GOVERNMENT PROCEDURES AND OPERATIONS

National Security and Artificial Intelligence

James A. Greene

SNOVA

Government Procedures and Operations

snova
New York

Government Procedures and Operations

COVID-19 in Nursing Homes: Impact, Transmission and Prevention
Fred C. Wallace (Editor)
2023. ISBN: 979-8-88697-673-1 (Hardcover)
2023. ISBN: 979-8-88697-683-0 (eBook)

The Significance of the COVID Pandemic in Nursing Homes
John S. Gill (Editor)
2023. ISBN: 979-8-88697-660-1 (Hardcover)
2023. ISBN: 979-8-88697-682-3 (eBook)

Putting an End to Surprise Medical Billing
Danial E. Tackett (Editor)
2022. ISBN: 979-8-88697-296-2 (Hardcover)
2022. ISBN: 979-8-88697-327-3 (eBook)

Abortion Rights, Access, and Legislative Response
Jorge P. Sandford (Editor)
2022. ISBN: 979-8-88697-258-0 (Softcover)
2022. ISBN: 979-8-88697-294-8 (eBook)

Review of Capitol Police Procedures During the Capitol Attack
Rafael B. Phillips (Editor)
2022. ISBN: 979-8-88697-260-3 (Hardcover)
2022. ISBN: 978-1-53617-478-6 (eBook)

More information about this series can be found at https://novapublishers.com/product-category/series/government-procedures-and-operations/

James A. Greene
Editor

National Security and Artificial Intelligence

snova
New York

NOTICE TO THE READER

The Publisher has taken reasonable care in the preparation of this book but makes no expressed or implied warranty of any kind and assumes no responsibility for any errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of information contained in this book. The Publisher shall not be liable for any special, consequential, or exemplary damages resulting, in whole or in part, from the readers' use of, or reliance upon, this material. Any parts of this book based on government reports are so indicated and copyright is claimed for those parts to the extent applicable to compilations of such works.

Independent verification should be sought for any data, advice or recommendations contained in this book. In addition, no responsibility is assumed by the Publisher for any injury and/or damage to persons or property arising from any methods, products, instructions, ideas or otherwise contained in this publication.

This publication is designed to provide accurate and authoritative information with regards to the subject matter covered herein. It is sold with the clear understanding that the Publisher is not engaged in rendering legal or any other professional services. If legal or any other expert assistance is required, the services of a competent person should be sought. FROM A DECLARATION OF PARTICIPANTS JOINTLY ADOPTED BY A COMMITTEE OF THE AMERICAN BAR ASSOCIATION AND A COMMITTEE OF PUBLISHERS.

Library of Congress Cataloging-in-Publication Data

ISBN: 979-8-88697-994-7(eBook)

Published by Nova Science Publishers, Inc. † New York

Contents

Preface

Artificial intelligence (AI) is a rapidly growing field of technology with potentially significant implications for national security. As such, the United States and other nations are developing AI applications for a range of military functions.

Chapter 1

Artificial Intelligence, China, Russia, and the Global Order[*]

Nicholas D. Wright

Preface: US Perspective

Given the wide-ranging implications for global competition, domestic political systems and daily life, US policymakers must prepare for the impacts of new artificial intelligence (AI)-related technologies. Anticipating AI's impacts on the global order requires US policy makers' awareness of certain key aspects of the AI-related technologies—and how those technologies will interact with the rapidly changing global system of human societies. One area that has received little in-depth examination to date is how AI-related technologies could affect countries' domestic political systems—whether authoritarian, liberal democratic, or a hybrid of the two—and how they might impact global competition between different regimes.

This work highlights several key areas where AI-related technologies have clear implications for globally integrated strategic planning and requirements development:

- Since 2012, new AI-related technologies have entered the real world with rapidly accelerating scale and speed. While the character of these technologies currently favors enhanced surveillance, it is limited by a need for extensive human involvement and the preparation of big-data platforms. This will likely dominate current efforts to incorporate AI into social governance, as we see now in China.
- AI may help enable a plausible competitor to liberal democracy allowing large and industrially sophisticated states to make their citizens rich while maintaining rigid control. China is now building core components of such a system of digital authoritarianism. Such systems are already being emulated in a global competition with liberal democracy.
- Russia has a different political regime than China. The Russian model is a hybrid that relies on a mix of less overt and often nontechnical mechanisms to manipulate online information flows. Competition for influence between digital liberal democracy and

[*] This is an edited, reformatted and augmented version of *Artificial Intelligence, China, Russia, and the Global Order*, a publication of the United States Airforce, dated October 8, 2019.

In: National Security and Artificial Intelligence
Editor: James A. Greene
ISBN: 979-8-88697-871-1

more authoritarian digital regimes will occur at many levels: international institutions (and norms), nation states, and corporations. The United States must adopt a multifaceted approach to influence with allies and crucial swing states. It must also carefully prevent unwanted escalation of this competition—as a number of contributors argue in this work, insecurity drives much of Chinese and Russian decision making.

- China's foreign policy decision making will not necessarily become more expansionist if its domestic regime becomes more authoritarian. Mapping out AI's effects on foreign policy choices requires mapping them out within the domestic ecosystem and content from which those choices emanate.
- Military dimensions of global competition will change with AI. Hackers become more prominent, and new crisis escalation risks emerge. Chinese domestic social governance systems that become ever more reliant on vast digital systems will be tempting targets for adversaries—a fact likely to prompt Chinese regime insecurity that may feed a spiraling security dilemma.

The emerging digital liberal democracy in the United States, digital hybrid regime in Russia, and digital authoritarian regime in China will each exert influences far beyond their physical borders. This competition for influence will likely prove a defining feature of the twenty-first-century global system. We must not be caught by surprise.

ALEXUS G. GRYNKEWICH
Major General

Preface: UK Perspective

In the 1990s, there was talk of a revolution in military affairs (RMA) resulting from the combination of improved sensors, digital communications, and precision-guided munitions. In retrospect this was both more and less of a revolution than supposed at the time. It was less of a revolution because the drivers of military conflict were not technological but lay in broader social, economic, and political factors. The new technologies made possible military operations that ran at a faster tempo and used weapons of greater lethality that allowed for greater discrimination. Military power could be directed against vital targets to achieve the optimum effects. It was soon discovered that enemies could limit the advantages these capabilities gave the United States and its allies by adopting guerrilla strategies based on ambushes and terrorism. However well suited they might be to fights between regular armies, their limitations became evident in struggles over "hearts and minds."

Yet, it was also more of a revolution than really understood in the 1990s. The RMA was then assumed to represent an advanced stage in a line of technological development that could be traced back to the 1960s when Gordon Moore first observed that the number of components per integrated circuit would double every two years. Yet, as we can now see, it was really only an interim stage. Over the past two decades, we have seen the arrival of smart phones putting data sets, imagery, navigation, and forms of communication into the hands of individuals that were once only specialist military tools. Forms of international connectivity have created new opportunities for productive and benign activities but also for mischief and malign influences. The kinetic aspects of conflict have now been joined by nonkinetic forms of struggle, including cyberattacks and information campaigns. These have moved the arena of conflict away from the field of battle to the essentials of everyday life and the state of public opinion.

Artificial intelligence (AI) now points to the next stage. The ability to gather data and interrogate it with scant human engagement now starts to set tests for whole societies: regarding the efficient exploitation of scarce resources on the one hand, and the ability of individuals to live free and fulfilled lives on the other. As this volume makes clear, the government of China is now embarking on a vast experiment in social control that aims to use AI to ensure that individuals are following the party line and rewards or punishes them according to how well they behave. Russia does not have the capacity or the political structures capable of following this example, though it has been a pacesetter in the use of cyber and information operations to undermine its foes (without actually starting a war).

It is worth recalling that the Cold War was decided not by force of arms but because the Soviet system imploded, having failed to deliver for its people and having lost legitimacy as a result of its repressive methods. The military balance of the time, and in particular the fear of nuclear war, maintained a stalemate so that instead of a hot war there was intense ideological competition. Liberal democracy posed a threat to authoritarian systems because it was seen to be better able to meet human needs, including free expression. But during the Cold War, the United States and its allies always led the ideological competition and over time demonstrated with relative ease the superiority of their political systems. As before, there are formidable reasons for both sides to avoid pushing any contest to open hostilities. This means that there is now a different form of ideological competition. This time it will be tougher, because China has invested heavily in the technologies of social control, and in particular in AI, while liberal democracy has lost some of its luster in unpopular wars and financial crises. The West has yet to work out how to cope with so much personal data being stored and analyzed by both private and state organizations. But liberal democracies must somehow demonstrate that it is possible to take advantage of the new technologies without losing sight of their core values.

Another difference from the Cold War is that China's economy depends on trade with the rest of the world. It has recently started to be viewed as an unreliable partner, for example by getting its technology into the critical systems of Western countries. This issue has acquired more salience because of growing concern over rather old-fashioned geopolitical issues, as China pushes to turn itself into the dominant regional power in the Asia-Pacific region. This takes us back to the question of how much the new technologies have influenced classical forms of military conflict. The answer will depend on how well AI is integrated into command systems, as well as the ability to disrupt enemy systems. In the new era of AI, when humans might be perplexed by what is going on in the machines on which they must depend, the strategies of disruption and disorientation that have been prominently in play in international affairs in recent years could well move to new levels and become more central than before to the conduct of conflict.

It is unwise to try to predict the future just by following trends or assuming that the structures of international economics and politics will continue to follow familiar patterns. The US network of alliances, for example, is currently under a lot of pressure. Anticipating the likely path of technological development may therefore be far less difficult than grasping the forms of its interaction with a changing context. The future is unpredictable because itwill be shaped by choices between options that are currently barely understood. The great value of this work is that it describes some of the big issues coming our way and urges us to stretch our imaginations when thinking about the challenges that will need to be faced.

SIR LAWRENCE FREEDMAN

Acronyms

AI	Artificial intelligence
AIDP	Artificial intelligence development plan
AR	Augmented reality
ARF	Advanced Research Foundation
BMI	Brain-machine interface
BRI	Belt Road Initiative
BRICS	Brazil, Russia, India, China, and South Africa
CCP	Chinese Communist Party
CECIC	China National Electronics Import & Export Corporation
CIA	Central Intelligence Agency
CMC JSD	Central Military Commission Joint Staff Department
CORA	Cyber Operational Resilience Alliance
DARPA	Defense Advanced Research Projects Agency
DOD	Department of Defense
EEG	Electroencephalogram
EU	European Union
FAANG	Facebook, Apple, Amazon, Netflix, and Google
GDP	Gross domestic product
ICT	information and communication technology
IoT	Internet of Things
ISP	Internet service provider
IT	Information technology
MIIT	Ministry of Industry and Information Technology
ML	Machine learning
MLP	Medium- and long-term plan
MOD	Ministry of Defense
NDRC	National Development and Reform Commission
OBOR	One Belt, One Road
PLA	People’s Liberation Army
PRC	People’s Republic of China
R&D	Research and development
RMA	Revolution in military affairs
S&T	Science and technology
SA	Situational awareness
SORM	System for Operative Investigative Activities
STEM	Science, technology, engineering, and mathematics

STES	Socio-technical-economic system
UNGA	UN General Assembly
VR	Virtual reality

Acknowledgments

The Editor would like to thank Mariah Yager for her hard work and insightful comments that helped make the project a reality. Sincerest thanks also to Hriar 'Doc' Cabayan, the Director of the Strategic Multilayer Assessment Branch at the Joint Staff (J-39), whose intellectual curiosity, rigour and dedication to the real world helped shape this chapter.

Introduction and Overview

Artificial intelligence (AI) and big data promise to help reshape the global order. For decades, most political observers believed that liberal democracy offered the only plausible future pathways for big, industrially sophisticated countries to make their citizens rich. Now, by allowing governments to monitor, understand, and control their citizens far more effectively than ever before, AI offers a plausible way for big, economically advanced countries to make their citizens rich while maintaining control over them—the first since the end of the Cold War. That may help fuel and shape renewed international competition between types of political regimes that are all becoming more "digital." Just as competition between liberal democratic, fascist, and communist social systems defined much of the twentieth century, how may the struggle between digital liberal democracy and digital authoritarianism define and shape the twenty-first?

The technical nature of AI's new advances particularly well suits all-encompassing surveillance and, as a consequence, authoritarianism. New forms of authoritarianism arose with previous waves of global authoritarian expansion: fascism in the 1920s or bureaucratic authoritarianism in the 1960s. China has begun constructing core components of a digital authoritarian state. America's liberal democratic political regime is turning digital, and so too is Russia's hybrid political regime that lies between democracy and authoritarianism.

Swing states from Asia to Africa, Europe, and Latin America must manage their own political regimes within the context of this global competition. Several like-minded countries have begun to buy or emulate Chinese systems. Russian techniques are diffusing. To be sure, competing models for domestic regimes must be seen within the broader strategic context—relative military or economic power also matter deeply—but, as in the twentieth century, it will likely prove a crucial dimension.

This work focuses on the emerging Chinese and Russian models and how they will interact with the global order. We bring together deep expertise on China, Russia, strategy, and technology—as well as artists to provide illuminating sidelights.

The key recommendation is that US policy makers must understand the potential for the new AI-related to technologies to affect domestic political regimes (authoritarian, hybrid, and democratic) that will compete for influence in the global order.

We recommend policy makers use the following three-pronged strategy to understand the challenge and develop global policy:

- US democracy must be kept robust as it adapts to these new technologies. It must respond to domestic threats (e.g., capture by a tech oligopoly or drift to a surveillance state) and external threats without becoming governed by a military–industrial complex. US digital democracy, if successful at home, will exert gravitational influence globally.
- The United States must exert influence effectively and manage potential escalation in the swing states (e.g., in Asia or Europe) and global systems (e.g., norms and institutions) that form the key terrain for competition among the digital regime types. Diplomatic, economic, informational, and commercial dimensions will be crucial, with allies and other states.
- The United States should push back on the digital authoritarian and digital hybrid heartlands but do so in ways that manage the significant risks of spiraling fear and animosity.

Overview of the Chapter

In the remainder of this Introduction we provide an overview for each of these six sections. We bring together leading experts on China, Russia, strategy, and artificial intelligence (AI), as well as artists.

- **Part I** examines the AI-related technologies and their implications for the global order. It provides a framework that describes how the technologies' effects on domestic political regimes may affect the global order. This helps structure the diverse contributions below.
- **Part II** describes specific aspects of the Chinese and Russian regimes in more detail.
- **Part III** examines specific aspects of the export and emulation of the Russian and Chinese models within a global competition for influence.
- **Part IV** explores how AI's potential implications for the Chinese domestic political regime may affect its foreign policy decision making.
- **Part V** examines specific military dimensions of AI, including in the Chinese and Russian contexts.
- **Part VI** takes a very different approach and provides thought-provoking new viewpoints from artists and perspectives from the humanities.

Part I. Artificial Intelligence, Domestic Political Regimes, and the Global Order

In Part I, Nicholas Wright provides an overarching analysis and framework, going all the way from the specific technical characteristics of the new technologies through to the global order.

Section 1 examines the artificial intelligence (AI)–related technologies and asks: what specifically is new? By *artificial intelligence* here we mean a constellation of new technologies: AI itself more narrowly defined (essentially giving computers behaviors that would be thought intelligent in humans), big data, machine learning, and digital things (e.g., the "Internet of

Things"). This constellation is bringing in a new technological epoch. Following a leap in AI research around 2012, we now have: *Automated systems learning directly from data to do tasks that are complicated*. The key leap is that AI's can now do much more-complicated tasks (e.g., AI can now do good facial recognition). Crucially, AI has particularly improved for tasks related to "perception"—e.g., perceiving images or speech or some kinds of patterns in big data—and these are the advances now being rapidly rolled out across diverse real-world uses.

Section 2 considers AI's bewildering profusion of implications for the global order and breaks them down into three more manageable bites. This work primarily focuses on the first area, which has received by far the least attention.

1) The first is how this new technology's potential impacts on *domestic political regimes* (e.g., authoritarian, hybrid, or liberal democratic) may affect competition among them in the *world order*. AI will help enable a plausible competitor to liberal democracy for big industrially sophisticated states to make their citizens rich and maintain rigid control: *digital authoritarianism*. China is building core components of such a system—which are already being exported and emulated in a global competition with liberal democracy.
2) An "*nth industrial revolution*": AI will radically change the means of production across economic and societal sectors, e.g., transport, healthcare, or the military.
3) The "singularity" and the sense of self: In the *singularity*, exponentially accelerating technological progress creates an AI that exceeds human intelligence and escapes our control, potentially destroying humanity or disrupting humans' conceptions of themselves.

Section 3 examines AI and domestic political regimes in more detail, and introduces three crucial cases: China, Russia, and the United States. A *domestic political regime* is a system of social organization that includes not only government and the institutions of the state but also the structures and processes by which these interact with broader society. Three broad types dominate globally today: authoritarian (e.g., China), liberal democratic (e.g., the United States), and hybrid regimes that fall somewhere in between (e.g., Russia). New variants of these regime types emerge in response to changing times. For instance, historically new forms of authoritarianism emerged in the 1920s (fascism) and 1960s (bureaucratic authoritarianism). We arguably now see "digital" variants of each regime type emerging: digital authoritarianism (e.g., China), digital hybrid regimes (e.g., Russia), and digital liberal democracies (e.g., the United States). However, the character of the new AI-related technologies (i.e., enhanced perception) best suits the augmentation of the surveillance, filtering, and prediction in digital authoritarianism, making that perhaps the largest departure of the three.

Section 4 discusses global competition and, in particular, the export and emulation of these alternative models for influence over swing states—as occurred in the twentieth century among liberal democratic, fascist, and communist regime types. The global competition for influence occurs through active promotion; export of control and surveillance systems, competition between Chinese and US tech titans, as well as battles over global norms and institutions. Swing states across Europe, Africa, Asia, and elsewhere are highly heterogeneous, and even within states, the elites and populations may disagree over the models' relative merits. Of course, the attractiveness or otherwise of the competing models is just one factor in the broader strategic context, as was the case between competing twentieth-century regime types. Finally, we also examine two further ways the AI-related technologies may affect global competition: firstly, how AI's potential impacts on domestic political regimes may affect foreign policy decision making, and second, military dimensions.

Part II. Digital Authoritarianism: Evolving Chinese and Russian Models

In Section 5, Jeffrey Ding provides an overview of China's artificial intelligence (AI) strategy. He first places it in the context of past science and technology plans, which helps analyze China's most important current policies and initiatives to further its AI-related industries. Next, he outlines how AI development intersects with multiple areas of China's national interests—and in particular its domestic social governance. He concludes by discussing the main barriers to China realizing its AI dream.

In Section 6, Samantha Hoffman describes how understanding developments in China‘s technology-enhanced authoritarianism requires placing them in context of the Chinese Communist Party's political control process known as "social management." The modern "grid management" system, the "Skynet" surveillance project, and "social credit system" are all conceptually linked to long-existing Leninist control processes.

In Section 7, Shazeda Ahmed describes the Chinese "credit city," in which local governments and tech companies share their data with one another to determine the degree of individuals' and businesses' trustworthiness. The value judgments that come out of assessing these data—in some instances, a numeric score or a verbal rating—become a basis for determining the benefits that a person or company can unlock. However, her research on the ground reveals the huge technical and administrative challenges that have yet to be overcome.

In Section 8, Jaclyn Kerr describes how Russia's innovative and experimental approach to information manipulation and control differs significantly from the more-often discussed Chinese "Great Firewall" system, as well as other approaches that emphasize systemic technical censorship. The Russian model relies on a mix of less overt, and often nontechnical, mechanisms to manipulate online information flows, narratives, and framings to shape public opinion without resort to universal censorship. This model for the domestic control of information not only fits Russia's own political system but also is likely to prove more resonant and easier to emulate in many other countries.

Part III. Export and Emulation of the Models in Global Competition

In Section 9, Valentin Weber provides a more granular view of how the Chinese model is being exported—by the government, state-owned companies, and private companies that make up China's security–industrial complex. This export has been successful in Africa, Asia, the Middle East, and South America. If the United States wants to maintain a strategic advantage in regions where China's construction of internet infrastructure and the installation of filtering/surveillance technology challenges America, then US policy makers require a global view of the underlying agents that drive exports. This will allow the United States to tailor policies that counter the diffusion of information controls.

In Section 10, Laura Steckman describes China's dual-pronged strategy to become the world's technology leader for AI. Its two primary pathways are: (1) establishing partnerships with nations, organizations, and other entities that demonstrate AI talent and (2) globally exporting its domestically developed AI-related technologies. These approaches raise questions for countries with different political and social structures or that remain wary of using these technologies to shape societies in ways that contradict national values and norms or, more profoundly, to assert control through mechanisms of digital authoritarianism.

In Section 11, Robert Morgus details the spread of Russia's model. Pervasive communications collection, absent oversight, and government cooption of industry—particularly internet service providers—to do their bidding characterizes Russian digital authoritarianism. Russia's digital authoritarianism is neither as well defined nor as technologically robust or reliant on AI as the Chinese model. The Russian government exports or encourages emulation its model of digital authoritarianism globally and in their near abroad through diplomatic, informational, and economic means.

In Section 12, James Lewis takes a skeptical look at ideas of AI and China's unstoppable rise. Judging any Chinese digital authoritarian model's potential attractiveness requires viewing it in strategic context—not only in the context of a more comprehensive view of what drives influence in the global system but also in the context of how such influence compares to that of China's major competitor: the United States. He outlines five factors that will limit the Chinese model's impact. Although AI ripped from its strategic context can seem powerful or even frightening, given strategic competence the United States will remain superior to China.

In Section 13, Chris Demchak posits that as AI-related technologies rise in criticality for the nations' future economic and political wellbeing, China now has the advantage in three of the four "horsemen" of AI conflict (scale, foreknowledge, and strategic coherence), leaving only a fourth (speed) to the Western democratic societies. To counter China's AI advantages, democratic societies need a new narrative that places their future as minority states in the global order who seek long-term survival—and also novel but practical organizational architecture to implement that vision. Militaries must also change, preparing to "fight" a constant war in AI-led military operations while collectively embedded in the community of democratic states.

Part IV. Artificial Intelligence and Domestic Impacts on China's Foreign Policy Decision Making

In Section 14, Benjamin Chang asks: How will domestic use of AI affect Chinese foreign policy, particularly with respect to US–China relations? Drawing on relevant threads of political science, he discusses two possible consequences: (1) significantly worsened US–China relations due to increased ideological friction and opacity and (2) increased Chinese assertiveness due to increased confidence and a smaller "winning coalition." Finally, he assesses implications for US policy.

In Section 15, Kacie Miura discusses the implications of increased internal control on China's international behavior. Although a small group of top leaders dictates foreign policy making in China, several key domestic factors constrain and complicate China's international behavior. These include regime insecurity, public opinion, factional competition, and bureaucratic discord. AI—if it improves the Chinese leadership's ability to monitor and control societal and elite actors—could presumably reduce the influence of these internal drivers of China's international behavior. This will allow China's leaders to more efficiently advance their aspirations for China's position in the world, regardless of whether they choose to do so through confrontational or cooperative foreign policies.

In Section 16, Rachel Esplin Odell explores the crucial interrelationship among China's regime insecurity, domestic authoritarianism, and foreign policy. Too often, Western narratives fail to perceive that a deep-seated insecurity about its ability to maintain power while reforming its economy drives the Chinese Communist Party's authoritarianism. Moreover, Western observers falsely assume China’s domestic authoritarianism infuses its international ambitions, leading China to challenge the existing liberal international order. Instead, Western governments recognized China’s foreign policy behaviors as largely status quo-supporting efforts to foster economic growth, they could craft more effective, positive-sum policies in response.

In Section 17, Rogier Creemers examines the international and foreign policy impact of China’s AI and big-data strategies. In the past few years, China has embarked upon an ambitious strategy to build up its capabilities in AI and big data. The primary aims for this agenda are domestic: transforming the government’s social management and governance abilities and creating new areas for economic growth. Nonetheless, this agenda also has an international impact, both in terms of foreign governments’ responses to China’s domestic strategy and the extent to which Chinese technologies are exported or become part of global cyber processes. This chapter reviews the development of this agenda and assesses its impact for China’s foreign policy.

Part V. Artificial Intelligence and Military Dimensions in International Competition

In Section 18, Martin Libicki argues artificial intelligence (AI) will change the character of warfare by making hacking more important and by changing hacking. Computer hacking may be understood as the search for vulnerabilities in opposing systems whose exploitation permit leverage: small efforts have great effect. Injecting AI into systems systematizes the hackers’ search for vulnerabilities. Moreover, AI also multiplies vulnerabilities. Systems can be trained on a corpus of expected environments, but if the other side generates edge cases that the defender failed to imagine, the receiver’s AI may exhibit behavior favorable to the hacker. In sum, as AI becomes more important, searching for such vulnerabilities will likely constitute a growing share of military activity.

In Section 19, Herbert Lin examines the risks of conflict escalation from AI-enabled military systems. He describes how AI may feed deliberate, inadvertent, accidental, or catalytic escalation. Today’s AI—in particular, machine learning—poses particular risks because the internal workings of all but the simplest machine-learning systems are for all practical purposes impossible for human beings to understand. It is thus easy for human users to ask such systems to perform outside the envelope of the data with which they were trained and for the user to receive no notification that the system is indeed being asked to perform in such a manner.

In Section 20, Elsa Kania examines AI in future Chinese command decision making. The Chinese People’s Liberation Army (PLA) is exploring the use of AI technologies to enhance future command decision making. In particular, the PLA seeks to overcome admitted deficiencies in its commanders’ capabilities and to leverage these technologies to achieve decision superiority in future “intelligentized” (智能化) warfare. Chinese military experts have examined the DARPA program Deep Green and are inspired by AlphaGo’s recent successes. The PLA’s apparent expectation that the future increases in the tempo of operations will outpace human cognition could result in a pragmatic decision to take humans “out of the loop” in certain operational environments. In others, the PLA also recognizes the importance of integrating and leveraging synergies among human and machine “hybrid” intelligence.

In Section 21, Lora Saalman gains insight into Chinese AI research using an illuminating case, Chinese efforts to integrate neural networks into its hypersonic platforms. Based on analysis of over 300 recent Chinese technical journal papers and articles issued by researchers at Chinese universities and military institutes, she uncovers two major trends. First, increasingly innovative and prolific research that demonstrates expanded domestic and international collaboration. Second, a quantitative and qualitative shift away from defensive countermeasures to offensive platforms, suggesting a trend from China's traditional stance of "active defense" toward a stronger, AI-enabled offense.

In Section 22, Samuel Bendett examines Russia's expanding AI development. The Russian government's increasing attention to developing AI-assisted and AI-facilitated technologies drives this expansion. Moscow's AI development still lags far behind nearest peer competitors like China and the United States. However, progress is evident. Specifically, the Russian military is investing heavily in creating the intellectual and physical infrastructure for AI development across its services. The government is also eager to expand debate and cooperation between the country's growing hi-tech private sector and expansive military–academic infrastructure.

Part VI. Artistic Perspectives and the Humanities

Section 23 is the short story "Infinite Bio-Intelligence in the World of Sparrows" by Eleonore Pauwels and Sarah Denton. *Nothing lives or dies without being monitored*. In a future where artificial intelligence, advanced genomics, and biotechnologies converge, we will constantly be aware of the biological evidence we unwittingly leave behind as we go about our daily lives. In this fictional, futuristic scenario, the authors attempt to convey the social, political, and ethical implications of deploying such technologies without regard for human rights. What is lost in the fray of the "Internet of Bodies," the ubiquitous bio-surveillance network, are the human stories that emerge from such a system. This is one such story.

Section 24 is the visual piece "Two Memos from the Future" by Lydia Kostopoulos. In efforts to look backward into the present, she has chosen futuristic scenarios to help us visualize the future in a way that technical reports do not. Predicting the future in an era of exponential change and rapid technological convergence is partly making an educated guess based on technological assessments and partly creative exploration of the status quo and imaginative alternatives. These scenarios are on the horizon in some form or another.

Section 25 is the short story "The Parade Cleaners" by Lt Col Jennifer Snow. The story explores one of the darker futures of a burgeoning surveillance state. We follow Chad, a security worker responsible for digitally patrolling the prestigious main thoroughfare. The short proposes some challenging questions for our growing global information culture and pushes the reader to consider "what if?" Are these potential technological calamities that could or are becoming real today? Who determines which people benefit and which people do not? The AI programmers? The government? The public? What is the future of free speech, public access, or upward mobility in an increasingly divided global infospace between authoritarian and libertarian ideals?

Section 26 is the essay "Beware the Jabberwocky: The AI Monsters Are Coming" by Natasha Bajema. Science fiction plays an important role in shaping our understanding of the implications of science and technology and helping us to cope with things to come. This artistic piece describes three AI monsters depicted in science fiction films as one day disrupting the global order and potentially destroying humanity: the automation monster, the supermachine monster, and the data monster. Fears about the implications of the automatic and supermachine monsters distract us from the scariest of them all. Below the surface of our daily lives, the data monster is stealthily assaulting our sense of truth, our right to privacy, and our freedoms.

Section 27 is the essay "Is China's AI Future the Snake in the Wine? Or Will Our Future Be FAANGed?" by Regina Joseph. China's urgent plan to dominate in AI is characterized in similar world-changing terms to Silicon Valley's. Both portrayals emphasize limitless opportunity, brilliance, and social good. However, a different potential lurks beneath. In the United States, younger generations seem to slowly recognize the bondage posed by addictive technologies—a fate prophesized by Aldous Huxley's *Ultimate Revolution*. In China, centralized control and soft coercion stymie public opposition to techno-nationalism, leading to an unchecked zeal for AI expansion that will adversely affect China, the United States, and beyond.

Part I. Artificial Intelligence, Domestic Political Regimes, and the Global Order

Section 1. The Technologies – What Specifically Is New?[1]

Abstract

This section discusses the new technologies:

- By *artificial intelligence* (AI) here we mean a constellation of new technologies: AI itself more narrowly defined, big data, machine learning, and digital things (e.g., the "Internet of Things").
- This constellation of technologies is bringing in a new technological epoch. Following a leap in AI research around 2012, we now have: *Automated systems learning directly from data to do tasks that are complicated*. The key change is that the task is now complicated (e.g., AI can now do good facial recognition).
- Crucially, AI particularly improved for tasks related to "perception"— e.g., perceiving images or speech, or some kinds of patterns in big data— and these are the advances now being rapidly rolled out across diverse real-world uses. AI also improved when choosing actions in tasks that are bounded enough to be very well described by vast amounts of data.
- AI's current technical limitations mean that its current incorporation into social governance must include extensive human involvement; and also, that setting up big data platforms will likely dominate current efforts. This is what we see now in China.
- AI adds a new layer to traditional "cyber."

What Are the AI-Related Technologies?[2]

By the term "AI" here we refer to a constellation of AI-related technologies (AI more narrowly defined, machine learning, big data and digital things) that together provide powerful, wide-ranging and new capabilities (Figure 1.1).[3] Together they enable a new industrial revolution,

[1] This section was written by Nicholas D. Wright, Intelligent Biology; Georgetown University; University College, London; New America.

[2] The author thanks Zeb Kurth-Nelson for insightful discussions on the technical aspects of AI research included here.

[3] We use a broad characterization here because the term *artificial intelligence* has come to refer to many significant things that are not captured by narrower definitions. An analogy is the term "rational," which means many things to many people.

taking the vast reams of data now produced by the computers and Internet of the preceding revolution—and turning it into useful data (Box 1.1).

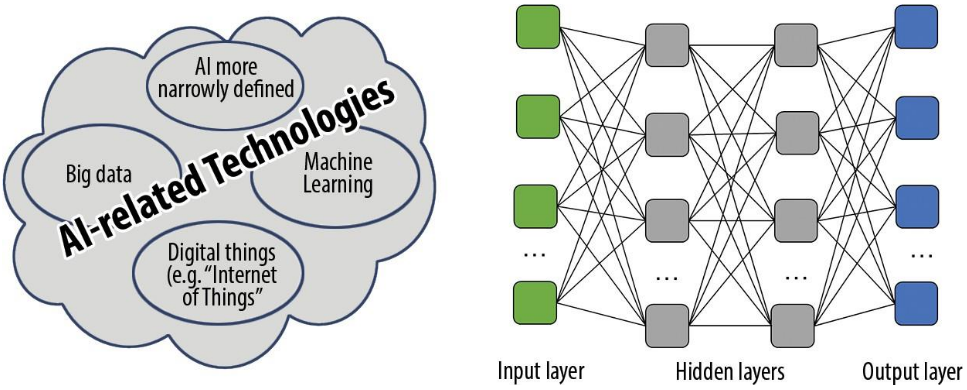

Figure 1.1. The left panel shows the constellation of AI-related technologies. The right panel illustrates "deep learning." Deep learning is one of many approaches to machine learning. It was inspired by the brain, and in particular the interconnecting of many neurons (Artificial Neural Networks). In deep learning, the key idea is that the neural networks have at least one "hidden layer" in the middle between inputs and outputs, whose "neurons" can take on different weights while learning about the task.

Box 1.1. Is this new "nth Industrial Revolution" really distinct from what went before?

This question is closely related to the question: why is AI different to "cyber?" An industrial revolution may be defined as "A general term for the process of the rapid onset of continued economic change and advancement through the application of industrial techniques to traditional forms of manufacture" (Lawrie, 1999). As was recently argued by perhaps the most prominent voice behind the idea that AI reflects a fourth industrial revolution:

> "There are three reasons why today's transformations represent not merely a prolongation of the Third Industrial Revolution but rather the arrival of a Fourth and distinct one: velocity, scope, and systems impact. The speed of current breakthroughs has no historical precedent. When compared with previous industrial revolutions, the Fourth is evolving at an exponential rather than a linear pace. Moreover, it is disrupting almost every industry in every country. And the breadth and depth of these changes herald the transformation of entire systems of production, management, and governance." (Schwab, 2015)

I discuss this "nth industrial revolution" in the next section.

None of the technologies is entirely new, but there have been big recent improvements (particularly from deep learning, see below) and together the constellation has revolutionary applications. Within the constellation of new technologies, four are crucial:[4]

- *"AI" more narrowly defined.*[5] One can describe AI as the analysis of data to model some aspect of the world, where inferences from these models are then used to predict

[4] This subsection's definitions draw in particular on (ICO, 2017).

[5] This narrower definition of AI itself is also highly debated, and is further sub-divided in various ways. For instance, one might contrast "general AI" that can apply its intelligence to many tasks, against an AI such as Siri that is

and anticipate possible future events. Importantly, AI programs do not simply analyze data in the way they were originally programmed. Instead they learn from data to respond intelligently to new data and adapt their outputs accordingly. AI is ultimately about "giving computers behaviors which would be thought intelligent in human beings" (ICO, 2017).

- *"Machine learning."* Many of the computational techniques related to AI are actually from a field called machine learning. This can be described as "…the set of techniques and tools that allow computers to 'think' by creating mathematical algorithms based on accumulated data." Arthur Samuel coined the phrase, in 1959, defining it as, "the ability to learn without being explicitly programmed." (McClelland, 2017). *Deep learning* is one method for machine learning—and it is improved deep learning that recently led to big advances in AI (Figure 1.1 right panel).
- *"Big data."* These are high-volume—as well as often high-velocity and high-variety—information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making. The massive recent increase in the amount of big data everywhere is new.
- *Digital things.* Things (e.g., smartphones, "Alexa," toasters, military drones, robots in factories) will increasingly be able to perceive (e.g., facial or speech recognition), decide, and act. The things may not be connected to the Internet but may be smart. The "Internet of Things" refers to the growing interconnectedness of things, and getting them onto the Internet.

Together these technologies are more than the sum of the parts. Firstly, consider "*big-data analytics*." One can think of big data as an asset that is hard to exploit, for which AI is a key to unlocking its value, and where machine learning is one technical mechanism for doing AI. When big data analytics is merged with things in the real world, we have *online-to-offline* merging.

What Was the New Big Improvement Around 2012?

The computer and Internet-related revolution made lots of data, and now this AI-related revolution turns that data into usable information. We are early in this new epoch. Around 2012 researchers made a large improvement in the quantity of big data that automated systems can analyze, which was sufficiently large that it essentially provided qualitatively new capabilities. After 2012 we have qualitatively new: *Automated systems learning directly from data to do tasks that are complicated.*

We can unpack this. "Automated systems" means AI programs themselves, not relying on humans. "Learning directly from data" means the way the AI doing the job does not depend on hard coding from humans. A task is something like facial recognition. That the task is now complicated is the key change, and how we measure complexity may be via comparisons to human performance, or previous AI performance. For instance, AI can now do good facial recognition. These basic advances were those leveraged to achieve Al-phaGo's victory over a top human in 2016.

programmed to essentially perform a single task (called "narrow AI," although not AI more narrowly defined in the sense we use in this volume that also includes "strong" or "general AI"). This also broadly corresponds to "strong AI" versus "weak AI."

Two papers signaled and illustrate this change:

1) Krizhevsky, Sutskever, and Hinton, (2012): This was a big breakthrough in perception. In a visual object recognition task, they trained a deep convolutional neural network to classify visual images. They trained the neural network on 1.2 million images—all labeled—from a huge and then new dataset called "Imagenet." They roughly halved the error rate of the previous state-of-the-art on the most challenging benchmark to date. Such AIs recently approached human-level performance on some object recognition benchmarks. This paper triggered huge interest in AI research, with some 33,000 Google Scholar citations in under six years.
2) Mnih et al., (2015): The AI learned to play a large range of classic "Atari" computer games, with essentially the only inputs being the pixels on the screen and the game score—and it achieved human or superhuman performance on many games. It had to deal with the huge perceptual challenge, and also control actions. They combined ideas from deep learning and reinforcement learning (i.e., learning from the rewards and punishments associated with previous events). Within the tightly bounded environment in each game, the AI could play vast numbers of times to learn from a huge dataset on each game environment.
3) Such advances were crucial for AlphaGo's famous 2016 victory over a world-class human go player. Go is a lot more difficult than chess. Within the tightly bounded environment of go, before beating world champion Lee Sedol, AlphaGo effectively learned from some 100 million or more games altogether (Lake, Ullman, Tenenbaum, and Gershman, 2017).

What led to this big change? There was no magic bullet. Instead, three factors combined:

1) Raw computer power increased.
2) Datasets for training became available. For instance, the advance by Krizhevsky et al. (2012) was possible because they had a huge dataset of millions of labeled images on which to train. The imagesets often need to be labeled, so the AI can learn.
3) Deep-learning algorithms were improved. It was not a single innovation, but instead multiple moderate improvements (e.g., "dropout" and "ReLUs" in the 2012 "Imagenet" advance).

Current Strengths and Weaknesses—and Where AI Is Going

These advances have been huge but not uniform, and it is important to understand the technical strengths and weaknesses. This helps understand both what we might expect to see in real-world applications—for instance in the construction of a surveillance state—and also where the research is likely to go.

Strengths

The new technology has two big new strengths:

1) First, one really huge new improvement in AI capabilities relates primarily to "perception," such as perceiving images or speech, or patterns in some types of big

data that humans may not be able to perceive. That is what the 2012 advance in classifying “Imagenet” pictures was all about in the preceding subsection.

Thus, now local devices such as smartphones, digital assistants, or cheap cameras in office lobbies can effectively monitor speech or faces—and indeed such technology is already widespread in the West and China. One can see why this is particularly good for surveillance, as discussed later in this work.

Moreover, being able to learn to perceive well also means that if you reverse those models you can be very good at producing images or audio. In a strategic context that may be useful for fooling others (e.g., “deepfakes”).

Databases of data, much of which may have originally been collected for other purposes, can also be examined for patterns—adding value to the “big data” that may just have been sitting there.

2) Second, AI also improved in choosing actions in tasks that are bounded enough to be very well described by vast amounts of data. Go or the Atari games above are a good example. A well-known real-world example is Google Deepmind training AI on data from Google’s datacenters, and so “more accurately predicting when the incoming compute load is likely to land,” which reduces power consumption for cooling (Burgess, 2016).

Current Limitations

However, there are two major limitations in the current AI technology. These help us know what we should expect if these new AI-related technologies were applied in domestic security.

1) *Huge amounts of data are needed to train the system, and this data often needs to be labeled* (e.g., this is a picture of a cat). The availability of a huge dataset of labeled images—“Imagenet” described above—was a crucial factor enabling the big leap in 2012. The algorithms cannot yet generalize well from learning in one environment to learning in another, and also they cannot learn things from just a few instances as humans often can.

As discussed in later sections, this is why it is so important in a surveillance state to add “ground truth” data (e.g., tax returns, criminal records or medical records) that acts like labels for your broader data (e.g., smartphone usage; Figure 3.3).[6] Often governments are the only parties with such data (e.g., tax returns) or they heavily regulate who can access data (e.g., medical records or genetic data). Without the ground truth data, just having tons of big data by itself will be a lot less useful.

Moreover, this greatly raises the value of having very detailed monitoring specific populations with extensive ground truth data, because you can then use that very detailed data to train your algorithms. For those working on Chinese surveillance, that would be one big advantage of the very heavy physical and online monitoring in Xinjiang province.

[6] One example would be training a program to predict tax payment (or avoidance) based on innumerable aspects of smartphone data. Or training a program to predict criminal acts, including those that may have political dimensions, from smartphone data. An analogy is given by recent reports of the Chinese company Smart Finance, which uses seemingly irrelevant smartphone data, such as the typing speed or battery charge levels, to predict individuals’ creditworthiness for loans, with reportedly high accuracy (Lee, 2018). One could also create links between such systems.

Further, this is a good reason why lots of humans will be needed in any AI system of surveillance for the foreseeable future—to do labeling. Indeed, the importance of cheap labor for labeling has even been touted as a key Chinese strength in AI more broadly (Yuan, 2018).

Making the datasets is a huge challenge. Creating the "Imagenet" labeled dataset was a *pre*condition of the leap made in 2012. Similarly, building big datasets that are in right form with the right type of labeling and so on should be the current major effort, if one were building an AI-enabled surveillance state now.

2) *Context is still very poorly understood by the systems*—that is, they lack common sense (e.g., is this likely to be a picture of a baby holding a toothbrush or a gun?). This is why human-machine teams and semi-automated systems are often the only way to harness the benefits of AI, by adding the human ability to add context.

   The challenge of context is another key reason why any plausible surveillance system will only be semi-automated for the foreseeable future—lots of humans would still be needed even if a system built with current cutting-edge AI technology worked perfectly.

Where Is AI Going in the Lab and at Scale in the Real-World?

Given the state of AI-related research, where might we expect the technologies to go over the next five years or so?

First, we might ask: where will the cutting-edge research go? Efforts to overcome the limitations above are perhaps the two hottest current research areas—and given the huge resources being spent to overcome them, this is where to expect potential research advances. The scientific literature is looking to augment deep-learning methods that learn from experience, for instance by adding more informed models of the world. Just as the human brain does, and as AlphaGo arguably began to do (Lake et al., 2017). To give a flavor of the Defense Advanced Research Project Agency's[7] goals with AI, they describe a focus on moving beyond what they call "second wave AI" of the type we have now—which is good at perception and learning but not at abstraction and reasoning—and toward "third wave AI." That third wave AI aims at "Contextual Adaptation [in which] Systems construct contextual explanatory models for classes of real-world phenomena."

But while that is the cutting-edge research, crucially we must also allow for time lags: not just from lab to real-world, but also to large-scale in the real-world. The Internet, for instance had certainly reached the real-world by the early 1990s—growing up then in London my family had an Internet connection—but many of the Internet's large-scale real-world impacts took another one decade or two to occur, such as Amazon or Facebook reaching their huge scale.

Thus, second we might ask: where are the AI-related technologies going in the real-world at scale? AI advances in visual or speech perception are now rolling out at huge scale in our smartphones and digital assistants. However, we are still working on how to usefully use all the information this produces, and how to work that into broader commercial or governmental systems. Perceiving patterns in big data that are hard for humans to perceive will almost certainly bring about great advances fields like medicine or predictive policing—but it is important to realize that such real-world applications are only in development ("The Promise and Perils of AI Medical Care," 2018) and fully operational systems are certainly not ready for

[7] Defense Advanced Research Project Agency, https://www.darpa.mil/about-us/darpa-perspective-on-ai.

rollout at huge scales. Driverless vehicles still require a lot more training data, and are now being deployed in very limited circumstances, such as a pilot taxi service in Phoenix or Tesla's partial driving assistance. Overall, for most AI-related technologies the next five years will likely see a lot of piloting to find out what works in the real-world, while building the crucial datasets and also assessing how the technologies can later be rolled out at scale. That is also what we would expect if one were building these powerful new technologies into a surveillance state, and for instance is what we see now in China.

How Do the New AI-Related Technologies Relate to Traditional "Cyber?"

AI adds new properties and a new layer of value to what can be done using the Information and Communication Technologies. It is a bit like an onion. As depicted in Figure 1.2, each new layer adds value to that beneath. The Internet increased the value of computers. As so many things have become computers and can communicate electronically this has generated huge amounts of data—big data. The new AI-related technologies help turn this big data into something useful. They add more value.

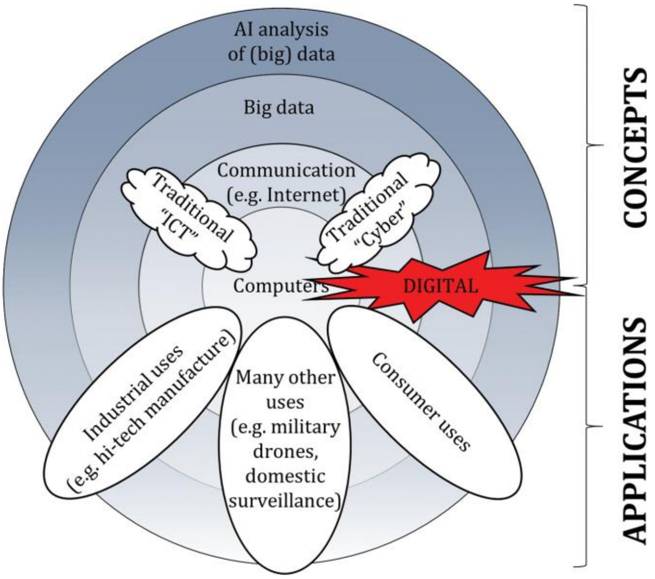
AI analysis of (big) data
Big data
Communication (e.g. Internet)
Traditional "ICT"
Traditional "Cyber"
Computers
DIGITAL
Industrial uses (e.g. hi-tech manufacture)
Many other uses (e.g. military drones, domestic surveillance)
Consumer uses
CONCEPTS
APPLICATIONS

Figure 1.2. The onion of digital technologies.

The prefix "cyber" essentially relates to computers and electronic communication.[8] Much of what happens with information will involve traditional issues in "cyber" rather than AI, so we do not need AI to tell us much about them. Computers still matter—hence the "chip wars" between the US and China ("Chip wars," 2018). Communications still matter—hence the battles over "5G" standards, over social media regulation within and between countries, as well as the global struggle for Internet governance. The challenges and opportunities of cyber will still be meaningful, it is just that there will also be other things from the new layer of the onion.

[8] The academic Thomas Rid notes "the increasing use of the word "cyber" as a noun among policy wonks or many a uniformed officer I've come to be highly distrustful of "nouners," as they all too often don't seem to appreciate the necessary technical details" p. ix (Rid, 2013). Acknowledging that the term is suboptimal, I use it here as the prefix at least does have meaning among scholars and the word among practitioners.

Finally, the different applications to which one can put the AI-related and other digital technologies speaks to an important question: *does China have a data advantage over the liberal democracies?* One can make two points here:

1) China does not have an advantage in terms of number of users if one includes the global userbases for US tech giants like Facebook or Google. But it does have an advantage in terms of integration of data across platforms[9] and also, most importantly, in terms of combining breadth of data with "ground truth" data (e.g., government data). Training AI depends on both quantity and quality. The liberal democracies *should not* compete.
2) The above comments relate to human user data, for example for consumer uses or domestic surveillance. However, a lot of important AI training will occur on other types of data such as from sophisticated machines. For example, Germany's AI strategy relies on that alternative aspect of data— which is harnessed from cyber-physical manufacturing processes and the industrial Internet of things—that plays to German industrial strengths. Many industrial, military and other applications will rely much more on that type of data than human user or consumer type data.

Conclusions

The advances in the AI-related technologies are very real. They greatly multiply the value derived from the preceding digital technologies, by turning data into usable information. Understanding the current technical strengths and weaknesses helps anticipate and understand its applications in the real world.

A big advance has been in perception; hence the ready application to surveillance and censorship. The big limitations are that the AI handles context poorly (hence human-machine teams are key), and that it requires vast amounts of well-labeled data (hence the importance of combining ground truth data—often only from government—with the breadth of data from myriad smart devices and sources). As the big research breakthrough only really began in 2012 we are in the early stages of rolling out many of the new AI capabilities, and so much of what happens now in the real world will essentially be piloting, or the building and preparing of good enough datasets. Indeed, even where AI technologies are being rolled out at massive scale—notably in commercial devices such as smartphones or digital assistants— while their outputs may be dual use for surveillance, usefully harnessing those outputs at societal scale is itself surely a massive additional IT program that requires careful building and piloting.

Finally, in terms of military uses or foreign policy decision making, these technical characteristics explain why AI's main uses have often been for more perceptual tasks, such as satellite image analysis—as is seen in the Chinese case (e.g., Elsa Kania, Section 20 this volume). With current technology, decision support will likely only work well with human-machine teams to provide contextual capabilities—and there are likely considerable risks in allowing many types of decisions to be made with humans "out of the loop."

[9] Kai-fu Lee's prominent recent book compared Chinese and US approaches to AI. He repeatedly refers to key Chinese apps that bundle together what is done separately in the West by companies like Google, Facebook or Uber (e.g., chapters 1 and 3 in Lee, 2018). An example is the "super app" WeChat, from the tech giant Tencent.

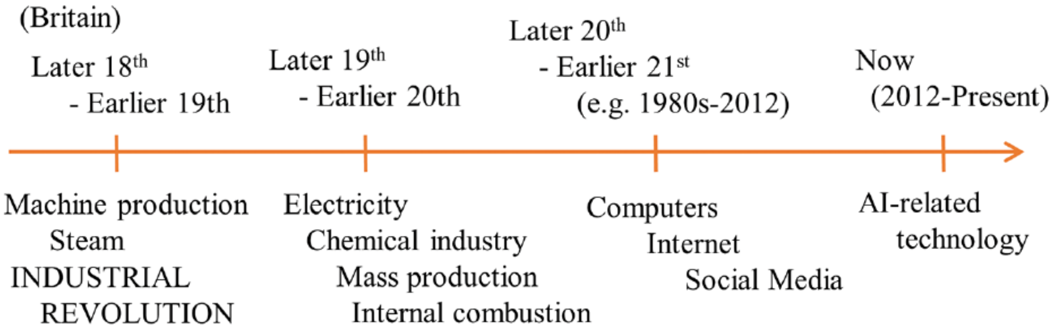

Figure 1.3. Timeline of industrial revolutions.

Section 2. Artificial Intelligence's Three Bundles of Challenges for the Global Order[10,11]

Abstract

Artificial intelligence (AI) raises a bewildering profusion of implications for the global order—which this section breaks down into three more manageable bites. This section primarily focuses on the first area, which has received by far the least attention.

1) First is how this new technology may impact domestic *political* regimes (e.g., authoritarian, hybrid, or liberal democratic) may affect competition between them in the *world order*. AI will help enable a plausible competitor to liberal democracy for big industrially sophisticated states to make their citizens rich and maintain rigid control: digital authoritarianism. China is building core components of such a system—which are being exported and emulated in a global competition with liberal democracy.
2) An "nth Industrial Revolution": AI will radically change the means of production across economic and societal sectors, e.g., transport, healthcare or the military.
3) The "singularity" and the sense of self: In the singularity, exponentially accelerating technological progress creates an AI that exceeds human intelligence and escapes our control, potentially destroying humanity or disrupting humans' conceptions of themselves.

Introduction

Everybody now seems to agree that AI seems important for everything. From what it means to be human; to the social impacts of laying off Uber drivers once cars drive themselves; to AI propaganda in politics; to the rise of the robots or a superintelligence exterminating humanity. But what does AI's bewildering profusion of implications mean for the global order? Anticipating AI's challenges for the global order requires breaking them down into more manageable bites—because failure in any one of these three distinct bundles of challenges I identify would be catastrophic. As Figure 2.1 shows, bewildering profusion of implications mean for the global order? Anticipating AI's challenges for the global order requires breaking

[10] This section was written by Nicholas D. Wright, Intelligent Biology; Georgetown University; University College, London; New America.
[11] This section draws in part on Wright, 2018.

them down into more manageable bites—because failure in any one of these three distinct bundles of challenges I identify would be catastrophic. As Figure 2.1 shows, they are: (1) Competing types of political regimes in the global order; (2) Change in the means of production across social sectors in an "nth Industrial Revolution"; and (3) The "singularity" and the sense of self.

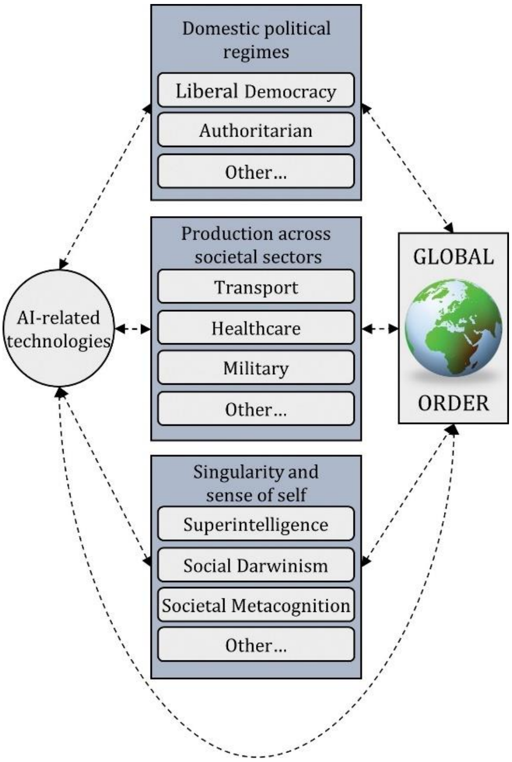

Figure 2.1. AI's impacts on the global order.

This volume focuses primarily on the first area—AI's impacts on competing domestic political regimes in the global order—because it is critical but has been least examined. This section also describes the other bundles of challenges for two reasons. Firstly, because global strategy must address all three areas. Each of the three bundles of challenges requires different thinking—and policies—at the level, of nations, of the UN, business and other stakeholders. Second, many who debate the potential importance and/or urgency of AI's impacts on the global order end up discussing different bundles of impacts, and so talk past one another. Here we put them in one space.

Competing Political Regimes in the Global Order

New technologies may affect the form and/or relative attractiveness of different types of *domestic political regimes*—e.g., authoritarian, liberal democratic or hybrids combining features of each—and this may affect competition between such regimes in the *global order*. A domestic political regime is a system of social organization that includes not only government and the institutions of the state, but also the structures and processes by which these interact with broader society. Competition between different types of domestic social system was a crucial feature of twentieth-century global politics. While liberal democracy's eventual triumph may now seem inevitable, fascist regimes in the interwar period and communist regimes for much longer were plausible paths forward for big, industrially sophisticated societies to make their citizens rich. Now the new AI-related technologies could crucially help reinvigorate the idea that more authoritarian regimes can make their citizens rich and maintain social control.

Section 3 and Part II of this volume examine digital authoritarian regimes in particular in more detail. Section 4 and Part III of this work examine export and emulation of such regimes in global competition.

Before moving on, however, it is important to note multiple reasons why domestic political regimes matter for the global order.

- First, as described above, they may offer competing visions of the future, and these may compete for influence within swing states in global competition. This is only one potential facet of influence between states, but it can be highly significant as it was in the twentieth century. This is particularly the case if large countries develop particular types of domestic regimes, such as Russia in the early twentieth century (i.e., the Soviet-style Communist regime) or potentially China during its rise in the twenty-first century.
- Second, aspects of domestic regimes, such as bureaucratic or domestic political audiences, can profoundly affect foreign policy decision making. Part IV of this work examines how the development of digital authoritarianism may affect Chinese foreign policy decision making.
- Third are ideas that some types of regime are inherently less, or more, problematic in the global system. Most prominent is the idea of "democratic peace theory," which identifies a correlation between domestic structure and the absence of war between democracies—a very prominent notion among scholars and indeed practitioners (Russett, Layne, Spiro, and Doyle, 1995).

Change in the Means of Production across Social Sectors

A second basket of challenges arise because AI and big data will radically change the means of production across many economic and societal sectors. There will be winners, losers and new ways of doing things, which will roil societies across the globe. Consider three sectors. One now classic example is transport: after Uber rolls out self-driving cars, where will all the unemployed drivers work (Edwards, 2017)? Another sector is the military. Drones and AI will likely contribute to a revolution in military affairs, which may be destabilizing (Horowitz, 2018). There may be arms races. A third example is the colossal health sector, accounting for some 18 percent of US GDP (CMS.gov, 2018), where AI promises to change how medical decisions are made and care delivered ("A revolution in health care is coming," 2018). One can point to essentially any social sector.

But not much so far suggests this will be bigger than other technological impacts, such as those contributing to the Industrial Revolution itself—or the internet's rise in the 1990s-2000s. Uber drivers being sacked isn't much different to the internet reducing retail jobs with Amazon's rise.[12] The airplane, steamship, machine gun or tank all revolutionized warfare; and

[12] See e.g., (Thompson, 2017) but note e.g., (Manne & Maclean, 2017).

so did the internet, with cyber now a military domain alongside land, sea, air and space. Potential change in healthcare is exciting but powerful human, regulatory and institutional factors make the health sector as nimble as a supertanker. One potential caveat is that the rapidity of these changes renders them different, but as Section 1 describes making many of the AI-related technologies work in the real-world at scale means overcoming numerous tough practical problems, which downloading a software update won't solve.

We might call this an "*n*th Industrial Revolution," as the popular term fourth Industrial Revolution has been around since the 1940s (Thornhill, 2018). These changes and their attendant disruptions will require management, just as welfare states were created and adapted to manage the social disruptions from industrialization. It requires sector-by-sector planning. Much will rely on relatively straightforward, although politically challenging, means such as welfare nets and retraining for the swathes of workers whose jobs become obsolete.

Changes in the means of production matter for the global order for multiple reasons:

- First, twentieth-century history illustrates how failure to manage domestic social dislocations, such as in interwar Germany, disrupts global order.
- Second, twentieth-century history also illustrates how new military technologies can affect the balance of power or strategic stability. Chapter 4 and Part V in this work examine military aspects.
- Third, how well different countries harness new technologies within their societies can affect the relative balance of power between them. An example is that while Britain dominated economically in the original Industrial Revolution, instead Germany and the US harnessed twentieth-century technologies equally well or even slightly better.
- Fourth, it is possible that the AI-related technologies may alter the inequalities in power between nations. For instance, if robotic manufacturing becomes highly effective, this may remove a key advantage that poor countries have traditionally had when developing—supplies of low-skilled workers for labor-intensive manufacturing such as in textiles. The AI-related technologies may also exacerbate inequalities between the developed economies, as we have seen to some extent with US tech giants totally unmatched in Europe or Japan.

The "Singularity" and the Sense of Self

The singularity is the single biggest concern for many AI scientists. The idea is that exponentially accelerating technological progress will create an AI that exceeds human intelligence and escapes our control ("What is the Singularity?," 2018). This superintelligence may then deliberately or inadvertently destroy humanity, or usher in an era of plenty for its human charges. As Henry Kissinger also describes, the catastrophic consequences may not only be physical but also apply to humans' conceptions of themselves (Kissinger, 2018). For him, the most important question is: "what will become of human conscious- if its own explanatory power is surpassed by AI, and societies are no longer able to interpret the world they inhabit in terms meaningful to them?" Given the rate of progress, the singularity may occur some point this century. But although clearly momentous, given that nobody knows when, if or how a possible singularity will occur, limits clearly exist on what can sensibly be said or planned for now. Previous existential technologies have emerged: nuclear weapons can obliterate humanity. Indeed, nuclear weapons provide a useful, although imperfect, analogy for global efforts to manage or prevent a singularity. Preventing nuclear war required careful management and luck, which we will need again. Preventing nuclear proliferation is tough, and despite considerable success we couldn't prevent North Korean nuclear weapons. Could one persuade Russian, Chinese or US leaders to stop AI programs viewed as vital for their security? Indeed, this is more concerning than Kissinger's further concerns about human understanding of our own nature. Human egocentrism is remarkably robust—if we can (despite wobbles) deal with Darwin telling us we're just hairless apes, we'll survive this new disclosure.

The bottom line is that, just like nuclear weapons, singularity-related issues will require managing within the international order as best we can, although our best will inevitably be grossly imperfect. The singularity potentially represents a qualitatively new challenge for humanity that we need to think through and discuss internationally. But plenty of other fish also need frying, and a lot sooner.

Conclusions

Global strategy must address all three bundles of challenges that AI presents for the global order. Most attention has been paid to the singularity and a new Industrial Revolution. Thus, this work focuses primarily on an equally crucial bundle of challenges for the global order, posed by AI's implications for domestic political regimes.

Box 2.1. What is the global order?[13]

Below I give my working definition and some other examples of definitions, so that the reader can see the concept's broad shape.

My working definition: The global order is a system covering the whole of human society that includes: (1) social institutions around which actors' expectations converge; and (2) the distribution of power among key subsystems in the global system, where these subsystems include states (e.g., the US or China), international subsystems (e.g., the global financial system or the UN), and important systems at other levels (e.g., regions below the level of the state, such as Catalonia; or systems above the level of the state, such as during the Cold War there were the liberal international system and the Communist international system).

That is, the global order is a system of systems. It involves material factors, subjective ideas/perceptions, path dependence (i.e., "history matters") and multiple levels.

A textbook definition: "World order is the distribution of power between and amongst states and other key actors, giving rise to a relatively stable pattern of relationships and behaviours." (Heywood, 2013, p. 422).

A prominent academic definition: "International regimes have been defined as social institutions around which actor expectations converge in a given area of international relations. Accordingly, as is true of any social institution, international regimes limit the discretion of their constituent units to decide and act on issues that fall within the regime's domain. And, as is also true of any social institution, ultimate expression in converging expectations and delimited gives international regimes an intersubjective quality." … "The

[13] Many different terms are used and discussed, such as *World Order*, *International Order*, *Liberal World Order*, or *New World Order*. I prefer the term *global system*. However, here I avoid including the word *system* to prevent confusion that may arise as this chapter also discusses systems at many other levels within the global order. For instance, domestic political regimes may be called systems (see below), while the digital social governance systems used and planned in China are in fact best thought of as *systems of systems*.