taiwanese government would restrict TSMC, the country’s largest employer, exporter and single largest contributor to GDP. In addition, Huawei is TSMC’s second-largest customer, after Apple. Discontinuing foundry services to Huawei thus would come at a very high cost to TSMC. Nevertheless, it is unclear what will happen if the US government decides to reduce the US content minimum to 10 percent or zero, as proposed by the US Commerce Department. In any case, while economic arguments matter, geopolitical considerations may prevail, especially for a small island like Taiwan, which is sandwiched between the world’s largest military powers. The victory of the Democratic Progressive Party in Taiwan’s January 2020 election may increase the importance of geopolitics.

The Technology War against Huawei also Hurts the US Semiconductor Industry

America’s technology export restrictions will no doubt damage Huawei. But in the long term, this may end up hurting US component suppliers even more. US suppliers will suffer, as Huawei responds to the US technology ban by redoubling efforts to become self-reliant and placing heightened emphasis on sourcing components from suppliers based in Europe and elsewhere in Asia.

According to Bill McClean, a leading industry observer:

> Chinese systems suppliers will choose an alternative to a part offered by a U.S. company if the alternative is at all viable. If, for example, Texas Instruments has a part that competes with parts from companies such as STMicroelectronics, NXP Semiconductors, or Renesas Electronics, the Chinese company is going to go with one of the other vendors…. They’re going to avoid U.S. suppliers any chance they get going forward, even if this thing is settled tomorrow. The fear of God has been put into them, and so they are going to be looking at South Korean companies, Japanese companies, and European companies, and if there is an alternative, they are going to pick it. I guarantee that. The trust is gone….The gloves are off now. Anything goes now because of what the U.S. government is doing to Huawei.” (quoted in McGrath and Jorgenson 2019)

102 Official announcement at TSMC’s technology symposium on May 23, 2019, as quoted in Patterson (2019).

103 Email to the author, dated June 11, 2019, from Taiwanese industry expert who has requested anonymity.

Baidu and Alibaba Enter the Game

Two other large Chinese AI companies have entered the AI chip race — Baidu in July 2018, and Alibaba in September 2018. Both companies are still lagging behind Huawei/HiSilicon, in the size and the depth of their AI chip R&D. Nevertheless, their efforts have added new innovative AI chip designs, which are taken seriously by international AI chip experts.

Baidu's Kunlun AI chips are designed both for the company's autonomous vehicles and for data centres and come in two versions — for algorithm training and for inference. Baidu has partnered with Intel to optimize Intel's Xeon processor to run Baidu's own machine-learning framework, called PaddlePaddle.[104] That partnership is now on hold due to the US technology restrictions. In any case, Baidu's PaddlePaddle struggles to gain ground against TensorFlow and PyTorch, the globally dominant AI deep learning frameworks. According to a recent GitHub comparison, Baidu's PaddlePaddle trails TensorFlow by a factor of eight, and its use is declining.[105]

Baidu's Kunlun AI chip offers 512 gigabytes per second memory bandwidth and supplies up to 260 Tera operations per second at 150 watts. In addition, the new chip allows Ernie, a pre-training model for natural language processing, to infer three times faster than the conventional GPU/FPGA-accelerating model. According to a recent report, "leveraging the chip's limit-pushing computing power and power efficiency, Baidu can effectively support a wide variety of functions including large-scale AI workloads, such as search ranking, speech recognition, image processing, natural language processing, autonomous driving, and deep learning platforms like PaddlePaddle" (Jessie Shen 2019).

Baidu's Kunlun chip uses Samsung's 14 nm process node for fabrication. Samsung's decision to provide foundry services will allow Baidu to provide advanced AI platforms for maximizing AI performance, while Samsung will be able to expand its foundry business into high-performance computing chips that are designed for cloud and edge computing. This diversification of China's overseas sources of advanced foundry services signals that the country no longer relies on Taiwan's TSMC as the only source.

Alibaba has formally established a semiconductor business operation to design AI chips. Alibaba subsidiary T-Head (also known as Pingtouge Semiconductor) collaborates with the parent company's Machine Intelligence Technology Lab, to design AI chips. Focused on customized AI chips (i.e., ASICs), different versions are reported to be under development for autonomous vehicles, smart city planning and logistics. In April 2019, Alibaba acquired C-Sky Microsystems, a successful IC design company, also based in Hangzhou, Alibaba's hometown.

On August 21, 2019, Alibaba announced a specialized AI-FPGA chip for speech synthesis, called Ouroboros.[106] Alibaba claims that the Ouroboros FPGA is the industry's first AI FPGA chip design dedicated to speech synthesis algorithms, which can increase the computational efficiency of speech generation algorithms by more than 100 times. A review by the Linley Newsletter provides the following assessment: "Alibaba's data centers have deployed an Ouroboros FPGA prototype running on Xilinx VU9P boards... Alibaba says that in simulations of a 16nm ASIC, Ouroboros can synthesize speech with 30ms latency, matching the performance of the WaveNet model in Google data centers" (Demler 2019).

China no longer just depends on Huawei in its efforts to compete in advanced AI chips. Both Alibaba and Baidu are emerging as serious competitors.

Alibaba has also developed a low-power consumption chip, which, according to Linley Gwennap (2019), is unlikely to match DeepMind's voice quality. More recently, Alibaba has introduced a high-performance AI chip named Hanguang 800. This new chip is optimized for inference operations, a market that is growing faster than AI chips for the training of algorithms. The Hanguang 800 is "achieving 10x the performance of Nvidia's fastest GPU on the popular ResNet-50 model" (ibid.). Alibaba has already developed accelerator cards and systems using the new chip, and it has tested them

104 PaddlePaddle (Parallel Distributed Deep Learning) is an easy-to-use, efficient, flexible and scalable deep learning platform, which was originally developed by Baidu scientists and engineers for the purpose of applying deep learning to many products at Baidu. See https://github.com/PaddlePaddle/Paddle.

105 See https://python.libhunt.com/compare-paddle-vs-tensorflow.

106 See www.firstxw.com/view/237098.html.

in its data centres. It will also allow Alibaba Cloud customers to access Hanguang-enabled systems. Mass production of the 7 nm process technology was scheduled for December 2019 at TSMC.

In short, China no longer just depends on Huawei in its efforts to compete in advanced AI chips. Both Alibaba and Baidu are emerging as serious competitors.

AI Chip Unicorns

According to a recent OECD report, China has seen a dramatic upsurge in AI start-up investment since 2016. "It now appears to be the second player globally in terms of the value of AI equity investments received. From just 3% in 2015, Chinese companies attracted 36% of global AI private equity investment in 2017. They maintained an average of 21% from 2011 through to mid-2018" (OECD 2019a).

But what happened specifically in AI chip design? When I first started to look at Chinese AI chip start-ups in 2017, I was soon dealing with a growing list of 30-something companies. The floodgates of government and venture money had been thrown open for start-up funding. This was very different from Silicon Valley, where, as one expert put it, "investors flee at the mention of chip investment" (Rowen 2017). The main drivers for these initial investments in Chinese AI chip start-ups was a perception of lower chip development costs in China, due to lower engineering salaries, and because fabrication did not initially require costly bleeding-edge process technologies below 20 nm.

This initial AI chip fever was concentrated around four clusters: Beijing, Shenzhen, Shanghai and Hangzhou. In each of these locations, local governments followed the direction laid out by AIDP, competing against each other in providing incentives and other support. In our interviews, we found that the founders of the more promising of those AI chip start-ups typically had studied at leading research universities in both China and the United States. They also sought to cultivate close contacts with both the central and the local governments. In essence, the typical start-up business model was focused on monetizing intimate knowledge of highly specialized leading-edge IC design techniques with a narrow focus on a handful of high-growth AI mass applications.

An important weakness of Chinese AI chip start-ups is that they have largely focused on inference, neglecting the training of algorithms, which is technologically much more demanding and remains dominated by US firms, especially Nvidia's GPUs. An additional weakness is that China's AI chip design companies have focused primarily on the rapidly growing demand for surveillance applications, especially for face identification, pedestrian tracking and crowd monitoring. This narrow specialization has set China's AI chip start-ups apart from the United States, where two-thirds of the start-ups are focused on cloud software and services, leveraging the easy development and powerful ecosystem effect of cloud services, provided by Amazon AWS[107] or Microsoft Azure,[108] the global industry leaders.

China's approach to "infant industry protection" in internet services has produced conflicting results.

Cloud-centric AI chip start-ups are still rare in China, perhaps reflecting the negative effects of the "Great Firewall."[109] By blocking access to Google, Facebook and other global digital platform leaders, the Great Firewall has enabled companies such as Tencent, Alibaba and Baidu to build up their own cloud services industry in China. But this has come at a heavy cost, delaying the development of China's infrastructure-as-a-service (IaaS) public cloud market. This is changing — China's IaaS market is expanding now at a faster pace than in other countries. After posting 86 percent year-on-year growth in 2018 to become worth $4.65 billion, China now has the world's second-largest public cloud IaaS market (CGTN 2019). However, Amazon, followed by Microsoft, Alibaba, Google and IBM, continues to dominate the global cloud services market (Gartner 2019). Of the more than 400

107 See https://aws.amazon.com/.

108 See https://azure.microsoft.com/en-us/.

109 According to Wikipedia, the Great Firewall of China "is the combination of legislative actions and technologies enforced by the People's Republic of China to regulate the Internet domestically. Its role in the Internet censorship in China is to block access to selected foreign websites and to slow down cross-border internet traffic. The effect includes: limiting access to foreign information sources, blocking foreign internet tools (e.g. Google search, Facebook, Twitter, Wikipedia, and others) and mobile apps, and requiring foreign companies to adapt to domestic regulations....The Great Firewall is operated under the 'Golden Shield Project' by the Bureau of Public Information and Network Security Supervision" (see https://en.wikipedia.org/wiki/Great_Firewall).

worldwide data centres, 147 are run by Amazon, Microsoft and Google (Hwang 2019, 39).[110]

In short, China's approach to "infant industry protection" in internet services has produced conflicting results. By blocking Google and Facebook from the Chinese market, the government has created an "online parallel universe" for Alibaba, Tencent and Baidu. This divided market for cloud services may now constrain China's efforts to broaden the specialization of Chinese AI chip start-ups beyond vision and speech recognition into cloud server applications.

China's Leading AI Chip Unicorns

China's most promising AI chip unicorns[111] have received massive support from the government as well as from venture capital investors. On the demand side, the government is also an important customer. As a result, some of these new AI chip players have rapidly developed their technological and management capabilities. As Huawei now faces increasing restrictions on access to US technology, the companies discussed below are likely to play an important role in strengthening China's AI chip ecosystem.

DeePhi

DeePhi was founded in 2016 by a group from Stanford and Tsinghua Universities, and has raised around $40 million in investment. The company specializes in neural networks, and to this end has two AI chip architectures that are pursuing both vision and speech applications. Aristotle is designed for video and image recognition, and uses convolutional neural networks, while Descartes is intended for speech recognition. The start-up was known for its expertise in model compression — the practice of taking the neural network model that interprets data and comes to a decision, and then compressing it into a smaller package of embedded devices. Crucially, DeePhi was good at doing this without the model losing accuracy.

In July 2018, the company was acquired by Xilinx, one of America's two global leaders in the FPGA market.[112] Xilinx's main motivation for acquiring DeePhi was to gain access to a team of experienced experts in this critically important technique. This created quite some excitement in China's AI chip community. Some observers argued that, in contrast to the usual narrative in the United States of China acquiring US technology, the acquisition of DeePhi by Xilinx provides an example of a leading US AI chip design company acquiring Chinese technology, as embodied in the intangible knowledge accumulated by DeePhi's Chinese engineers. However, there are reasons to be skeptical. According to Sun Yongjie, a prominent Chinese technology blogger, DeePhi's deep learning processors are entirely built on Xilinx FPGA frameworks: "If DeePhi Tech ever broke away from Xilinx's FPGA platform, it would be completely cut off from all sustenance" (quoted in Laskai and Toner 2019).

An interesting question for further research will be to examine how the regulations of the Export Control Reform Act (ECRA) of 2018 will deal with technology exports originating from the China-based research team of Xilinx that is made up of top-rated Chinese engineers.

Cambricon

Beijing-based Cambricon Technologies was founded in 2016 by Chen Tianshi, a professor at the Institute of Computing Technology at the Chinese Academy of Sciences. The company started out as a supplier of specialized AI chip design to HiSilicon, by licensing its Cambricon-1A design to HiSilicon's Kirin 970 chip. However, in 2018, Cambricon lost Huawei's partnership. During the Huawei Connect conference in October 2018, Huawei introduced its own Da Vinci AI architecture and new Ascend family of chips, Ascend 310 and Ascend 910, both of which were developed by Huawei itself.[113] According to Zhi Jun, rotating chairman of Huawei, "Cambrico's IP is quite good, but it is insufficient as we require chips that can support all AI application scenarios. We want to have a system that can provide both massive computing capacity and extremely low-power consumption" (quoted in Huang 2018).

Despite this setback, Cambricon has raised several hundreds of millions of US dollars in a June 2018

110 According to this study, Amazon, Microsoft and Google are leveraging big data analytics and cloud services to raise entry barriers for latecomers.

111 The selection is based on information received during our interviews in China, and from phone and email interviews with international AI chip experts.

112 See www.xilinx.com/about/company-overview.html.

113 See earlier discussion of Huawei's AI chip development.

series B funding round, increasing its valuation to $2.5 billion. The funding was led by SDIC Venture Capital, a subsidiary of state-owned investment holding group China Reform Holdings Corp Ltd.[114] Cambricon's long list of big-name investors includes Alibaba, Lenovo, CICC Capital, CITIC Securities Goldstone Investment Fund, TCL Capital and a Chinese Academy of Sciences fund. This funding round made Cambricon a "unicorn."

In addition to the Cambricon-1A chip, Cambricon designs core processor chips (called MLU100) for intelligent cloud servers, intelligent terminals and smart robots. Cambricon chips are fabricated at TSMC, the MLU100 uses "slightly-behind-the-leading-edge" 16 nm technology. Like Huawei, Alibaba and Baidu, Cambricon's MLU100 thus critically depends on secure access to TSMC foundry services.

MIT's prestigious Lincoln Laboratory has included Cambricon's cloud server chips in the 2019 Survey of AI chips (Reuther et al. 2019). This has added credibility to Cambricon's AI chip design business. China's main server companies, Lenovo and Sugon, are using Cambricon's chips.

Note, however, the Linley Group's rather critical assessment:

> The company expects more than one billion devices to use its AI designs by 2020. For comparison, we project cumulative shipments of all AI accelerators...[AI chips in the terminology of this report]...by 2020 will be 1.9 billion units, most in smartphones; to gain a majority share of these shipments, the startup would need to win large foreign smartphone customers (e.g., Apple, Qualcomm). Cambricon has developed an AI chip for security cameras and is developing chips for autonomous driving and for data centers. The latter two markets consume few units, but security cameras are a big market: China alone plans to deploy 400 million new units by 2020. Most of these cameras, however, probably won't have an integrated AI chip, instead relying on remote intelligence.
>
> Cambricon claims its AI design is 20 times more power efficient than standard GPUs such as Nvidia's Kepler, a 2013 design, but even Nvidia has delivered a 20x improvement since then. The startup has withheld performance and power specifications for its products, making it difficult to assess their competitiveness. But I believe a true world leader would boldly disclose its capabilities. Simply announcing an ambitious goal and throwing billions of dollars at it doesn't ensure success. (Gwennap 2018)

The latter statement highlights a troubling constraint to any meaningful evaluation of the often quite adventurous claims of AI chip start-ups in China, but also in Silicon Valley. There is still no reliable common benchmark scheme for evaluating the performance of specialized AI chips.[115] As one industry insider puts it: "I'm honestly getting a bit exhausted seeing Chinese and the western media (and western analysts) make claims of 'ground breaking' and 'novel' Chinese advances when it's really again just something that is trying to mimic and then slightly improve upon pre-existing innovation."[116]

In fact, a lot of the foreign commentary on China's efforts in AI chip design is about hype — either boosting it or debunking it. The more significant point for analysis, though, is that China's convergence to the tech frontier means that it has now the ability to mimic or make small improvements. Hence, what the above analyst really is saying is that China has now reached the technology frontier. That is pretty impressive.

Horizon Robotics

Horizon Robotics (HR) is a deep learning AI chip design house, founded in 2017 by Yu Kai, who led the self-driving car project at Baidu. HR's co-founding team combines capabilities in algorithms, software, hardware, and chip development. With 220 people, HR seeks to develop AI chips across a broad front for vision systems, including smart home, gaming, voice-vision integration and self-driving cars. It also has adopted a tight hardware-software integration strategy for more complete and efficient

114 See www.sdic.com.cn/en/about/A0201index_1.htm.

115 While officially MLPerf (https://mlperf.org/) is supposed to provide a standard, its efforts are still at an early stage.

116 Semiconductor industry expert, who requested anonymity.

solutions. Over time, it has increasingly focused on imaging and facial recognition algorithms.

In 2015, HR introduced its BPU (brain processing unit) architecture for inference acceleration.[117] HR claims that its Sunrise 1.0 processor is superior to Nvidia's TX1 in power consumption and imaging performance. The company's Journey 1.0 processor is fabricated by TSMC, while its Journey 2.0 processor is based on Intel's FPGA. Feedback from experts who were interviewed did not question these claims.

The response from investors provides an indirect confirmation of HR's achievements. Since 2015, the company has attracted substantial venture capital investment from Morningside Venture Capital, Hillhouse Capital, Sequoia Capital, GSR Ventures, Linear Venture, Innovation Works, ZhenFund, Wu Capital, Tsing Capital and Vertex Ventures, as well as from Yuri Milner, a venture capitalist from Silicon Valley. In February 2019, HR had raised $600 million, led by South Korean chip maker SK Hynix Inc. and several "first-tier" Chinese automakers, while other participants included China Oceanwide Capital, CITIC Securities' One-Belt-One-Road Fund, and Minsheng Capital (Reuters 2019). And in November 2019, HR raised up to $1 billion in a funding round that will value it at between $3 billion and $4 billion (Lucas 2018).

Finally, expert interviews emphasize that, in response to the intensifying US technology restrictions, companies like HR are now benefiting from increasing government support. Ironically, the US technology war against China has provided a powerful catalyst for Chinese AI chip companies to move ahead with their own designs, no matter the costs.

Yuntian Lifei/Intellivision

Yuntian Lifei was founded in 2014 by a team of doctorate degree holders who had returned to Shenzhen from the United States. The company is focused on edge-based AI chips for visual recognition and security monitoring systems, a rapidly growing market in China. The company claims that its AI processors were developed in-house, in addition to supporting software and big data analytics. It is reported that the company cooperates with Huawei on an AI vision platform.

117 See https://autonomous-driving.org/wp-content/uploads/2018/06/Horizon-Robotics-Embedded-AI-Processors-for-Autonomous.pdf.

Its most recent processor seems to have been fabricated by TSMC at 22 nm process technology.

Suiyuan Technology

Founded in March 2018, the company has R&D centres in Beijing and Shanghai. As mentioned earlier in the report,[118] a Tencent-led funding of $500 million is reported to be focused on cloud-based AI chips that aim to compete with the GPUs of Nvidia, the global industry leader. Zhao Lidong, the company's CEO, has previously worked at AMD China. He says that the company will follow the Chinese government's Development Plan on the New Generation of Artificial Intelligence, and aims to become the AI chip solution and technology leader in the Chinese market (Tian 2018). [119]

Is China's AI Chip Investment Fever Cooling Down?

In January 2019, Chris Rowen observed: "It feels like valuations in China for equivalent quality teams is perhaps 2x that of Silicon Valley (and Silicon Valley is perhaps 2x the rest of the world). It is definitely a case of too much money chasing too few experienced, credible AI startup teams in China."[120] It seems that the AI chip investment fever that reached its climax in 2018 is now cooling down.

One possible explanation could be that chip development costs in China are no longer as attractive as before, as Chinese engineering salaries are rising. In addition, AI chips developed in China have become more expensive, as they now use increasingly complex chip designs and leading-edge fabrication process technology.

A closer look reveals a split picture. New venture capital investments both from China and the United States increasingly are directed toward AI chip start-ups for cloud server applications, in line with the accelerating growth in this AI applications market. This contrasts with a slowdown in venture capital investments into the markets for vision and speech recognition. The latter markets not only may have been saturated with funds, but for US venture capital funds, these surveillance-related markets are also politically

118 See the section "China's Three Separate AI Development Trajectories" (p. 19).

119 As discussed earlier in the section "Contrasting America's and China's Different AI Development Trajectories."

120 Chris Rowen, email to the author, January 6, 2019.

risky, in light of widespread criticism in the United States. In short, China's AI chip investment fever seems to continue, albeit now with a new focus on cloud computing AI applications.

To summarize, our field research has shown that China has made some progress in its efforts to improve its access to AI chips in the face of intensifying US technology export restrictions. Huawei, as well as Alibaba and Baidu, have demonstrated a capacity to design technologically sophisticated AI chips as part of an increasingly integrated AI ecosystem. In a very short period of time, these three companies have developed a substantial body of new technological and management capabilities. Some of these capabilities are now also emerging in a small group of AI chip start-up companies. Our interviews show that China's response to US restrictions in AI chips is now drawing on a better understanding of the strengths and weaknesses of global industry leaders from the United States and from other advanced countries.

There is no doubt, however, that all these efforts are still very much work in progress. It remains unclear how fast China can leverage its newly developed technological capabilities to reduce its dependence on advanced specialized semiconductors for AI. This raises a fundamental question that will be discussed in the conclusions: in light of the headwinds that China is facing, what are its strategic options to gradually catch up with global industry leaders in AI chips?

Conclusions

Key Findings

A basic objective of this report has been to fact check the assumptions underpinning America's claim that China is about to forge ahead as an AI technology leader. Focusing on AI chips, our research finds that China still has a long way to go before it could credibly threaten America's leadership. Such fears are not grounded in reality, but reflect a general China-bashing mode in Washington, DC, that is driven primarily by ideology and geopolitical considerations.

A similar assessment of China's persistent gap in AI technology can be found in a recent testimony before the US-China Economic and Security Review Commission: "China is not poised to overtake the U.S. in the technology domain of AI; rather, the U.S. maintains structural advantages in the quality of S&T inputs and outputs, the fundamental layers of the AI value chain, and key subdomains of AI" (Ding 2019, 1). Our findings are also supported by a report from a think tank closely linked to the US IT industry: "Overall, the United States currently leads in AI, with China rapidly catching up, and the European Union behind both. The United States leads in four of the six categories of metrics this report examines (talent, research, development, and hardware)...[while]...China leads in two (adoption and data)" (Castro, McLaughlin and Chivot 2019).[121]

Developing a robust domestic AI chip industry is arguably China's most immediate and difficult AI challenge. China struggles to ensure secure access to AI chips. This was the case even before the outbreak of the technology war. Now that China faces increasingly tight US technology restrictions, these entry barriers have become much more severe.

China's AI chip industry is mounting a vigorous catch-up campaign whose technological achievements are well received by experts. Nevertheless, China continues to lag considerably behind the United States in the depth of its domestic AI chip value chain, and even more so in terms of its international reach. Islands of technological excellence continue to coexist with deeply entrenched structural weaknesses in China's emerging AI chip industry. It is hard to see how China could become a threat to the United States any time soon. The real question is how fast China can converge to the technological frontier in AI chip development despite US restrictions.

US pressure is indeed seeking to slow down China's rise as a peer competitor in AI chips. We saw that the unintended negative consequences of such enforced delays are severe, both for the United States and for China. In any case, it is certainly unrealistic to assume that America could effectively block China's efforts to forge ahead in this critically important building block of the AI industry. China will be able to move forward in AI chips, as it can draw on significant government-supported investments, a large engineering pool and dozens of semiconductor fabs under construction or on the books.

Decoupling Prospects

A central question, of course, is how far the decoupling will progress between the American and Chinese production and innovation systems. Irrespective of the ups and downs of the official US-China trade negotiations, it is clear that the heating up of technology warfare is already transforming long-established trade and investment patterns in the IT industry. Especially in the AI industry, at least some partial decoupling is under way between these two most important competitors.

In fact, both the United States and China are driving decoupling. China was an early mover. As my CIGI colleague Paul Blustein (2019, 245-46) puts it, "Thanks to China's Great Firewall, a 'splinternet' already divides the world's netizens into those who use Google, Facebook and Twitter versus those who use WeChat, Weibo and Alibaba. The US crackdown has accentuated the probability that something akin to a Digital Iron Curtain is in the offing, which would mean companies in a variety of tech sectors — telecommunications, artificial intelligence, semiconductors, and so on — operating in separate ecosystems, with some obliged to choose between China and the United States."

121 The Center for Data Innovation is affiliated with the Information Technology and Innovation Foundation, a think tank representing major US IT companies.

In both countries, decoupling was driven by powerful players: "For security hawks in the U.S., decoupling with China — first in trade and then in technology and finance — would reduce China's economic growth potential significantly, and thus contain its power. From the perspective of their Chinese counterparts, excessive dependence on the U.S. market and technology also presents unacceptable risks to sovereignty and national security" (Pei 2019, 1).

While decoupling is real, it remains porous and meandering, and hence may leave room for pragmatic counterstrategies.

An insightful analysis of the US-China technology war concludes that "the US has gone into full decoupling mode with China," giving rise to "a Balkanized technology world, with systems used by China and its satellites failing to communicate with systems in the West" (Stevenson-Yang 2019, 1). As mentioned at the beginning of the report, war-gaming US-China technology competition has been going on for quite some time. On Huawei, the Defense Innovation Board has extensively prepared 5G competition scenarios (Medin and Louie 2019). On supercomputing (especially exascale computing), relevant scenario research is being conducted at the National Security Agency and, more operationally, at the Office of Advanced Simulation and Computing, part of the National Nuclear Security Administration at the US Department of Energy (Thibodeau 2017). As for AI, DARPA's new five-year, $2 billion plan seeks to outdo China (DARPA 2018). A massive push in decoupling thus is the main goal of US technology export restrictions codified in ECRA, FIRRMA and CFIUS.

Thus far, however, relocation by leading US semiconductor firms of their affiliates out of China has remained limited. Too much is at stake for these companies that critically depend on the vast Chinese market. In the United States, powerful interest groups, both in industry and in leading research universities, represent formidable voices in the fight against decoupling. In December 2019, US technology companies rebuffed a Trump administration request that they pledge to stop sourcing supplies from Huawei and other Chinese IT companies (Stacey 2019). In addition, US technology companies have engaged teams of top attorneys to figure out ways to bypass US technology restrictions, some of which have been quite hastily prepared. Even the US Defense Department seems to have second thoughts. As described at the beginning of the report, in January 2020, the DoD decided to block a further extension of technology restrictions against China.

My gut feeling is that, given the considerable lobbying power of the US semiconductor industry, US technology export restrictions may be implemented incrementally, rather than in a massively destructive big push.[122] In other words, while decoupling is real, it remains porous and meandering, and hence may leave room for pragmatic counterstrategies.[123]

But decoupling is not just shaped by US actors. Take Huawei, the company at the centre of US technology warfare. Huawei is under considerable pressure to search for alternative suppliers, in order to be prepared in case the Commerce Department technology blockade would be fully implemented. The US blockade is already having negative implications for US companies, as they are now viewed as potentially unreliable suppliers. Some industry experts argue that a new supply chain for Huawei and other Chinese IT firms may gradually take shape across Asia as Chinese companies are seeking to wean themselves off of US suppliers of crucial components.[124] For instance, Huawei reportedly is redesigning its base stations to accommodate slightly lower-performance components for local deployments; Huawei is also relying more on Taiwanese suppliers, with a new ecosystem possibly taking shape to serve mainly Chinese IT firm (*Digitimes* 2019). This, as discussed, would assume that US exterritoriality pressure on Taiwan will not block such new supply sources.

Since January 2020, China is now facing additional pressure. The outbreak of COVID-19 has dramatically increased China's challenge, adding to de facto decoupling of China from the global IT value chain.[125] On January 23, 2020, the Chinese government closed off Wuhan (an important

122 An example may be the unexpected decision by CFIUS, mentioned at the beginning of this study, to clear the acquisition of Beijing OmniVision Technologies Co., Ltd. by Shanghai Will Semiconductor Co.

123 See further below under "Implications for China's Future AI Chip Development."

124 Interview with industry experts based in Asia who have requested anonymity.

125 On the impact of the coronavirus crisis on the GVC of the information technology industry, see Hille, McMorrow and Liu (2020).

semiconductor industry cluster) and other cities in Hubei province in an effort to quarantine the epicentre of the COVID-19 outbreak.[126] On February 2, 2020, the US government started implementing stringent travel restrictions that are seriously disrupting trade and knowledge exchange in the global IT industry.[127] In addition, most global air carriers have suspended flights to and from China.[128] In an industry that is unmatched in its dependence on international trade and knowledge exchange, it is clear that enormous damage will be caused by such fundamental disruptions. An indicator of how deeply semiconductor GVCs might be affected is the announcement on February 4, 2020, that the all-important meeting for the global semiconductor industry, the SEMICON / FPD China 2020 trade fair in Shanghai will be postponed indefinitely: "In accordance with government guidelines and to protect the health and safety of exhibitors and guests, we regret to inform you of the postponement of SEMICON/FPD China 2020 and related events originally scheduled for March 18–20, 2020," SEMI has announced.[129]

In short, it is impossible at this stage to predict how much and for how long the coronavirus epidemic might further add to the decoupling of the GVC linkages between the United States and China that are critical for China's development of its AI chip industry.

Implications for China's Future AI Chip Development

Faced with rising US technology restrictions and the GVC disruptions caused by COVID-19, China's secure access to leading-edge AI chips is under a very real threat. China's leadership believes that a robust domestic AI chip industry is needed, if China wants to sustain its still highly fragile achievements in commercial AI applications. Even more than the ZTE shock in 2018,[130] the massive campaign against Huawei will likely force China to accelerate its drive for self-reliance. From a US perspective, it is somewhat ironic that US restrictions on technology exports may actually strengthen China's resolve to strengthen its domestic capabilities.

In January 2019, President Xi announced a new 15-year Science and Technology Innovation Plan to speed up its indigenous innovation campaign.[131] Under the headline "Let everyone focus on innovation," Xi used a talk at the Tianjin Binhai-Zhongguancun Collaborative Innovation Exhibition Center to highlight China's achievements in the "Tianhe" series of supercomputers; the Feiteng chips, a series of ARM microprocessors designed and introduced by Phytium, a leading Chinese IC design company.[132] Xi also highlighted Huawei's Kirin AI chips, AI-enabled smart robots and drone cluster intelligent control systems. Talking to a large crowd of R&D managers from leading Chinese IT companies, Xi emphasized that "high-quality development depends on innovation, and our country's further development depends on independent innovation."[133]

China's leadership believes that, as the United States seeks to block China's progress in AI chips, it has little choice but to develop an almost completely integrated domestic value chain. China continues to advance its foundry industry with huge investments in new fabs and technology, despite trade tensions and a slowdown in the IC market (Lapedus 2019). According to data from SEMI's 2018 World Fab Forecast report, China has the most fab projects in the world, with 30 new facilities in construction or on the drawing board (Dieseldorff and Liu 2018). Of those, 13 fabs are targeted for the foundry market, according to SEMI. The remaining facilities are geared toward LEDs, memory and other technologies.

In addition, China's push to increase vertical integration now is targeted to expand across all stages of the semiconductor industry value chain, including production equipment. At the September 2019 China Semiconductor Industry Summit, the

126 See https://en.wikipedia.org/wiki/2020_Hubei_lockdowns#Elsewhere_in_China.

127 Entry into the United States is temporarily denied to foreign nationals who visited China in the 14 days prior to their arrival to the United States. Restrictions also apply to US citizens who have been in China's Hubei province, the epicentre of the COVID-19 outbreak, in the two weeks prior to their return to the United States. Upon their return, those citizens will be subject to a mandatory quarantine of up to 14 days. US citizens returning from the rest of mainland China in the 14 days prior will undergo health screenings at selected ports of entry and face up to 14 days of self-monitored quarantine (Andone 2020).

128 See Toh and Riley (2020).

129 As reported in Chen and Chan (2020).

130 In April 2018, the US Commerce Department imposed a seven-year components ban on ZTE, China's second-largest telecommunications equipment maker, which was later withdrawn (Delaney 2018).

131 See www.xinhuanet.com/politics/2019-01/18/c_1124006805.htm.

132 See https://en.wikichip.org/wiki/phytium/feiteng.

133 See www.xinhuanet.com/politics/2019-01/18/c_1124006805.htm.

Big Fund's manager revealed that Phase II will focus on etching machines and film, test and cleaning equipment. "The goal is to build an independent, self-sufficient and controllable industrial chain for the Chinese IC industry" (Liu 2019).

In China, the slogan of self-reliance may be attractive as a tool for mass mobilization. After all, "'catch up and surpass'...abbreviated as 'ganchao' in Chinese, has long been one of the Chinese Communist Party's defining goals....In the minds of China's leaders, from Mao Zedong to Xi Jinping, technological progress is not only a means to economic and military prowess but also an ideological end in itself — offering final proof of China's restoration as a great power after decades of struggle" (Gewirtz 2019).

Within China's semiconductor research community, indigenous innovation is widely believed to be the missing link to China's otherwise successful catching up through global knowledge sourcing. Take Ni Guangnan, a well-known academician at the Chinese Academy of Engineering and a senior researcher at the Institute of Computing Technology of the Chinese Academy of Sciences.[134] In a widely circulated interview, Ni argued that "the development of ...[China's]...integrated circuit industry requires long-term ideological preparation and investment. We cannot expect to get returns in a few years. It may take another decade or two to really develop the integrated circuit industry. We must have the determination to make up for the shortcomings in the industry and persevere steadfastly."[135]

In our interviews, voices within China's semiconductor industry suggested a pragmatic approach to indigenous innovation that is informed by the dynamics of competition in this industry. Most interviewees agreed with the proposition of this report that decoupling is real, but porous and meandering, rather than cast in iron. Under such circumstances, a pragmatic Chinese approach to indigenous innovation may consist of three building blocks:

→ Vertical integration across all stages of the semiconductor industry value chain, including production equipment, should be the overriding long-term goal. However, two complementary and supporting short-term strategies are needed in order to improve the chances of success of the long-term strategy.

→ One complementary strategy would focus on "slightly behind-the-leading-edge" chip architectures and process nodes.[136] This approach would be good enough for diffusing AI technologies across China's manufacturing and service industries. It is argued that in a "porous trade decoupling" scenario, such a pragmatic approach could build enough capabilities until the time when global trade can grow again. A key proposition is that the big changes, discussed earlier, in chip design may open up new opportunities for Chinese firms to move gradually forward in at least some niches of the AI chip market.

→ Another complementary strategy would be a hybrid approach to leapfrogging, which might work under two conditions:

  - if leapfrogging is focused on specific niche markets that are of critical importance for China's development, i.e., 5G telecom, or the Smart Grid; and

  - if such leapfrogging is supported by strategic partnerships with leading foreign semiconductor companies.

Most of our interviewees agreed that, in order to improve self-reliance, China would first need to conduct in-depth empirical research on what precisely are China's major challenges in AI chips. Such an analysis would need to explore: the access to top experienced talent; the persistent dependence on leading-edge GPUs, CPUs and FPGAs; the dependence on software tools and frameworks; whether access to leading-edge fab capacity is a serious bottleneck; the role that China should play in AI chip-related standardization; and the role of entry barriers due to oligopolistic market structures.

134 A former chief scientist of Legend, the predecessor of Lenovo, Ni lost a power struggle with Liu Chuanzhi (the company's founder) when he was arguing for the indigenous development of a PC hardware card that allowed the input of Chinese characters.

135 See www.jqknews.com/news/202860-Chivalrous_Island_Interview_with_Ni_Guangnan_China_Core_Design_Progress_Making_is_Short_Board.html. See also Lingzhi and Jie (2018).

136 For a detailed analysis of such a strategy, see Part Three: Upgrading Prospects — Economic Reasons for a Bottom-up, Market-led Industrial Policy in Ernst (2015).

In order to make self-reliance work, China also needs to adjust its approach to international technology competition. The report has demonstrated that global knowledge sourcing through international cooperation in science and technology remains critical, if China wants secure access to key AI technologies. After all, a defining characteristic of AI is that AI researchers publish their algorithms, codes, chip designs, data and results in real time and to an international readership on open web platforms such as arXiv and GitHub, where others, irrespective of nationality, can find and use the research (O'Meara 2019). Techno-nationalism would destroy AI's global knowledge-sharing culture. This is true both for the "America First" doctrine, and for China's attempts to claim sovereignty over its cyberspace through the Great Firewall and the Cybersecurity Law.

China's leadership thus needs to reconsider the notion that the country can only progress in AI if it pursues a zero-sum competition policy in its relationship with the United States and other advanced countries. There is a widespread perception in the international community that China's government over the last few years has overplayed its hand with a bevy of new stringent, and often quite ambiguous, regulations. This perception extends well beyond the United States, and is widely shared across Europe and Canada, as well as in Japan, Korea and Taiwan. It is time for China to acknowledge that it needs to provide safeguards to foreign companies against forced technology transfer through policies such as compulsory licensing, cyber security standards and certification, and restrictive government procurement policies. Paraphrasing Blustein (2019, 266), China would clearly benefit if efforts to "Make the WTO Great Again" would enable trade to "be pretty free, pretty fair and pretty reciprocal — with the big bonus of a rule of law system that fosters stability and predictability."

On the domestic front, this report highlights a number of priority issues that need to be addressed if China wants to broaden its achievements in AI.[137] Most importantly, it is necessary to overcome the still-prevailing fragmentation of diverse policies that reflect deeply engrained interagency rivalries. Knowledge sharing between academic research and industry needs to be drastically strengthened in order to overcome the entrenched fragmentations across China's innovation system. By the same token, China needs to establish effective policies to incentivize companies to conduct basic AI research, and to improve R&D productivity.

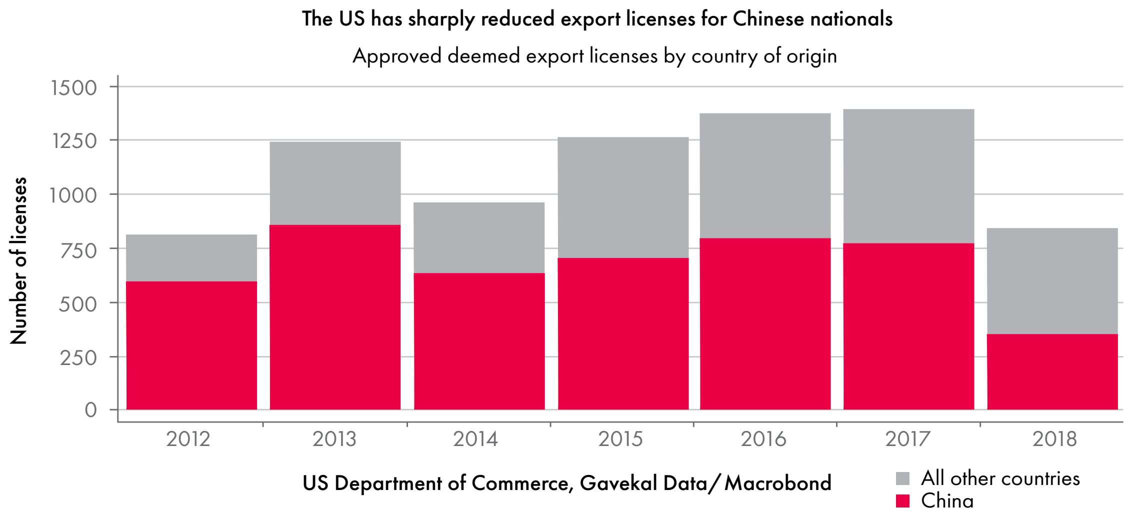
To conclude, efforts to upgrade China's AI chip industry amidst technology war require a unified strategy that combines innovation, trade, competition and standardization policies.

Acknowledgements

The study owes much to extremely helpful comments and criticisms from anonymous reviewers, and from Dan Ciuriak, Doug Lippoldt, Raul Rojas, Ron Wilson, Chris Rowen, Linley Gwennap, Bill McClean, Mark Li, Michel Girard, Michael G. Plummer, Paul Triolo, Jimmy Goodrich, Julian Snelder, Arthur Kroeber, Dan Wang, Anne Stevenson-Yang, Jan-Peter Kleinhans, Boy Luethje, John Ravenhill, Anton Malkin, Alex He, Xue Lan, Wang Ping, Gu Shulin, Chen Tain-Jy, Chen Shin-Horng, Colley Hwang, Liang Zheng, Yin Shouyi, Yu Zhen, Ning Zou, Gu Xiongli, Hou Junjun, Zhang Yaokun, Yu Hanzhi and Yi Zhao. I am immensely grateful to Rohinton Medhora, CIGI's president, and to Bob Fay, director, Global Economy, for their encouragement and support. Many thanks also to Heather McNorgan, program manager, and to Jennifer Thiel, program assistant, for helping me implement this project. I am very grateful to Jennifer Goyder for editing the report and for her kindness and patience.

137 Issues related to AI standardization and data governance are addressed in a separate paper (Ernst 2019).

Appendix 1: US Reduction in Export Licenses for Chinese Nationals



US university labs and US companies will find it more difficult to recruit and retain Chinese talent, which would quite drastically constrain the diffusion of AI technologies.

Source: Gavekal. Reproduced with permission.

Appendix 2: List of AI-related Emerging Technologies Considered to Be Essential to US National Security, Proposed by the US Commerce Department's BIS

1. AI and machine-learning technology, such as:

- Neural networks and deep learning (e.g., brain modelling, time series prediction, classification);
- Evolution and genetic computation (e.g., genetic algorithms, genetic programming);
- Reinforcement learning;
- Computer vision (e.g., object recognition, image understanding);
- Expert systems (e.g., decision support systems, teaching systems);
- Speech and audio processing (e.g., speech recognition and production);
- Natural language processing (e.g., machine translation);
- Planning (e.g., scheduling, game playing);
- Audio and video manipulation technologies (e.g., voice cloning, deepfakes);
- AI cloud technologies; or
- AI chipsets

2. Microprocessor technology, such as:

- Systems-on-Chip (SoC); or
- Stacked Memory on Chip

3. Advanced computing technology, such as:

- Memory-centric logic

4. Data analytics technology, such as:

- Visualization;
- Automated analysis algorithms; or
- Context-aware computing

5. Quantum information and sensing technology, such as:

- Quantum computing;
- Quantum encryption; or
- Quantum sensing

6. Robotics, such as:

- Micro-drone and micro-robotic systems;
- Swarming technology;
- Self-assembling robots;
- Molecular robotics;
- Robot compliers; or
- Smart Dust.

7. Brain-computer interfaces, such as:

- Neural-controlled interfaces;
- Mind-machine interfaces;
- Direct neural interfaces; or
- Brain-machine interfaces

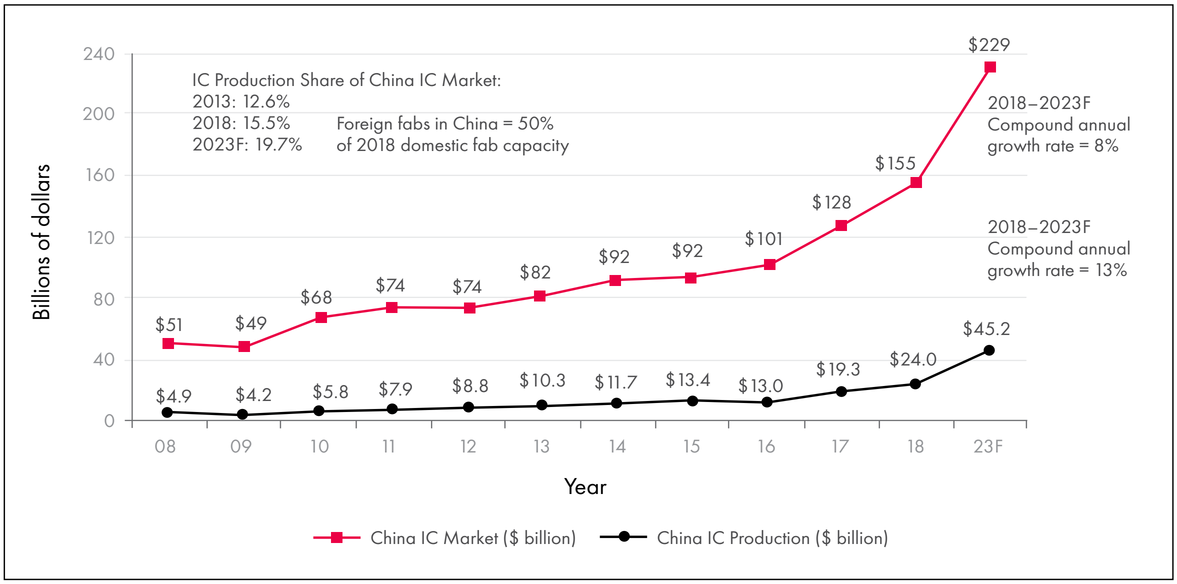
8. Advanced surveillance technologies, such as:

- Faceprint and voiceprint technologies

Source: Excerpts from a list, circulated by the US Commerce Department's BIS, published at www.globaltradeandsanctionslaw.com/u-s-commerce-department-releases-proposed-rulemaking-for-export-controls-on-emerging-technologies/#more-898.

Note: Out of the 14 "emerging technologies" identified by the US Commerce Department's BIS, eight technology families are relevant to the development of AI. Note that the US Commerce Department is expected to issue a separate rule making for proposed controls on "foundational technologies."

Appendix 3: China IC Market versus China IC Production Trends



Source: www.icinsights.com/news/bulletins/Can-We-Believe-The-Hype-About-Chinas-Domestic-IC-Production-Plans/. Reproduced with permission.

Works Cited

Allen, Gregory C. 2019. "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security." Center for a New American Security, February 6. www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

Andone, Dakin. 2020. "US travel restrictions go into effect to combat coronavirus spread." CNN, February 3. www.cnn.com/2020/02/02/us/coronavirus-us-travel-restrictions/index.html.

Bailey, Brian. 2018. "AI Market Ramps everywhere." *Semiconductor Engineering*, December 20.

———. 2019. "Power Is Limiting Machine Learning Deployments." *Semiconductor Engineering*, July 25. https://semiengineering.com/power-limitations-of-machine-learning/.

Baldwin, Richard E. 2014. "Multilateralising 21st Century Regionalism." Global Forum on Trade Reconciling Regionalism and Multilateralism in a Post-Bali World, Paris, France, OECD, February 11-12.

Barton, Dominic, Jonathan Woetzel, Jeongmin Seong and Qinzheng Tian. 2017. *Artificial Intelligence: Implications for China.* McKinsey Global Institute. April.

Batra, Gaurav, Zach Jacobson, Siddarth Madhav, Andrea Queirolo and Nick Santhanam. 2018. *Artificial-intelligence hardware: New opportunities for semiconductor companies.* McKinsey & Company. www.mckinsey.com/industries/semiconductors/our-insights/artificial-intelligence-hardware-new-opportunities-for-semiconductor-companies.

Beijing Innovation Center for Future Chips and Tsinghua University. 2018. *White Paper on AI Chip Technologies.* www.080910t.com/downloads/AI%20Chip%202018%20EN.pdf.

Bekkers, R. N. A., B. Verspagen and J. M. Smits. 2002. "Intellectual property rights and standardization. The case of GSM." *Telecommunications Policy* 26 (3-4): 171–88.

Blustein, Paul. 2019. *Schism: China, America and the Fracturing of the Global Trading System.* Waterloo, ON: CIGI.

Boeing. 2019. "Boeing Reports Record 2018 Results and Provides 2019 Guidance." Press Release, January 30. https://investors.boeing.com/investors/investor-news/press-release-details/2019/Boeing-Reports-Record-2018-Results-and-Provides-2019-Guidance/default.aspx.

Bonvillian, William B. 2018. "DARPA and its ARPA-E and IARPA clones: a unique innovation organization model." *Industrial and Corporate Change* 27 (5): 897–914.

Brandt, Loren and Thomas G. Rawski. 2019. "Policy, Regulation and Innovation in China's Electricity and Telecom Industries." https://brandt.economics.utoronto.ca/wp-content/uploads/2019/11/201910_Intro_Brandt-and-Rawski_2019.pdf.

Braverman, Burt. 2018. "Congress Enacts the Export Controls Act of 2018, Extending Controls to Emerging and Foundational Technologies." Davis Wright Tremaine LLP, September 26. www.dwt.com/insights/2018/09/congress-enacts-the-export-controls-act-of-2018-ex.

Broadman, Harry G. 2019. "As CFIUS Turns 45 Years Old, U.S. Regulation of Foreign Investment Is the Strictest Among Advanced Countries." *Forbes*, November 30. www.forbes.com/sites/harrybroadman/2019/11/30/as-cfius-turns-45-years-old-us-regulation-of-foreign-investment-is-the-strictest-among-advanced-countries/#5ac997495ecf.

Brynjolfsson, Erik, Daniel Rock and Chad Syverson. 2017. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics." Paper presented at NBER conference, Toronto, ON, September 13-14.

Castro, Daniel, Michael McLaughlin and Eline Chivot. 2019. "Who Is Winning the AI Race: China, the EU or the United States?" Washington, DC: Center for Data Innovation. www.datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/.

CB Insights. 2019. "The Increasingly Crowded AI Unicorn Club." February 14. www.cbinsights.com/research/ai-unicorn-club/.

CGTN. 2019. "China's public cloud IaaS market grows to $4.65 bln in 2018: report." CGTN, July 22. https://news.cgtn.com/news/2019-07-22/China-s-public-cloud-IaaS-market-grows-to-4-65-bln-in-2018-report-IwP8QYOGuA/index.html.

Chen, Monica and Rodney Chan. 2020. *Digitimes*, February 4.

CISTP. 2018. *China AI Development Report 2018.* Beijing, China: Tsinghua University.

Ciuriak, Dan. 2019a. "A Trade War Fuelled by Technology." CIGI Opinion, January 11. www.cigionline.org/articles/trade-war-fuelled-technology.

———. 2019b. "Economics of AI/ML and Big Data in the Data-Driven Economy: Implications for Canada's Innovation Strategy." May 17. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3362083.

———. 2019c. "Geoeconomic Disruption: A Comment on the Causes, Consequences and Policy Responses." Notes for the workshop The Shape of Geoeconomics to Come, Cato Institute, Washington, DC, March 4.

Cockburn, Iain M., Rebecca Henderson and Scott Stern. 2017. "The Impact of Artificial Intelligence on Innovation." Paper presented at NBER Toronto Conference, September 13-14.

Congressional Research Service. 2019. "CFIUS Reform: Foreign Investment National Security Reviews." October 4. https://fas.org/sgp/crs/natsec/IF10952.pdf.

Contreras, Jorge L. 2015. "A Brief History of FRAND: Analyzing Current Debates in Standard Setting and Antitrust Through a Historical Lens." *Antitrust Law Journal* (80).

Correa, Francesco D. 2017. *Artificial Intelligence and Exponential Technologies: Business Models Evolution and New Investment Opportunities.* Cham, Switzerland: Springer.

Crockett, Zachary. 2019. "How much money can you make on Amazon Mechanical Turk?" *The Hustle*, July 13. https://thehustle.co/making-money-on-amazon-mechanical-turk/.

DARPA. 2018. "DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies." September 7. www.darpa.mil/news-events/2018-09-07.

Delaney, Robert. 2018. "US slaps China's ZTE with 7-year components ban for breaching terms of sanctions settlement." *South China Morning Post*, April 16. www.scmp.com/business/companies/article/2142002/us-slaps-zte-seven-year-components-ban-breaching-terms-sanctions.

Demler, Mike. 2018. "Alibaba Looks to C-Sky for CPU IP." Linley Newsletter, August 21. www.linleygroup.com/newsletters/newsletter_detail.php?num=5901&year=2018&tag=3.

———. 2019. "Alibaba's Ouroboros is a Fast Talker." Linley Newsletter, August 29. www.linleygroup.com/newsletters/newsletter_detail.php?num=6058&year=2019&tag=3.

Deng, Iris. 2019. "Building China's own chip industry will be a costly 10-year marathon, former Intel China MD says." *South China Morning Post*, May 29. www.scmp.com/tech/science-research/article/3012140/building-chinas-own-chip-industry-will-be-costly-10-year.

Development Solutions for Europe Ltd. 2018. *China's "1+N" funding strategy for Artificial Intelligence.* December. http://chinainnovationfunding.eu/wp-content/uploads/2018/12/2-Chinas-1N-funding-strategy-for-Artificial-Intelligence.pdf.

Dieseldorff, Christian G. and Eugenia Liu. 2018. "SEMI FabView Update: More Investments in Semiconductor Fabs." *SEMI Blog*, September 20. https://blog.semi.org/business-markets/semi-fabview-update-more-investments-in-semiconductor-fabs.

Digitimes. 2019. "Highlights of the day: Huawei continues 5G deployments." *Digitimes*, June 11.

Ding, Jeffrey. 2019. Testimony before the US-China Economic and Security Review Commission Hearing on Technology, Trade, and Military-Civil Fusion: China's Pursuit of Artificial Intelligence, New Materials, and New Energy, 7 June, www.uscc.gov/sites/default/files/2019-10/June%207,%20 2019%20Hearing%20Transcript.pdf.

Dyson, George. 2012. *Turing's Cathedral: The Origins of the Digital Universe.* New York, NY: Vintage Books.

Editors. 1984. "Extraterritorial Application of United States Law: The Case of Export Controls." *University of Pennsylvania Law Review* 132. https://scholarship.law.upenn.edu/penn_law_review/vol132/iss2/4.

Erickson, Simon. 2019. "This Stock Could Become the Next NVIDIA." The Motley Fool, February 21. www.fool.com/investing/2019/02/21/stock-become-next-nvidia-xilinx-fpga.aspx.

Ernst, Dieter. 2009. *A New Geography of Knowledge in the Electronics Industry? Asia's Role in Global Innovation Networks.* Policy Studies No. 54. August. Honolulu, HI: East-West Center.

———. 2011. "China's Fragmented Innovation System." Paper presented at the MINERVA/IGCC Symposium "What is DARPA in Chinese? The Nature and Prospects for Radical Defense R&D and Innovation in China," June.

———. 2015. *From Catching Up to Forging Ahead: China's Policies for Semiconductors.* East-West Center Special Study, September. www.eastwestcenter.org/node/35320.

———. 2016. "China's Bold Strategy for Semiconductors—Zero-Sum Game or Catalyst for Cooperation?" East-West Center Working Papers: Innovation and Economic Growth Series, No. 9, September. www.eastwestcenter.org/node/35798.

———. 2017. *China's Standard-Essential Patents Challenge: From Latecomer to (Almost) Equal Player?* CIGI Special Report. Waterloo, ON: CIGI. www.cigionline.org/publications/chinas-standard-essential-patents-challenge-latecomer-almost-equal-player.

———. 2018a. "The Information Technology Agreement, Manufacturing, and Innovation — China's and India's Contrasting Experiences." In *Megaregionalism 2.0: Trade and Innovation within Global Networks*, edited by Dieter Ernst and Michael G. Plummer, 361–88. World Scientific Studies in International Economics. Vol. 67. Hackensack, NJ: World Scientific.

———. 2018b. "Beyond Value Capture — Exploring Innovation Gains from Global Networks." In *Megaregionalism 2.0: Trade and Innovation within Global Networks*, edited by Dieter Ernst and Michael G. Plummer, 55–89. World Scientific Studies in International Economics. Vol. 67. Hackensack, NJ: World Scientific

———. 2019. "Technology war — new challenges for the governance of standardization." Paper presented at the 4th International Symposium on Standardization and Governance, Hunan University, Changsha, China, November 14-15.

———. Forthcoming. "New Challenges for AI Standardization and Data Governance."

Fisher, Lawrence M. 2017. "Siri, Who Is Terry Winograd?" *strategy+business*, January 3. www.strategy-business.com/article/Siri-Who-Is-Terry-Winograd?gko=8513c.

Forbes. 2018. "The Rise In Computing Power: Why Ubiquitous Artificial Intelligence Is Now A Reality." July 17. www.forbes.com/sites/intelai/2018/07/17/the-rise-in-computing-power-why-ubiquitous-artificial-intelligence-is-now-a-reality/#69169efb1d3f.

Fuchs, Erica. 2009. "Cloning DARPA Successfully." *Issues in Science and Technology* 26 (1).

Gartner. 2019. "Gartner Says Worldwide IaaS Public Cloud Services Market Grew 31.3% in 2018." Press Release, July 29. www.gartner.com/en/newsroom/press-releases/2019-07-29-gartner-says-worldwide-iaas-public-cloud-services-market-grew-31point3-percent-in-2018.

Gavekal. 2019. "The US-China Trade and Technology Mess." Audio and transcript, May 29. https://research.gavekal.com/content/audio-transcript-%E2%80%94-us-china-trade-and-technology-mess.

Germano, Sara. 2019. "Huawei Strikes German 5G Deal Despite Political Pushback." *The Wall Street Journal*, December 11.

Gershgorn, Dave. 2017. "The data that transformed AI research — and possibly the world." *Quartz*, July 26. https://qz.com/1034972/the-data-that-changed-the-direction-of-airesearch-and-possibly-the-world/.

Gewirtz, Julian Baird. 2019. "The Roots of Xi Jinping's Ambition to 'Catch Up and Surpass.' *Foreign Policy*, August 27.

Gu, Shulin. 1999. *China's Industrial Technology: Market Reform and Organizational Change*. A United Nations University Study. London and New York: Routledge.

Gwennap, Linley. 2018. "China's AI Dream: CPU Transition Opens Doors for New Vendors." *Microprocessor Report*, April 30.

———. 2019. "Alibaba Takes AI Performance Lead." Linley Newsletter, November 28. www.linleygroup.com/newsletters/newsletter_detail.php?num=6096&year=2019&tag=3.

Hille, Kathrin, Ryan McMorrow and Qianer Liu. 2020. "Coronavirus shakes centre of world's tech supply chain." *Financial Times*, February 5.

Horowitz, Michael, Elsa B. Kania, Gregory C. Allen and Paul Scharre. 2018. "Strategic Competition in an Era of Artificial Intelligence." Center for a New American Security, July 25. www.cnas.org/publications/reports/strategic-competition-in-an-era-of-artificial-intelligence.

Howell, Thomas R., William A. Noellert, Janet H. MacLaughlin and Alan Wm. Wolff. 1988. *The Microelectronics Race: The Impact of Government Policy on International Competition.* Boulder, CO: Westview Press.

Hruska, Joel. 2019. "Chinese Foundry SMIC Begins 14nm Production." *Extreme Tech*, August 19. www.extremetech.com/computing/296802-chinese-foundry-smic-begins-14nm-production.

Huang, Yuanpu. 2018. "Cambricon Technologies, Potential Challenger Of Nvidia, Still Has A Long Way to Go." November 12. https://equalocean.com/ai/20181112-cambricon-technologies-still-has-a-long-way-to-go.

Huawei. 2019. "Huawei launches Ascend 910, the world's most powerful AI processor, and MindSpore, an all-scenario AI computing framework." Press Release, August 23. www.huawei.com/en/press-events/news/2019/8/huawei-ascend-910-most-powerful-ai-processor.

Hwang, Colley. 2019. *Asian Edge: On the Frontline of the ICT World.* Taipei, Tawain: *Digitimes*.

Hwang, Tim. 2018. "Computational Power and the Social Impact of Artificial Intelligence." March 23. https://arxiv.org/abs/1803.08971.

Kharpal, Arjun. 2019. "Huawei tops $100 billion revenue for first time despite political headwinds." CNBC, March 28. www.cnbc.com/2019/03/29/huawei-earnings-full-year-2018.html.

Kline, Ronald R. 2011. "Cybernetics, Automata Studies and the Dartmouth Conference on Artificial Intelligence." *IEEE Annals of the History of Computing*, October–December. IEEE Computer Society.

Knight, Will. 2016. "What Marvin Minsky Still Means for AI." *MIT Technology Review*, January 26. www.technologyreview.com/s/546116/what-marvin-minsky-still-means-for-ai/.

———. 2017. "China's AI Awakening." *MIT Technology Review*, October 10. www.technologyreview.com/s/609038/chinas-ai-awakening/.

Kraemer, Kenneth L., Greg Linden and Jason Dedrick. 2011. "Capturing Value in Global Networks: Apple's iPad and iPhone." July. http://economiadeservicos.com/wp-content/uploads/2017/04/value_ipad_iphone.pdf.

Kratsios, Michael. 2019. "The White House: Accelerating America's Leadership in Artificial Intelligence." Talk at the Center for a New American Security, February 28, www.cnas.org/publications/podcast/the-white-house-accelerating-americas-leadership-in-artificial-intelligence.

Kroeber, Arthur and Dan Wang. 2020 "A Tough Ask On Trade, Trouble Brewing On Tech." Gavekal Dragonomics, January 16. https://research.gavekal.com/article/tough-ask-trade-trouble-brewing-tech.

Kynge, James and Nic Fildes. 2020. "Huawei: the indispensable telecoms company." *Financial Times*, January 31.

Lamm, Ben. 2019. "Defense industry must take a lead on AI as tech firms waver." *The Hill,* January 19. https://thehill.com/opinion/national-security/426051-defense-industry-must-take-lead-on-ai-as-tech-firms-waiver.

Lapedus, Mark. 2019. "China's Foundry Biz Takes Big Leap Forward." Semiconductor Engineering, January 28. https://semiengineering.com/chinas-foundry-biz-takes-big-leap/.

Lardy, Nicholas. 2019. *The State Strikes Back:. The End of Economic Reform in China?* January. Washington, DC: Peterson Institute for International Economics.

Laskai, Lorand and Helen Toner. 2019. "Can China Grow Its Own AI Tech Base?" *New America*, November 4. www.newamerica.org/cybersecurity-initiative/digichina/blog/can-china-grow-its-own-ai-tech-base/.

Lee, Kai-Fu. 2018. *AI Superpowers: China, Silicon Valley, and the New World Order*. Boston, MA: Houghton Mifflin Harcourt.

Leonard, Jenny and Ian King. 2019. "Tech Shudders as U.S. Weighs New Limits on Huawei Sales." *Bloomberg*, December 18.

Li, Lauly and Ting-Fang Cheng. 2019. "US hits at Huawei innovation with blacklist of R&D centers." *Nikkei Asia Review*, August 21.

Lingzhi, Fan and Shan Jie. 2018. "79-year-old Chinese academic spearheads efforts to break US high-tech monopoly." *Global Times*, May 3.

Liu, Luffy. 2019. "China's 'Big Fund' Phase II Aims at IC Self-Sufficiency." *EETimes*, October 30.

Lucas, Louise. 2018. "Chinese AI chipmaker Horizon Robotics raises up to $1bn." *Financial Times*, November 26.

Ma, Kwan-Chen. 2019. "Happy New Year? Nvidia." Seeking Alpha, March 6. https://seekingalpha.com/article/4246760-happy-new-year-nvidia.

Marcus, Gary. 2018. "Deep Learning: A Critical Appraisal." https://arxiv.org/abs/1801.00631.

Markoff, John. 2012. "For Web Images, Creating New Technology to Seek and Find." *The New York Times*, November 19.

———. 2015. *Machines of Loving Grace: The Quest for Common Ground between Humans and Robots*. New York, NY: Harper Collins.

McCarthy, J., M. L. Minsky, N. Rochester and C. E. Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." August. http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf.

McGrath, Dylan and Barb Jorgenson. 2019. "Huawei Ban May Prove to Haunt U.S." *EET India*, June 11.

Medin, Milo and Gilman Louie. 2019. *The 5G Ecosystem: Risks & Opportunities for DoD*. Defense Innovation Board. April 3.

Merritt, Rick. 2018. "AI Flood Drives Chips to the Edge." *EETimes*, July 11.

———. 2019. "Huawei: Covering all the Bases in AI." *EETimes Asia*, August 22.

Nakazawa, Katsuji. 2019. "The new Long March — Xi's 15-year battle plan with the US." *Nikkei Asian Review*, May 30.

National Security Commission on Artificial Intelligence. 2019. *Interim Report*. November.

Nelson, Mei. 2019. "Debating China's AI Path: 'Alternative Routes,' or 'Overtaking on the Curve'?" *New America* (blog), November 12. www.newamerica.org/cybersecurity-initiative/digichina/blog/debating-chinas-ai-path-alternative-routes-or-overtaking-on-the-curve/.

Nenni, Daniel. 2019. "Can We Believe the Hype about China's Domestic IC Production Plans?" SemiWiki, June 13. https://semiwiki.com/forum/index.php?threads/can-we-believe-the-hype-about-china%E2%80%99s-domestic-ic-production-plans.11398/.

Newell, A., J. C. Shaw and H. A. Simon. 1958. "Elements of a theory of human problem solving." *Psychological Review* 65 (3): 151–66.

Newell, A. and H. A. Simon. 1976. "Computer Science as Empirical Inquiry: Symbols and Search." *Communications of the ACM* 19 (3): 113–26.

Ng, Andrew. 2016. "What Artificial Intelligence Can and Can't Do Right Now." *Harvard Business Review*, November 9. https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now.

Nilsson, Nils. 2010. *The Quest for Artificial Intelligence: A History of Ideas and Achievements.* Cambridge, UK: Cambridge University Press.

Norvig, Peter. 2003. *Artificial Intelligence: A Modern Approach*. 2nd ed. Upper Saddle River, NJ: Prentice Hall.

OECD. 2019a. *Artificial Intelligence in Society*. Paris, France: OECD.

———. 2019b. "Measuring distortions in international markets: The semiconductor value chain." OECD Trade Policy Papers No. 234. December 12. Paris, France: OECD Publishing http://dx.doi.org/10.1787/8fe4491d-en.

Okoshi, Yuki. 2019. "China overtakes US in AI patent rankings." *Nikkei Asia Review*, March 10.

O'Meara, Sarah. 2019. "AI researchers in China want to keep the global-sharing culture alive." *Nature*, May 29.

Patterson, Alan. 2019. "Huawei Catches a Break with TSMC." *EETimes Asia*, May 28.

Pei, Minxin. 2019. "No Easy Road for US in Decoupling." *Foreign Policy*, December 6.

Pohlmann, Tim and Knut Blind. 2002. *Landscaping Study on Standard Essential Pantes (SEPs).* Berlin, Germany: IPlytics. www.iplytics.com/wp-content/uploads/2017/04/Pohlmann_IPlytics_2017_EU-report_landscaping-SEPs.pdf.

Potember, Richard. 2017. *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD.* JSR-16-Task-003, January. https://fas.org/irp/agency/dod/jason/ai-dod.pdf.

Press, Gil. 2017. "6 Reasons Why China Will Lead in AI." *Forbes*, November 5.

PwC. 2019. "Opportunities for the global semiconductor market: Growing market share by embracing AI." April 3. www.pwc.com/gx/en/industries/tmt/publications/assets/pwc-semiconductor-report-2019.pdf.

Qu, Tracy. 2019. "Can China catch up in chip design?" Tech in Asia, September 9. www.techinasia.com/china-catch-up-chip-design.

Ren, Shuli. 2019. "Low R&D Spending May Be China's Achilles' Heel." Bloomberg, August 12.

Reuters. 2019. "Chinese AI chip maker Horizon Robotics raises $600 million from SK Hynix, others." Reuters, February 27.

Reuther, Albert, Peter Michaleas, Michael Jones, Vijay Gadepally, Siddharth Samsi and Jeremy Kepner. 2019. "Survey and Benchmarking of Machine Learning Accelerators." arXiv, August 29. https://arxiv.org/abs/1908.11348.

Romer, Paul M. 1994. "The Origins of Endogenous Growth." *Journal of Economic Perspectives* 8 (1).

Rotenberg, Ziv. 2016. "Where next for university IP commercialisation in China?" *Intellectual Asset Management* (79).

Rowen, Chris. 2017. "Deep Learning Startups in China: Report from the Leading Edge." *Cognite Blog*, July 7. www.cogniteventures.com/2017/07/07/deep-learning-startups-in-china-report-from-the-leading-edge/.

Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. 2015. "ImageNet Large Scale Visual Recognition Challenge." arXiv, January 30. https://arxiv.org/pdf/1409.0575.pdf.

Sacks, Samm. 2019. Testimony before the House Foreign Affairs Committee during a hearing on "Smart Competition: Adapting U.S. Strategy Toward China at 40 Years." *New America*, May 8. www.newamerica.org/cybersecurity-initiative/digichina/blog/samm-sacks-testifies-house-foreign-affairs-committee-smart-competition-china/.

Schulte, Drew J. 2019. "Keys to Successful AI Patents in the U.S. and Europe." Haley Guiliano, March 11. www.hglaw.com/news-insights/successful-ai-patents-us-europe/.

SEMI. 2019. *China IC Ecosystem Report, 2019 Edition.* September 10. www.semi.org/en/news-resources/market-data/china-ic-ecosystem.

Semiconductor Industry Association. 2019. *2019 SIA Databook*. Washington, DC.

Shambaugh, David. 2019. "U.S.-China Decoupling: How Feasible, How Desirable?" December 6. www.chinausfocus.com/foreign-policy/us-china-decoupling-how-feasible-how-desirable.

Shen, Jessie. 2019. "Baidu, Samsung team up for AI chips." *Digitimes*, December 19.

———. 2020. "SMIC scaling up 14nm chip output." *Digitimes*, January 16.

Shen, Jill. 2019. "Huawei's new venture capital firm highlights aim to create an ecosystem." TechNode, April 26. https://technode.com/2019/04/26/huawei-new-vc-hubble/.

Shilov, Anton. "SMIC Begins Volume Production of 14 nm FinFET Chips: China's First FinFET Line." AnandTech, November 14. www.anandtech.com/show/15105/smic-begins-volume-production-of-14-nm-finfet-chips-chinas-first-finfet-line.

Sperling, Ed. 2018. "Big Changes for Mainstream Chip Architectures." Semiconductor Engineering, August 23. https://semiengineering.com/big-changes-for-mainstream-chip-architectures/.

Stacey, Kiran. 2019. "US tech groups rebuff Trump's new anti-Huawei push." *Financial Times*, December 16.

Stevenson-Yang, Anne. 2019. "What does it mean to shun Huawei? The Silicon wars" *JCAP China Primary Insight*, June 3.

Thibodeau, Patrick. 2017. "NSA, DOE say China's supercomputing advances put U.S. at risk." *Computerworld*, March 15. www.computerworld.com/article/3180984/spy-agency-doe-see-china-nearing-supercomputing-leadership.html.

Tian, Robert. 2018. "Suiyuan Technology Gets $50M to Build Cloud-Based Deep Learning Chips." Synced, August 7. https://syncedreview.com/2018/08/07/suiyuan-technology-gets-50m-to-build-cloud-based-deep-learning-chips/.

Toh, Michelle and Charles Riley. 2020. "These airlines have suspended flights to and from China." CNN, February 6. www.cnn.com/2020/02/05/business/china-flights-travel-coronavirus-outbreak/index.html.

US Department of the Treasury. 2019. "Fact Sheet: Proposed CFIUS Regulations to Implement FIRRMA." Office of Public Affairs, September 17. https://home.treasury.gov/system/files/206/Proposed-FIRRMA-Regulations-FACT-SHEET.pdf.

Wang, Dan. 2020. "Huawei Comes Out Ahead, For Now." Gavekal Dragonomics, January 30.

Waters, Richard. 2019. "World's biggest chip created to meet demands of AI." *Financial Times*, August 19.

Winograd, Terry and Fernando Flores. 1986. *Understanding Computers and Cognition: A New Foundation for Design.* Norwood, NJ: Ablex Publishing Corporation.

Wolff, Alan Wm. 2011. "China's Indigenous Innovation Policy." Testimony before the US- China Economic and Security Review Commission Hearing on China's Intellectual Property Rights and Indigenous Innovation Policy. Washington, DC, May 4.

WIPO. 2019. *Technology Trends 2019: Artificial Intelligence.* Geneva, Switzerland: WIPO.

Wu, Mark. 2016. "The China, Inc. Challenge to Global Trade Governance." *Harvard International Law Journal* 57 (2).

Yeung, Karen. 2019 "China nudges top tech firms from Tencent to Alibaba to give new life to struggling state giants." *South China Morning Post*, August 4.

Yu, Eileen. 2019. "China top internet firms pump 45% more into R&D." ZDNet, August 18. www.zdnet.com/article/china-top-internet-firms-pump-45-more-into-r-d/.

Zekun, Wang. 2019. "Top 30 Players of Chinese Artificial Intelligence Chip Industry." Equal Ocean, April 8.

Zhang, Sarah. 2017. "China's Artificial-Intelligence Boom." *The Atlantic*, February 16. www.theatlantic.com/technology/archive/2017/02/china-artificial-intelligence/516615/.

Zhu, Jun, Tiejun Huang, Wenguang Chen and Wen Gao. 2018. "The Future of Artificial Intelligence in China." *Communications of the ACM* 61 (11): 44-45. https://cacm.acm.org/magazines/2018/11/232209-the-future-of-artificial-intelligence-in-china/fulltext.

Zou, Julian. 2019. "Finally, Some Good News from CFIUS: OmniVision." *Global Trade & Sanctions Law* (blog), April 19. www.globaltradeandsanctionslaw.com/finally-some-good-news-from-cfius-omnivision/.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York, NY: PublicAffairs.

Centre for International
Governance Innovation

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org

RECYCLED
Paper made from
recycled material
FSC® C023070