



Artificial Intelligence and International Relations Theories

Bhaso Ndzendze · Tshilidzi Marwala

palgrave
macmillan

Artificial Intelligence and International Relations Theories

Bhaso Ndzendze · Tshilidzi Marwala

Artificial Intelligence and International Relations Theories

palgrave
macmillan

Bhaso Ndzendze
University of Johannesburg
Johannesburg, South Africa

Tshilidzi Marwala
University of Johannesburg
Johannesburg, South Africa

ISBN 978-981-19-4876-3 ISBN 978-981-19-4877-0 (eBook)
<https://doi.org/10.1007/978-981-19-4877-0>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Palgrave Macmillan imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

PREFACE

Given the nature of the advances being made in artificial intelligence (AI) and its manifestation in areas relevant to international relations (IR), and the impact these are having on a number of its underlining assumptions, its theoretical paradigms may be in need of updating. To these ends, this book investigates the implications of the deepening interface between AI and various aspects of IR. We examine and, where appropriate, we update liberalism, realism, dependency, hegemonic stability theory, the English School, constructivism, postcolonial theory, feminism, and green theory. Updates are made to the economic interdependence-peace thesis, the democratic peace thesis, and hegemonic stability while also revisiting accounts of deterrence and dependency. Extant literature of AI in IR are otherwise integrated or further developed. The book is rich in empirical case studies, which substantiate or otherwise dismiss the hypotheses advanced by traditional and critical theories, using falsifiable methods and open access data so as to allow further corroboration by subsequent scholarship. The book is a contribution to the literature on contemporary technology-based geopolitics, IR theory, and their attendant transformations of the nature of international interaction, including the forms of global inequality, power dynamics, conflict, and cooperation. Minor parts

of the book (in Chapters 5, 9, and 10) have been previously published and are partially reproduced here with permission from the editors of the outlets.

Johannesburg, South Africa
Johannesburg, South Africa

Bhaso Ndzendze
Tshilidzi Marwala

CONTENTS

1	Introduction	1
2	Theory in International Relations	15
3	Artificial Intelligence and International Relations	33
4	Realism and Artificial Intelligence	55
5	Liberalism and Artificial Intelligence	73
6	Hegemonic Stability Theory and Artificial Intelligence	87
7	Dependency and Technology in the 4IR	105
8	The English School and Artificial Intelligence	133
9	Critical IR Theories and Artificial Intelligence: Constructivism, Postcolonialism, Feminism, and Green Theory	141
10	Conclusion	157
	Index	163

ABBREVIATIONS

4GW	Fourth Generation Warfare
4IR	Fourth Industrial Revolution
5G	Fifth Generation Connectivity
AfDB	African Development Bank
AI	Artificial Intelligence
AIIB	Asian Infrastructure Investment Bank
AU	African Union
BERD	Business Expenditure on Research and Development
BRI	Belt and Road Initiative
CIA	Central Intelligence Agency
DFH	Demand-Following Hypothesis
DTI	Development of Telecommunications Infrastructure
EU	European Union
FBH	Feedback Hypothesis
FDI	Foreign Direct Investment
GDP	Gross Domestic Product
GERD	Government Expenditure on Research and Development
ICT	Information and Communications Technology
IDEA	Institute for Democracy and Electoral Assistance
IoT	Internet of Things
IPE	International Political Economy
IR	International Relations
IW	Information Warfare
LDC	Less Developed Country
MNC	Multinational Corporation
MNE	Multinational Enterprise

NATO	North Atlantic Treaty Organization
NLH	Neutrality Hypothesis
NPT	Non-Proliferation Treaty
NTI	Nuclear Threat Initiative
OECD	Organization of Economic Cooperation and Development
OFDI	Outward Foreign Direct Investment
OSS	Open Source Software
PPP	Purchasing Power Parity
R&D	Research and Development
SDO	Standards Development Organizations
SME	Small and Medium Enterprise
STEM	Science, Technology, Engineering and Mathematics
STI	Science, Technology and Innovation
SLH	Supply-Leading Hypothesis
UKUSA	UK, US and Australia
WEF	World Economic Forum
WSIS	World Summit on the Information Society
WWI	World War I
WWII	World War II

LIST OF FIGURES

Fig. 2.1	Methods in articles by leading and regional IR journals surveyed (<i>By authors. Data sourced from</i> EJIR, IA, IO, JASIA, JIRD, and SAJIA <i>journal archives</i>)	24
Fig. 3.1	Taxonomy of AI (<i>Source</i> By authors)	34
Fig. 4.1	Overview of some major thinkers in the three realist strands (<i>Source</i> Authors)	60
Fig. 5.1	Proposed interaction of democratic peace thesis (constitutional constraints variant) and AI (e.g. DeepFakes) (<i>Source</i> Authors. See also Artificial Intelligence and Emerging Technologies in International Relations [Ndzendze and Marwala 2021a, b])	80
Fig. 5.2	Proposed automation-trade interdependence status quo disruption (<i>Source</i> Authors. See also Artificial Intelligence and Emerging Technologies in International Relations [Ndzendze and Marwala 2021a, b])	83
Fig. 6.1	Status quo power military expenditure and war outbreak correlation	92
Fig. 6.2	Status quo power AI military expenditure and war outbreak correlation (<i>Source</i> Authors)	98
Fig. 7.1	Denmark (<i>Source</i> Authors)	113
Fig. 7.2	Finland (<i>Source</i> Authors)	114
Fig. 7.3	Hong Kong, SAR (<i>Source</i> Authors)	114
Fig. 7.4	Luxembourg (<i>Source</i> Authors)	115
Fig. 7.5	Netherlands (<i>Source</i> Authors)	116
Fig. 7.6	Norway (<i>Source</i> Authors)	116

Fig. 7.7	Singapore (<i>Source</i> Authors)	117
Fig. 7.8	Sweden (<i>Source</i> Authors)	118
Fig. 7.9	Switzerland (<i>Source</i> Authors)	118
Fig. 7.10	United Kingdom (<i>Source</i> Authors)	119
Fig. 7.11	Bolivia (<i>Source</i> Authors)	120
Fig. 7.12	Burundi (<i>Source</i> Authors)	120
Fig. 7.13	Chad (<i>Source</i> Authors)	121
Fig. 7.14	Ethiopia (<i>Source</i> Authors)	122
Fig. 7.15	Lesotho (<i>Source</i> Authors)	122
Fig. 7.16	Mauritania (<i>Source</i> Authors)	123
Fig. 7.17	Nepal (<i>Source</i> Authors)	124
Fig. 7.18	Rwanda (<i>Source</i> Authors)	124
Fig. 7.19	Tanzania (<i>Source</i> Authors)	125
Fig. 7.20	Zimbabwe (<i>Source</i> Authors)	126
Fig. 7.21	Generalised incidence summary, 2009–2017 (Data sourced from WEF. Chart and calculations by authors)	128
Fig. 10.1	Countries with ongoing interstate violence past and future, 1970–2040 (<i>Source</i> RAND ‘Conflict Trends and Conflict Drivers: An Empirical Assessment of Historical Conflict Patterns and Future Conflict Projections’ [2015] https://www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1063/RAND_RR1063.pdf)	159

LIST OF TABLES

Table 4.1	Major states' AI strategies and their discourses	62
Table 5.1	Regime types' niches and setbacks in AI development	84
Table 6.1	Summary of contemporary AI military applications and developer nations and corporations	100
Table 7.1	High scorers' raw data, 2009–2017 (core/core-periphery)	112
Table 7.2	Low scorers' raw data, 2009–2017 (periphery)	112
Table 7.3	Ranking of correlations scores	127



Introduction

If there is anything that can be agreed upon, and there are few even among specialists, it is that artificial intelligence (AI) is a complex phenomenon, rendered all the more so by its integration into complex human institutions and modes of interaction. There is perhaps no more complicated and layered an interaction as international relations (IR). The interaction of these two human-engineered spheres requires a re-look at the underlying sets of assumptions that inform our analysis and predictions for IR. Theory is central to IR. It is how scholars generalise and predict outcomes at a global level (both territorially and in terms of universal knowledge generation). Scholars of IR are primarily concerned with the international system and how it shapes, and is shaped by, states, and sometimes non-state actors, as they pursue their interests. States vary in size and capabilities. They also vary in their intentions. Both of these attributes, capabilities, and intentions, are ever-evolving elements. Size is more than a function of geography and population. Intentions are a function of historical milestones and the interlinked networks of interests. In sum, no station is permanent, and neither are alliances or conditions of enmity. A substantial part of this change is due to technology and its impact on the trajectories of nations. Technology, being differentially distributed and accessed at different stages by states, shapes the capabilities of these actors, and thus their perception of what they have licence to do (intentions).

As a discipline of study, IR is comprised of a number of fields, most prominently international security (IS), international political economy (IPE), international organisation (IO), and foreign policy (FP) analysis, all four of which are undergirded by theory. Understanding AI's impact on these, therefore, requires an understanding of its impact on theory. We have two classes of theories in IR: mainstream (realism and liberalism) and critical (economic structuralist theories such as dependency theory which are derived from Marxist philosophy; postcolonial theory; constructivism; feminism; and green theory).

The section which follows consists of a survey of the literature, giving an overview of recent trends in scholarship on IR and technology, and thus the theoretical updating necessitated by these. This is then followed by a summary of the book's aims and findings.

1.1 LITERATURE: IR SCHOLARSHIP, THEORY, AND AI

IR scholarship has been variously updated by major events, upheavals, and processes, including the two world wars, the Cold War, decolonisation, globalisation, 9/11, and more recently thinking around climate change. Similarly, it is observed that the scholarship has many theories which accumulate and at worst undergo lulls and revivals but not elimination on account of the variety of phenomena and events it attempts to explain (war, patterns of FDI and trade, regional integration, and international conventions amongst many others). Noting that new theories and sub-theories of IR have been brought on by particular phenomena (such as the balance of power by wars in early modern Europe, liberalism for post-WWII order, postcolonial theory by the independence waves, and constructivism by the late Cold War, as well as green theory by ecological changes) the question posed by the book is whether the advent of the ubiquity of AI in IR necessitates a new theory, or whether extant IR theories can be utilised in explaining and generating insights on AI in IR. This is a question ultimately for the future, but we think that these theories have considerable utility, with some modifications which we demonstrate. It is also made clear that the book is not aimed at comparative theory testing (CTT), and for the advocacy of one theory over another, but rather to assess the variety of ways in which these theories cohere with AI and how each does not. Thus each theory is given sufficient description, with concise information made on its context of emergence, the gaps it closes and the seminal texts it stands on, and in turn the relevance of AI is integrated into the frameworks it sets out.

AI has been in the works at least since the 1950s. In recent decades, however, it has found increasing application in commercial, administrative, and other procedures and interactions. Moreover, its current capabilities, already substantial, are projected to grow substantially in the coming years and decades. By one estimate, artificial intelligence is predicted to constitute some \$15 trillion to the global economy by 2030 (Kastner 2021). The onset of COVID-19 has enhanced the uptake of AI through direct and indirect use in track-and-tracing, remote work, digital infrastructure and, through contingency, the growth of fifth generation (5G) connectivity among many other reasons. This is crucial to consider; indeed a major booster of the prevalence of AI is the growth of related technologies such as the Internet of Things (IoT) and blockchain in addition to 5G itself. This convergence is the impetus for the fourth industrial revolution (4IR [see Schwab 2016; Marwala 2020]) and is mired in geopolitics (Ndzendze 2020; Ndzendze and Marwala 2021). This explosion in the uptake of these emerging technologies has led to vistas of new literature as different scholars seek to understand the implications for their areas of study, as well as to make use of the methodological opportunities presented by AI for their field-specific research (for example, there has emerged the “digital humanities”). Adams et al. (2021), for example, examine the 4IR’s implications for South Africa’s human rights. On the other hand, Njotini and Mpedi (2021) consider its impact on legal education, while Gleason (2018) does so for the higher education industry as a whole. Other notable writings include health (Mazibuko-Makena 2019), impact on economies and livelihoods (Mazibuko-Makena and Kraemer-Mbula 2020), and for trade policy, particularly manufactured exports (Cilliers 2019). DeepMind, Google’s AI lab, has also used AI to improve the study of ancient history. Yannis Assael, a DeepMind research scientist, in 2019 published a paper with Oxford University historian Thea Sommerschild on a deep learning model called Pythia, designed to *fill in the gaps* in missing text from ancient Greek inscriptions (Assael et al. 2019; Minsky 2020). Matthew Gentzkow has written on media and artificial intelligence, in which he described the emergence of automated journalism (Gentzkow 2018).¹ On the other hand, Nah et al. (2020) consider the implications for communication studies.

¹ For example, “On March 17, 2014, a magnitude 4.4 earthquake shook southern California. The first story about the quake on the LA Times’ website—a brief, factual account posted within minutes—was written entirely by an algorithm.” See also Oremus,

As an emerging area of study, AI in IR has been benefitted by some excellent foundational texts. Many of these books are empirical issue-by-issue analyses of AI in IR (e.g., comparative assessments of submarines in WWI vis-à-vis drones today, applications in combatting drug smuggling, the emergence and leadership role of China, and implications for methodology). Henry Kissinger, one of the leading thinkers in the field and former US secretary of state, wrote an article on artificial intelligence (‘How the Enlightenment Ends’) which appeared in the June 2018 edition *The Atlantic*. In it, he warned that “philosophically, intellectually—in every way—human society is unprepared for the rise of artificial intelligence” (Kissinger 2018). Kissinger’s article puts the emergence of AI in a longer historical context of technological evolution, especially in the storage, operationalisation, and transmission of information since the dawn of the printing press:

Heretofore, the technological advance that most altered the course of modern history was the invention of the printing press in the 15th century, which allowed the search for empirical knowledge to supplant liturgical doctrine, and the Age of Reason to gradually supersede the Age of Religion. Individual insight and scientific knowledge replaced faith as the principal criterion of human consciousness. Information was stored and systematized in expanding libraries. The Age of Reason originated the thoughts and actions that shaped the contemporary world order. (Kissinger 2018)

Kissinger, proceeds to note an emergent danger with the increasing complexity and capabilities of AI:

But that order is now in upheaval amid a new, even more sweeping technological revolution whose consequences we have failed to fully reckon with, and whose culmination may be a world relying on machines powered by data and algorithms and ungoverned by ethical or philosophical norms. (Kissinger 2018)²

Will. 2014 (March 17). “The first news report on the L.A. earthquake was written by a robot,” *Slate*.

² The trouble, as Kissinger sees it, begins with Internet culture: “The digital world’s emphasis on speed inhibits reflection; its incentive empowers the radical over the thoughtful; its values are shaped by subgroup consensus, not by introspection. For all

More specifically, Kissinger’s concern stems from the black-box nature of AI (see also Chapter 3):

Paradoxically, as the world becomes more transparent, it will also become increasingly mysterious... How will we manage AI, improve it, or at the very least prevent it from doing harm, culminating in the most ominous concern: that AI, by mastering certain competencies more rapidly and definitively than humans, could over time diminish human competence and the human condition itself as it turns it into data. (Kissinger 2018)

In April 2021, Ariel Kastner, lead researcher for the World Economic Forum (WEF)’s Geopolitical Agenda program summed up seven views prevalent in the literature on how technology will shape geopolitics. These views are:

- ‘We need to agree on norms and rules’
- ‘We may see a further erosion of interconnection’
- Tech companies as ‘the battleground for geopolitical influence’
- ‘Democracies need data sharing, ‘common standards,’ ‘technological infrastructure’
- ‘We must address challenges ‘jointly and across borders’
- ‘We must work together to address both the vast benefits and the enormous risks of data’
- ‘The basis of unnecessary and dangerous geostrategic conflict’.

A Chatham House report edited by Cummings et al. (2018) titled *Artificial Intelligence and International Affairs* similarly anticipates disruption. This report, like numerous other works which have come in its wake, considers the effects of AI in IR on a sectoral basis (particularly war, human security, and economic implications). Specifically, Heather M. Roff (2018), in her chapter, makes the following two points. Firstly, AI can have positive results for human security, but this will first require understanding its potential roles so that its positives can be enhanced. Secondly, this requires us to understand the distinction between using data and knowledge. As we will uncover in Chapter 2 of this book, this second point is at the heart of contemporary debates among AI scholars and

its achievements, it runs the risk of turning on itself as its impositions overwhelm its conveniences.”

developers. Kucier (2018) notes the deepening of wealth inequality at the behest of automation, and the implications that this will have for both domestic and international politics. What seems to be emerging is that AI is reinforcing the owners of capital and disempowering the working class and this can potentially lead to a massive increase in inequality.

On the other hand, a National Bureau of Economic Research (NBER) article by Brynjolfsson et al. (2018) finds that the introduction of an AI-based machine translation system (eBay Machine Translation) has significantly increased international trade on a major trading platform (eBay's international trade), by some 17.5% in terms of exports to Spanish-speaking countries in Latin America (control group, non-Spanish, countries such as Brazil and European Union member states saw no difference, showing a strong likely causal effect). In sum, their findings seem to indicate “[that] language barriers significantly hinder trade, and that AI has already begun to improve economic efficiency in at least one domain” (Brynjolfsson et al. 2018: 2).

Other notable works confronting the contemporary and future impact of technologies in IR include the following. *Technology and International Relations: Challenges for the 21st Century* by Bhaskar Balakrishnan (2017), which examines the current and future role of diplomacy as an equalising factor in the transfer and acquisition of science and technology in general. *Future War: Preparing for the New Global Battlefield* by Robert H. Latiff (2017) looks at the transformative aspects of emerging technologies, including AI, and their effects on military ethos, capabilities, and civil-military relations. *Philosophy and Theory of Artificial Intelligence* edited by Vincent C. Müller (2018) grounds AI in various facets of philosophical questions with a particular focus on ethics. The interdisciplinary volume titled *The Political Economy of Robots: Prospects for Prosperity and Peace in the 21st Century* edited by Ryan Kiggins (2018) deals comprehensively with the effects of robots on IPE, including on trade, norms, “state-sanctioned robot violence,” and institutions among others. The book *Technology and Agency in International Relations* edited by Marijn Hoijtink and Matthias Leese (2019) turns to governance and knowledge among others (focusing on satellite imagery, digital payments, and drone systems). Finally, *Artificial Intelligence and International Politics* by Valerie M. Hudson (2020) and *Artificial Intelligence and Emerging Technologies in International Relations* by Ndzendze and Marwala (2021) study the specific roles of AI in international relations.

In this book, we close two crucial gaps. Firstly, we place particular emphasis on AI as a specific technology of knowledge and action. Secondly, this book presents a comprehensive treatment of how AI impacts or interacts with all established IR theories. To contribute to the field, this volume presents a theory-by-theory assessment and seeks to comprehensively articulate the implications of the growing ubiquity of AI in international relations for the central assumptions and mechanisms of IR's major theories. Given the above texts, our volume will harness and integrate the strengths of the existing literature, while also greatly improving on some of their shortcomings. Crucial to this is a review that is truly international in scope (i.e., taking stock of AI-related developments in every region, including Africa, Asia, Europe, Latin America, the Middle East, North and South America, which are presently under-covered or not fully integrated in the current literature) and making contributions based on empirical findings on theory testing and bottom-up typology-building for theoretical dialogue in the age of AI.

1.2 SUMMARY OF THE BOOK

Scholars of international relations, at a minimum, seek to describe the world empirically and present models for the future. Over the past several decades, as the field has become its own area of specialisation, two powerful paradigms and several critical ones have come to the fore. The former two, Realism and Liberalism, have internal debates over methods and patterns. Marxism, constructivism, postcolonial theory, green theory, and feminism—the critical theories—have vast differences among each other but are unified by their critical and metanalytical focus which sees them deconstruct the underlying assumptions in the field. Throughout, however, there is a commitment to the possibility of objective observation to inform analyses.

Realism seeks to focus on the relationship between states. Liberalism seeks to understand relations among states, but also among states and non-state actors that have come about as a result of progress and globalisation. Marxism presents a model of the world in which there are exploitative power asymmetries between the economic haves and have-nots based on observations about capitalism. Constructivism is concerned with the ways in which the world is represented and understood in ways aligned with power interests. Postcolonial theory, on the other hand,

questions the extent to which there is true sovereignty enjoyed by countries which are former possessions, and notes the historically determined (path dependency) patterns of exploitation between former metropolises and colonies. The Feminist paradigm is based on the gendered nature of such occurrences as conflict and increased economic activity. Additionally, the latter two have germinated out of dissatisfaction with assumptions of universality by Western male scholars. The aim is to distil a more accurate picture of the world by shedding light on those aspects of international affairs which are outside the scope, and interest, of the mainstream paradigms.

To be sure, different theories can end up independently reaching similar or related conclusions; whereas a theory can, when taken to its fuller logical conclusions, yields the opposite results its originators intended for it. Over the decades of IR theorising, there have been numerous world-changing developments, including the Second World War, the Cold War, the attainment of independence in Africa, Asia, and the Caribbean, the fall of the USSR, the rise of China.

Artificial intelligence is poised to be one such world-transforming development.

We survey the impact in the following five spheres of IR. On diplomacy: States have responded to the opportunities and threats presented by AI. They have set up structures to mitigate the threats and accentuate the opportunities while noting AI as a functional mechanism. This has been manifest in the introduction of embassies to Silicon Valley to engage with non-state actors (companies), such as those by Denmark and the UK. Non-state actors, increasingly accumulating their own leverage, have sought to frame international regulations. IR theory has a task to understand and explicate the implications of these. Realists, being state-centric, will have to contend with the prominence of private corporations, and the civil society organisations seeking to hold them accountable. All the while, they will be advantaged by the regulatory environment being set up by states. On the other hand, liberals will find coherence with the initiation of these regulations by non-state actors, and the preponderance of private sector-derived AI over decisions by states. AI presents unique dilemmas through its shaping of the information landscape. On war: The conduct of war decisions, and combat itself, are increasingly made through the aid of AI. Here, too, non-state actors have played a role and are leading in the research and development of war-applicable AI. On the other hand, the incentives for peace are bound to be revised by the onset

of changes to the global economic structure at the behest of automation in commercial processes. On trade: International commerce is being transformed at a rapid pace by the onset of digital trade. Here AI has been used to inform market-related decisions. With changes in the structure of the global economy, we anticipate some knock-on effects on states. Do market fundamentals as advanced by liberalism still hold or is it time, as some have argued with growing ferocity in recent years, for a re-writing of the social contract so that there are some measures of protectionism? Given also that AI has been the subject of geopolitical dispute, we are seeing free trade scaled back. In turn, this renders the countries which are democratic paradoxically more likely to impose restrictions on technology from certain regions of the world. With liberalism failing to provide an account of this, we identify the value of thinking through the lenses of the English School. On developmental considerations: AI's economic facets, including deployments in production, have already become significant factors of consideration in development issues. This includes FDI and trade flows, as discussed, but from the perspective of growing inequality and reproduction of contemporary structures of asymmetry. As some countries pass key turning points before others, dependency theory's insights will be both deployed but also need reconsideration in light of the AI-era pathways. The benefits of AI, however, are not evenly spread even among the first countries to automate. On inequality (and bias): Being reliant on existing datasets, AI is characterised by bias along the lines of gender, race, and geographical location. Numerous IR theories take a critical stance and can be relied on to explicate the causes behind these and how the world of IR will look as IR becomes more integrated into AI. The critical theories have done much to elucidate the sometimes ill-equipped nature of AI to be inclusive. At the same time, however, they also have much potential to further exploit by way of integrating one another's insights. Indeed this holds for many of the theories studied in the book as a whole.

Evidently, AI has implications for the *raison d'être* of every major and emerging theory in IR. Each theory should have something to say about its effects on processes and outcomes. Will these developments, then, lead to a fundamental rethink of these theories and IR theory itself? To some extent. If the debate seen in the social sciences regarding AI is an indicator, there will be some scholars who will seek to reconceptualise the theories, and others who will downplay its significance. Moreover, these will be individuals writing within the same theoretical paradigm.

The world of IR paradigms is composed of vast, lively internal debates. There is also a lack of clarity on the extent and nature of AI in world affairs, partly because the “international” and the domestic are no longer strictly defined. It is perfectly to be expected, then, that some will see nothing notable—or paradigm-altering—in AI-linked developments. It is also to be expected that others will see major significance as the world enters a phase with AI ubiquity.

Chapter 2 provides an overview of international relations theory. It does so through an in-depth discussion of the nature and purpose of theorising in the field. The chapter provides an analysis into the history and current dynamics in IR theory. This is done through a review of the great debates that defined the early discipline, the post-Cold War crisis as well as changes and patterns in the methods that IR theorists use to engage with these paradigms, as well as in the meta-analysis of their field. Crucial to this chapter is the elemental nature of theory to the field of international relations, and of actors and circumstances to that theorising. The proliferation of several theoretical paradigms has resulted in ongoing debates, partially as a result of disagreements over the relative roles of the actors and of the inadequacy of the more traditional theories. This has given rise to theoretical dialogue, with scholars engaging the notion of borrowing from two or more theories in the construction of typological theories, as well as simplistic hypothesis testing. The emergence of AI in IR is set to complicate this analysis, while also promising to simplify it in some respects. Particularly, while AI will lead to methodological sophistication, its presence as an actor by the main units of analysis, it will result in unforeseen complexity. In Chapter 3, we provide an overview of AI, delving into its history, contemporary capabilities, and future directions. The chapter begins with an overview of the nature and history of artificial intelligence, with a focus on its evolution, core tenets, trends (including “AI winters” and resurgences and the state of affairs in present terms). It then examines the relevance and timeliness of researching AI in social science in general and IR in particular, given the implications that AI carries on account of its applications. These lead to an interface with society and operate across national boundaries. In light of these, the third section examines recent and contemporary implementations of AI in international relations, while the fourth section gives a more specific analysis of AI’s actual and potential implications of AI for trade, war, diplomacy, and AI’s relevance to FDI. The fourth section looks at AI

as a potential unit of analysis in IR (utilising debates in Ethics and Law on liability) and how it interacts with the Waltzian concept of images (levels of analysis); asking whether it is ultimately only a tool as some have argued, or another level of its own, or both in certain configurations.

In Chapters 4–9 we respectively study AI through the prisms of realism, liberalism, hegemonic stability, dependency theory, the English school, and three critical theories. Realism (Chapter 4) is the most established theoretical paradigm in International Relations scholarship. It has arguably shaped the foreign policies of successive governments throughout the world, most prominently the superpowers of the ages, from Sparta to eighteenth-century France and England to modern-day America. But this approach to understanding global affairs is not without internal divisions. Indeed, the story of realism has been the evolution of its tools of analysis from a focus on human nature to structural factors. This structural realism is also characterised by debates among those who argue for a defensive posture that should limit a state’s power capabilities only to satisfy its security needs, on the one hand, and another offensive approach which argues that power competition is the inevitable, and logical, action of states. A new variant (neoclassical realism) seeks to incorporate the role of domestic factors and capabilities to enrich our understanding of states’ varied responses to system-level developments. We present a model for measuring an “AI balance of power” in this chapter. In our analysis, of liberalism (Chapter 5) we present a useful set of lenses for understanding AI’s economic and transnational economic facets. One of the emerging patterns of the past few years is that the onset of geopolitical tensions over AI, which has led to the desire to curb access to “foreign AI,” thereby leading to democracies being marginally less liberal. More fundamentally, AI is changing liberal assumptions of interdependence and domestic audience participation. In sum, AI will make states less dependent on one another—and AI will cause domestic participation (the key to the democratic peace thesis) to operate under ambiguity and without the ability to discern fact from fiction while also being subjected to manipulation both domestic and foreign.

References to hegemony and emerging applications of AI and the perceived race for AI leadership between the United States and the PRC have been made by leading IR scholars and theorists. But in so doing, they have mainly been in passing. In other words, they have not been in a systematic and theory-generating or theory-testing method with implications for hegemonic stability theory (HST). Chapter 6 seeks to close

this gap. Situating the social dynamics of the next number of decades as being headed for greater ubiquity of cyber-physical interfacing with AI, in particular, the chapter argues that this necessitates an extension of the understanding of hegemony to include the future shaping of AI ethics, as well as a new operationalisation of HST that may link the maintenance of peace and outbreaks of conflict to *changes* in the global distribution of AI.

Given the pre-existing economic and technological advantages of some nations over others (in some respects an outcome of the first three industrial revolutions), it is worth examining the extent to which global capital is moving across these similar patterns as the world enters the fourth industrial revolution, and thereby fortifying pre-existing economic parities (a pattern which would be akin to the so-called Matthew effect and in keeping with the insights made by dependency theory in terms of technology flows). Findings in Chapter 7 have implications for how we are to understand dependency in the age of AI, with a window of opportunity noted for LDCs in the coming decade. Chapter 8 posits the contributions of the English School in a world of growing AI ubiquity and AI geopolitics. One of the critical contributions of the theory is the notion of international society. With growing geopolitics around AI, the English School can be a valuable lens for understanding the contingent nature of fault lines and divergences, with apparent mutual side-lining between China and Russia one side (themselves with their own subtle differences) and the United States and its allies on the other. In essence, the definition of who belongs and who does not belong to “mainstream” international society (or world community) is playing out in the digital realm as well in the era of AI. Chapter 9 assesses the predictive power of critical theories of international relations in the era of AI. These theories—constructivism, postcolonial theory, feminism, and green theory—differ in their focus, but they all demonstrate the inequitable nature of AI’s distribution within states and in transnational settings. Moreover, many of their assertions have anticipated the rollout of these technologies along the lines of gender and race, and to have ecological casualties. Their critical stance, combined with their mid-range nature (as opposed to being paradigms) and engagement with non-state actors, means they are closer to policy intervention through activism. In Chapter 10, we conclude the book with an evaluation of the findings made in individual chapters and what they might mean for the field as a whole. We thus present the emerging areas of opportunity in theoretical scholarship and pedagogy of IR.

BIBLIOGRAPHY

- Adams, R., Michael Gastrow, Nokuthula Olorunju, Mark Gaffley, Yuri Ramkissoon, Shanelle Van Der Berg, Fadlah Adams, and Tseliso Thipanyane. 2021. *Human Rights and the Fourth Industrial Revolution in South Africa*. Johannesburg: HSRC Press.
- Assael, Yannis, Thea Sommerschild, and Jonathan Prag. 2019. Restoring Ancient Text Using Deep Learning: A Case Study on Greek Epigraphy. DeepMind website. <https://deepmind.com/research/publications/Restoring-ancient-text-using-deep-learning-a-case-study-on-Greek-epigraphy>.
- Balakrishnan, Bhaskar. 2017. *Technology and International Relations: Challenges for the 21st Century*. New Delhi: Vij Books India.
- Brynjolfsson, Erik, Xiang Hui, and Meng Liu. 2018. Does Machine Translation Affect International Trade? Evidence from a Large Digital Platform. NBER Working Paper No. 24917. https://www.nber.org/system/files/working_papers/w24917/w24917.pdf.
- Cilliers, J. 2019. *Made in Africa: Manufacturing and the Fourth Industrial Revolution*. Pretoria: Institute for Strategic Studies.
- Cummings, M.L. 2018. Artificial Intelligence and the Future of Warfare. In *Artificial Intelligence and International Affairs: Disruption Anticipated*, eds. M. L. Cummings, M. Roff, Heather, Kenneth Cukier, Jacob Parakilas, and Bryce Hannah, 7–18. Chatham House. <https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial-intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf>.
- Cummings, M.L., Heather M. Roff, Kenneth Cukier, Jacob Parakilas, and Hannah Bryce (eds.). 2018. *Artificial Intelligence and International Affairs: Disruption Anticipated*. Chatham House. <https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-14-artificial-intelligence-international-affairs-cummings-roff-cukier-parakilas-bryce.pdf>.
- Gentzkow, Matthew. 2018. Media and Artificial Intelligence. TNIT Working Paper. <https://www.tse-fr.eu/publications/media-and-artificial-intelligence>.
- Gleason, Nancy W. 2018. *Higher Education in the Era of the Fourth Industrial Revolution*. Singapore: Palgrave Macmillan.
- Hoijsink, Marijn, and Matthias Leese (eds.). 2019. *Technology and Agency in International Relations Emerging Technologies, Ethics and International Affairs*. Oxfordshire: Routledge.
- Hudson, Valerie M. 2020. *Artificial Intelligence and International Politics*. Oxfordshire: Routledge.
- Kastner, Ariel. 2021. 7 Views on How Technology will Shape Geopolitics. *World Economic Forum*, April 7. <https://www.weforum.org/agenda/2021/04/seven-business-leaders-on-how-technology-will-shape-geopolitics/>.
- Giggins, Ryan (ed.). 2018. *The Political Economy of Robots: Prospects for Prosperity and Peace in the 21st Century*. New York: Palgrave Macmillan.

- Kissinger, H. 2018. How the Enlightenment Ends. *The Atlantic*. <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>.
- Kucier, K. 2018. The Economic Implications of Artificial Intelligence. In *Artificial Intelligence and International Affairs: Disruption Anticipated*, eds. M.L. Cummings, Heather M. Roff, Kenneth Cukier, Jacob Parakilas, and Hannah Bryce, 29–42. London: Chatham House.
- Latiff, Robert H. 2017. *Future War: Preparing for the New Global Battlefield*. New York: Knopf.
- Marwala, T. 2020. *Closing the Gap: The Fourth Industrial Revolution in Africa*. Johannesburg: Macmillan.
- Mazibuko-Makena, Z. 2019. *Epidemics and the Health of African Nations*. Johannesburg: Mapungubwe Institute for Strategic Reflection.
- Mazibuko-Makena, Z., and E. Kraemer-Mbula. 2020. *Leap 4.0: African Perspectives on the Fourth Industrial Revolution*. Johannesburg: MISTRA.
- McCarthy, D. 2015. *Power, Information Technology, and International Relations Theory: The Power and Politics of US Foreign Policy and the Internet*. New York: Palgrave Macmillan.
- Minsky, Carly. 2020. How AI Helps Historians Solve Ancient Puzzles. *Financial Times*, July 1. <https://www.ft.com/content/2b72ed2c-907b-11ea-bc44-dbf6756c871a>.
- Müller, Vincent C (ed.). 2018. *Philosophy and Theory of Artificial Intelligence*. New York: Palgrave Macmillan.
- Nah, Seungahn, Jasmine McNealy, Jang Hyun Kim, and Jungseock Joo. 2020. Communicating Artificial Intelligence (AI): Theory, Research, and Practice. *Communication Studies* 71 (3): 369–372.
- Ndzendze, Bhaso. 2020. ‘Huawei-Google Clash is an Opportunity for African Search Engines.’ *Business Maverick*. <https://www.dailymaverick.co.za/article/2020-07-28-huawei-google-clash-is-an-opportunity-for-african-search-engines/>.
- Ndzendze, Bhaso, and Tshilidzi Marwala. 2021. *Artificial Intelligence and Emerging Technologies in International Relations*. Singapore: World Scientific Press.
- Njotini, M., and L.G. Mpedi. 2021. *Law and Industry 4.0: Selected Perspectives on a New Scholarship of Teaching and Learning*. Cape Town: LexisNexis.
- Oremus, Will. 2014. *The First News Report on the L.A. Earthquake was Written by a Robot*. *Slate*, March 17.
- Roff, H.M. 2018. Advancing Human Security Through Artificial Intelligence. In *Artificial Intelligence and International Affairs: Disruption Anticipated*, eds. M.L. Cummings, Heather M. Roff, Kenneth Cukier, Jacob Parakilas, and Hannah Bryce, 19–28. London: Chatham House.
- Schwab, K. 2016. *The Fourth Industrial Revolution*. New York: Penguin.
- Szkarłat, Monika, and Katarzyna Mojska (eds.). 2016. *New Technologies as a Factor of International Relations*. Cambridge: Cambridge Scholars Publishing.



Theory in International Relations

2.1 INTRODUCTION

Scholars of IR frame the mainstay of their discussions in theoretical terms. They generate and test theoretical accounts of their universe; a world of states and non-state actors interacting with each other in one form or another. The theories differ in their basic assumptions of what exists, what should be analysed and the conclusions they reach about the world. But they are brought together by a common interest in describing international politics with accuracy. There is widespread disagreement as a result. There is disagreement, that is, on what exists (or what the most important level of analysis is between individuals or states in the international system), what causes certain outcomes, and on what ought to be done (and, indeed, whether anything should be done in the first place). This is not to imply purely compartmentalised thinking, though there is enough of it to go around; disagreement exists not only between theories but is alive and well within theoretical paradigms. That is, scholars within a theoretical tradition may agree on the fundamentals. Crucially, however, they still find a substantial degree upon which to agree, including on methods, the application of their theories, the implications of certain events, and the scope for borrowing or collaborating across the theoretical aisle. This is the basis for the numerous journals which serve as platforms for discussion, convergence, and divergence. Occasionally, many

axiomatically disparate theoretical accounts may end up reaching similar conclusions about the world.

This chapter unpacks the nature of theory in the field of IR with the aim of evaluating its readiness to account for the rise of artificial intelligence. To do so, we discuss the nature and purpose of theory in IR in Sect. 2.2; we then turn to the major debates in the history of the field in Sect. 2.3. Section 2.4 discusses the theory in the post-Cold War and post-9/11 eras and the crises these events appeared to present for IR thinking, as well as the evolution which came about in their wake. In Sect. 2.5 we turn to methods and their role in theory in this field. In Sect. 2.6 we discuss the concept of theoretical dialogue in IR. Section 2.7 examines criticisms of mainstream IR theory, and we conclude with an overview of the opportunities presented by the presence and growth of AI in IR.

2.2 THE NATURE AND PURPOSE OF THEORY IN INTERNATIONAL RELATIONS

As seen, International Relations is a field dominated by theory. Theories are ubiquitous. Hoffman writes that “even the most hardened, detail-oriented, empirically grounded analyst or diplomat is a theorist, if only implicitly, with ideas about how actors behave and how the world works” (Hoffmann 2003: 37). What is meant by this? In other words, what is a theory in our field? At a minimum, theory is a systematic explanation of events in the world. It is able to offer general accounts of the world through consistent patterns. In this way, theory is deduced from observation and systematic analysis of actors in specific circumstances. Theories, therefore, are accounts of patterns in international affairs. These have mainly to do with conflict and cooperation in what is considered an anarchic world (i.e., states themselves being the highest level of authority and accountable only to themselves and thereby choosing only to cooperate voluntarily with other states and international organisations and laws). This makes trade and war the two main preoccupations of IR theory, flanked by questions around voluntary adherence to international laws seen in the international community (in light of the underlying anarchy), and international cooperation on a contingent or long-standing issues. On the main, there are two mainstream IR theories with centuries-old origins, and about six theories that have come to the fore in the past five decades. The first two are realism and liberalism (hegemonic stability theory containing elements of both). The latter six are Marxism

(with associated theories such as dependency theory), postcolonial theory, constructivism, feminism, the English school, and green theory.

As the oldest theory in IR, realism can trace its origins back to ancient Greek thought. Specifically, the ideas of the historian and military general Thucydides, who wrote the history of the decades-long conflict between his native city-state of Athens, an emerging power at the time, and the powerful Sparta, are looked to as having laid the groundwork for this perspective for the analysis of international relations in the fifth century BCE. His recognition of the role of power, security thinking, and self-interest form the core of the assumptions upon which realism rests. While diverging on many important aspects in the form of classical, structural (and in turn defensive and offensive branches), and neoclassical realism, these various strands have common assumptions about the international order. For example, all realists assume that the nation-state is the principal actor in international relations, as opposed to non-state actors; in turn, realists are the most insistent among the theories that the world is in a state of anarchy in which the nation-state is not compelled by any binding higher body. By contrast, some liberals, and in turn the English School, consider the world to be less so. Contradicting all these, dependency theorists see the world as a single global economy that is characterised by the specialisation of labour and inequality. As a consequence of their staunchly anarchic view of the world, structural realists, the dominant perspective within realism today, argue that the international arena can be said to be the manifestation of the individual pursuits of their self-interest by all states. Structural realist theory is in two broad strands, defensive and offensive realism. The source of divergence stems from different readings of the intentions of state actors towards the outside world. One branch of the realist school of thought argues that the state is principally concerned with ensuring its own security and the other asserts that the general anarchy in the international arena encourages states to pursue limitless power.

Liberalism, originally termed idealism in its immediate post-WWI inception, is a paradigm of politics which as a theory of international relations, argues that the world can, and should be, characterised by greater cooperation and increased trade. The insistence on what *should be* rather than a restricted focus on describing what is, is termed “normativism” and is shared by liberalism and the critical theories. This makes realism the only exclusively positivist (i.e., purely descriptive) theory in IR. Unlike critical theories such as postcolonial, dependency, and feminism theories,

on the other hand, liberals advocate for reform rather than a fundamental rethinking of international relations. In other words, they are inclined towards working within the established system and retaining its positive facets rather than engineering a total overhaul.

Adherents of liberal theory argue that increases in trade can lead to interdependency and thereby make military confrontation undesirable and too costly to seriously contemplate. Like realists, liberal thinkers assume the natural position of rational self-interest on the part of their actors. Where they differ is in the unitary nature of the state. Liberals persist that there is a role for non-state actors such as individuals, corporations, and intergovernmental organizations and that it would be impossible to even attempt understanding the world without accounting for these. Indeed, liberals encourage the proliferation of these in the era of globalisation, *insisting* that they facilitate the interdependent world system needed to curtail war and foster economic development on a universal scale. This insistence should stand out in another, more fundamental sense; unlike realists, liberal thinkers have a bottom-up view of the state. They see the state as being the product of a collection of individual interests who decide to enter into organised membership (citizenship) within a state that is run by a government. Liberalism is a sophisticated paradigm that has also made penetrating observations about the world. One of these is the liberal democratic peace thesis. This describes the historical lack of any wars between or among liberal democracies. It is one of the main points that liberals use to argue for the adoption of liberal principles on the domestic and international scenes. However, it is not universally accepted and neither are its mechanisms, including among liberals themselves (we shall return to this in Chapter 5).

For its part, constructivism is a theory of international relations which assumes that the world is in the mind of people and not an objective reality as realists, and even liberals, persist. Adherents argue that the world is essentially a set of social constructs. For example, while the physical territory of Eastern Europe is the same, the entity which was once called the Soviet Union was dissolved in the 1990s and so the territories were renamed and with that the attitudes towards, and perception of, the region on the map changed. Additionally, the concept of anarchy, which refers to the absence of any structural form approaching a “world government” other than a litany of voluntary principles and organisations which states become party to, is treated by realists and liberals as

a natural outcome of this constellated pursuit of individual national interests by the various state and non-state actors. Constructivists on the other hand view this as a socially constructed phenomenon. Constructivists note that the construction of the “international” as a philosophical conception requires shared (or imposed) conceptions of global governance. In the absence of a such global government, this is the regulatory framework brought about by the set of voluntary international institutions, mechanisms, norms, values, agreements, and procedures which most states in the international system abide by, this is a powerful explanation. The implications of this can range from trade (mediated by the World Trade Organisation), conduct in war (guided by the Geneva Conventions), to ideas of justice (imposed by the International Criminal Court) among others. Integral to this paradigm, then, is the idea that events and institutions not only occur but are also given meaning to a substantial degree that is not delved into by the other mainstream theories.

One of the common features that transcend theoretical paradigms in IR is the use of multiple levels of analysis (i.e., individual, state and the international system). In a 2015 article, Temby makes the argument that levels of analysis are relational; each is defined by its effect upon the others, while being itself affected by them. This view expands them from being ontological entities alone and allows them to be methodological tools (Temby 2015: 721). More specifically, Temby shows that “the methodological issue of which levels of analysis a researcher employs is separate from the ontological issue of whether the theoretical lens is atomistic (reductionist) or holistic at any given level. One implication of this definition is that researchers need not view their ontological commitments as overly methodologically constraining.” Temby’s work addresses the phenomenon of multiple social structures which can co-exist within the same level of analysis (Temby 2015: 721). Below we turn to the debates which birthed and continue to define the field and its theories.

2.3 THEORETICAL DEBATES OF THE TWENTIETH CENTURY

Though its cannon dates back millennia, the field of international relations as an independent academic area of study and specialisation is relatively recent. The Department of International Politics at Aberystwyth University in Wales was the first such department. It was established in 1919, with the aid of a considerable endowment of £20,000 donated by David

Davies. The department was established as a memorial to students who were killed or wounded during the course of the war. Davies was inspired by a global vision, against the backdrop of the global war, that was aimed at “repairing the shattered family of nations and, more ambitiously, to redeem the claims of men and women in a great global commonwealth.” The proposed League of Nations featured prominently in Davies’ vision, as did its principal mover, US president Woodrow Wilson. Thus, the world’s first chair in international politics, similarly located at Aberystwyth was named in honour of President Wilson. The first Woodrow Wilson Professor was Sir Alfred Zimmern. An author of multiple works on international affairs—including the essay collections *Nationality and Government* (1919), *Europe in Convalescence* (1922) and *The League of Nations and the Rule of Law 1918–1935* (1936)—was among the pioneering generation of scholars who identified the importance of understanding international affairs through personal contacts and grass-roots democracy, as well as the facilitating role of global civil society in fostering understanding among different nations. But Zimmern’s ideas, which were regarded as the most notable of the interwar “idealists,” were put under scrutiny by E. H. Carr. Carr was by then the fourth Woodrow Wilson Professor, and is the most celebrated. His worldview, “realism” as it came to be called, was articulated in his book, *The Twenty Year’s Crisis* published in September of 1939, just as WWII was breaking out, gave harsh criticism towards what it termed a liberal “utopianism” and placed emphasis instead on the importance of power. For some, this went against the original intent of Davies’ original aims with the chair (Cazacu 2013: 2).

In this way, the history of the field is understood as having undergone debates in the twentieth century. According to the traditional historiography of the discipline, the damage caused by the Great War and the overwhelming determination by key players to prevent another such catastrophic conflict led to the inception of a field and a school of thought that is today called “idealism.” In this telling, idealists had a naïve sense of optimism towards the prospect of ending war perpetually, and had a belief in the role of supranational entities to achieve this end. They were also confident in the attainability of a harmony of interests among states despite the anarchic nature of global politics. This “utopian” phase in the history of the field was then surpassed by a decidedly more “realistic” one in the wake of geopolitical developments such as the Japanese invasion of

northern China (Manchuria) and the League of Nations' failure to act against it (Villanueva 2012: 1).

It is during this period that the IR traditional historiography affirms that realist scholars, such as E. H. Carr and Hans Morgenthau, had an intellectual confrontation with 'idealists' (i.e. what is conventionally called in the discipline the 'first great debate') and successfully crushed their utopian ideas by stressing that they had neglected the issue of power in real world events. (Villanueva 2012: 1)

But this historiography has come under increasing reconsideration by historians of the field. Problematising the conventional chronology, Osiander is able to demonstrate that the early twentieth-century thinkers were already engaged with many of the ideas that would eventually fall under the uniform umbrella of realism from 1939. Earlier Vasquez (1983: 13) had asserted that "realism did not come at them in 1939 as a bolt from the blue, as a consequence, as it is sometimes suggested, of some kind of Kuhnian scientific anomaly" (Vasquez 1983: 13).¹ To the contrary, it was argued, earlier thinkers were "thoroughly aware" of ideas which were characterised through the terminology of "international anarchy," "state sovereignty," "endemic war," and the "balance of power" ([see Osiander 1998: 414; Wilson 2012a: 32]). A sizeable portion of their writing was aimed in opposition to it. For this reason, Osiander argues that the often-referenced "first great debate" between idealists and realists did not occur in the way that the name implies. Rather, there was dialogue (Osiander 1998: 415). In *International Relations and the First Great Debate* (edited by Schmidt), Peter Wilson's chapter, further contends that as "a cohesive and certainly self-conscious school of thought, an "idealist" or "utopian" paradigm never actually existed" (2012a, b: 16). Ashworth concurs and views the presentation of the interwar years as dominated by an idealism as "a mere device of realism" intended to disqualify the contributions of "liberal internationalism" (2012a, b: 60–61). The second key argument of Schmidt's (2012) book is that the epoch of "first great debate" did not actually experience "any meaningful intellectual exchange" between realists and the so-called

¹ "Kuhnian" derives from Thomas Kuhn, in whose *Structure of Scientific Revolutions* he discussed the societal determinants of acceptance of new theories.

idealists (1). If accurate, this is a fundamental update to Osiander's (1998) idea of the debate being a dialogue instead.

In Wilson's argument, therefore, as a teaching device the idea of a first great debate as an episode has no value as it has no historical basis (2012a, b: 16). Contrary to the idea of an "exclusive intellectual confrontation" between realists on one hand and idealists on the other on issues considered by these two schools of thought, Wilson observed that "the period produced interesting theoretical writings on several issues important in the discipline that are frequently outside the topics covered by these theories such as class-based analyses" (Wilson, 2012a, b: 28). Ashworth (2012: 60) is similarly adamant that there is no historical evidence for a contest between realists and idealists but instead establishes that in the 1920s and 1930s there was an internal debate within the discipline on whether the war was endemic to capitalism (60). In the final chapter of the book, Schmidt (2012: 92) contends that the first great debate was in effect "little more than a disciplinary myth." Nonetheless, Schmidt advances some evidence for an intellectual dispute between realists and idealists. But, crucially, he argues that this took place only following WWII, and did not take place in the interwar years and the 1940s as understood in the mainstream (94). The focus on the Anglo-American sphere, also leads to a lack of attention to the German world at this time. Specifically, Villanueva (2012: 3) asserts, the significant strides and insights developed by what was later termed the Frankfurt School, formed in 1923 as the Institute for Social Research is worth a historical review. The School generated penetrating critiques of the "positivism" in the social sciences, including IR, and thereby preceded the emergence of later constructivism as it would later emerge in the 1980s (Villanueva 2012: 3). All these demonstrate the need for continual re-engagement with the field. We turn to this below, in light of the crises brought on by the end of the Cold War and the 9/11 attacks.

2.4 IR THEORY IN THE POST-COLD WAR AND POST-9/11 ERAS

For many reasons, the end of the Cold war was unexpected, and the manner that it did end left realism at a loss. In particular, the USSR dissolved away somewhat peaceably (at least towards the US) and gave way to a globalised economy. This led to a period of liberal triumphalism, in which the proponents of this theory are deemed to have overstated the

onset of a new world order in which liberalism's ideas would not find a conceptual or practical challenge (see Chapter 5). Yet this too did not last, as the 9/11 attacks launched a new period of protracted conflict in which power once again flowed from capabilities rather than ideas of progress. Writing in 2003, Hoffman (37) observed that:

The international relations discipline has had a difficult time providing concepts and frameworks that facilitate understanding or even addressing the discontinuous and non-linear change that rare events like the end of the Cold War and the terrorist attacks of 9/11 potentially engender. (Hoffmann 2003: 37)

Hoffman still finds, however, that this shortcoming is not indicative of whether the theory is relevant, but rather presents the question of *which* theory is the appropriate one for a given situation. The end of the Cold War and the 9/11 attacks were two seminal events, which, in rapid succession, demonstrated the existence of non-linearity to scholars within an ever-evolving science (Hoffmann 2003: 37–38). Thus, IR scholars have come to recognise the need to incorporate and anticipate change and disruption in a way the two traditional theories previously could not. This also enables the theorist to be of substantial policy input given that practitioners operate within an ever-changing landscape (Hoffman 2003: 39). Another recent event which seems to evince a new era in the international system is the retreat of the United States from Afghanistan, followed by a rapid Taliban takeover of the country in August 2021. Does this indicate a new era in which US power is going to be increasingly challenged? Moreover, might it also vindicate the long-held view that technological advantage (as possessed by the US vis-à-vis the Taliban) is not the key factor in military outcomes? Perhaps, or perhaps not. Scholars will ponder these questions for decades. But in so doing they will have to utilise cogent and convincing methods of analysis. We turn to an overview of methods in IR.

2.5 METHODOLOGY IN INTERNATIONAL RELATIONS

Theorists use a multiplicity of tools to conduct their inquiry. Scholars carry out their research with a view to test existing theories or generate new ones. As such, the methods utilised in each study are determined by the parameters, scope and aims of the research. These fall within the broad

qualitative-quantitative domains which define most of the social sciences. Increasingly, scholars have made use of mixed methods in their studies. Our review of papers using either method in leading (high impact factor) IR journals (*European Journal of International Relations* [EJIR], *International Affairs* [IA], *International Organization* [IO], *Journal of East Asian Security and International Affairs* [JASIA], *Journal of International Relations and Development* [JIRD], and the *South African Journal of International Affairs* [SAJIA]) in 2020 showcases the prevalence of qualitative methods over quantitative ones. For EJIR and IO, the leading methods were qualitative followed by mixed methods, with quantitative methods in last place. SAJIA primarily consisted of qualitative methods, followed by mixed methods, and had no quantitative methods. For IA and JIRD, the most common methods were qualitative, followed by quantitative, and then mixed methods. These are represented in Fig. 2.1.

Research in IR has come to largely take the form of case studies or within-case analysis, with growing scepticism around the major theories. As Bennett (2013: 459) observes, “Theorizing under the rubric of paradigmatic “isms” has made important conceptual contributions to International Relations, but the organisation of the subfield around these isms is based on flawed readings of the philosophy of science and has run its course.” Case studies are usually designed to test the causes behind the occurrence of a phenomenon among different polities (country, regional

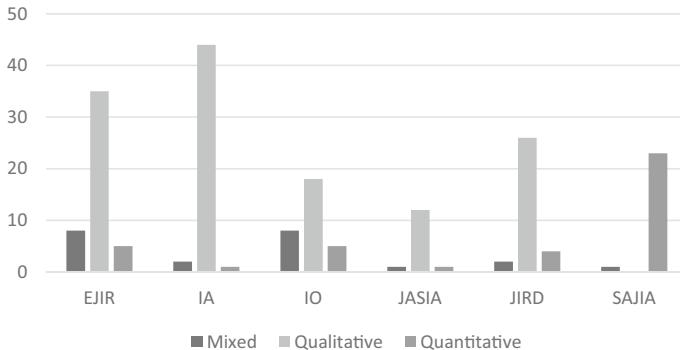


Fig. 2.1 Methods in articles by leading and regional IR journals surveyed (By authors. Data sourced from EJIR, IA, IO, JASIA, JIRD, and SAJIA journal archives)

or international levels) or some other environment. Within-case analysis on the other hand is usually interested in the longitudinal changes within a single environment. This is especially useful for the analysis of outliers (i.e., cases which cannot be examined in comparison to others due to their rarity). There are many such outliers in IR. Some examples may include the use of nuclear weapons in combat (this was only done once, in Hiroshima and Nagasaki in 1945 during World War II) or the lack of colonisation in Africa (only Ethiopia escaped this fate and as such is an outlier country among African countries insofar as this characteristic is concerned). Thus scholars can develop a method to determine some cause and effect relationship of a set of variables within the case in question to reveal answers about the evolution of conditions from a starting point in time to the present point of interest, while they may examine different case studies in order to investigate the causal relevance of a particular variable. The use of different cases also brings about a greater causal connection as it provides scope for falsification. Falsification takes place when the presence of a variable in one context leads to an outcome and its absence in another leads to no outcome. Falsification also holds when the presence of a variable leads to no outcome and its absence leads to an outcome. Scholars may be interested in either of these uses of falsification depending on the causal relationship they are asserting or disproving. For each of these, the researcher may use the appropriate framework.

For quantitative studies, simplistic hypothesis testing is a mainstay. These studies include the use of numerical data, determined with specific variables. Like their counterparts in the rest of the social sciences (psychology, economics, and sociology), IR scholars with a quantitative bent make use of common statistical methods such as regression analysis to determine causal relationships in large N studies—to study, among others, conflict, economic development trajectories, international organisations, and other transnational dynamics such as organised crime, migration and climate change.

In 1966, Rummel wrote *Factorial Analysis* with the aim of popularising numerically enriched studies in IR. (It is to be noted that the use of quantitative techniques has been mostly associated with IR scholarship in the US or US-based journals). Robert K. Yin (2009: 14) defines case study research as “an empirical enquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident.” For within-case analysis, researchers commonly make use of

process tracing. This is a method which involves investigating, over a specified period of time, the changes which occur following the introduction or removal of a specific variable. Process tracing, for example, may be used to investigate the processes which follow the ratification of a treaty within a domestic context of a country. This is commonly done by scholars interested in the domestication of international norms and standards.

Andrew Bennett and Colin Elman (2010) have observed that qualitative research methods currently have “an almost unprecedented popularity and vitality...in the international relations sub-field,” such that they are now “indisputably prominent, if not pre-eminent” (2010: 499). Our Fig. 2.1. attests to this. In their view, this is owed to the numerous advantages that case study methodologies present for those interested in studying “complex and relatively unstructured and infrequent phenomena that lie at the heart of the subfield” (Bennett and Elman 2007: 171). The use of mixed methods, though difficult, can lead to possibly greater causal explanation. It has been observed that while quantitative studies can effectively demonstrate correlation, they cannot by themselves deduce causality. Thus, qualitative analysis is needed to provide a deeper analysis of the apparent relationship. On the other hand, qualitative analysis does not have the level of scale seen in quantitative studies. This, in addition to the burgeoning popularity of inter- and trans-disciplinary work, may explain the growing use of mixed methods in IR, including theory testing and generation.

Methods in IR also can also be used to provide a platform for engagement between theories. A theory commonly arises as a means of dismissing another following findings from a research undertaking. They may subsume those of others or cross-pollinate their ideas through theoretical sequencing and typology-building. We turn to the history of such theoretical dialogue and discuss the practical outcomes in the next section.

2.6 THEORETICAL DIALOGUE IN IR

In their 2003 *Comparative Political Studies* article Jupille, Caporaso, and Checkel (19) propose that theories in political science have a sufficient degree of overlap that ought to enable conducive interoperability. There are our main ways in which conversation takes place between two or more theories. These are competitive theory testing (CTT), sequencing (such that one theory explains one stage of a phenomenon and yields to another

theory for another), incorporation (otherwise termed subsumption), and domains of application (Jupille et al. 2003: 19). Of these, CTT consists of an attempt at refuting one theory at the expense of others upon testing the same basic problem and determining each theory's causal validity. In related fashion, subsumption consists of a stronger theory absorbing those that are less powerful (sometimes through criteria established by the powerful theory of the set) (Jupille et al. 2003: 20). Strong tests are tests within the CTT method which require high thresholds for confirmation, so that a theory will have overcome difficult standards and be that much more satisfactory in its displacing the alternative explanation (Jupille et al. 2003: 22). In this way, theories are granted respective "home turfs," which in turn fit in the larger picture (Jupille et al. 2003: 21). Through sequencing, every theory's claims are specified and the results, if successful, lead to an additive theory that contributes more comprehensively than the individual theories (Jupille et al. 2003: 22). This requires precision in understanding and defining the scope conditions of each theory (Jupille et al. 2003: 22). In our estimation, theoretical dialogue is going to be increasingly crucial in the era of AI, as scholars attempt to grapple with various aspects of it. As an all-encompassing general-purpose technology (GPT), AI needs a renewed eclecticism. Understanding state-level incidents will be to the detriment of effects on the climate, which in turn would require synchronised understanding of colonial legacies, gender, while also factoring in regime types. To be sure, eclecticism in IR has a history, and has enjoyed largely mixed results. IR theory, as seen in its geographical blind spots, is still refining itself. We review one such major blind spot in the next section and what it means for the future of AI studies in IR.

2.7 META-THEORY IN IR

One of the main critiques of the field are the racial undertones in IR theoretical thinking put forth by numerous works. For his part, W. E. B. Du Bois argued that Western IR cannot explain Africa's continental processes, or shape its future (1946). At the root of this, he argued, was a fundamental lack of interest in the continent. Ali Mazrui on the other hand observes the continued domination of the continent by its former colonial administrators, including in the education systems that its countries have in place which he regards as "mechanisms for further westernization" (1980: 46). In his 1980 lectures titled "The African Condition: A

Political Diagnosis,” he pointed out the inadequacy of these lenses for thinking about the economic development and security of the continent. More directly towards the study of IR, Neuman (1998: 2) points out that mainstream IR theory is deeply Eurocentric, having been founded in the West to understand what happened there and not moving away from such a focus even as other regions gained independence and greater importance in world affairs. More directly, Henderson (2015) adds that IR’s cannon has been defined by a consistent dismissal of the developing world. A similar observation is made by Engel and Olsen (2005: 6) when they assert that Africa has “had strikingly little impact on IR theory.” This leads them to conclude that “by and large, empirical research on Africa, within either IR or comparative politics, has dealt with specific problems not general theories” (Engel and Olsen 2005: 6). In his 2015 work *African Realism?*, Errol A. Henderson observes that:

Today, it is no less true that as one studies a field, its scholarship rarely engages Africa as a site to test IR theory, nor does it seem to engage – much less incorporate – developments in Africa in ways that inform IR theory. (Henderson 2015: 2)

Henderson also highlights the varying degrees of overt overlooking of Africa by seminal scholars such as Waltz (1979), Huntington (1993), Collier and Hoeffler (1998).² For example, Waltz (1979), the founder of structural realism, insists that a theory of world politics should have as its focus the major powers and should not be preoccupied with explaining the foreign relations of less powerful states—despite these being the overwhelming majority of humanity and sovereign states in the world system (Henderson 2015: 2). Waltz thought it “ridiculous” to construct a theory of international politics based on countries outside the Eurasian and North American mainland.

Given the field’s struggle with engaging with “other” parts of the world and feminine perspectives (see Chapter 9), is the infusion of AI a distraction before these have been adequately accounted for? (Parallel

² Though more mid-range than exemplary of a paradigm, Henderson argues that Collier and Hoeffler’s (1998) work on Africa consists of nothing more than an “econometric glance.” The greed-grievance framework, by which civil wars on the continent are understood to be driven by the appetites of political entrepreneurs than genuine struggles for political participation, offers no analysis of the complexities underlying these conflicts, and advanced to subsume both international and civil wars.

arguments have been made regarding the adoption of the 4IR by poorer countries that are still lagging behind in electricity supply, manufacturing capacity, basic infrastructure, and advanced technology markets [e.g., Sutherland 2020]). In our view, that is not so. Precisely because AI emerges into a world defined by these inequalities along issues of gender and geography, in addition to IR itself being an imperfect instrument, AI needs to be integrated (however imperfectly) and mainstreamed as soon as possible in IR theory. AI will continue to improve and with it its influence in domestic and transnational processes. Moreover, the developing world, which has a passive and only consumer role towards AI, is no more impervious to AI's effects, be they positive or negative. It is therefore timely to consider the technology both through mainstream and critical theoretic lenses.

2.8 CONCLUSION

It can be observed from the foregoing chapter that IR theory evolves through dialogue and through events. One such event in our times is the emergence of AI. Will the theories which exist live up to this new world, or will new paradigms be needed? If the current theoretical paradigms suffice, will some require amendment? These are the fundamental questions for this book. It springs from a gap in the field which this book seeks to close: the limited engagement with AI by IR scholars. In Chapter 3, we provide an overview of AI. We provide a description of its characteristics, functions, and capabilities (including proven and conceptual). We also undertake a historical review of AI's emergence, through successive epochs and trajectories in different states. Invariably, these are linked to international politics, and it is for this reason that many of its new horizons as well as its dashed hopes were conditioned by the Cold War climate.

BIBLIOGRAPHY

- Ashworth, Luke. 2012. 4. Did the Realist-Idealist Great Debate Really Happen? In *International Relations and the First Great Debate*, ed. Brian Schmidt, 60–77. London: Routledge.
- Bennett, Andrew. 2013. The Mother of All ISMS: Causal Mechanisms and Structured Pluralism in International Relations Theory. *European Journal of International Relations* 19 (3): 459–481.

- Bennett, Andrew, and Colin Elman. 2007. Case Study Methods in the International Relations Subfield. *Comparative Political Studies* 40 (2): 170–195.
- Bennett, Andrew, and Colin Elman. 2010. Case Study Methods. In *The Oxford Handbook of International Relations*, eds. C. Reus-Smit and D. Snidal, 499–519. Oxford: Oxford University Press.
- Cazacu, Andrei. 2013. International Relations: A Brief Journal Analysis, 2008–2012. University of Groningen. <https://www.rug.nl/let/studeren-bij-ons/master/rema/mhir/international-relations-cazacu.pdf>.
- Collier, P., and A. Hoeffler. 1998. On Economic Causes of Civil War. *Oxford Economic Papers* 50 (4): 563–573.
- Du Bois, W.E.B. 1946. *The World and Africa: An Inquiry Into the Part Which Africa Has Played in World History*. New York: Viking Press.
- Engel, U., and G.R. Olsen. 2005. *Africa and the North: Between Globalization and Marginalization*. London and New York: Routledge.
- Henderson, E.A. 2015. *African Realism? International Relations Theory and Africa's Wars in the Postcolonial Era*. Lanham: Rowman and Littlefield.
- Hoffmann, Matthew J. 2003. Constructing a Complex World: The Frontiers of International Relations Theory and Foreign Policy-making. *Asian Journal of Political Science* 11 (2): 37–57.
- Huntington, S.P. 1993. The Clash of Civilizations? *Foreign Affairs* 72 (3): 22–49.
- Jupille, J., J.A. Caporaso, and J.T. Checkel. 2003. Integrating Institutions: Rationalism, Constructivism, and the Study of the European Union. *Comparative Political Studies* 36 (1–2): 7–40.
- Mazrui, A. 1980. *The African Condition*. London: Heinemann.
- Neuman, S.G., ed. 1998. *International Relations Theory and the Third World*. New York: St. Martin's Press.
- Osiander, Andreas. 1998. Rereading Early Twentieth-Century IR Theory: Idealism Revisited. *International Studies Quarterly* 42 (3): 409–432.
- Schmidt, Brian, ed. 2012. *International Relations and the First Great Debate*. London: Routledge.
- Sutherland, Ewan. 2020. The Fourth Industrial Revolution—The Case of South Africa. *Politikon* 47 (2): 233–252.
- Temby, Owen. 2015. What are Levels of Analysis and What Do They Contribute to International Relations Theory? *Cambridge Review of International Affairs* 28 (4): 721–742.
- Vasquez, J.A. 1983. *The Power of Power Politics: A Critique*. New Brunswick: Rutgers University Press.
- Villanueva, Ricardo. 2012. Review: International Relations and the First Great Debate. *The Kelvingrove Review* 13 (1): 1–4.
- Waltz, K. 1979. *Theory of International Politics*. New York: McGraw-Hill.

- Wilson, Peter. 2012a. 2. The Myth of the ‘First Great Debate.’ In *International Relations and the First Great Debate*, ed. Brian Schmidt, 16–32. London: Routledge.
- Wilson, Peter. 2012b. Where are We Now in the Debate About the First Great Debate? LSE Research Online. [http://eprints.lse.ac.uk/41819/1/Where%20are%20we%20now%20\(LSERO\).pdf](http://eprints.lse.ac.uk/41819/1/Where%20are%20we%20now%20(LSERO).pdf). Accessed 28 January 2021.
- Yin, Robert K. 2009. *Case Study Research: Design and Methods* (4th Ed.). Thousand Oaks, CA: Sage.
- Zimmern, A. 1919. *Nationality & Government: With Other Wartime Essays*. London: Chatto and Windus.
- Zimmern, A. 1922. *Europe in Convalescence*. New York and London: G.P. Putnam’s Sons.
- Zimmern, A. 1936. *The League of Nations and the Rule of Law 1918–1935*. London and New York: Macmillan Co.



Artificial Intelligence and International Relations

3.1 INTRODUCTION

For centuries now, the goal of creating machines as intelligent as humans has been an elusive one. But it has become possible and, for some, realised in recent decades due to the unprecedented digital platforms and computational power of our time (Mitchell 2019: 3; Marwala 2021). AI stands at the crossroads of multiple fields rather than representing a single research and commercial area unto itself. In general, AI denotes “machines that respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment, and intention” (West 2018). Russell Stuart and Peter Norvig, in their book *Artificial Intelligence: A Modern Approach* (2009), identify four general aims of AI: thinking humanly, thinking rationally, acting humanly, and acting rationally. These functions involve multiple overlapping sets of theories and methodologies, including natural language processing (NLP), machine learning (ML), and robotics (see also Wang 2018). It is commonly held that AI has its origins in cybernetics, which was the “study of control and communication in the animal and the machine” (Wang 2018). In Fig. 3.1, we present an overview of the branches of AI.

Today’s AI developers and scholars establish two basic and related criteria to establish what AI is; intentionality and adaptability. By intentionality, AI experts are referring to AI algorithms being designed for

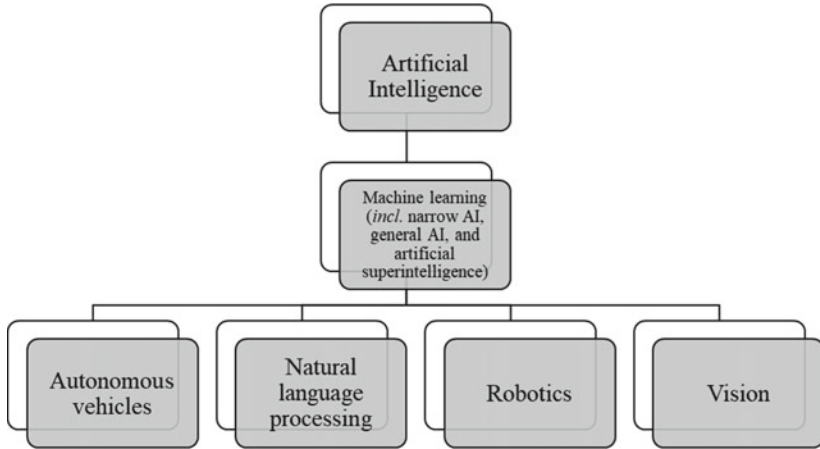


Fig. 3.1 Taxonomy of AI (*Source* By authors)

an ability to make use of sensors, digital data, and remote inputs as well as synthesise information from different sources and thereby analyse data in real-time to act on the basis of these insights (West 2018). It is for this reason that AI often is designed in tandem with ML and data analytics, which inform AI’s ability for intelligent decision-making. ML is able to take in data and search for underlying trends. If it identifies relevant data, such data can be used for analytics and understanding specific issues (West 2018). Moreover, this ML function is what differentiates AI from ordinary computer programming as it gives machines the capacity to correct themselves (i.e., to *learn*) through trial-and-error, in a similar fashion to the cognitive processes of humans (Dai and Shen 2018). This is particularly the case with deep learning (DL), also known as deep neural networks. It is through DL that AI is able to “learn” from its so-called experiences.

By adaptability, AI developers are referring to AI’s ability to learn and change its behaviour while simultaneously compiling new information and making real-time decisions. Thus, the most efficient AI are those which at a minimum adjust their behaviour as circumstances continuously shift (West 2018). In the real world, this involves “alterations in financial situations, road conditions, environmental considerations, or military circumstances” (West 2018). AI must to integrate such changes in its

algorithms in order to be able to decide in an adaptable way (West 2018). For example, when autonomous vehicles “warn” other automobiles on the road about traffic congestion, they are using their “experiences” and that of other vehicles to navigate efficiently in the absence of human decision, and the entire body of their achieved “experience” becomes immediately accessed by other similar automobiles. This also highlights AI’s function as the bedrock of the internet of things (IoT), which is an emerging technology based on sensors, cameras, and algorithms. This finds application in the so-called smart factories, which, for example, enable production to operate autonomously in different parts of the world through a central node. Applied to robotics, it indicates instances wherein, by autonomously accessing data not initially fed into it (i.e., tapping the experiences of other, otherwise located AI), robots can take on new functions not initially programmed into them. This is enabled by the advent of DL.

Section 3.2 of this chapter provides an overview of AI, while Sect. 3.3 traces the evolution of AI across different leading countries, particularly the US and China. The fourth Sect. 3.4 concludes the chapter.

3.2 OVERVIEW OF AI

In the popular imagination, AI is often conflated with robots. But though common, their integration is not always the case. That is, not all robots are capable of AI, and not all AI is integrated with the physical world through robots. At a basic level, an artificially intelligent agent is any software which functions in a digital (online) platform. This allows it to produce analyses or actions within this cyber world. When combined with a physical interface, these commonly become integrated with robotics. Robots are located in the real world, interacting with physical objects and buttressed by artificially created intelligence. The decision-making process of these robots is best understood as one of “sense-plan-act.” This means that a robot first has to detect and scan its environment, plan its actions accordingly, and then execute those actions (Bartneck et al. 2021: 12). While robots have traditionally been used as specialised production machines, with the advent of AI they are gaining broad functions outside the factory floor. For example, an autonomous vehicle is a type of robot that operates in the open world, learning and deciding as it does. Restaurants have also introduced intelligent robots to perform the functions of servers. Increasingly, moreover, the cyber and the physical

are being incrementally connected. Through the advent of the Internet of Things (IoT), there has been an upsurge in the ubiquity of sensing systems, through sensing devices implanted on real-world objects, which allows for the cyber control of the physical world (Andreu-Perez et al. 2017: 1). AI is the major growth industry of what is regarded as the fourth industrial revolution. For example, it is expected that the AI market will pass the US\$3 trillion mark by the year 2024 due to the current pace and rate of government and private sector investment. Moreover, due to their growing levels of accuracy, AI systems are expected to perform an increasingly wide range of functions in various sectors such as finance, pharmaceuticals, energy, teaching and learning, and transport among others (Moloi and Marwala 2020; Andreu-Perez et al. 2017: 1).

AI still faces many challenges, and human performance still greatly outperforms it in many respects. For example, AI systems are not yet fully able to generalise across learned concepts. While some specialised classifiers may be trained to work efficiently on related problems, their performance does tend to dip when the data is drawn from different sources (Bartneck et al. 2021: 14); for example, face recognition classifiers can generate good results for faces that are facing straight, but struggle with the same faces when they change to profile (side) views. Put differently, AI systems lack the necessary robustness to deal with a *changing*, world, which is both dynamic and ultimately unpredictable (Bartneck et al. 2021: 14). As a result of this,

AI currently works best in constrained environments, but has trouble with open worlds, poorly defined problems, and abstractions. Constrained environments include simulated environments and environments in which prior data accurately reflects future challenges. The real world, however, is open in the sense that new challenges arise constantly. (Bartneck et al. 2021: 7)

AI systems currently have only a specific set of abilities when it comes to by analogy, and when it comes to reasoning from one setting or situation to another. Because of this, they need to be taught new lessons even for closely related problems. They still lack general reasoning (or common sense), especially when problems are poorly defined (Bartneck et al. 2021: 7). This means that artificial general intelligence (AGI), which is still theoretical at this point and entails the capability of AI to perform with the same level of abstraction as human beings without any of the shortfalls of machine thinking, is still not possible. Given the states of the fields

involved, experts suggest that such realities are decades or even centuries away. We can thus deduce that it will not be possible for AI to assume complete responsibility for the variety of functions that entail statesmanship as these are such complex and consequential tasks as governance, diplomacy, and military affairs. Developing and operationalising context for these, moreover, has proven to be more challenging than the otherwise simpler process of automating repetitive tasks. The key hurdle in this regard stems from the difficulty of defining objects which the AI has not previously encountered. This is against the impressive performance of human agents who have been able to identify danger from intuition when “something didn’t seem right” (Kumar and Batarseh 2020).

3.3 THE HISTORY OF AI

AI has gradually grown from its meek origins to a mega-industry with global impact projected to affect every sphere of work and life in industrialised societies and global processes on a planetary scale in the near future. Crucially, the origins of AI were tied to international politics:

The dawn of automata theory can be traced back to World War II with what [were] known as the “codebreakers”. The amount of operations required to decode the German trigrams of the Enigma machine, without knowing the rotor’s position, proved to be too challenging to be solved manually. The inclusion of automata theory in computing conceived the first logical machines to account for operations such as generating, codifying, storing and using information. Indeed, these four tasks are the basic operations of information processing performed by humans. (Andreu-Perez et al. 2017: 3)

In the 1950s, researchers put forth a link between the fields of automata theory and neuroscience, thereby making the case the first conceptualisation of an “artificial neuron.” By 1958, with financial support from the US Navy, this paved the way for the formulation of the first computational intelligence algorithm. The developers called it a “perceptron” (Mitchell 2019). This proposal sparked the interest of the leading scientists of the day, including most prominently John von Neumann (the pioneer of modern computing) and paved the way for the “connectivism movement” (Andreu-Perez et al. 2017: 3). To Alan Turing, Von Neumann and other pioneers of the field, there were self-evident and strong similarities, at

least by analogy, between the way the human brain functioned and the computers they sought to create. Indeed, they were adamant that human intelligence could be imitated in computer programs (Mitchell 2019: 3).

AI historians locate its modern official founding in a small workshop held in 1956 at Dartmouth College organised by John McCarthy, then a 28-year-old professor with a background in both psychology and computer science (then called “automata theory”), having worked briefly at Bell Labs and IBM. The workshop, held over two months, included the participation of Marvin Minsky, Claude Shannon, and Nathan Rochester. McCarthy coined the term “artificial intelligence” in order to set it apart from the field of cybernetics (which was more focused on the combination of man and machine). Why did he go for this term? As he put it, “I had to call it something, so I called it “Artificial Intelligence”” (in Mitchell 2019: 4). Interestingly, in their proposal to the Rockefeller Foundation, the group provided a list of topics they intended to study, and which are the demarcations of the field of AI to this day. These are:

- Natural-language processing
- Neural networks
- Machine learning
- Abstract concepts and reasoning
- Creativity

Thus, at Dartmouth, “the field was named, and its general goals were outlined” (Mitchell 2019: 5). These early participants were optimistic. Minsky, who later founded the MIT AI Lab, forecasted that “the problems of creating artificial intelligence will be systematically solved” within a generation (Minsky 1967: 2).¹ Herbert Simon, one of these pioneers, predicated that machines would be capable within twenty years “of doing any work that man can do” (Simon 1965: 96).² McCarthy established the Stanford Artificial Intelligence Project in the early 1960s with the goal of building “a fully intelligent machine” within a decade. These capabilities

¹ Minsky, Marvin. 1967. *Computation: Finite and Infinite Machines*. Upper Saddle River, NJ: Prentice-Hall.

² Simon, HA. 1965. *The Shape of Automation for Men and Management*. New York: Harper & Row.

are yet to be obtained, however, and the goal of a fully intelligent machine is still in pursuit (see Marwala 2021).

In 1971, in one of the first major public investments in AI, the Defense Advanced Research Projects Agency (DARPA), funded a consortium of leading labs in the area of speech recognition. The agency, under the auspices of the US Department of Defense, had been established in 1958 following the successful launch of Sputnik by the Soviet Union and was one of a wide array of initiatives by the US government to regain the advantage over the USSR in the space race and the broader scientific rivalry. The speech recognition project was aimed at the mammoth task of creating a speech recognition system that had a perfect vocabulary. The lab fell short of these objectives, however. By the middle of that decade, the field underwent a lull due to these shortcomings. Funding dried up, and criticisms of the field grew. In the wake of these disappointing results, DARPA withdrew its funding for speech recognition. Such financial scarcity, along with the Lighthill report,³ marked the first winter of AI (Andreu-Perez et al. 2017: 4).

In the 1980s, AI programmes known as expert systems were taken on board by companies. From this point on, the goal of knowledge acquisition was the defining focus of most AI research. There was also a resurgence in government funding, not only in the United States but in Japan as well. Starting in 1982, the government of Japan, through its Ministry of International Trade and Industry (MITI), focused on its fifth generation computer systems (FGCS) initiative intended to provide a computational platform for AI (Andreu-Perez et al. 2017: 4).⁴ The following decade saw the emergence of new concepts. In the 1990s, the notion of intelligent agents came about; these were AI programmed systems that perceive their environment and can undertake actions to maximise their probability of success in real-time. This concept of the agent led, for the first time, to the idea of intelligent units working

³ See Lighthill, James. 1973. *Artificial Intelligence: A General Survey*. London: Science Research Council. The report gave a devastating criticism of the failure of AI developers to deliver on their promises.

⁴ This was a ten-year initiative, partially motivated by Japanese government's aim to refute the notion that the country was only a modifier and consumer of foreign developed technologies. See Shapiro, Ehud Y. 1983. "The fifth generation project—a trip report," *Communications of the ACM* 26(9): 637-641.

independently as well as collaboratively to achieve a common objective. This drew direct inspiration from the structure of human society, which is based on individuals, with individual minds and independently acquired and independently-processed knowledge, working together in large groups and accomplishing more than the sum of their collective efforts. By the late 1990s, statistical learning, including probability, fuzzy logic, and frequentist statistics were brought into AI in order to deal with uncertain decisions (Andreu-Perez et al. 2017: 4).

During this phase, high-speed trains based on fuzzy logic systems were deployed alongside numerous other industrial applications such as reactor control in power plants and factories. Some additions were also made to household devices.

These were different from the expert systems in 1980s; the modelling of the inference system for the task, achieved through learning, gave rise to the field of Machine Learning. Nevertheless, although machine reasoning exhibited good performance, there was still an engineering requirement to digest the input space into a new source, so that intelligence could reason more effectively. (Andreu-Perez et al. 2017: 4–5)

Since 2000, a new renaissance for the connectionism paradigm has emerged with the advent of Big Data. Big Data itself is propelled by the rapid growth of the Internet and mobile communication that has taken place since the early 2000s, which have accumulated large amounts of digital information. Once again, neural networks are being given serious consideration, especially in the role they could play in improving perceptual intelligence (Andreu-Perez et al. 2017: 4–5). New milestones were also reached in the area of computer vision. This led to increases in visual perception and thus the capabilities of intelligent agents when it comes to executing complex activities that had a visual pattern recognition component. Perez and colleagues note that “these paved the way to new AI challenges such as, speech recognition, natural language processing, and self-driving cars” (Andreu-Perez et al. 2017: 4–5). Since 2010, a branch of AI known collectively as deep learning (or deep neural networks) has come to dominate the field. DL, however, is one among many methods in the aim of realising machines with intelligence.

The history of AI has not been linear, and much less so across societies. Given the largely Western focus of the mainstream history of AI, below we turn to consider the history of AI in the Chinese context. This is necessary

in light of the huge role that country is playing and is set to further play in this field. Such a study is also illuminating in terms of uncovering the political and ideological currents which underpin AI development.

3.3.1 *History of AI in China*

As seen, the 1950s was a period in which scientists and innovators of various sorts in the western world were enthused by the new ideas of computing, their potential applications, and parallels with human cognition (Wang 2018). In the PRC, however, there was no formal research on cybernetics or AI. This stemmed from two main reasons; the first was the focus on basic economic reconstruction from the damage caused by the Second World War (which in some ways had begun two years earlier in China), the second was the dominance of political ideology, and with it the influence of the Soviet Union upon China. On 1 October 1949, the PRC was declared by the Communist Party of China, having just defeated the Kuomintang forces, who then retreated to the island of Taiwan. The following year, on February 14th, the PRC “and the Soviet Union signed the Sino-Soviet Treaty of Friendship, Alliance and Mutual Assistance.”⁵ This was accompanied by a loan of US\$300-million by the USSR to China.

The Soviet Union also transferred technical and scientific knowledge to China in the 1950s, and thereby shaping Beijing’s science and technology policies. Soviet approaches towards science and technology were at this time guided by an ideological outlook, including in the area of cybernetics. In short, cybernetics was seen as “[a] bourgeois and reactionary pseudoscience” that stemmed from the capitalist West. For example, in the 1954 edition of the *Short Philosophical Dictionary* (*Kratkiy Filosofskiy Slovar*), the ‘Cybernetics’ entry defined the field as one that was “a reactionary pseudo-science that emerged in the USA after the World War II and became widespread in capitalist countries.” Owing to Soviet influence, therefore, no serious research on cybernetics was conducted in China for a while. This seemed poised to change in 1956. In that year, the *Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967)* was formulated. This is reportedly the first long-term plan for scientific and technological development in China

⁵ http://wangjieshu.com/2018/10/17/history_of_ai_in_china/.

which saw a role for “new technologies,” including, also, atomic energy, radio electronics, and ultrasonics, were identified as requiring special and concerted attention because of their lack in the PRC at that time (Wang 2018). There were nine core areas, among which (for the first time in an official policy document) cybernetics was included, notably nuclear and semiconductor physics, among others. Thus from 1956, the China’s scientists began to increasingly pay attention to cybernetics.

But politics once again interfered. In 1957, the Chinese government, under the direction of Communist Party chairman Mao Zedong, initiated the Anti-Rightist Campaign (ARC). As part of this campaign, which was in effect a purge of individuals who were suspected of holding ideas not fully in line with CPC ideology, intellectuals were especially targeted. Among those most affected were scientists and university professors, as well as writers and artists. Between July 1957, when the ARC began, and late 1958 “between 550,000 and a million intellectuals” were given the label of Rightist and thus rendered incapable of carrying on with their work. The result was not only a long-term economic recession but the campaign also “severely tarnished the public image of Chinese intellectuals and killed their motivations, thus hindered the development of science and technology” (Wang 2018). From 1958, the Chinese government also initiated the Great Leap Forward. It was a policy intended to grow China into a globally leading country in various industries, including agriculture and steel production. However, it led to widespread famine and the deaths of tens of millions. This famine, and 1966–1976 Cultural Revolution (with similarities to the ARC) which followed, had a dampening effect on the rollout of the Twelve-Year Plan. For many historians, it came to exist only in name (Wang 2018). All the while, in the United States the seeds were being sown for many forward-looking ideas such as the graphical user interface and eventually the Internet.

Alongside the Sino-Soviet split of the 1960s, the two former allies experienced an ideological and geopolitical schism. Thus while cybernetics and AI were embraced in the Soviet Union, the Chinese scientific community became reluctant towards them. The trend continued for a sizeable portion of the 1970s (until 1978), with a Chinese issue of a translated Soviet article on AI and how machines learn prefaced with a highly critical statement by the editor, who said the article “shows these revisionists [i.e., the Soviet Union] have gone too far on the road that betrays Marxism-Leninism” (Wang 2018). A new study published in the same issue of the Chinese journal *Selective Translation: Philosophy of the Natural Science*

from Foreign Countries titled “Does AI Exist?” acknowledged the potential held by the technology, particularly in aiding thinking. It also rejected the degree of emphasis placed on it by the Soviet Union:

Can humans build intelligence? No. The term ‘artificial intelligence’ is easy for idealists to exploit. If humans can build ‘intelligence,’ then, in the future, something more intelligent than humans will emerge. ... Some Soviet revisionist academicians are vigorously propagating AI. This behavior fully exposes their ugly face as traitors. (1974: 26)

Furthermore, within China itself, discourse on artificial intelligence also reportedly came to be a useful accusation against those who held different ideas in the ideological terrain. Notably, an article in 1976 in the journal *Philosophical Trends* argued that “we must work against Deng Xiaoping, ... walking our own way in the war against those reactionary trends such as ‘image recognition’ and ‘artificial intelligence’” (in Wang 2018). The Cultural Revolution ended in that same year, however, and the Gang of Four soon lost their power. Soon developments would take shape in science and technology. The year 1978, in particular, is looked at as marking a turning point. In March of that year, the National Science Conference took place in Beijing under the auspices of the Central Committee of the CPC and the State Council with the aim of restoring and enhancing science and technology in the country. In the conference, Deng Xiaoping delivered a speech, in which he declared that “science and technology constitute a primary productive force” (in Wang 2018). The *National Science and Technology Development Plan (1978–1985)* was also established at the conference, proposing over one hundred major research projects. Among these was an “Intelligent Simulation and Intelligent Control System,” which received particular emphasis by the Chinese Academy of Sciences and the Ministry of Education to be a key research project. Still, in 1978, the Chinese Association of Automation (CAA), which had originally been formed in 1961 and suspended operation at the behest of the Cultural Revolution, was reactivated. Soon it reported numerous achievements in research, including an optical character recognition systems and cybernetics. For many, this showed that concerted studies on cybernetics and artificial intelligence had begun to take on a life of their own in the PRC (Wang 2018). This research flourished because of changes in ideological outlook as the Reform and Opening Up Policy of Deng relegated the idea that political ideology came first, and in so

doing changed the nature of Chinese society, “and broke the chain for scientific and technological innovation” (Wang 2018).

In 1982, the inaugural issue of *The Journal of Artificial Intelligence* was published. This was to be first dedicated academic journal the first specialised academic journal in the discipline in the country. This signalled official government recognition of the importance of the discipline. It is reported that in 1984, upon seeing a chess-playing computer program designed by young children, Deng Xiaoping, by then the Chairman of the State Military Commission, stated that the mainstreaming of computers should be targeted at the younger generations, especially children. More works were published on the subject of AI, including in the prominent *People’s Daily*. In 1987, the first academic book on an AI topic was published at Tsinghua University, and in the following year, another monograph on robotics was published (Wang 2018). Another year later, the first intelligent computer in China was showcased in Xi’an Jiaotong University on 26 April 1989, and a Bridge-Playing Computer (described by *China Daily* as “The New Achievement in AI”) on 13 November 1990.

Today, China’s government sees AI as a significant part of its goal of enhancing its scientific/technological leverage. The *New Generation AI Development Plan* published in 2017 lists some specific and broad goals that the country has for its role in the development of AI. The document also envisions a role for the private sector; as a result more companies have been integrating AI-related technologies into their products and services, while academia has linked its AI research with both. Foreign competition continues to spur the government on. 2017 was a significant year for two reasons in China’s AI landscape—not only was the *AI Development Plan* published in July but two months prior, in May of that year, AlphaGo, a computer program developed by Google, beat China’s (and the world’s) top player of the Chinese game Go “in a Sputnik-like moment that spurred China into a concerted, state-directed effort to catch up in artificial intelligence” (Dai and Shen 2018).

China’s continued AI development is a result of three factors whose common denominator is the proactive role of its government: reverse engineering, market isolation, and the development of tech ecosystems. The first bandwagons on global innovation, the second and third foster it from within China itself. China’s technology development trajectory was also aided by its backwardness when it initiated its Reform and Opening Up policy in 1978. Unlike the Soviet Union, it had been less

industrially developed and thus, it faced fewer obstacles to plug-and-play with foreign technologies. In turn, from the Chinese point of view, what the United States now terms technology theft “began as technical co-operation 40 years ago, when the two countries committed to a closer relationship in the form of, for example, joint ventures or equity ownership, subject to governmental licences” (Bin 2020). That these “agreements” are being evoked as examples of coerced transfer of technology are therefore seen by China as “an act of American duplicity” (Bin 2020).⁶ The pro-Chinese stance, therefore, argues that China is an underdeveloped country trying to catch up developing countries, including the United States in the eighteenth and nineteenth^h centuries, necessarily do by imitation through reverse engineering (Bin 2020). Indeed, using the example of Taiwan which is historically a part of China and societally analogous in many respects, Alford’s (1995) book *Intellectual Property* partially concludes by noting that China will enact protection for intellectual property once the country has developed sufficiently such that it has no need to steal, and has a stake of its own in terms of protecting its own innovations. By many estimates, China has reached such a moment in the 2020s; it is leading in specific areas of AI research and implementation, and Huawei has been the first to develop 5G. This technology is deemed as essential to the 4IR as it will lead to the rapid transfer of information, as well as increased capacity for the data demands that will define it (through network slicing). Today, Asia accounts for most of the new patents in the world. And within Asia, however, China stands out:

Asia has strengthened its position as the region with the greatest activity in patent filings. Offices located in Asia received two-thirds (66.8%) of all applications filed worldwide in 2018 – a considerable increase from 50.8% in 2008 – primarily driven by growth in China. Offices located in North America accounted for just under one-fifth (19%) of the 2018 world total, while those in Europe accounted for just over one-tenth (10.9%). The combined share of offices located in Africa, Latin America and the Caribbean, and Oceania was 3.3%. (WIPO, 2019)

⁶ According to Jason Fang and Michael Walsh (2018, April 29) “China requires foreign firms to start joint ventures with local companies in order to gain access to its lucrative market. Those companies are also often required to share their intellectual property with the local Chinese company they partner with. In the end, Chinese firms end up gaining the advanced technology of foreign companies without needing to compete with them in the Chinese market, thanks to heavy government subsidies.”.

Among China's new patents, AI stands out. Between 2016 and 2018, China more than doubled the number of companies in the *Nikkei* ranking of the Top 50 Patent Filers for artificial intelligence from 19 to 38. Prior to this, there were only eight in the previous three-year span. On the other hand, the United States experienced a trend in the opposite direction: moving from 19 companies to 12 (Okoshi 2019). In particular, the Chinese search giant Baidu has been soaring, moving from 11 to 4th place. By field specialty, China's growth is most observed in AI patents related to e-commerce, data searches, and language processing:

Filings for image processing technology were particularly strong at 16,000, more than quadruple the filings in the US. The country is a leader in facial recognition technology that can match a person's identity from an image of their face caught on video. (Okoshi, 2019)

Some 95% of China's new AI patents are submitted to the China National Intellectual Property Administration, but many inventors have also been undertaking overseas filings. Being a relatively new field, AI patents have been noticeably more easy to register compared to other fields of invention: the average success rate is 70%, but it is 90% when it comes to the ML subset (Okoshi 2019). While indicative of lowering of standards, it is also part of the country's strategy as the government of China may be accepting as many patents on AI as possible to get ahead of other countries. This may be working, too. While the six decades between 1950 and 2016 saw the US, and then Japan, dominate the number of applications, the PRC has surpassed both countries as of 2017 (Okoshi 2019). The Chinese government has approached its overall AI strategy, widely regarded as the most ambitious in the world, through a combination of talent cultivation and research financing, both of which find expression in its tech ecosystems, which are also very entrepreneurial (Westerheide 2020). In parallel to this, its government has worked efficiently with digital companies such as Baidu, Alibaba, and Tencent to orient them towards AI (Westerheide 2020). As part of this, funds are made available. Carter and Crumpler (2019) observe that this support allows Chinese AI companies and entrepreneurs to outcompete international competition across various sectors, including smart consumer services, transportation, and health (Carter and Crumpler 2019: 4). At the level of municipalities, cities and provinces, various sub-national governments actively compete for new AI clusters:

The implementation of the national strategy varies greatly from region to region. While cities such as Tianjin and Shanghai have already launched multi-billion-dollar AI city Venture Capital funds and had entire districts and islands built for new AI companies, other provinces are still in the process of learning and development. (Westerheide 2020)

After all, it has been observed that China's various incentives (e.g., government investment, tax credits and subsidies meant to bolster its AI developers) are disbursed by municipal governments, "driven by local cadres competing for advancement in the CCP by advancing Chinese AI" (Carter and Crumpler 2019: 5). Finally, the country has also not only benefitted by who enters its special economic zones but also who does not get access to its market. As a result, companies such as Baidu, WeChat, Alibaba have emerged due to the restrictions placed on Google, Facebook, Amazon, and other western firms operating in similar spaces. These have been benefitted through phased western financing and expertise, which got surpassed by Chinese innovation buoyed by a competitive but closed-off market of up to a billion consumers.

Thus, the US and China are the two leading countries on AI. Our chapter on hegemony will deliberate on the long-term consequences of these two AI giants as they interact on the global stage. Below we look at AI research in social science.

3.4 AI RESEARCH IN SOCIAL SCIENCE

Advances in AI mean social change. As such, social scientists have always observed developments in this arena. Today, "along with the unprecedented enthusiasm of AI, there are also fears about the impact of the technology on our society" (Andreu-Perez et al. 2017: 1). For example, equity has been an issue. For some, the dissemination and development of AI needs to be cognizant of ethical requisites so that the whole of society benefits from AI's advancement. But given that debates still abound between deontological and utilitarian approaches, this is not a straightforward phenomenon:

There are many moral approaches that Siri could be designed to take. Siri could take a consequentialist approach. This is the idea that the value of an action depends on the consequences it has. The best known version of consequentialism is the classical utilitarianism of Jeremy Bentham and John

Stuart Mill. Alternatively, Siri might be designed to take a deontological approach. A deontologist like Immanuel Kant might prioritise duty over happiness. Kant might advise Chris that lying is wrong. He has a duty not to lie so he should tell the truth about his weight, even if this would decrease his chances of getting a date. (Bartneck et al. 2021: 7)

Social science researchers have thus undertaken research on AI with interest in the implications for privacy, jobs, legal responsibility, civil rights, and military use (Andreu-Perez et al. 2017: 1). As with AI, “the relevance of international relations theory hinges on the ability of theorists to predict or at least understand the seminal events of any era” (Hoffmann 2003: 37). This book is a contribution to this research as it more directly links AI to its dominant theories, and their ethical systems and assumptions. IR as a discipline stands at the crossroads of political science, philosophy, and economics and thus ought to lead in considering the implications of AI on a global scale.

3.5 AI AND INTERNATIONAL RELATIONS PROCESSES

AI systems have begun shaping international relations processes. We note two manners in which this occurs: explicit (or direct) and implicit (or indirect). This distinction is necessary given the pervasiveness of AI such that it is evident or embedded in systems that do not appear to have an obvious AI component to them. In the more explicit dimensions, we have seen states and non-state actors make use of AI in ways that have transnational implications. These include military decision-making and deployment, and manufacturing and trade facilitation. On the indirect dimension, state and non-state actors make use of AI-enhanced processes in risk calculation, which in turn informs foreign direct investment flows and communications, including by terrorist organisations or soft power initiatives on social media (which are enabled for microtargeting through algorithms). We determine the role of AI through the crucial component of the presence of automation (represented by the functions performed by non-human systems without active human oversight), either in analysis or decision-making or both. Automation is not without conceptual dispute:

What is considered “autonomous” in a vehicle is different to what is considered “autonomous” in an [*sic*] weapon. In bioethics autonomy refers to the ability of humans to make up their own minds about what treatment to accept or refuse. In Kantian ethics⁷ autonomy refers to the ability of humans to decide what to do with their lives and what moral rules to live by. (Bartneck et al. 2021: 7)

Even in the face of these debates, many recognise the transformative nature of AI for IR processes. In her contribution to the edited volume *Remote Warfare: Interdisciplinary Perspectives*, Jennifer Gibson presents the story of a Yemeni imam who fell victim to a US automated drone deployment in 2016 and was killed, ironically just days after preaching against al-Qaeda to his community. In recognition of such incidents, and not wanting to be complicit in future killings, Google’s employees took a stand against the company in 2018 by openly protesting its role in Project Maven. The project is a US government-run initiative intended to integrate AI into military operations and sought Google’s unparalleled data analysis systems. Google’s role, its executives argued, would have been confined to the automation of drone footage analysis. But, aware of the difficulty of isolating between civilians and terrorists, the Google employees protested the suggestion. As Gibson (2021) argues,

For many communities in places like Pakistan and Yemen, computers are already making life and death decisions. Massive amounts of signals intelligence are being run through algorithms that make decisions as to who is ‘suspicious’ and who ‘isn’t.’ For populations with a drone flying overhead, those decisions can be deadly.

Because of its applications in robotics and the internet of things, AI is increasingly recognised as a key technology in the next generation of technology, termed the fourth industrial revolution. According to the World Intellectual Property Organization (WIPO), artificial intelligence is among the fastest growing sectors in the five-year span between 2015 and 2020. It has seen the most growth in new patents and investments, from governments and even more from the private sector.

⁷ In reference to the deontological system of ethics which came in the wake of German philosopher Immanuel Kant in the eighteenth century. According to the deontological tradition (as opposed to its utilitarian rival), the basis for human actions ought to be their ethical value in themselves and not as ends.

A 2019 report by the Carnegie Endowment for International Peace (2019) puts it in no uncertain times: “Given AI’s economic and geopolitical significance, they’re also increasingly being considered strategic assets.” In recognition of AI’s power and promise, many states have proactively initiated transnational AI agreements, at both bilateral and multilateral levels. Today’s global AI landscape is dominated by traditional players in the international system, namely China, the EU, and the US. China’s and Canada’s AI plans, both published in 2017, were soon succeeded by those of European (Danish Finnish, French, and Italian) strategies, alongside that of the EU itself, as well as by India, Japan, Singapore, and South Korea in Asia. Other notable players are Mexico, the UK, and South Africa. Many of these states represent the majority of the hyper-scale data centres in the world (the US alone constitutes 40%, followed by China, Japan and the UK who collectively form the next 20%). AI developers and other specialists can play a crucial role in shaping countries’ science diplomacies; but AI is becoming such a general technology that it can longer be confined to this type of diplomacy alone. By the same token, the definition of diplomacy has grown to now encompass state representation to leading technology companies. Similarly, governments are eager to acquire AI developers, and to avoid brain drains of their own.

3.6 AI AS A UNIT OF ANALYSIS IN IR?

Alan Turing proposed what is now termed the “Turing test.” According to this test’s standard, a machine can be said to pass when it can successfully convince a human that it is not a machine but another human as well. To be sure, AI has been able to defeat people in complex games through “skills” attained from unsupervised learning. Despite claims of numerous software systems passing the Turing test, these have been somewhat controversial and contested (Bartneck et al. 2021: 14). There are numerous interfaces between AI and IR processes, with implications for the field. We identify at least four: trade, war, diplomacy, and AI’s relevance to FDI flows. In all these, does AI emerge as a unit of analysis in IR? In our analysis, AI currently hinges on human control at the individual level of analysis, and thus does not constitute a fourth image (at least yet). This is the stance taken in the theories as analysed here.

3.7 CONCLUSION

Chapters 2 and 3 should demonstrate that both IR and AI have conceptual splinters and vibrant debates (moreover, these persist within theoretical paradigms, as we shall see). Within IR, there is divergence in methods used, questions asked, and even basic units of examination. There arise numerous sub-disciplines, including International Political Economy, War Studies, Transnational Issues, as well as qualitative-quantitative splits within these. Within AI, on the other hand, the splinter has mainly been between those more interested in a scientific approach and those with a practical one. Scientific researchers are interested in investigating the mechanisms by which biological/natural intelligence can be embedded in computers (Mitchell 2019: 7). Proponents of a more practical AI, on the other hand, have a more pragmatic goal: “to create computer programs that perform tasks as well as or better than humans, without worrying about whether these programs are actually *thinking* in the way humans think” (Mitchell 2019: 7; italics in original). More fundamentally, the scientifically minded AI developers and mathematicians have promoted mathematical logic and deductive reasoning as “the language of rational thought.” On the other hand, the practically minded researchers champion inductive methods by which programs can extract patterns from data and use probabilities to close the gap where there are uncertainties. To this day, “each approach has generated its own panoply of principles and techniques, fortified by specialty conferences and journals, with little unification among the sub-specialities” (Mitchell 2019: 7). This has resulted in an anarchy of methods. As we shall uncover, the same is true for IR theories. This book puts forth a common issue for all to consider. This is carried out in the chapters which follow.

BIBLIOGRAPHY

- Alford, William P. 1995. *To Steal a Book is an Elegant Offense: Intellectual Property Law in Chinese Civilization*. Stanford, California: Stanford University Press.
- Andreu-Perez, Javier, Fani Deligianni, Daniele Ravi, and Guang-Zhong Yang. 2017. Artificial Intelligence and Robotics. Cornell University. <https://arxiv.org/abs/1803.10813>.

- Baidu. 2020 (March 11). How Baidu is Bringing AI to the Fight Against Coronavirus. *MIT Technology Review*. <https://www.technologyreview.com/2020/03/11/905366/how-baidu-is-bringing-ai-to-the-fight-against-coronavirus/>. Accessed 19 April 2020.
- Bartneck, Christoph, Christoph Lütge, and Alan Wagner Sean. Welsh. 2021. *An Introduction to Ethics in Robotics and AI*. Cham, Switzerland: Springer.
- Bin, Gu. 2020 (14 January). Americans are Wrong to Paint China as an Intellectual Property Thief. *Financial Times*. <https://www.ft.com/content/26903a94-3617-11ea-ac3c-f68c10993b04>. Accessed 14 April 2020.
- Caboz, Jay. 2019 (September 13). 4 Ways Artificial Intelligence is Transforming African Industries and Communities—Including by Helping to Diagnose Eye Disease. *Business Insider SA*. <https://www.businessinsider.co.za/5-ways-scientists-are-teaching-robots-to-transform-the-african-industries-they-work-in-from-depression-chatbots-to-diagnosing-eye-diseases-2019-8>. Accessed 14 April 2020.
- Carnegie Endowment for International Peace. 2019. What the Machine Learning Value Chain Means for Geopolitics. *Carnegie Endowment for International Peace*. https://carnegieendowment.org/files/7-1-19_Stanton_etal_Machine_Learning.pdf.
- Carter, William A., and Crumpler William D. 2019. China's National Strategy for AI. In *Smart Money on Chinese Advances in AI*, 4–6, Washington, DC: Center for Strategic and International Studies.
- Chan, Elaine. 2018 (September 10). 'Made in China 2025': Is Beijing's Plan for Hi-tech Dominance as Big a Threat as the West Thinks it is? *South China Morning Post*. <https://www.scmp.com/business/china-business/article/2163601/made-china-2025-beijings-plan-hi-tech-dominance-big-threat>. Accessed 18 April 2020.
- Crawford, Emily. 2019. Made in China 2025: The Industrial Plan that China Doesn't Want Anyone Talking About. *PBS*. <https://www.pbs.org/wgbh/frontline/article/made-in-china-2025-the-industrial-plan-that-china-doesnt-want-anyone-talking-about/>. Accessed 18 April 2020.
- Dai, Sarah, and Shen Alice. 2018 (October 1). 'Made in China 2025': China has a Competitive AI Game Plan but Success will Need Cooperation. *South China Morning Post*. <https://www.scmp.com/tech/article/2166177/made-china-2025-china-has-competitive-ai-game-plan-success-will-need>. Accessed 18 April 2020.
- De Klerk, Aphiwe. 2020 (February 13). Sona: Ramaphosa Announces new University for Ekurhuleni. *Times Live*. <https://www.timeslive.co.za/politics/2020-02-13-sona-ramaphosa-announces-new-university-for-ekurhuleni/>. Accessed 18 April 2020.
- Deng, Iris. 2018 (17 July). China's AI Industry Gets the Most Funding, but Lags the US in Key Talent, Says Tsinghua. *South*

- China Morning Post*. <https://www.scmp.com/tech/china-tech/article/2155600/chinas-ai-industry-gets-most-funding-lags-us-key-talent-says>. Accessed 18 April 2020.
- Department of International Cooperation Ministry of Science and Technology (MOST), People's Republic of China. 2017. *Next Generation Artificial Intelligence Development Plan*. Beijing. <http://fi.china-embassy.org/eng/kxjs/P020171025789108009001.pdf>. Accessed 18 April 2020.
- Ding, Jeffrey. 2018. *Deciphering China's AI Dream: The Context, Components, Capabilities, and Consequences of China's Strategy to Lead the World in AI*. Oxford: Future of Humanity Institute.
- Dutton, Tim. 2018 (June 28). An Overview of National AI Strategies. *Medium*. <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6cdf>. Accessed 14 April 2020.
- Fang, Jason, and Walsh Michael. 2018 (April 2019). Made in China 2025: Beijing's Manufacturing Blueprint and Why the world is Concerned. *ABC Australia*. <https://www.abc.net.au/news/2018-04-29/why-is-made-in-china-2025-making-people-angry/9702374>. Accessed 18 April 2020.
- Gibson, J. 2021. Death by Data: Drones, Kill Lists and Algorithms. In *Remote Warfare: Interdisciplinary Perspectives*, eds. Alasdair McKay, Abigail Watson, and Megan Karlsjøj-Pedersen, 187–198. Bristol: E-International Relations Publishing.
- Hamilton, Isobel Asher. 2020 (February 7). Chinese Tech Giant Baidu has made a Maps App that Shows the Location of Coronavirus Patients. *Business Insider US*. <https://www.businessinsider.sg/baidu-map-app-coronavirus-cases-location-2020-2>. Accessed 18 April 2020.
- Hoffmann, M.J. 2003. Constructing a Complex World: The Frontiers of International Relations Theory and Foreign Policy-Making. *Asian Journal of Political Science* 11 (2): 37–57.
- Kumar, A., and F.A. Batarseh. 2020. The Use of Robots and Artificial Intelligence in War. *LSE Blog*. <https://blogs.lse.ac.uk/businessreview/2020/02/17/the-use-of-robots-and-artificial-intelligence-in-war/>.
- Marwala, Tshilidzi. 2020 (March 6). South Africa must have a Stake in Artificial Intelligence Technology. *Mail and Guardian*. <https://mg.co.za/article/2020-03-06-south-africa-must-have-a-stake-in-artificial-intelligence-technology/>. Accessed 18 April 2020.
- Marwala, Tshilidzi. 2021. *Rational Machines and Artificial Intelligence*. Cambridge, MA: Academic Press.
- Minsky, M.L. 1967. *Computation: Finite and Infinite Machines*. Englewood Cliffs: Princeton-Hall.
- Mitchell, Melanie. 2019. *Artificial Intelligence*. London: Pelican.
- Nd zendze, Bhaso. 2020 (February 13). SA's Star Role in the 4IR. *The Star*.

- Norvig, Peter, and Stuart J. Russell. 2009. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice-Hall.
- Okoshi, Y. 2019. China Overtakes US in AI Patent Rankings. *Nikkei Asia*. <https://asia.nikkei.com/Business/Business-trends/China-overtakes-US-in-AI-patent-rankings>.
- Pham, Sherisse. 2018 (March 23). How much has the US Lost from China's IP Theft? *CNN*. <https://money.cnn.com/2018/03/23/technology/china-us-trump-tariffs-ip-theft/index.html>. Accessed 14 April 2020.
- Simon, H.A. 1965. *Administrative Behavior*. New York: The Free Press.
- VAO. 2013. IP Theft Costs US \$300 Billion Per Year: Report. *VOA*. <https://www.voanews.com/archive/ip-theft-costs-us-300-billion-year-report>. Accessed 14 April 2020.
- Wang, Jieshu. 2018. The Early History of Artificial Intelligence in China (1950s–1980s). Annual Meeting of the Society for the History of Technology (SHOT), St. Louis, MO.
- Wei, Lingling, and Davis Bob. 2018. China Prepares Policy to Increase Access for Foreign Companies. *Wall Street Journal*. <https://www.wsj.com/articles/china-is-preparing-to-increase-access-for-foreign-companies-11544622331>. Accessed 18 April 2020.
- West, Darrell M. 2018. What is Artificial Intelligence? *Brookings Institution*. <https://www.brookings.edu/research/what-is-artificial-intelligence/>. Accessed 14 April 2020.
- West, Darrell M., and Allen John R. 2018. How Artificial Intelligence is Transforming the World. *Brookings Institution*. <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/>. Accessed 14 April 2020.
- Westerheide, Fabian J.G. 2020 (January 14). China—The First Artificial Intelligence Superpower. *Forbes*. <https://www.forbes.com/sites/cognitiveworld/2020/01/14/china-artificial-intelligence-superpower/#490350112f05>. Accessed 14 April 2020.
- WIPO. 2019. *World Intellectual Property Report 2019*. World Intellectual Property Organization. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_944_2019.pdf.
- Yonay, Guy. 1995. Intellectual Property in Chinese Civilization. *Harvard Journal of Law & Technology* 8 (2): 537–541.



Realism and Artificial Intelligence

4.1 INTRODUCTION

The seventeenth-century English thinker Thomas Hobbes enjoys the rare legacy of being a key thinker in two distinct political traditions which are diametrically opposed. In addition, being one of the foundational thinkers of (domestic) liberalism (see Chapter 5 of this book), his work is integral to the germination of the paradigm of realism through his conceptualization of the notion of anarchy and advancement of the ideal of the strong state (or ‘leviathan’). Critical to his realism, however, were the adoption of notions of progress and ingenuity. His magnum opus *Leviathan* (1651) opens with an account of governance-enabling and man-made creations that draw their derivations from nature which they seek to mimic. He writes:

“Nature (the Art by which God hath made and governs the World) is by the *Art* of man, as in many other things, so in this also imitated, that it can make an Artificial Animal. For seeing life is but a motion of Limbs, the beginning whereof is in some principall [*sic*] part within; why may we not say, that all *Automata* (Engines that move themselves by springs and wheels doth a watch) have an artificial life?” (Hobbes 1651: 81).

Ample more inventions have emerged since Hobbes’ era, which preceded the Industrial Revolution in England by a century. Present-day changes and advancements in technology have made this process even more

intensified. To follow Hobbes' logic further, today's rational machines, powered by AI, are capable of making decisions at a more efficient rate than their human counterparts (Marwala 2021). How much of this will shift the international system and how leaders craft their states' response to these on behalf of their states? What levers will states uncover or enhance as a result of AI? That is the focus of this chapter. We ask these questions through the prism of realist theory.

Section 4.2 gives a comprehensive overview of realist theory, including its origins, the arguments it advances, and the axioms upon which these are based. The section is closed with an analysis of the internal debates within the overarching realist paradigm. Section 4.3 is a discussion of AI and realism, in which we argue that realist theory causes those who espouse it to consider AI in comparative and zero-sum terms, as well as in its relevance to the war. This section also consists of a discussion of what we term an AI balance of power, and proceeds to an examination of the possible elements of AI-induced information warfare. We argue that it would occur in a manner akin to mutually assured destruction in the wake of automation (particularly second-strike capabilities). The concluding Section 4.4 provides an overview of how realism may be limited or counterargued in its examinations of AI in IR, and thus the caveats it may need to concede.

4.2 CONCEPTUALISING REALISM: ORIGINS, CLAIMS, AND INTERNAL DEBATES

Realism is the oldest theoretical paradigm within IR. It draws its intellectual origins to the fifth century BCE, with the work of Thucydides, the *History of the Peloponnesian War*, being seen as the tradition's founding text on account of its emphasis on power differences in the initiation and conduct of the war between Sparta and Athens and their respective alliances. Thucydides placed the war's origins on an emerging empire's power aspirations and the reaction (essentially insecurity) that this evokes from the status quo power of the day. He also gave an account of the impossibility of neutrality for smaller states in a war between major powers. Through the subsequent millennia—the next major book in the conventional realist cannon would be written nearly two thousand years later during the Renaissance (Niccolò Machiavelli's *The Prince* in 1532)—realism has evolved to become more sophisticated and undergone splinters. Realists, however, are unified by their permanent focus on the

importance of power in the international system and an appreciation of its centrality to states' capacity for action. In Mearsheimer's description, realists hold firmly that power is the sole "currency" of international politics (2013: 77). Though layered and differently understood, at a minimum, power is defined economically and militarily in relative terms. But why do states want power? Here realists essentially differ between classical realist and structural/neorealist answers. For classical realists, the focus is on human nature, while the latter focus on reasons that are rooted in the system or structure itself and the incentives they present. In the classical sense, for example, in the views of Thucydides (and subsequent commentators on his work), the driver of international conflict are the divergent ambitions of states rooted in the human capacities for pride and fear. In turn, Machiavelli's philosophy was heavily shaped by the political nuances of his day. Living in an era of the emergence of the modern-day nation-state, Machiavelli's work sought to give statesmen (particularly the leader of Florence at the time, Lorenzo de' Medici) practical advice on how to govern, and an exposition of the philosophical grounding for such manner of governing. In this sense, "the main foundations of every state,[...] are good laws and good arms; and because you cannot have good laws without good arms, and where there are good ones, good laws inevitably follow" (Machiavelli 1532). Thus, he advised rulers to break commitments made in treaties if necessary in their pursuit of power and national interest. For structural realists, on the other hand, the answer is instead to be found in the international system, which they claim compels states to pursue power. But within the structural realist tradition, there are internal debates, particularly between defensive and offensive realists. Both these notions rely on core assumptions made by either sub-tradition. We delve into these in turn in the two sub-sections which follow. What follows these is an application, and evaluation, of such assumptions in the wake of AI in international relations.

4.2.1 *Structural Realist Debates*

For structural realists, states are very alert to the international balance of power and compete with one another to gain an advantage at the direct expense of each other. They care deeply about the balance of power because they are left with no choice but to do so by the structure of the international system for states that want to survive (as they are assumed to by realists). For these realists, then, there is an inherent power

competition. This sometimes causes states to fight one another. Observes Mearsheimer:

In a system where there is no higher authority that sits above the great powers, and where there is no guarantee that no one will not attack another, it makes eminently good sense for each state to be powerful enough protect itself in the event it is attacked. (Mearsheimer 2013: 78)

Unlike liberalism, especially through its democratic peace theorists (see Chapter 5), structural realism intentionally pays no attention to differences in regime types and domestic institutions and cultures. The international system creates, realists argue, basic incentives for all great powers (the only unit and type of state realists are interested in). Accordingly, it nearly does not matter whether a state is a democracy or an autocracy when we consider its conduct outwardly. This focus on the outcome rather than the process means that these theorists approach states as though they are black boxes: “they are assumed to be alike, save for the fact that some states are more or less powerful than others” (Mearsheimer 2013: 78).

Structural realists, however, do not agree on everything. The different branches take on two key strands: defensive and offensive realism. Defensive realists make the assertion that structural factors limit the extent of power which can be gained by a given state, and this in turn ensures amelioration of security competition (Mearsheimer 2013: 77). Defensive realists such as Kenneth Waltz (1983) think that it is misguided for any state to seek maximisation of their share of world power, because of the likelihood of a backlash for wanting too much of it. As such, the pursuit of power is deemed a foolhardy one. The opposite view is taken by offensive realists, however. Mearsheimer (2001), for example, argues that “it makes good strategic sense for states to gain as much power as possible and, if the circumstances are right, to pursue hegemony.” In advancing this argument, offensive realists are not claiming conquest and domination for their own sake. Rather, they recognise that possessing overwhelming power is the surest method for guaranteeing the survival of one’s state. On the other hand, classical realists conceive of power as an end in itself rather than a means for survival in an anarchic world. How will power be defined in an AI-dominated world? To answer this question, we briefly discuss the five foundational assumptions of structural realism, which in turn accounts for why states want power:

1. Powerful states are the main, nearly exclusive actors in the international political system. The system in turn is characterised by sovereignty, leading to anarchy.
2. States have military capabilities that can cause harm to other states. Such capabilities are not even distributed; they vary among states.
3. States have no way of being certain with regards to the intentions of other states. Specifically, states want to know whether other states are revisionist or status quo in orientation. However, it is not always possible to know other states' intentions with a high degree of certainty. Even if they can be (through speeches and policy documents—assuming they can be relied upon), future intentions cannot be truly known.
4. The objective of each individual state is perpetuating its own survival in territorial terms as well as political self-determination.
5. States are rational actors. Miscalculation is possible due to point number-3 in this list (Mearsheimer 2013: 80).

On the first count, great powers fear and mistrust each other. Secondly, states operate in a self-help world; even alliances are not as central as self-interest; there is a positive correlation between a state's relative power. The more of it that it has, the lower the likelihood that it will be put under attack. This drives great powers to search for opportunities to turn the balance of power overwhelmingly in their favour (Mearsheimer 2013: 80). Every country in the system knows this, and operates within this logic, thereby leading to a jostle for power. The structure of the system pushes every great power to act in this way. Why is it not possible for all the major powers to become content and pursue peace, and thus the status quo? Here offensive structural realists argue that it is not possible for them to be certain about the current and future intentions of other states:

In an anarchic system, where there is no ultimate arbiter, states that want to survive have little choice but to assume the worst about the intentions of other states and to compete with them. This is the tragedy great power politics. (Mearsheimer 2013: 80)

This is the security dilemma articulated by Hertz in 1950, characterised by perpetual security competition. Whereas defensive realists argue that it is strategically more sensible to strive for only an appropriate amount

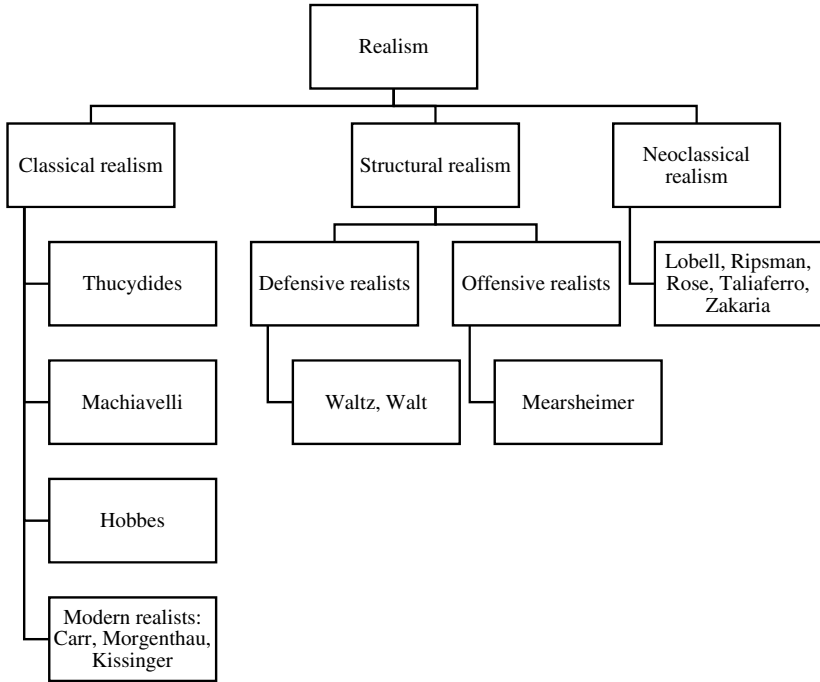


Fig. 4.1 Overview of some major thinkers in the three realist strands (*Source Authors*)

of power despite the incentives presented by the system (Waltz 1983: 40), their offensive counterparts are of the view that a state must always search for opportunities for greater power and must do precisely so whenever this is feasible. In this way, then, their ultimate goal ought to be hegemony because that is the only way to guarantee their survival (Fig. 4.1).

4.3 REALISM IN THE AGE OF AI

What can realism (or *realisms*) elucidate about the AI landscape? In the wake of AI, security and power considerations are the most evident areas in which realism would be expected to respond—and be responsive

towards. Given its focus on powerful states as the primary actors, realism's role will be to illuminate how these states incorporate AI, whether domestically developed or acquired, into their arsenals. Moreover, realism will have to define what a major state looks like in the age of AI. In other words, will AI be significant and ubiquitous enough to necessitate a reconceptualization of the practice of power? And will its rollout fundamentally alter the current balance of power (if we may be said to be operating under one)?

As discussed in Chapter 3, AI has security applications which have been paid close attention by powerful governments, including the US, Russia, and China. Some security applications include prediction, analysis, and simulation. An overview of these states' AI strategies also indicates the extent to which they see AI as a powerful tool for gaining geostrategic advantage. A perusal of three such states (China, Russia and the US) as summarised in Table 4.1, demonstrates this recognition. By extension, a perusal of these countries' understanding of the global AI scene demonstrates how much they deem themselves as being behind the others—this is the AI security dilemma in practice. True to the expectations of realism, these states are concerned with their sovereignty, though, at least explicitly, not directly with their survival. They thus may see AI as a mid-range threat; serious and advantage-generating enough to merit advantage but not powerful enough to yet present extermination. It is also noteworthy that these states are nuclear-armed and may not be as confronted by the immediacy of attack or invasion or annexation. It is also the case that the major disagreements are to do with regimes currently in power and not with the acquisition of one another's territories.

One difficulty in the way forward is the inherent disagreement, even among realists, of who the major states of today are. For many, the United States would seem to be the unrivalled superpower (based on its unparalleled military budget, economy and overseas bases), while for some, it had been the sole superpower following the end of the Cold War but has begun to decline in comparison to China, which is catching up and Russia gaining greater power status (even if still materially harangued) (Brands 2016). Still for others, we are in a multipolar phase with the emergence of numerous states and regional leaders outside the orbit or control of the United States (Dong 2020). For its part, the United States has conducted an introspective study known as the National Security Commission on Artificial Intelligence (NSCAI). The commission consisted of a “bipartisan commission of 15 technologists, national security professionals,

Table 4.1 Major states' AI strategies and their discourses

<i>AI strategy title</i>	<i>AI and IR discourse</i>	<i>Extract</i>
China's New Generation Artificial Intelligence Development Plan (2017)	An economic game-changer with implications for development and security in which China must obtain first-mover advantage	"Follow the coordinated development law for economic and national defence construction; promote two-way conversion and application for military and civilian scientific and technological achievements and co-construction and sharing of military and civilian innovation resources; form an all-element, multi-domain, highly efficient new pattern of civil-military integration"
Decree of the President of the Russian Federation on the Development of Artificial Intelligence in the Russian Federation (2019)	AI as economically viable and growing; Russia as advantaged but behind; Russia as needing to catch up with a dominant minority on the globe	"The Russian Federation has considerable potential for becoming an international leader in the development and use of artificial intelligence technologies... Taking into account the current situation on the global artificial intelligence market and medium-range forecasts for its development, the implementation of the Strategy at hand is a necessary condition for the Russian Federation's entry into the group of world leaders in the field of the development and introduction of artificial intelligence technologies, and consequently, for the country's technological independence and competitiveness"

(continued)

Table 4.1 (continued)

<i>AI strategy title</i>	<i>AI and IR discourse</i>	<i>Extract</i>
US's National Security Commission on Artificial Intelligence (2021)	AI being nefariously used by adversaries against the US; US losing likely to lose competitive edge if not proactive enough. Need for an expanded definition of security	“Simultaneously, AI is deepening the threat posed by cyber attacks and disinformation campaigns that Russia, China, and others are using to infiltrate our society, steal our data, and interfere in our democracy. The limited uses of AI-enabled attacks to date represent the tip of the iceberg. Meanwhile, global crises exemplified by the COVID-19 pandemic and climate change highlight the need to expand our conception of national security and find innovative AI-enabled solutions”

Source Authors

business executives, and academic leaders” and was chaired by former Google CEO and Chairman Eric Schmidt. In its report, published in 2021, the NSCAI determines that:

America is not prepared to defend or compete in the AI era. This is the tough reality we must face. And it is this reality that demands comprehensive, whole-of-nation action. Our final report presents a strategy to defend against AI threats, responsibly employ AI for national security, and win the broader technology competition for the sake of our prosperity, security, and welfare. The U.S. government cannot do this alone. It needs committed partners in industry, academia, and civil society. And America needs to enlist its oldest allies and new partners to build a safer and freer world for the AI era. (Schmidt et al. 2020: 7)

By recognising the impotence of the government outside the cooperation and assistance of non-state actors such as industry, academia, and civil society, the above statement is an affront to structural realism (with the

possible exception of neoclassical realism, which is still outside the mainstream and rejected by many realists). China's approach (in addition to admitting its own self-perceived shortcomings¹) is also cognizant of the limits of the state and pays heed to the role of the private sector:

Follow the rules of the market, remain oriented toward application, highlight companies' choices on the technological line and primary role in the development of commercial product standards, accelerate the commercialization of AI technology and results, and create a competitive advantage. Grasp well the division of labor between government and the market, better take advantage of the government in planning and guidance, policy support, security and guarding, market regulation, environmental construction, the formulation of ethical regulations, etc." (Government of the People's Republic of China 2017)

It is also possible, however, that the stance of the multisectoral NSCAI is indicative of a national government being led by a powerful private sector rather than the other way around. In his farewell address, US President Dwight Eisenhower referred to what he termed a "military-industrial complex" through which the private arms manufacturing corporations influence the government through lobbying. Taken to its logical conclusion, this phenomenon may be responsible for the instigation of conflicts for the purposes of manufacturing and selling ammunition. In the age of AI, there may be such an undertone; it reads as an "AI-military-industrial complex" when the leaders of AI-producing corporations and the AI research community in academia suggest that the United States is not keeping up, despite being the world leader in all major metrics. For offensive realists, this might be in order; the core aim should be expanding the gap and obtaining absolute hegemony in AI. Yet for liberals and other theories cognizant of the domestic sphere, the policy process itself matters

¹ The document states that "there is still a gap between China's overall level of development of AI relative to that of developed countries—lacking major original results in the basic theory, core algorithms, key equipment, high-end chips, major products and systems, foundational materials, components, software and interfaces, etc. Scientific research institutions and enterprises do not yet possess international influence upon ecological cycles and supply chain, lacking a systematic research and development layout; cutting-edge talent for AI is far from meeting demand. Adapting to the development of AI requires the urgent improvement of basic infrastructure, policies and regulations, and standards systems" (Government of China, 2017).

a great deal. It is worthwhile to examine the discourses presented and whence they emerge. In other words, where offensive realists see the hegemonic US eager to maintain its position, others might well see a state held captive by its industrial (and academic) elite.

4.3.1 *AI Balance of Power: Towards a Neoclassical Turn*

We have already seen that realists pay heed to power and power alone as the currency of international politics among nations. In the realist tradition this power stems from the material capabilities that states control. The balance of power is quantifiable: it is the tangible military possessions of states, and the differences therein. This means paying attention to such physical items as armoured divisions, fighter jets, nuclear weapons, and aircraft carriers among others (Mearsheimer 2013: 78). However, there is a second kind of power which is brought forth by the emergent Neoclassical school of realism as led by Gideon Rose, Norrin M. Ripsman, Jeffrey W. Taliaferro, Steven E. Lobell, and Fareed Zakaria among others. Neoclassical realism seeks to present a theory of foreign policy and in so doing brings in domestic, and non-state, actors into the realist explanatory paradigm. For neoclassical realists such as Ripsman et al. (2016), both structural realism, with its emphasis on the international system, and liberalism, with its focus on *Innenpolitik* (domestic politics), are in need of correction. Critiquing structural realism, they state that “only in rare circumstances does the international system provide clear information to states about the external constraints and opportunities they face” (Ripsman et al. 2016: 9). This has implications for how power is measured and balanced. For example, the 1930s was a period of debate about whether Germany and Japan posed a threat to the United Kingdom before the former two ultimately revealed themselves to be in 1939. Such debate is mirrored in today’s discussions about China and whether it will pose a threat to the United States or in fact undergo a peaceful rise such as observed by post-WWII Japan and Germany. On the other hand, an exclusive focus on *Innenpolitik* leads to obfuscation about the choices faced by states in crafting their foreign policies. These neoclassical realists assert that only during elections or other threats to regime survival do domestic variables exclusively set the tone of a state’s foreign policy. They draw attention to the distinction between national power (i.e., technological, economic and human resources in the hands of the population) as opposed to state power (i.e., economic and military power which is

readily in the hands of the government) and highlight the prerequisites for translating the former into the latter as it should not be assumed to be a seamless process. Aaron Friedberg's (2000) *In the Shadow of the Garrison State: America's Anti-Statism and Its Cold War Grand Strategy* explores this theme in the context of the Cold War, during which the United States failed to optimally take advantage of its endowments and translate these into state power. In this line of thought, power is not just military potency but also resides in the socio-economic base to be found within the domestic sphere. This is termed "latent power" and is based on the wealth and population size of a given state (Mearsheimer 2013: 78). These are not primarily military in function but are rapidly convertible and can thus lead to an advantage in times of war; they are the raw potential that a state has in relation to its competitor states (Mearsheimer 2013: 78).

In the age of AI, these can be incorporated accordingly so that instead of domestic wealth more generally, we measure funding of AI research and development specifically. And instead of the general populace, we are interested in the relative number of AI developers and engineers per population. We present the following model for measuring the "AI balance of power" which can be considered alongside other traditional markers such as military expenditure, nuclear stockpiles, total enlisted soldiers, etc. The AI balance of power ought to be focused on outputs and therefore should consider not only the averaged innovation indicators of a given country (e.g., WEF scores), but also AI patents registered within the preceding period of interest; the total share of AI technologies among its exports, and these markers may be subtracted from those of the rival state in comparison. We present this below as:

$$BoP_{AI} = (\mu_{innovation(Y1-Y2)} + P_{AI} + X_{Tech(Y1-Y2)} - R) \quad (4.1)$$

where $\mu_{innovation(Y1-Y2)}$ refers to average innovation score from year to year, P_{AI} refers to total AI patents with industrial and military applications that are exclusive to the state in question and have been deployed in the preceding 10 years, $X_{Tech(Y1-Y2)}$ refers to the total share of technology exports of the given country, accounting for change between the current and previous year, and where R refers to these exact markers for the "opposing" country. In this way, if the score is positive, it indicates that the state being compared is currently lagging behind the rival state, while a negative score indicates a gap in favour of the benchmark state.

For example, if a state and its rival have the following indicators compared to a rival with an R score of 60,

$$BoP_{AI} = (7 + 20 + 50) - R \quad (4.2)$$

It can be deduced that the benchmark state is currently leading by a margin of 17. Of course debates abound whether technology has any parsimonious impact on war outcomes, but the figure ought to represent an estimation of the relative standing of each country in relation to another such that we can deduce which has an “upper hand.”

4.3.2 *Information Warfare as Mutually Assured Destruction*

Waltz begins his paper “The Spread of Nuclear Weapons: More May Better” (1983) with a breakdown of the conventional view of nuclear weapons dissemination in three key points. The first is the widespread belief at the time that the world runs the risk of greater danger in the wake of nuclear weapons proliferation. Therefore, more nuclear weapons were perceived as bad. The second is that the regime type of the nuclear state in question matters, as it influences its outlook (particularly its sense of responsibility) towards its stockpiles, and its administrative efficiency for managing them. Finally, there is the conventional view, at least at the time, that if states are long-standing rivals, they are likely to use the weapons in their confrontation. He subsequently counters this logic with the following observation: “The prevalence of peace, together with the fighting of circumscribed wars, indicates a high ability of the post-war international system to absorb changes and to contain conflicts and hostility.” The most notable changes following WWII are the shifts from a world defined by multipolarity to one of bipolarity (centred on Moscow and Washington), as well as the introduction of nuclear weapons. For realists, this combination was essential, with different emphases by different scholars on the primary variable (bipolarity or nuclear weaponry). Placing emphasis on the role of nuclear weapons is a thesis in the realist paradigm termed mutually assured destruction (MAD). It argues that the cost of a possible nuclear attack poses such a threat that no nuclear possessing state is likely to initiate a nuclear war on another. The United States was the first country to develop a working nuclear arsenal and was the first to deploy them, doing so in Japan in 1945. The United States was soon followed by the Soviet Union in 1949, with reports that it managed to do

so with secrets stolen from the US.² The UK followed suit in 1952, as did France in 1960 and the People's Republic of China in 1964. These are the only UN-recognised “nuclear weapons states” as they developed these capabilities prior to the opening of the Non-Proliferation Treaty (NPT) for signature in July of 1968 and entry into force 2 years later. However, some countries abstained from it and were to later develop nuclear capabilities; India (1974), apartheid South Africa (1970s), India (1974), and Pakistan (1998). North Korea (2006) is so far the only country to exit the NPT and subsequently develop nuclear weapons. While Israel's nuclear program has been shrouded in secrecy (even being denied by its leaders) but it is believed to have begun testing in 1966 (two years before the NPT). In the late 1980s, South Africa became the first country to voluntarily dismantle its nuclear arsenal. Others were to follow suit, however, as concerns over the fall of the USSR led to some of its former republics inheriting sizeable portions of its nuclear stockpiles. Many observers and experts in the early 1990s, were of the view that neither these states nor Russia itself could guarantee their security. Thus under the Lisbon Protocol (1992), three former Soviet Republics (Belarus, Kazakhstan, and Ukraine), Russia, and the United States became parties under the START Treaty. Under the treaty, the three republics agreed to transfer their weapons to Russia, who in turn committed to a reciprocal reduction in stockpiles with the US.

The possession of nuclear weapons by the United States and the Soviet Union is said to have been the principal reason behind why the Cold War never materialised into a conventional war (Wilson 2008: 421). Scholars of the MAD emphasise the importance of second-strike capabilities; that is, upon being attacked, a state should still be in a position to retaliate with equal or greater force on the initial attacker. Some of these appear to have an overlap with the logic behind escalations in information warfare capabilities. Assessments such as the below, which deploy the concepts of “first strikes” customarily reserved to works on nuclear capabilities, demonstrate as much:

Information warfare of [a] disruptive variant is particularly challenging for our understanding of conflict because it blurs the peace–war boundary.

² “Klaus Fuchs, a German-born British physicist who was involved in the Manhattan Project, was later convicted of passing secret information on the theory and design of atomic bombs to the Soviet government” (*Encyclopedia Britannica* 2019).

Given the technological potential for intrusion, the temptation to pre-emptively disrupt in order to ‘prepare the battlefield’ before conventional hostilities or a crisis begins, or to incapacitate an adversary’s war-making system by causing a complete or partial loss of function, is strong. A pre-battle information suppression operation might shatter an enemy’s will to fight but does this first strike constitute a use of force? (Bishop and Goldman 2003: 121)

This would appear to be the case as information attacks are intended for sending enough messages to convince the adversary that they should stop any further fighting or, alternatively, to completely diminish the communications systems of the rival such that it is unable to counterattack (Bishop and Goldman 2003: 121). Like nuclear attacks, information warfare is able to target both civilian and military key points (Bishop and Goldman 2003: 122). In light of the long-standing use of information as a tactic, we are seeing a ratcheting up of these pre-existing capabilities (see also Chapter 5). This paves the way for an intensified form of information war that is nonetheless rooted in the policy rationale posited by Carl Von Clausewitz (i.e., war as continuation of politics by other means) and whose insight remains unchallenged (1831).

4.4 CONCLUSION

This chapter has discussed realism as a theoretical paradigm—a school of thought composed of three distinct strands which signify the theory’s evolution over time. The classical realists located their analyses on human nature, while the structural realists have developed more sophisticated models to depict how the international system presents states with incentives for security-minded thinking. Less mainstream in adoption has been neoclassical realism, with a focus on domestic factors. The emergence and proliferation of AI with military applications will inevitably impact realist thinking. In our analysis, based on existing and conceptual AI, we have determined three areas in which realism will interact with AI. Firstly, AI will become part of the balance-of-power rationale. As part of the latent power arsenal, modernised and comparatively more capable AI will be increasingly necessary and will inform how states interact with the international system. Secondly, technological advancements have shifted the conduct of war, including the very nature of conventional battles among the major powers (realism’s only species of focus)—AI will accelerate the

process. This means the definition of “defensive” and “offensive” will need revisiting and definition for the era of AI. Thirdly, the focus on only “major powers” requires rethinking as well. As seen in Chapter 3, moreover, AI is relevant to international processes in ways other than its military offering. Thus, if we rely solely on realism to give an account of an AI-dominated world, we may struggle to get a sufficiently comprehensive understanding of it. A major drawback for realism will be to define a world in which AI is churned out by non-state actors, major corporations chief among them, and a world in which AI is created by multiple countries in conjunction with one another and with attempts at universal sets of ethics to guide the development, deployment, and management of AI. A theoretical paradigm that can take cognizance of these actors is needed. We thus turn next to such a paradigm, liberalism, in the following chapter.

BIBLIOGRAPHY

- Bishop, Matt, and Emily Goldman. 2003. The Strategy and Tactics of Information Warfare. *Contemporary Security Policy* 24 (1): 113–139.
- Brands, Hal. 2016. *Making the Unipolar Moment: U.S. Foreign Policy and the Rise of the Post-Cold War Order*. New York City: Cornell University Press.
- Dong, Wang. 2020. *Re-globalization: When China Meets the World Again*. New York City: Routledge.
- Encyclopædia Britannica*. 2019. Nuclear Proliferation. *Encyclopædia Britannica*. <https://www.britannica.com/topic/nuclear-proliferation>.
- Friedberg, Aaron. 2000. *In the Shadow of the Garrison State: America’s Anti-Statism and its Cold War Grand Strategy*. Princeton, NJ: Princeton University Press.
- Government of the People’s Republic of China. 2017. *A New Generation Artificial Intelligence Development Plan*. Beijing: Government Press.
- Hobbes, T. 1651 [2017]. *The Leviathan*. London: Penguin.
- Machiavelli, N. 1532 [2011]. *The Prince*. London: HarperCollins.
- Marwala, T. 2021. *Rational Machines and Artificial Intelligence*. London: Academic Press.
- Mearsheimer, J.J. 2001. *The Tragedy of Great Power Politics*. New York: W.W. Norton & Company.
- Mearsheimer, John J. 2013. Structural Realism. In *International Relations Theories: Discipline and Diversity*, eds. Tim Dunne, Milja Kurki, and Steve Smith, pp. 77–93. New York: Oxford University Press.
- National Security Commission on Artificial Intelligence. 2021. *Final Report*. Washington, DC: National Security Commission on Artificial Intelligence.

- Ripsman, Norrin M., Jeffrey W. Taliaferro, and Steven E. Lobell. 2016. *Neoclassical Realist Theory of International Politics*. New York: Oxford University Press.
- Schmidt, E., R. Work, S. Catz, E. Horvitz, S. Chien, A. Jassy, M. Clyburn, G. Louie, C. Darby, W. Mark, K. Ford, J. Matheny, J. Griffiths, K. McFarland, and A. Moore. 2020. *Final Report: National Security Commission on Artificial Intelligence*. Washington, DC: The National Security Commission on Artificial Intelligence.
- Von Clausewitz, Carl. 1949 [1831]. *On War*. London: Routledge and Keagan Paul.
- Waltz, Kenneth. 1983. Nuclear Weapons: More May be Better. *Adelphi Papers*.
- Wilson, Ward. 2008. The Myth of Nuclear Deterrence. *Nonproliferation Review* 15 (3): 421–439.



Liberalism and Artificial Intelligence

5.1 INTRODUCTION

Though considerably more recent than realism, liberalism has a pedigree which stretches centuries and remains as relevant today as ever. Initially a set of descriptions of the ideal society, it has—in the post-Westphalian world—evolved into a complex explanation of world politics. In its wake and under the influence of liberal-leaning statesmen and stateswomen, lasting institutions and processes of global cooperation have been established. Moreover, unlike its rival paradigm of realism, liberalism is able to offer explanations of our modern world in which non-state actors are shaping global politics. Characterised by a firm belief in transparency, norms, law, and the possibility of political, economic and social progress, this theoretical paradigm offers powerful critiques of the emerging world of AI domination. The paradigm has not been immune from some blunders. For example, the belief in democratic values has led to predictions that appeared to downplay the prospect of different forms of political order which would co-exist with, and even challenge, liberal democracy in the post-Cold War era. Most notably Francis Fukuyama, in his essay “End of history?” wrote that “The triumph of the West, of the Western *idea*, is evident first of all in the total exhaustion of viable systematic alternatives to Western liberalism” (Fukuyama 1989: 3). In our analysis, however, liberalism presents a useful set of lenses for understanding AI’s economic and transnational economic facets. These include an expansion

of the democratic peace thesis and the concept of economic interdependence as a prerequisite for peace. Faced with rising AI protectionism, moreover, avowedly liberal democratic states risk becoming like the autocracies from which they seek to shield themselves and their populations. Indeed, a state which props up barriers to “foreign” AI technologies (as the US, India, and other democracies have done with regard to Chinese apps and devices), particularly in times of peace, in part forgoes the claim to being a true bastion of liberalism. This is one among many paradoxes this theory will be invited to unravel in this chapter as AI gains greater prominence in economics, society and politics.

Section 5.2 gives a comprehensive overview of liberal theory including its origins, the arguments it advances and the assumptions upon which these are based. Section 5.3 is an overview of AI and its impact on liberalism. Section 5.4 updates the liberal notion of democratic peace, while Sect. 5.5 updates the concept of interdependence. The penultimate Sect. 5.6 provides a brief discussion of regime types and their influence on AI development. Section 5.7 is an overview of some apparent shortcomings of liberalism in the wake of AI.

5.2 LIBERALISM: ORIGINS, CLAIMS, AND INTERNAL DEBATES

Liberalism’s most recognised originators and propagators are Thomas Hobbes, John Locke, Thomas Paine, and John Stuart Mill, Montesquieu and Jean-Jacques Rousseau. These philosophers wrote between the seventeenth and nineteenth centuries. Their body of work is vast. We note four key innovations of liberalism. Firstly, according to Hobbes and Locke, the state comes about as a result of a compromise people enter into with one another as individuals to avoid anarchy. This phenomenon was put most eloquently by Hobbes when he asserted that in the state’s absence life would be “solitary, poor, nasty, brutish and short” (1651 [2017]). For Locke specifically, it was crucial that the state be arranged in order to grant equality particularly when it came to religious belief and practice. Secondly, Rousseau argued that people create society, and in turn that society creates a social contract with those whom it appoints to rule over it. In other words, the citizenry forgoes some freedoms, in the service of its greater self-interest. These freedoms are entrusted to governments, who are in turn accountable to a population that also reserves a right to revolution when the government becomes autocratic (Rousseau 1763).

Thirdly, in relation to the prior point, liberals such as Mill argued that certain freedoms are not to be violated. Key among these is the freedom of speech and the free press. Finally, Montesquieu (1748) argued that there was a necessary separation of powers in order to restrict their exercise and contain any potential excess. Thus in liberal democratic states, legislatures make the laws, the judiciary interprets them, and the executive enforces (and is restrained through) them.

As a theory of international relations, liberalism argues that the world should be characterised by cooperation and increased trade. Adherents argue that increases in trade lead to interconnectedness that makes military confrontation undesirable and virtually too costly to take place. Liberal institutionalism argues that international institutions should be created to facilitate these processes. The eighteenth-century German philosopher Immanuel Kant articulated in his 1795 essay “Perpetual Peace: A Philosophical Sketch” a theory for a worldwide federation, or government, which would bring about eternal peace. Key to this, he argued, was the uptake of the republican form of government and in turn international cooperation. The United Nations is seen as the closest practical manifestation of this idea; having been conceived following WWII. The UN’s precursor, the League of Nations, had been put into place in the wake of WWI but ultimately failed in its goal of preventing another major conflict outbreak. One of the key factors in the failure of the League of Nations was possibly the absence of US membership and the unwillingness of major member states to act against Italy when it invaded Abyssinia and Japan when it annexed northern China. This hesitance to act was a display of impotence by these major states, among them France and the United Kingdom, as well as the failure of the concept of collective security. Seemingly, then, there can be a wide gap between liberal ideas and actual practice.

Moreover, not all liberal ideas are without criticism. One of the most criticised liberal variations is that of neoliberalism. This is the name given to a set of economic and political principles which include free trade, privatisation and “rolling back” of state involvement in the industry. In practice, this led to the World Bank and IMF enforcing “structural adjustment programs” in poorer countries. This, for example, led to decreases in social spending (including in health) and support for state-owned enterprises. At the same time, these smaller countries were required to open up their markets to large multinational corporations (MNCs). The system of embedded liberalism—the idea of free markets co-existing with the

welfare state—had enjoyed widespread support from the 1940s to the 1970s. By the 1980s, however, it gained considerable criticism in Britain and later the United States where two leaders (Margaret Thatcher and Ronald Reagan) were elected. Their avowed aims were to roll back much of the welfare spending in their respective countries and cut corporate taxes and subsequently export elements of this model to other parts of the world through the Bretton Woods institutions. In the United Kingdom, it also coincided with the privatisation of many state-owned enterprises. As a part of consequence of these ideas, the world underwent a rapid pace of globalisation. Key to this was technological innovation, including new milestones in AI as discussed in Chapter 3. What can the ideas of Hobbes, Locke, Montesquieu, Rousseau, and Mill say to us in today’s world of AI saturation? We discuss this in the following section.

5.3 AI AND LIBERALISM

There are a number of issues on which liberalism falls short in the wake of AI. Most evidently, the concept of a free market has proved too idealistic in light of the unseemly relationship between governments and big corporations to benefit the latter. Indeed, the past and present of AI development cannot be separated from the role played by funding ties to the military in research and development. Secondly, the faith in the market self-regulating to find the ideal price is also to be problematised in today’s world. Demand and supply are no longer a general phenomenon. Research shows that AI can funnel prices to individual consumers on online platforms (Marwala and Hurwitz 2017). As a result, we can no longer speak of a single market. Rather, the market has become segmented and individualised. Algorithms are able to predict, based on individual-level data, an individual’s disposable income and present them an ideal individual price.

Liberalism’s history and the above-highlighted shortcomings notwithstanding, the theory still has considerable explanatory value. The rise of AI has coincided with the decline of countries classified as liberal democracies in the world. Democracy and artificial intelligence appear to be having a negative correlation with one another. The more AI has become diffused, the fewer countries have qualified as free societies.¹

¹ <https://www.dailymaverick.co.za/article/2021-03-22-artificial-intelligence-and-emerging-technologies-are-powerful-tools-but-can-be-bad-for-democracy/>.

Freedom House observes that there has been democratic backsliding in every region in the 2010s. The year 2020 marked the fifteenth consecutive year of decline in the number of countries classified as “free,” a reversal of the trajectory seen from the early 1990s until 2005. Today, fewer than 20% of people in the world live in countries that reach this classification (Ndzendze and Marwala 2021a, b). Matters could hardly have been helped by COVID-19. The Economist Intelligence Unit (EIU) opens its 2020 report with the observation that “across the world in 2020, citizens experienced the biggest rollback of individual freedoms ever undertaken by governments during peacetime (and perhaps even in wartime).” For many states, COVID-19 has led to a rationale for greater surveillance, leading to a growth in the ubiquity of emerging technologies with application in tracking and tracing. The long-term effect of this will be “a question for the future, but it contains the ingredients for tech-based authoritarianism”.² With widespread recognition of the role of technology in lowering the costs of mobilisation by citizen and authorities alike, regimes have taken steps to deny access to digital platforms, as well as divide mobilising citizens by using bots and misinformation on those platforms.

5.4 DEMOCRATIC PEACE THESIS AND AI

The democratic peace thesis is particularly affected by the changes which have occurred and are set to further occur in the wake of AI within states. A powerful subset argument within this theory argues that domestic political factors explain the complete absence of military violence between democracies; indeed no two democracies have ever had a conventional military conflict (Marwala and Lagazio 2011: 3–5). Thus, whereas structural realism stipulates that states undertake military decisions based on evaluations of military capabilities in an environment of anarchy and information asymmetry, this variation of the liberal paradigm argues that the key determinant is whether the states involved are democratic or not. Despite the universal agreement on the tendency of democracies to not go to war with one another, it is still a matter of debate as to *how* this is achieved.

² <https://www.dailymaverick.co.za/article/2021-03-22-artificial-intelligence-and-emerging-technologies-are-powerful-tools-but-can-be-bad-for-democracy/>.

Democratic peace mainly takes two forms as a result: the monadic argument and the dyadic argument (Russett 1993: 164). These two propositions are set apart mainly by the different levels of emphasis they each place on the regime type of the targeted state within a potentially warring dyad (pair of states). The monadic proposition argues that the regime type of targeted states does not matter and that there are inherent variables within democracies in themselves that make them less likely to initiate wars with one another. The dyadic proposition considers the regime type of the targeted state to be of significant importance in such outcomes.

The monadic proposition asserts that there is a correlation between the level of democracy which characterises a state and its disposition towards other states, regardless of their regime type (Rummel 1995: 457). In support of this argument, Rummel (1995: 457) argues that democracies are the regime type that is the most passive above all others. Morgan and Schwebach (1992: 305) have likewise found that countries which are democratic are unlikely, or less likely, to escalate arising disputes into wars. According to this structural (or institutional) model, democratic states are able to maintain peace with one another due to the checks and balances which are codified into their constitutions and restrain the decision-makers. Also crucial to this is the existence of a vibrant civil society. All these converge to make the idea of a war-mongering democratic state highly unlikely (Bueno de Mesquita and Lalman 1992). Moreover, there is a long time lag between the decision to initiate war and when such a decision may be executed: leaders who gain power through democratic means are subject to separation of powers which, according to Russett (1993: 92), leads them to a path of caution in foreign policy conduct, thereby diminishing the likelihood of escalation. These institutional constraints in addition to regular elections in which leaders are forced to face appraisal for their policies and track record over the preceding period, mean that a reckless foreign policy or war initiation is likely to face removal from office (Bueno de Mesquita and Lalman 1992).

According to the dyadic account, democratic states are not necessarily less bellicose, but they are restrained from going to war with one another for reasons to do with the form of interaction they are forced into by virtue of their shared system (Doyle 1986: 1161). Contrary to the monadic proposition, and arguing that it misses the causal explanation as to why democracies are less war-prone when it comes to each

other (seeing as they regularly go into war with other regime types), Owen puts it as follows: “States sharing republican norms may be more willing to bargain, compromise” well as “[fulfil] contracts than states without these norms” (Owen 1994: 93–98). On the other hand, hostility towards states that are not democratic is a higher probability because it is easier to rally public support within democracies for military actions. The asserted reason for this is that non-democratic regimes are capable of being portrayed as illegitimate within their own states and, if they are especially autocratic, as being at war with their own populations (Doyle 1986: 1161). In this regard, liberal states are determined by some liberal scholars to be more peaceful in their orientation towards fellow liberal democracies because they ostensibly know themselves to be “reasonable, predictable, and trustworthy because they are governed by their citizens’ true interests, which harmonise with all individuals’ true interests around the world” (Owen 1994: 95) they, therefore, have a better understanding (contrary to realism, see Chapter 4) which sees all states as being in a state of information asymmetry regardless of regime type. This has led some liberal scholars to argue for a qualified typology of democratic regimes because some regimes may carry out procedures that seem democratic but may in fact be ruled over by governments that do not adhere to liberal values and practices (Owen 1994: 89). These regimes are termed “illiberal democracies.” Thus, even when war threats are made, illiberal leaders would not be able to “rally the public to fight,” as well as make these leaders “fear that an unpopular war would lead to their ouster in the next election” (Owen 1994: 89).

The monadic position is the most impacted by the advent of AI, though the dyadic literature itself requires some updating. The monadic position relies on the flow of common, nearly universally agreed-upon information. The tools of information warfare have a tendency to target “enemy population beliefs, enemy leadership beliefs, and the economic and political information systems upon which society relies to function” (Bishop and Goldman 2003: 119). More fundamentally, in a world where people are not the main fighting force anymore—or where, when deployed, they play a supportive role to drones—to what extent do they still need to be consulted on war initiation? As such, what becomes of DPT assumptions, which understand civil society and the populace as involved, active stakeholders? General Latiff of the United States observes already that “in the absence of clear and unambiguous public involvement, the military will respond to events in the way it deems most

appropriate” (Latiff 2017: 10). In a group that tends to be homogeneous, the literature indicates that there may be a consensus on war as a policy than on other diplomatic routes. Nevertheless, the growth of communicative technologies paved the way for increased public participation. This will likely operate on the basis of information that will be even more susceptible to manipulation, however, such as when the George W. Bush administration sought to convince the US population—and use as a justification to Congress—the claim that Saddam Hussein’s Iraq had weapons of mass destruction and, moreover, was likely to share these with terrorists. As Bishop and Goldman (2003: 115) put it, “we have entered an age where information is not only an adjunct to conventional military and business operations, but has become a key arena of conflict and competition.” This was in 2003. The situation has become substantially more complex in the era of AI. Figure 5.1 represents our model of how misinformation will distort the framework of democratic peace, through manipulation (among others by DeepFakes) of the population such that they do not confidently know what is real and what is only apparent.

A wave of global wave protests in 2019 that resulted in the resignations of the national leaders of Bolivia, Iraq, and Puerto Rico and a rollback of a piece of legislation in Hong Kong put on display that the public is not as complacent as is sometimes portrayed, and at the centre of the coordination methods of the various campaigners was social media.

Dyadic democratic peace also needs to explicate whether the growing political populism, which has led to different types of democracies with varying attitudes towards the global order and pluralism, can still be said

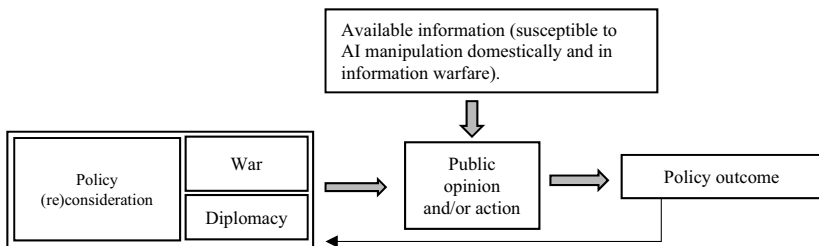


Fig. 5.1 Proposed interaction of democratic peace thesis (constitutional constraints variant) and AI (e.g. DeepFakes) (Source Authors. See also Artificial Intelligence and Emerging Technologies in International Relations [Ndzendze and Marwala 2021a, b])

to subscribe to the same set of principles such that war between or among them is unthinkable. The role of AI in playing up issues that further stratify voters, and in turn elect ideologically divisive governments, has consequences for the international order which are yet to be evaluated.

5.5 THEORETICAL UPDATING: ECONOMIC INTERDEPENDENCE

The second theory we consider to be relevant and impacted by the emergence of AI is the concept of interdependence. So far two trends in the literature have formed but have not been connected, this chapter is the first study to further take this up: one has focused on the job losses and structural changes to whole economies that are materialising. The second has seen AI gaining greater application in war, despite current limitations (stemming from object recognition difficulties that make them unable to be completely independent in battle as indicated in Chapter 3). We argue for a diminished interdependence, which will escalate tensions due to lowered opportunity cost among states.

As early as the eighteenth century, Montesquieu (who in his 1748 work, *The Spirit of the Laws*, argued that “movable wealth” encouraged “peace between and within states”), scholars have asserted the view that “peace is a positive externality of global commerce” (Gartzke et al. 2001: 392). In Book 20 of the same work, Montesquieu (1748: 346) also further stated that “peace is the natural effect of trade.” He makes the argument that two nations who conduct trade with one another are tied into a state of reciprocal dependence, “for if one has an interest in buying, the other has an interest in selling; and thus their union is founded on their mutual necessities.” This is borne out by empirical studies in recent years. Similar to the democratic peace thesis, therefore, contemporary debates seem to be around the mechanisms by which trade interdependence causes peace. Papayoanou (1999) contends that:

Economic linkages act as signals of resolve and credibility: because domestic economic actors in status quo states only support conflicts that protect their interests, these states are more easily constrained from balancing against revisionist states with which they share economic relations. (Papayoanou 1999: 16)

Trade interdependence may also have an indirect impact through reconfiguring the preferences of states in such a way that they are incentivised to no longer see gains from competition with one another (Solingen 1998: 9). For Solingen (1998: 9), the key mechanism is the linkage of domestic actors with international commercial interests. Stated simply, domestic elites who have internationalist preferences act as motivated agents for cross-national bonds, and thereby facilitate interdependence with foreign countries. These become mutually reinforcing. For Gartzke et al. (2001: 391), on the other hand, the prospect of disruption to economic linkages to capital leads to an occasional deterrence of small contests between states, while also not creating enough of an opportunity cost for larger militarised disputes. They thus offer an additional mechanism for how interdependence leads to peace and argue instead that interdependence presents means which are not military but are serious and costly enough for communicating signals of resolve for commitment to a warpath by one state towards another (Gartzke et al. 2001: 392).

The rise of advanced automation and market intelligence, at the behest of AI, is in the process of re-making the trade patterns which have defined the post-Cold War era of globalisation, and global value chains are undergoing transformation (Charalambous et al. 2019: 1) which will likely reduce the peace dividend of the present status quo for the majority of the globe. In this way, then, we propose that it will reduce the opportunity cost for initiating conflicts for those states who will be “left behind” (see Fig. 5.2). This section therefore draws a connection which has not been noted; we hypothesise that economic transformation in the wake of AI will disrupt the economic interdependence and thus the basis for interstate peace which has held since the emergence and growth of the present age of globalisation. This is represented in Fig. 5.2.

Moreover, market intelligence could be used to predict and fulfil consumption patterns of the population of a rival state’s ally and thereby possibly isolate it from its reliable network of allies. At the very least, this would split the considerations of the ally state (itself accountable to an economically rational public) and result in a slower decision-making process. And that would be the intended consequence.

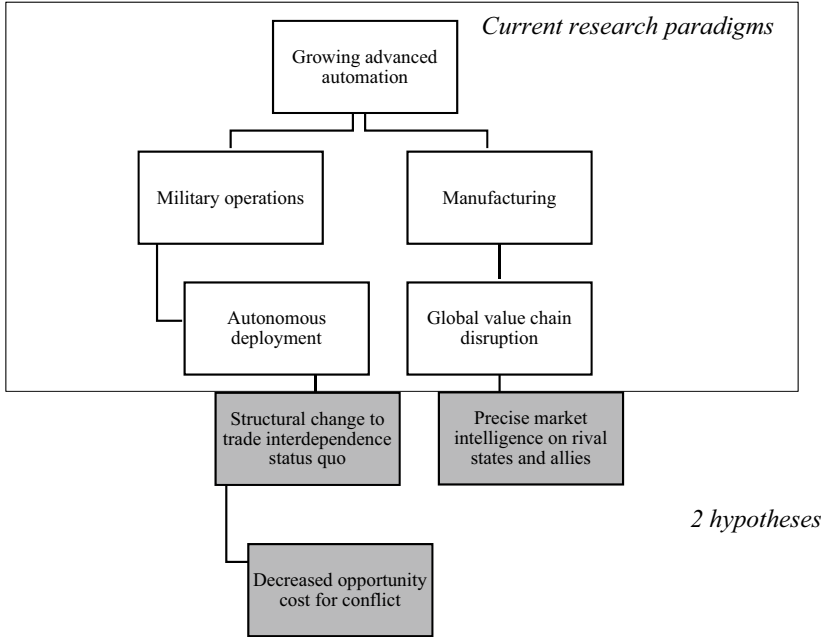


Fig. 5.2 Proposed automation-trade interdependence status quo disruption (*Source* Authors. See also Artificial Intelligence and Emerging Technologies in International Relations [Ndzendze and Marwala 2021a, b])

5.6 REGIME TYPE AND AI NICHES

Different types of regimes have already demonstrated different niches when it comes to AI. Essentially, all have advantages and setbacks. These are evident in the output seen in AI by some of the major states, regardless of regime type. This disregards it as a differentiating factor. Table 5.1 demonstrates that democratic states have the advantage of collaborative and open societies based on meritocracy which leads to the best engineers working on developing AI programs. On the other hand, authoritarian governments have few to no regulations on privacy, and thus developers working in these countries have the advantage of huge data troves. However, both face challenges that stem from overregulation for democracies and inconducive cultures and incentives in authoritarian states.

Table 5.1 Regime types' niches and setbacks in AI development

<i>Regime type</i>	<i>Niche</i>	<i>Setbacks</i>
Democratic	Collaborative, open societies typically meritocratic	Restrictive, regulatory environments on data acquisition and experimentation
Authoritarian	Fewer privacy regulations on government-aligned developers and easier data accumulation	Conservative, hierarchical culture
Hybrid	Combination of democratic and authoritarian advantages	Incoherent, unpredictable environment for development

Source Authors

Can democracies be able to maintain an exclusive form of collaboration in the era of AI? Somewhat paradoxically, to do so would be to abrogate an essential component of liberal democracy; they will be less open and be more restrictive towards different types of states. That is to say: the more exclusionary they are in the pursuit of safeguarding liberal values, the more they will have to rely on anti-liberal methods. This will render democracies less liberal, and perhaps even less democratic as they have to stand in the way of their own citizens' enshrined rights to consumption and procurement of technologies. This ideology- and survival-driven protectionism is nothing essentially new in modern international politics, but it was always propped up during times of war (Jefferson's Embargo), economic downturn (Smoot-Hawley Tariffs, for example) or under cases of Cold War politics (the Cuban embargo by the United States is the last vestige of this). The onset of AI presents new rationales, and new levers in addition to a new platform. This also raises new questions. Principally, what does a full-fledged AI embargo look like? Is it possible? And given the interconnectedness of data systems and data infrastructure, what are the full implications? Might democracies have become too inexorably linked with non-democracies to render them distinguishable? There are no easy answers to any of these, but the chief effect of this chapter should be to demonstrate liberalism's role in the era of AI.

5.7 CONCLUSION

Broadly speaking, technology and politics are two manifestations of the same long-term endeavour, namely humanity's desire to tame the world and make it more predictable. In other words, they are both human-made creations that are intended to be mechanisms for efficiency and management of scarcity. Liberalism highlights the importance of distribution and decentralisation of power. The deployment of AI should serve human ends and as such ought to be subject to checks and balances that reflect a form of updated social contract. Crucially, the players are no longer society and the state, as in Rousseau's era, but also powerful corporations.

BIBLIOGRAPHY

- Bishop, Matt, and Emily Goldman. 2003. The Strategy and Tactics of Information Warfare. *Contemporary Security Policy* 24 (1): 113–139.
- Bueno de Mesquita, Bruce, and David Lalman. 1992. Domestic Opposition and Foreign War. *American Political Science Review* 84 (1): 747–766.
- Charalambous, Eleftherios, Robert Feldmann, Gérard Richter, and Christoph Schmitz. 2019. *AI in Production: A Game Changer for Manufacturers with Heavy Assets*, March. McKinsey and Company. Available at: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ai-in-production-a-game-changer-for-manufacturers-with-heavy-assets>. Accessed 3 Apr 2020.
- Doyle, Michael W. 1986. Liberalism and World Politics. *American Political Science Review* 80 (1): 1151–1169.
- Fukuyama, F. 1989. The End of History? *The National Interest* 16 (1): 3–18.
- Gartzke, Erik, Quan Li, and Charles Boehmer. 2001. Investing in the Peace: Economic Interdependence and International Conflict. *International Organization* 55 (2): 391–438.
- Hobbes, T. 1651 [2017]. *The Leviathan*. London: Penguin.
- Latiff, Robert H. 2017. *Future War: Preparing for the New Global Battlefield*. New York: Knopf.
- Marwala, Tshilidzi, and Evan Hurwitz. 2017. *Artificial Intelligence and Economic Theories: Skynet in the Market*. Heidelberg: Springer.
- Marwala, Tshilidzi, and Monica Lagazio. 2011. *Militarized Conflict Modelling Using Computational Intelligence*. London: Springer-Verlag.
- Montesquieu, Charles-Louis. 1748 [2001]. *The Spirit of the Laws*, trans. Thomas Nugent. Ontario: Botache Books. <https://socialsciences.mcmaster.ca/econ/ugcm/3ll3/montesquieu/spiritoflaws.pdf>.

- Morgan, C.T., and V. Schwebach. 1992. Take Two Democracies and Call Me in the Morning: A Prescription for Peace? *International Interactions* 19 (4): 305–320.
- Ndzendze, Bhaso, and Tshilidzi Marwala. 2021a. Liberal Thought Has a Place in the Era of Artificial Intelligence. *Thought Leader*, January 28. <https://thoughtleader.co.za/liberal-thought-has-a-place-in-the-era-of-artificial-intelligence/>.
- Ndzendze, Bhaso, and Tshilidzi Marwala. 2021b. Artificial Intelligence And Emerging Technologies Are Powerful Tools – But Can Be Bad for Democracy. *Daily Maverick*, March 22. <https://www.dailymaverick.co.za/article/2021b-03-22-artificial-intelligence-and-emerging-technologies-are-powerful-tools-but-can-be-bad-for-democracy/>.
- Owen, John M. 1994. How Liberalism Produces Democratic Peace. *International Security* 19 (2): 87–125.
- Papayoanou, Paul A. 1999. *Power Ties: Economic Interdependence, Balancing, and War*. Ann Arbor: University of Michigan Press.
- Rousseau, J. 1763 [2018]. *The Social Contract*. Hertfordshire: Wordsworth.
- Rummel, Rudolph J. 1995. Democracies Are Less Warlike Than Other Regimes. *European Journal of International Relations* 1 (1): 457–479.
- Russett, Bruce M. 1993. *Grasping the Democratic Peace: Principles for a Post-Cold War*. Princeton, NJ: Princeton University Press.
- Solingen, Etel. 1998. *Regional Orders at Century's Dawn: Global and Domestic Influences on Grand Strategy*. Princeton, NJ: Princeton University Press.
- Von Clausewitz, Karl. 1949 [1832]. *On War*. London: Routledge and Keagan Paul.



Hegemonic Stability Theory and Artificial Intelligence

6.1 INTRODUCTION

China's Sputnik moment—such is how Kai-Fu Lee (2018: 1) characterises the three-day showdown in 2017 between Ke Jie, China's esteemed world champion in the ancient Chinese game of Go with Google's AI-powered AlphaGo. In that set of games, the grandmaster lost all matches to the machine.¹ Read at face value, the analogy by Lee, put forth in his book *AI Superpowers: China, Silicon Valley, and the New World Order*, and variations of it by other scholars, nominally raises a great alarm. Such comparisons may conjure up scenarios of superpowers at loggerheads, with unmeasurable power in their possession. Yet read in the fuller context of the Cold War, the USSR parallel is an optimistic one. The full story of the Cold War is not only about the intense rivalry between Washington and Moscow but also its lack of confrontation in direct war. The root of this aggregate pacificism (resorting to proxy wars in other parts of the world and espionage) is still debated by IR scholars, with MAD being one explanation among many (see Chapter 4). Indeed, one of the near casualties of the Cold War was realist theory itself, before being reinstated by the events of September 11th in 2001.

¹ Moreover, AlphaGo was playing not on the basis of pre-programmed instructions, but continual adaption through ML.

One of the explanatory factors attributed to the Cold War's trajectory is the widening and irreversible quantitative and qualitative military gap in favour of the United States over the USSR which began to be clearest from the late 1970s (having been the subject of debate in the 1950s and 1960s, though historical analysis shows that the United States had always had the advantage, despite claims to the contrary by political parties when they were out of power). In other words, whereas the early Cold War had been defined by a *growing* USSR that at times appeared capable of posing a credible threat to the United States (most prominently during the Cuban Missile Crisis in 1962 which sought to close the qualitative gap, if not the quantitative one, by placing nuclear weapons some 90 miles from the coast of Florida), the middle and late period was characterised by visibly diminishing Soviet capabilities on many fronts, including financially, technologically and militarily. The disastrous invasion and occupation of Afghanistan in 1979 best signifies this in retrospect. This culminated with the collapse of the Berlin Wall in 1989. Subsequent decades have seen growing financial globalisation, with many former Soviet satellites being drawn into the US-led order described as the "Washington Consensus." In this way, scholars who adhere to the hegemonic stability theory (HST) argue that the emergence of a single superpower is concomitant with durable peace and commercial activity. Historically, this was provided by Pax Romana (especially in the Mediterranean, and the various regions of Afro-Eurasia surrounding it), Pax Sinica (seen in the early Silk Road), Pax Mongolica under the Mongol Empire of Genghis Khan (under whom for the first and only time in its history, the entire Silk Route was ruled by a single power),² and Pax Britannia (whose navy was the most powerful in the world between the sixteenth and nineteenth centuries). Most recently some scholars have noted a "Pax Americana" with US presence in the Middle East and the Pacific (especially the crucial Straits of Malacca and the Straits of Hormuz). The hegemonic school contends, in line with the balance of power theories of which it is a variant that the root of the conflict lies not in unilateral power, but in the transformation or change in its equilibrium (Kennedy 1989). This is nominally understood as the emergence of a rising power

² UNESCO. 2020. "The Silk Route of the Mongols The Coming of Ghengis Khan," UNESCO. Available at: <https://en.unesco.org/silkroad/sites/silkroad/files/knowledgebank-article/the%20silk%20route%20of%20the%20mongols.pdf> (Last accessed: 22 April 2020).

in the face of a status quo superpower. As Thucydides famously put it: “it was the rise of Athens and the fear that this instilled in Sparta that made war inevitable.” This phrase has enjoyed a renaissance in recent years, due in large part to Graham Allison’s (2017) bestselling account of the US-China rivalry and the prospects of it falling into this “Thucydides trap.” Though not formally an IR scholar himself, Kai-Fu Lee likewise makes a considered case for caution, noting that conflict between the two is not inevitable but that cooperation is not a foregone conclusion either: recognising that “businesses and scholars have turned China into a bona fide superpower, the only true national counterweight to the United States in this emerging technology,” he therefore notes that the “how these two countries choose to compete or cooperate in AI will have dramatic implications for global economies and governance” (Lee 2018: x).

References to hegemony and the AI race between the United States and China have been made by leading IR theorists and practitioners (see Allison 2017, 2019; Kissinger 2018), but they have done so either in passing, on contingent issues, or (quite commonly) from the point of view of their respective countries. In other words, the literature has not been in a systematic and theory-generating or theory-testing fashion aiming to highlight implications for the theory most concerned with hegemony and technology; hegemonic stability theory. This chapter seeks to bridge this gap. Situating the social dynamics of the next number of decades as being headed for a near-ubiquity of cyber-physical interfacing (Xing and Marwala 2018: 7), the chapter therefore extends this to the global sphere, with particular relevance for the future of the balance of power equilibria related to AI. Our analysis necessitates an extension of the understanding of hegemony to include the future shaping of AI ethics, as well as a new operationalisation of HST as a whole. In sum, utilising the insights of hegemonic stability scholarship, we note that the theory can be restated as follows: sustained or increased relative economic size by the state with the most comparatively advanced militarily applicable AI leads to diminished prospects of conflict within its geographic proximity and spheres of influence.

Section 6.2 of this chapter reviews the assertions made by adherents of HST and the internal debates among them. Section 6.3 gives an account of HST-relevant developments in AI (and future directions), particularly in military affairs. Section 6.4 provides conceptual updates on HST due to the rise of AI. Section 6.5 concludes with potential shortcomings in

HST accounting for an AI-dominated military sphere and areas for further research needed in closing these gaps.

6.2 HST: DEFINITIONS, DEBATES, AND HISTORIES

As seen in Chapters 4 and 5, while liberalism puts forward democracy and commerce to bring about world peace, more realist theories are of the view that a stable international order can only be attained through optimal power distribution. One scholar argues that “when power is distributed unevenly, it is often the result of hegemony, defined as a situation in which a single powerful state controls or dominates the lesser states in the system” (Gilpin 1988: 29). Political hegemony is difficult to define, but it has been characterised by Goldstein as at least the ability of a state “to dominate the world militarily” (Goldstein 2012: 281). For three generations of scholarship, it has been put forward by a litany of scholars that there needs to be a state that acts as a hegemon if anarchy is to be made stable for trade to prosper (Ogunnubi and Okeke-Uzodike 2016: 110). HST is essentially the assertion that a single country can dominate a world order and create conditions for a more open and stable commercial environment than would be the case in a bipolar or multipolar order (Ogunnubi and Okeke-Uzodike 2016: 110). Earlier Kindleberger (1973: 28) has asserted that a peaceful international order needs there to be a state that is able and willing to set the standards and rules of the game for other countries, which are left little choice but to follow its lead.

Hegemony relies on two main determinants: military power and economic power. These concepts are not interchangeable (and certainly one does not easily convert into the other), but they do reinforce one another. While economic power arises first, a hegemon needs to have clear military capabilities (Sachse 1989: 7). Most importantly, a hegemon must have the demonstrable ability to halt other states using their military power to close off or interrupt trade (Sachse 1989: 7). More bluntly, Stephen Krasner had earlier stated that during a considerable gap between the capabilities of a hegemon and other states, the former may utilise its military power to force the latter into adopting a system of open trade (1976: 322). This is an important component in what Prys (2010) calls the “the three ‘Ps’ of being a hegemon: the ability to provide economic incentives, project military power and generate perception of these by other nations (2010: 2). However, Krasner emphasises that force alone is inefficient for changing economic policies, and may lead to growing

resistance by middle-tier powers, thereby creating instability (Krasner 1976: 322). Moreover, Keohane argues that the contemporary era poses a dilemma for hegemonic states: they cannot easily resort to military force in the pursuit of their economic aims because that may mean losing their cooperation in other areas of strategic importance (Keohane 1984: 40). Nonetheless, each period is defined by powerful states that dominate the global order. In 1978, Modelski proposed that the global order (or “world system”) has gone through a number of hegemonic cycles. Each of these has lasted about a century (Modelski 1978: 217). According to Modelski, a new cycle comes about in the wake of a global war involves all the hegemonic aspirants:

The result of such a global war is the emergence of one world power that is preponderant and thus able to dominate the system and maintain systemic order. Ultimately, the dominant power loses ground relative to competitors, and eventually the system again disintegrates, resulting in global war. (Modelski 1978: 217)

Accordingly, these have been the cycles:

the period of domination by Portugal from 1494 through 1576-1580; domination by The Netherlands (United Provinces) from 1609 through 1672-1678; a first period of British domination from 1713 through the late 1700s; a second period of British domination from 1815 through 1939; and a period of U.S. domination beginning in 1945. (Modelski 1978: 217)

Taking on board the assumptions of HST scholarship, below we review how the presence or absence of a hegemon predicts for interstate war as an outcome. We also review the implications that flow from this for gauging hegemony in the era of AI.

6.3 CORRELATING HEGEMONY, DECLINE, AND WAR

The empirical strategy for testing for hegemonic stability is rooted in the hypothesis that wars have come about when a hegemonic system underwent dissolution leading to a “system disequilibrium” that would only be resolved through conflict (Sachse 1989: 112). We also note Gilpin’s argument that relative distributions of power of time are the key predictor

of hegemonic stability. Against these two rules, we denote that the independent variable must be a change in relative capabilities over time. This can be gauged through changes in GDP over a period of time, which in turn affects military budgets. Accordingly, no conflicts should break out in periods wherein the hegemon is experiencing higher growth than other states in the system. Thus, the operationalisation of HST can be stated as: Sustained or increased relative economic relative size by the state with the largest comparative GDP in a defined geographical space (global or regionally) does not coincide with an outbreak of an interstate conflict within the region. (For example, in a historical examination of British power and international conflict; Spiezio [1970] identified a correlation between British decline and the outbreak of wars). A generalised version of such events is represented in Fig. 6.1 below; accordingly, we should expect an increased likelihood of conflict in instances when the comparative gap (i.e., not the actual figure by itself) in the military expenditure of the status quo superpower declines considerably. In Fig. 6.1, these are the hypothetical time points 4, 7, and 10 in the time series. Importantly, the periods following the wars may be characterised by the resumption of status of hegemony (such that there is a continuum of hegemony, but successive nation-states).

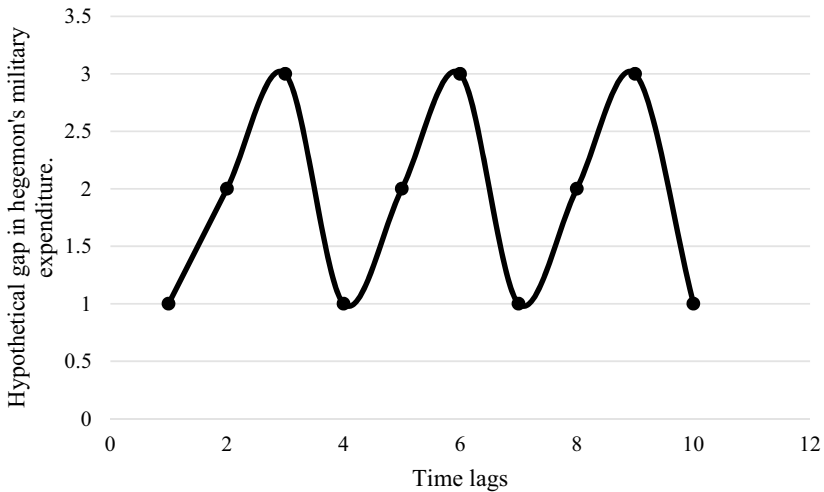


Fig. 6.1 Status quo power military expenditure and war outbreak correlation

The next section reviews developments in AI and their relevance for HST.

6.4 HST AND AI

If, as is assumed in numerous indices, states are compared in terms of their technological capabilities in infrastructure and general ICT, it is just as valid that they compare themselves in terms of technology's implications for their security capabilities, and war. Indeed, in our own times, Cummings (2017: 1) notes that "a metaphorical arms race is in progress in the commercial sphere of autonomous systems development." Vinod Anand (1999: 137) writing for the journal *Strategic Analysis* already at the tail end of the twentieth century, also detailed that technological improvements have always been used for the production of the tools of warfare. As such, the modern age (i.e., since the French Revolution according to his timeline) has seen systematic research in science bolster the development of new innovations for military application. In turn, the tactics of war and strategic thinking are modernised. Further contemporary analysis by Kumar and Batarseh (2020) blatantly predicts that "robots that are quicker, stronger, and more accurate will determine who the victor is." Such robotic warfare is yet to be seen in practice so far, though military expenditure in research and development has continued to grow among militarily leading countries. Indeed, in the US' total annual military budget of US\$600 billion in 2016, some US\$200 billion was set aside for research and development as well as procurement of new weapons systems. The United States, which has the largest budget, is followed by China, setting the scene for intense geostrategic rivalry with commercial nuances and far-reaching implications (Smith and Browne 2019: 253). Both these and other smaller countries, noticeably Israel and South Korea, see close cooperation between the private sectors and governments with a revolving door between innovative corporations and their departments of defence (Senor and Singer 2009).

6.5 AI IN WAR: APPLICABILITY AND LIMITATIONS

At its most basic and general form, AI must be able to replicate human cognitive patterns and mimic them. As a result, AI programmers are able to create machines that perform similarly or reasonably well on activities that human brains are able to do; these include production, as well

as analysis and prediction. The latter two are the most directly relevant to military matters and thus the focus of this section. AI machines are autonomous and adaptable, and increasingly so in the wake of machine learning, quantum computing, soft computing, and access to Big Data. In this manner, and in many other ways, AI can be dangerous. But the opposite is also true as AI can be beneficial to society. Artificial general intelligence, which is still theoretical at this point, is the capability of AI to execute the same cognition as human beings, without any of the shortfalls of machine thinking and thus assume complete responsibility for a variety of functions in governance, diplomacy and military affairs. Experts suggest that such realities are decades or even centuries away given the states of the fields involved.

Already, artificially intelligent systems are active in analytical roles, where they are involved in sorting over large datasets, where they then reach conclusions on the basis of pattern-recognition; “these are precisely the ‘dull’ tasks (of the ‘dull, dirty and dangerous’ formulation) that are generally regarded as the highest priority for automation”. The use of robotics and various other forms of technologies involved in information-gathering and reconnaissance in warfare has been in place since at least WWII. A high number of these early robots (perhaps most notably the American “Aphrodite” drones in 1944) are deemed to have either been not sufficiently effective or to have only been useful for specialised operations, with much of the heavy lifting ending up being done by conventional weapons and manned operations. The use of robotics for military operations enjoyed a renaissance and serious consideration and deployment in the 1990s, with the (MQB-1) “Predator” drone being used by the CIA in the Balkans between July and November of 1995, and Afghanistan as part of search efforts to find Osama bin Laden in September 2000, a year before the 9/11 attacks in New York and the Pentagon, and was further escalated after the attacks.

One of the advantages offered by drones and robotics is distance and remoteness in operations. Rather than meticulously controlling them from close-up radio signals, drones can be controlled through satellites from any location (Kumar and Batarseh 2020). However, despite the notable progression in making robots ever more intelligent, autonomous robots are still found lacking “the flexibility to react appropriately in unforeseen situations” (Nitsch 2013; in Kumar and Batarseh 2020). Thus, contrary to popular perceptions, the deployment of autonomous robots on battlefields means soldiers are acquiring more responsibilities,

not fewer. As such, “soldiers will be expected to perform normal military tasks and use robotic assets as per mission requirements” (Kumar and Batarseh 2020).

Indeed, the process of adoption is rather opportunity cost-based, with the pervading consensus being that in order to enjoy fully autonomous implementation, AI will need to be shown to have similar or greater efficacy as their human counterparts on the battlefield. Put more practically by is the idea that if policymakers already have confidence in the abilities of human analysts on strategic functions such as command and control, there is no basis for turning over these decisions entirely to as yet unproven systems. Overall, then, save by some extraordinary developments, deployment “will not simply be a case of handing over the keys or flipping a switch. There will be no ‘artificial analysts’ ready to simply take on human roles” and instead AI will undergo incremental pairing with human analysts on specific tasks.

Context greatly aids this process, as recent studies have shown that having context (also termed situational awareness [SA] in militaries) is crucial for intelligent computational agents. However, developing and operationalising context for these has proven to be more challenging than the otherwise simpler process of automating repetitive tasks. The key hurdle in this regard stems from the difficulty of defining objects which an AI has not previously encountered. This is against the impressive performance of human agents who have been able to identify danger from intuition, when something did not seem or feel quite right (Kumar and Batarseh 2020). On the other hand, AI continues relying on data from past experiences to be able to perceive subtle but important features such as differences in soil textures or missing equipment. Thus, if an AI agent can offer complimentary assistance to soldiers in noticing dangers through contextual data, they can be more effective and useful in battle.

The second role for AI agents is prediction. Placed in this role, AI systems present an opportunity to policymakers to anticipate future events. One such modality as identified by is in modelling complex negotiations. Additionally, they posit that these can be used in monitoring compliance to international commitments and increasing the capacities of negotiators, in machine learning methods may be used to forecast other players’ positions and tactics. This has not been performed so far, however, as limitations persist. In practice, “while predictive algorithms have been demonstrated with some success in some capacities, they are not yet necessarily more accurate than their human equivalents”.

On the other hand, scholars have used AI to model the prediction of interstate conflicts. In their 2011 book, *Militarized Modelling Using Computational Intelligence*, Marwala and Lagazio developed an AI that considered seven variables that drive interstate conflicts. One of these was the relative difference in the degree of militarisation between the two states. In their model, if two countries have parity in their degree of militarisation, there is a decreased probability of war between them if they are mutually aware of this fact (i.e., one does not see itself as likely to win a confrontation). However, what may be at play is the process of bolstering its military capabilities by one state in the perception of its rival. This problem, termed the security dilemma,³ is also at the heart of HST. Given the inevitability of economic transformation (i.e., a decline of one superpower and the emergence of another, perhaps more revolutionary power), the prospect of a “hegemonic war” becomes increasingly likely (Kennedy 1987; Allison 2017). In this regard, some theoretical revisions on account of the rapid emergence of AI and other relevant technologies in military affairs in the wake of AI and its growing ubiquity (and inequitable distribution) must be conducted. This is turned to in the section below.

6.6 TOWARDS AI-BASED HST THEORETICAL UPDATING?

The foregoing sections have reviewed power dynamics in AI possession, while also noting the applicability of AI in war. This section seeks to detail the implications of these for hegemonic stability. Perhaps the most relevant manifestation of AI’s relevance to HST is the AI race between PRC and United States. To begin with, following the 19th National Congress of the Chinese Communist Party, CCP General Secretary and Chinese President, Xi Jinping, made the announcement that the modernisation of the People’s Liberation Army’s (PLA) would be complete by the year 2035 (McCauley 2018). In this regard, the PLA has well noted the development of AI as “an imperative capability in the future of warfare” in its endeavour for modernisation (McCauley, 2018). This is a major undertaking given how far the PLA has to go. Nonetheless, recent investment trends show the centrality of AI:

³ This is a model which predicts that readiness for conflict can become a self-fulfilling prophecy because the intentions of each state are effectively unknown by others. It is a classic example of the prisoner’s dilemma.

To date, the PLA has pursued research, development, and testing for multiple military applications of AI, including, but not limited to: intelligent and autonomous unmanned systems, such as swarm intelligence; AI-enabled data fusion, information processing, and intelligence analysis; applications in simulation, war-gaming, and training; the use of AI for defense, offense, and command in information warfare; and intelligent support to command decision-making. (Kania 2017: 21)

Additionally, the two countries have overlapping interests and potential clashes (and therefore immediate potential pretexts for military engagement) including the US's chain of allies in East Asia such as Japan (with whom China has a territorial dispute) and Taiwan (regarded by China as a renegade province), and US navigation activities in the South China Sea (regarded by Beijing as part of China's exclusive economic zone and by the United States and much of the international community as free waters). Additionally, competition over international economic presence has led to US suspicion over the PRC's Belt and Road Initiative as a reach for hegemony by the PRC. It is perhaps because of this that the US trade war has not only been characterised by the reciprocal imposition of trade tariffs by the two countries but has also had a technology component. On May 15, 2019, for example, then US President Donald Trump issued an executive order that banned the export of US information technology to designated foreign countries on the grounds of national security. Among the companies most affected was Huawei, which also got banned from making use of Google's Android operating system (OS), in addition to early bans on Huawei technology in US defence-related sectors (and by its Five Eyes intelligence allies). This could be read as an attempt by the United States to prevent the emergence of a revolutionary AI superpower. Moreover, this has been rendered somewhat ineffective as the PRC is no longer a mere importer of technology, but also a leading innovator. By 2017, the year of "China's sputnik moment," the country had already overtaken the United States in terms of funding for AI start-ups (accounting for 48% of all AI venture funding globally [Lee 2018: 5]). The year 2017 is noticeable for these two reasons. Herein perhaps may lie the insightful nature of HST; the disturbance of the AI expenditure equilibrium and the trade war which followed it may be rightly indicative of hegemonic competition. The coincidence of AI-related expenditure as a proportion of national GDP and tensions with the United States may give a window into the necessary revision of HST. This is visualised in

Fig. 6.2; accordingly, we should expect an increased likelihood of conflict in instances when the comparative gap in the AI military expenditure of the status quo superpower declines considerably.

In Fig. 6.2, these are the hypothetical time points 4, 7, and 10 in the time series. Importantly, the periods following the wars may be characterised by the resumption of the status of hegemony (such that there is a continuum of hegemony, but successive nation-states, as in Fig. 6.1). However, the onset of more advanced ML-based AI also brings a qualitative difference that may remake the military capabilities paradigm, such that it is difficult to determine the hegemon. As seen, though, for the foreseeable future completely autonomous AI is unlikely to come about, and thus the military expenditure (inclusive of AI and non-AI expenditure) is likely to be a viable proxy, just as much as accumulation and modernisation of nuclear stockpiles have been concomitant with gross military expenditure for most nuclear states. These make up the top four of the top ten (i.e., the US, China, Russia, and India) military budgets, and the top six overall (with the addition of France and the UK at number six and eight respectively). Therefore the critical link between advancement in militarised AI and industrial AI should reasonably lead us to

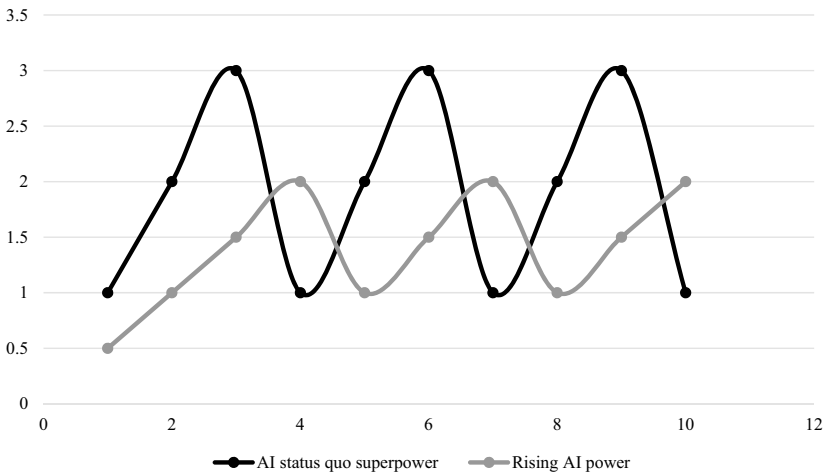


Fig. 6.2 Status quo power AI military expenditure and war outbreak correlation (*Source* Authors)

expect that the state with the largest productive output and economy should also be the one with the largest and most advanced militarily applicable AI. Furthermore, as intelligence (in the covert information-gathering/spying sense), an important aspect and application of AI, also lends itself to perception management and is important to the power projection of countries. Proven capacity for intelligence protection may become a more pronounced and salient factor in coalition formation and the depths of security cooperation as states seek to closely guard intellectual property. This may be behind US efforts at coercing its Five Eyes allies (Australia, Canada, New Zealand, and the UK) from using Chinese technology, and trading in advanced technology with the country since doing so would be to generate reliance on the technological infrastructure of a rival.

Partly because of this, then, AI data ownership and power reproduction are considered important elements of hegemony. As a new phenomenon, AI ubiquity is likely to be defined and regulated by standards and ethics determined by early adopters, and more likely originators. Therefore, it is no coincidence that the publication of AI policies has been correlated with states in possession of AI technologies and leadership in AI patents. Thus, in February 2020, the Pentagon unveiled the five “ethical principles” for AI in warfare, which assume and necessitate the separation of powers through a wide enough pool of responsible expertise with decision-making powers. The first of those proposed was the centrality and responsibility of human judgement in AI warfare. The second was making sure that AI is not biased. The third is ensuring a deep and widespread understanding of how AI works. The fourth is ensuring that the AI systems are always functional and reliable. The final, related to the first, is that humans can control AI to avoid unintended harm. China’s own AI strategy consists not only of increased expenditure but also of defining AI regulatory frameworks and ethics. The 2017-released *New Generation Artificial Intelligence Development Plan* (NGAIDP) outlines a step-by-step process for China to become the world’s leading AI power. In particular, by 2025, NGAIDP calls for China to not only compete in but lead the global AI industry, achieving “leading research results” and spearheading the development of laws, regulations, and ethics for AI (PRC 2017; see also Carter and Crumpler 2019: 4–5).

6.7 CONCLUSION

Through examination of AI applications in warfare (and their limits), we have given an account of AI's implications for HST. In particular, we have emphasised the importance of the inequitable distribution of AI military capabilities, and the congruence of AI into the HST framework. The limitations and lack of full autonomous deployment of AI in contemporary times, therefore, render the GDP/military expenditure framework so far adequate. The fuller implications of AI for HST are yet to manifest themselves. Table 6.1 demonstrates US leadership in proven military applications of AI.

By region, North America leads in military AI research and development (at 45% globally), followed by Europe (20%), the Asia Pacific (30%), the rest of the world accounts for 5%. Also noticeable from the table, however, should also be the number of private companies. This represents the fact that currently, the private sector outweighs governments in terms of investments in R&D on military AI. Noticeably,

Table 6.1 Summary of contemporary AI military applications and developer nations and corporations

<i>AI technology</i>	<i>Developer/origins</i>	<i>Military application(s)</i>
Robotic surgical systems	United States (IBM and Veterans Administration)	Providing remote surgical support and evacuation activities in war zones
Unmanned aerial vehicles	US (General Atomics, Lockheed Martin, Northrop Grumman), China (China Aerospace Science and Technology Corporation) and Israel (leading exporter of small drones)	Surveillance and reconnaissance
Logistics and transportation	United States (IBM and General Dynamics Land Systems [Stryker])	Lowering human cost and increasing efficiency of operations
Target recognition	United States (e.g., DARPA's Target Recognition and Adaption in Contested Environments [TRACE])	Improving ability of systems to identify the position of targets through probability-based prediction of target behaviour

Source Authors. Information adapted from Singh and Gulhane (2018), Wells (2018), and Cummings (2019)

key companies identified in (relatively) open datasets are mainly US-based: Lockheed Martin, Raytheon, Northrop Grumman, IBM, General Dynamics, NVIDIA, Harris Corporation, Charles River Analytics, SAIC, and Leidos. These are followed by the UK and France, represented by BAE Systems and Thales Group respectively. As noted by Davenport (2019), in recent years “the combined research and development spending at the U.S.-headquartered companies Google, Apple, Facebook, IBM, Microsoft and Amazon was \$54 billion,” while United States Defense Department research and development into AI “doesn’t even come close,” with *Bloomberg* reporting that “the Defense Department plans to allocate \$4 billion towards artificial intelligence and machine learning R&D activities in fiscal [year] 2020” which was only about 0.5% of the total budget of the department and less than 10% of the large corporations. Elsewhere, Cummings coheres with this, noting that “the global defence industry is falling behind its commercial counterparts in terms of technology innovation, with the gap only widening as the best and brightest engineers move to the commercial sphere” (2017: 1). Thus, while contending with the issue raised in the preceding section—of how to account for the qualitative-quantitative aspect of military AI—HST has this state and non-state actor equilibrium to consider; states (likely the AI hegemon in this sense) remain the prime movers in the framing of the international regulatory environment within which AI is to operate on a ubiquitous scale. However, given the insights made on the democratic peace thesis and audience costs in Chapter 5 (on Liberalism and AI), implementation and adherence to these will be based on domestic actors keeping their governments accountable to the extent that they can. Ultimately, one of the questions HST will have to answer is where to locate domestic and transnational non-state actors and the cultural/institutional contexts within which they emerge and act. This may require some theoretical dialogue with theories more abreast with such types of actors, particularly neoliberalism and the critical theories, while also taking stock of its realist roots. Inequality is produced by politics and for some theorists it is an engineered outcome. This is considered in the two forthcoming chapters (Chapters 7 and 9). We begin with dependency theory in the next chapter.

BIBLIOGRAPHY

- Allison, Graham. 2017. *Destined for War: Can China and America Escape the Thucydides Trap?* Boston, MA: Houghton Mifflin Harcourt.
- Allison, Graham. 2019. Is China Beating America to AI Supremacy? *The National Interest*. Available at: <https://nationalinterest.org/feature/china-beating-america-ai-supremacy-106861>. Accessed 15 May 2020.
- Anand, V. 1999. Impact of Technology on Conduct of Warfare. *Strategic Analysis* 23 (1): 137–150.
- Carter, William A., and William D. Crumpler. 2019. China's National Strategy for AI. In *Smart Money on Chinese Advances in AI*, 4–6. Washington, DC: Center for Strategic and International Studies.
- Cummings, M. 2017. *Artificial Intelligence and the Future of Warfare*. Policy Commons. Available at: <https://policycommons.net/artifacts/613666/artificialintelligence-and-the-future-of-warfare/1593718/>.
- Cummings, M.L. 2019. *Artificial Intelligence and the Future of Warfare*. Chatham House. Available at: <https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf>. Accessed 3 Jan 2020.
- Davenport, T.H. 2019. China is Catching Up to the US on Artificial Intelligence Research. *The Conversation*. Available at: <https://theconversation.com/china-is-catching-up-to-the-us-on-artificial-intelligence-research-112119>.
- Gilpin, R. 1988. The Theory of Hegemonic War. *Journal of Interdisciplinary History* 18 (4): 591–613.
- Goldstein, J.S. 2012. *Winning the War on War: The Decline of Armed Conflict Worldwide*. New York: Plume.
- Kania, E.B. 2017. *Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power*. Center for a New American Security.
- Kennedy, P. 1987. *The Rise and Fall of the Great Powers*. London: Unwin Hyman.
- Kennedy, Paul. 1989. *The Rise and Fall of the Great Powers*. New York: Random House.
- Keohane, R.O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, NJ: Princeton University Press.
- Kindleberger, C.P. 1973. *The World in Depression 1929–1939*. Berkeley: University of California Press.
- Kissinger, Henry. 2018. How the Enlightenment Ends. *The Atlantic*. Available at: <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>. Accessed 30 Apr 2020.
- Krasner, S. D. 1976. State Power and the Structure of International Trade. *World Politics* 28 (3): 317–347.

- Kumar, Abhinav, and Feras A. Batarseh. 2020. The Use of Robots and Artificial Intelligence in War. *London School of Economics Business Review*, February 17. Available at: <https://blogs.lse.ac.uk/businessreview/2020/02/17/the-use-of-robots-and-artificial-intelligence-in-war/>. Accessed 3 Apr 2020.
- Lee, Kai-Fu. 2018. *AI Superpowers: China, Silicon Valley and the New World Order*. Boston, MA: Houghton Mifflin Harcourt.
- McCauley, Kevin. 2018. The PLA Accelerates Modernization Plans. *The Jamestown Foundation* 18 (1): 1–5.
- Modelski, George. 1978. *The Long Cycle of Global Politics and the Nation-State*. Cambridge: Cambridge University Press.
- Ogunnubi, O., and Okeke-Uzodike, U. 2016. Can Nigeria be Africa's hegemon?. *African Security Review* 25 (2): 110–128.
- Prys, M. 2010. *Regions, Hegemony and Power: India and South Africa Compared*. PhD Dissertation, University of Oxford.
- Sachse, V.E. 1989. *Hegemonic Stability Theory: An Examination*. PhD Thesis, Louisiana State University and Agricultural & Mechanical College.
- Senor, Dan, and Saul Singer. 2009. *Start-Up Nation: The Story of Israel's Economic Miracle*. New York: Twelve Books.
- Singh, Tejaswi, and Amit Gulhane. 2018. 8 Key Military Applications for Artificial Intelligence in 2018. *Market Research*. Available at: <https://blog.marketresearch.com/8-key-military-applications-for-artificial-intelligence-in-2018>. Accessed 14 May 2020.
- Smith, Brad, and Carol Ann Browne. 2019. *Tools and Weapons: The Promise and Peril of the Digital Age*. London: Hodder and Stoughton.
- Wells, Sarah. 2018. IBM Watson Health and the VA Extends Partnership in Cancer Research. *Tech Crunch*. Available at: <https://techcrunch.com/2018/07/19/ibm-watson-health-and-the-va-extends-partnership-in-cancer-research/>. Accessed 14 May 2020.
- Xing, Bo, and Tshilidzi Marwala. 2018. *Smart Maintenance for Human-Robot Interaction: An Intelligent Search Algorithmic Perspective*. Cham, Switzerland: Springer.



Dependency and Technology in the 4IR

7.1 INTRODUCTION

A growing body of literature, flanked by a litany of governmental policy documents the world over, notes the global economy to be entering the 4IR. The defining changes to be brought on by this revolution are not emerging in a historical void. Indeed, the present global economy is underlined by inequality among the countries of the world. In many ways, the first three industrial revolutions shaped today's economic gaps among countries. Thus, technological readiness and capacity for innovation, which are tracked by the World Economic Forum (WEF) through a specialised index and are among the factors that arguably determine the retention and attraction of new FDI, could be important factors for new inflows. Conversely, however, FDI itself could play a role in their presence and maintenance. Although poor and technologically lagging states can hypothetically affect a rapid turnaround in their fortunes around in this regard, stories of late developers managing to do this are a rarity. Thus countries that lack technological readiness could be seeing less FDI because they do not have technology and innovation to a satisfactory degree for international investors. Conversely, higher FDI-netting countries could be experiencing the same or more FDI influxes on a year-to-year basis because they have high levels of technology and innovation (as predicted by dependency theory, where inherent/historic advantages

accrue further advantages, with much the same pattern for the historically disadvantaged). Furthermore, technological readiness factors, as measured by the WEF, are one out of many factors for FDI attraction. Furthermore, the absorptiveness of a country towards certain technologies is determined by its placement in the “catch-up” stage of its economy. For example, the Yokota–Tomohara model shows that the level of technology transfer spillovers (particularly between low- and high-technology transfers) varies depending on domestic skills levels, with countries that have skilled labour scarcities having technological spillovers only in low-tech industries; while high levels of skills seem to correlate with transfers in more high-tech areas (Yokota and Tomohara 2010: 5). But this model was developed in 2010, before the proposed timeframe in this chapter, during which the 4IR has become in vogue. This carves out a niche for the present study. Moreover, this chapter notes that these differences play themselves out unevenly on the global scale, with distinct winners and losers (Piketty 2014); noticeably, seven out of ten of the lowest scoring states in the WEF’s index at the onset of the study were on the African continent. This geographically linked typological framing presents the potential for specific proposals for policy measures and deepened analysis in further studies.

Section 7.2 is a background on the fourth industrial revolution. It is also a review of the literature on this revolution and its underlying inequality. Section 7.3 provides the conceptual framework of the WEF’s GCI, extracting the key concepts tested. These are in turn operationalised in Sect. 7.4, which is a description of the methodology used in the chapter. Section 7.5 presents the data sourced from the WEF reports and the World Bank for FDI influxes into the ten best scoring and the ten lowest scoring countries per the WEF reports over the 2009–2017 timeframe. The chapter then provides an analysis of the data findings in Sect. 7.6. The chapter concludes in Sect. 7.8 by noting the insights made and explores areas for further research in future studies.

7.2 LITERATURE

Various scholars have problematised and investigated the determinants of technology transfer. In his study, Griozard (2009), researches the determining factors behind technology imports using data from across 80 countries over the period between 1970 and 1985 in order to find out the role played by various factors in influencing technology importation. In

the study, findings showed that domestic investment, FDI, and quality of IP systems have a positive correlation with technology transfer (Groizard 2009: 1526). This demonstrates that countries without pre-existing financial means and technologies can have a positively reinforcing circular flow; therefore, poorer countries can still take measures to increase their probabilities of success in a capital-driven global economy. Groizard's results show that investment in human capital and skilling bode well for tech transfers from imports. Additionally, there is evidence that "the role of FDI in fostering technology imports is higher the larger the protection of IPRs [intellectual property rights]" (Groizard 2009: 1526).

The same author shows that there is some cross-country evidence that demonstrates that a state's openness to trade leads to technology adoption, as seen in their total factor productivity, as well as larger investment in computers or rates of adoption of specific innovations (pp. 1527–1528). As discussed in numerous works (see Ndzendze and Marwala 2021), domestic demand and consumption are also key factors in the degree to which technologies will take root. There are, however, other means through which technologies move across borders. The most prominent is foreign direct investment. In recent years, for example, African governments have made it a condition of Chinese FDI that operations are largely staffed by their citizens, along with commitments to skills transfers on certain technologies by the investors and management of the foreign firms.

Given the pre-existing economic and technological advantages that different states have, as a result of the first, second, and third industrial revolutions (see Shafaeddin 1998; Inikori 2002), this chapter seeks to review whether capital is following this pattern in the nascent fourth industrial era. If this is the case, it would demonstrate a pattern akin to the so-called dependency as it would perpetuate the global specialisation of labour that favours some countries over others.

Specifically, this chapter investigated the effect of improvements in 4IR indicators in countries that are in different stages of industrialisation as a predictor of FDI inflows. A paper by Yokota and Tomohara shows that the degree of technological spillovers is dependent upon country characteristics, particularly the level of skills a country's population generally has. As a result, countries with fewer skilled citizens experience transfers only in low-tech industries, while those with an abundance of skilled labour see transfers in higher technologies more easily (Yokota and Tomohara 2010: 5). The implications for AI are apparent here. The technology requires

very advanced education in computing, software development, research and market development. Additionally, the demand for infrastructure such as high-powered data centres act as a major barrier to entry.

7.3 CONCEPTUAL FRAMEWORK

Dependency theory can be understood as a means of understanding the global system, both in political and economic terms, as imbalanced towards the countries of the global North (for shorthand, the members of the OECD are usually taken to represent a comprehensive list of these). Dependency theory is thus a method of understanding how political and economic relations of the countries on the periphery have evolved and come to be embedded in the international system, which is also capitalist (Conway and Heynen 2014: 111). At its inception scholars (e.g., Raul Prebisch, Hans Singer, Celso Furtado, and Osvaldo Sunkel) argued that Latin America's (and the broader global South's) state of being marginalised and underdeveloped was due to two reasons: the apparent "specialisation of labour" among states, *and* inequality within states.

Accordingly, this theory determines that there are four different types of states (Ghosh 2001): the centre-core (CC) which are the wealthiest and enjoy the most direct influence over global institutions. Below these, tallied by their descending levels of power, are the periphery core (PC), the centre-periphery (CP), and the periphery of the periphery (PP).

The perception that the countries of the so-called Third World obtained true independence and sovereignty after the formal end of colonialism is deemed to be a fallacy (Conway and Heynen 2014: 113). There persists today an unequal competition that still persists "[as] an extremely powerful, dependency relationship in globalisation's transformative, disciplinary, and destructive influences" (Conway and Heynen 2006). Jha (2012: 8) adds that patterns of exploitation are being seen in greater earnest following the attainment of independence by the global South. Dependency theory's take on the role of technology is perhaps best summed up by Cardoso (1979) when he discusses the effects of the inequality produced by the nature of the pathways endemic to the capitalist system: "there is a financial and technological penetration by the developed capitalist centers of the countries of the periphery and semi-periphery" and this "produces an unbalanced economic structure both within the peripheral societies and between them and the centers [which] leads to limitations on self-sustained growth in the periphery" (Cardoso

1979; Tausch and Heshmati 2009: 5). The question is how much this will persist in the 4IR. There are instruments for tracking this on a year-to-year basis. The most prominent among them is reviewed below, as we lay the groundwork for examining dependency in the era of technology saturation.

7.3.1 *Towards Testing Dependency in the 4IR*

The *WEF Competitiveness Index* reports, which have been churned out since 2004, seek to pin down the determinants of economic growth to a set of 12 pillars. These include: “Institutions; Infrastructure; Macroeconomic environment; Health and primary education; Higher education and training; Goods market efficiency; Labor market efficiency; Financial market development; Technological readiness; Market size; Business sophistication; and Innovation.”

Each pillar is scored out of 10 based on weighted averages of its constituent sub-pillars that are based on answers to survey findings that are then averaged. Moreover, each test is flanked by additional empirical investigations that take stock of trends over a preceding five-year period and compare such changes to results from the survey to ensure coherence with those in other indicators that are congruent with the factors being compared (Schwab 2018; Appendix C: 339). The ninth pillar measures “Technological readiness” (which in recent years accounts for AI). This pillar seeks to measure “the agility with which an economy adopts existing technologies to enhance the productivity of its industries” (WEF 2009: 5). This stems from a recognition of the indispensable role of technology for competitiveness in the modern global economy. The WEF observes that ICT specifically has come to be a general-purpose technology in contemporary times such that it is the bedrock of many other economic sectors. Therefore, access to, and regulation of, ICTs determines the overall trajectory of economies (WEF 2009: 5).

When it comes to technological readiness, the main area of focus is access to these by the firms operating within the given country (WEF 2009). The WEF itself recognises that “the level of technology available to firms in a country needs to be distinguished from the country’s ability to innovate and expand the frontiers of knowledge” (WEF 2009). In recognition of this, the WEF has separate indices for technological readiness and capacity for innovation. The latter constitutes pillar 12. The pillar is of particular use for gauging and comparing “maturing economies”

as these are at the frontier of new knowledge such that the adoption of externally sourced technologies becomes of marginal consequence to their economic competitiveness (WEF 2009: 7).

7.4 METHODS

In order to determine whether there is a correlation between FDI influx and a single selected 4IR readiness-related indicator over an eight-year period (2009–2017), we make use of case studies that consist of the ten highest scoring and ten lower scoring countries per indicator ($N = 20$), and these are traced for in the subsequent eight years. Data is sourced from the World Bank for FDI (dependent variable) and the WEF for the GCI pillars (independent variables) over the 2009–2017 timeline.

The chapter seeks to infer whether there is a causal relationship between the various countries' scores on the "Technological Readiness" pillar and subsequent growths or declines in new FDI. In particular, the chapter adopts as an independent variable the changes in scores in the WEF's GCI's tracing of Technological Readiness over the 2009–2017 period. This score is represented as a figure ranging from 0 to an upper limit of 7 (best scorer). Importantly, this score may or may not change in each consecutive year. On the other hand, the dependent variable for this chapter is the subsequent growth or decline in FDI influx on an annual basis. Specifically, this is operationalised as an increase or decrease in inbound foreign direct investment stocks in monetary terms (US dollars). In operational terms, both figures are converted into percentages of themselves to make them comparable, with the WEF score's percentile figure multiplied by 10 so as to not distort the figures since it is on an upper limit of 7, whereas growth in FDI is virtually limitless. Correlations in changes between the score and new FDI for each country are calculated as Pearson r scores. We should thus expect figures between 1 and -1 , with figures close to 1 indicating mutual movement, and thus a reduction in technological readiness not being associated with new FDI. On the other hand, figures close to -1 indicate a lack of a relationship. The countries being studied here are Sweden, Netherlands, Switzerland, Denmark, Luxembourg, Singapore, Norway, the United Kingdom, Hong Kong (special administrative region [SAR]), and Finland among the high scorers (top ten), whereas among the low scorers we study Bolivia, Burundi,

Chad, Ethiopia, Lesotho, Mauritania, Nepal, Rwanda,¹ Tanzania, and Zimbabwe. The conditions for case selection that resulted in these twenty countries' inclusion necessitated a comparative analysis of the ten highest-ranking states in the index and ten of the lowest-ranked states as of 2009.

With the exception of Singapore and Hong Kong, all the top scorers are located in Western Europe, whereas eight of the low scorers (with the exception of Bolivia and Nepal) are in Africa. In terms of hypotheses, the chapter will test the following working propositions:

Proposition 1: Increases in the Technological Readiness indicator in a given year directly correlates with an increase in FDI influx.

In order to further note any differences among the two country sets (groups), we test this second proposition:

Proposition 2: Overall, there is a difference in FDI influx for low and high scoring countries in the 4IR readiness indicator per year.

These will provide insights from incidences where, particularly, all countries, regardless of grouping, observe growth in 4IR markers and FDI but the high scorers have a disproportionate growth in new FDI. Finally, interested in the degree to which there is a global “Matthew effect” (i.e., of global financial flows disfavouring low scorers regardless of their improvement in scores), we, therefore, test the following third proposition:

Proposition 3: Regardless of improvement in 4IR readiness, low FDI scorers experience no new FDI influxes.

The forthcoming section lays out the dataset to be utilised in the study. It is followed by an analytical section, which is a case-by-case analysis of the correlation scores, and a subsequent discussion section, which derives a hypotheses-related discussion based on the findings (Tables 7.1 and 7.2).

¹ Rwanda was added given infrequent availability of data on Angola noted by the WEF report authors themselves. The East African country was thus chosen on a stratified basis due to its apparent growth in recent years, as well as the generally assumed role of technology and FDI in this regard.

Table 7.1 High scorers' raw data, 2009–2017 (core/core-periphery)

	2009	2010	2011	2012	2013	2014	2015	2016	2017
Sweden	6.2	6.12	6.9	6.29	6.22	6.19	6.24	6.29	6.3
Netherlands	6	5.99	6.1	5.98	5.97	6	6.1	6.18	6.34
Switzerland	6	5.6	6.7	6.02	5.93	5.97	6.31	6.41	6.39
Denmark	5.9	5.62	6.2	6.17	6.05	6.1	6.11	6.1	6.09
Luxembourg	5.9	6.11	6	6.21	6.19	6.36	6.42	6.4	6.46
Singapore	5.9	5.35	6.3	6.1	6.01	6.09	6.2	6.14	6.09
Norway	5.8	5.56	6.1	5.78	6.08	6.12	6.14	6.17	6.12
United Kingdom	5.8	5.58	6.5	6	6.06	6.28	6.3	6.33	6.33
Hong Kong	5.7	5.96	6.1	6.16	6.03	6.1	6.13	6.21	6.17
Finland	5.6	5.17	5.7	5.92	5.89	5.97	5.98	5.97	5.98

Data extracted from WEF reports, 2009–2017

Table 7.2 Low scorers' raw data, 2009–2017 (periphery)

	2009	2010	2011	2012	2013	2014	2015	2016	2017
Bolivia	2.34	2.65	3.7	2.73	2.77	2.79	2.89	2.96	2.853
Burundi	2.31	2.31	3.1	2.22	2.2	2.1	2.1	2.96	2.11
Chad	2.39	2.28	3.2	2.23	2.09	2.09	2.05	1.93	1.96
Ethiopia	2.29	2.51	2.4	2.5	2.47	2.46	2.46	2.43	2.36
Lesotho	2.64	2.59	2.6	2.5	2.45	2.37	2.67	2.67	2.58
Mauritania	2.81	2.55	2.5	2.75	2.71	2.71	2.68	2.32	2.24
Nepal	2.21	2.5	4.2	2.63	2.55	2.61	2.62	2.56	2.77
Rwanda	3.09	3.09	3.05	3.04	3.1	3.14	3.12	3.25	3.2
Tanzania	2.61	2.59	2.7	2.77	2.7	2.51	2.46	2.59	2.64
Zimbabwe	2.29	2.48	2.7	2.83	2.98	2.95	2.79	2.73	2.74

Data extracted from WEF reports, 2009–2017

7.5 DATA

7.6 ANALYSIS

7.6.1 Core/Peripheral Core Countries

In terms of the Pearson r correlation test, Denmark observed a 0.0539 correlation score, indicating a weak positive correlation between change in the availability of the latest technology and new FDI. Positive correlations are noted in the 2010–2012 period, after which Denmark's 4IR indicators

observed no growth, but still saw new FDI in 2014 and 2016, but also saw a decline in FDI in comparative terms in 2015 and 2017 (Fig. 7.1).

Over the 2009 to 2017 period, Finland observed a correlation score of 0.018, demonstrating a weak positive correlation between changes in the availability of the latest technologies and new FDI. Years of positive correlation are noted in the 2010–2013 period, with a negative correlation in 2014 and a lack of co-variation in 2016–2017; with, respectively, a decline and a growth in FDI but with no changes for the 4IR indicator score (Fig. 7.2).

In the period studied Hong Kong observed a Pearson r correlation score of 0.3966, higher than Denmark and Finland's scores, but was still a weak positive correlation. Years of positive correlation are noted in the entire time series, except for 2016 (Fig. 7.3).

Luxembourg's Pearson r correlation score between change in the availability of the latest technology and new FDI over the 2009–2017 period was 0.4454, higher than Denmark, Finland, and Hong Kong studied so far but was still a weak correlation. Years of positive correlation were noted from 2010 to 2015, with negative correlations in 2016 and 2017 (Fig. 7.4).

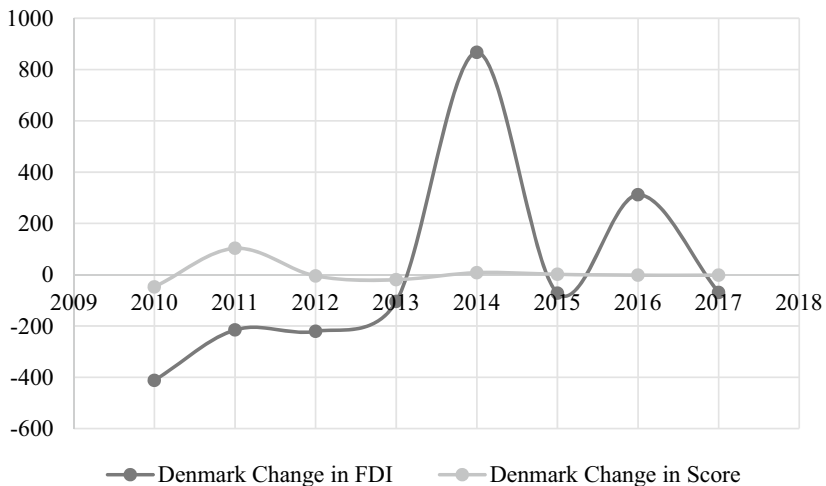


Fig. 7.1 Denmark (Source Authors)

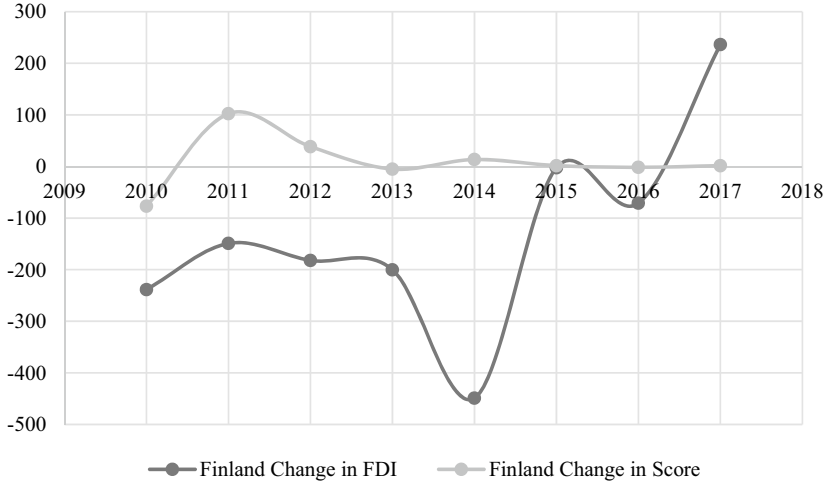


Fig. 7.2 Finland (*Source* Authors)

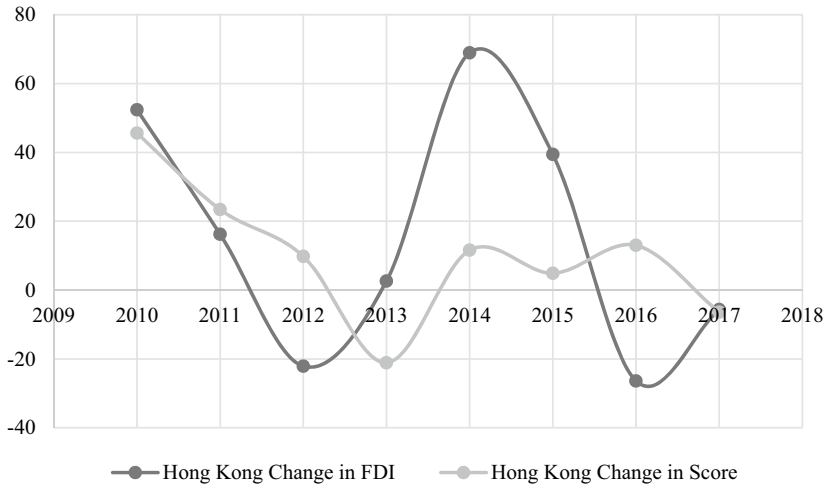


Fig. 7.3 Hong Kong, SAR (*Source* Authors)

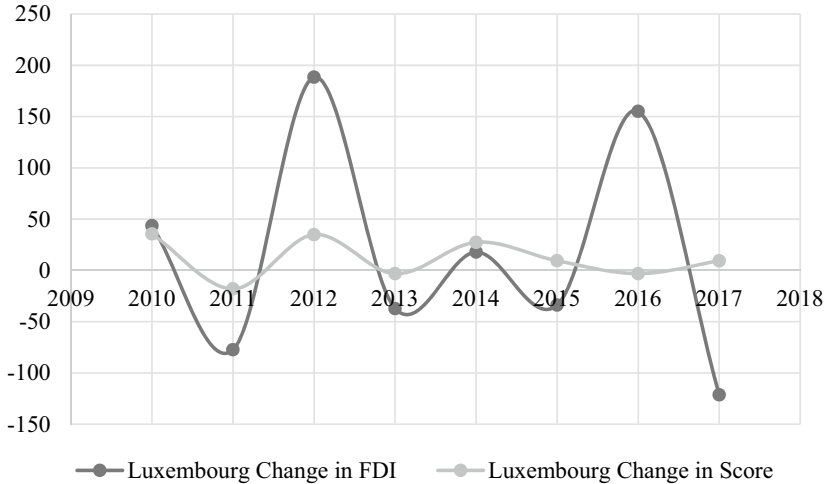


Fig. 7.4 Luxembourg (*Source* Authors)

For the period under study, the Pearson r score is 0.2849 for the Netherlands, indicating a weak positive correlation between the two variables. The positive correlations were observed in 2010–2013, and once again in 2015, whilst inverse correlations were noted in 2014–2017 (Fig. 7.5).

In the 2009–2017 period, Norway saw a weak negative correlation score of -0.3355 , one of only three countries among the high scorers (alongside Singapore and the UK) to observe such a negative correlation. Movements in the independent variable (change in technological readiness) coincided with a negative score in new FDI in 2011 and 2013–2014, whilst declines in its technological readiness score still correlated with new FDI in 2010, 2012, and 2016, and 2017 (Fig. 7.6).

In the period under study, Singapore observed a moderate negative Pearson r correlation score of -0.6677 . Singapore had observed negative growth in technological readiness in 2012 and 2016, during which it still saw new FDI. On the other hand, the country saw growths in its technological readiness score in 2011 and 2013–2014, but still saw comparative declines in new FDI. Thus, co-variation was only observed in 2015, 2016, and 2017 (Fig. 7.7).

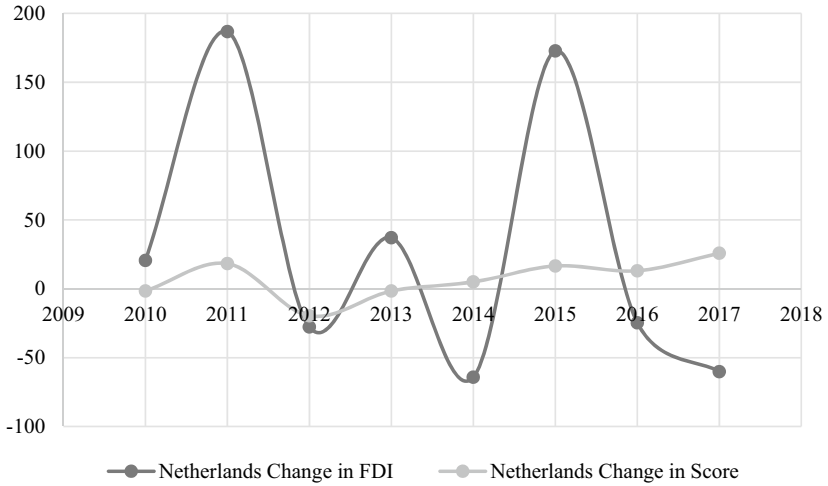


Fig. 7.5 Netherlands (Source Authors)

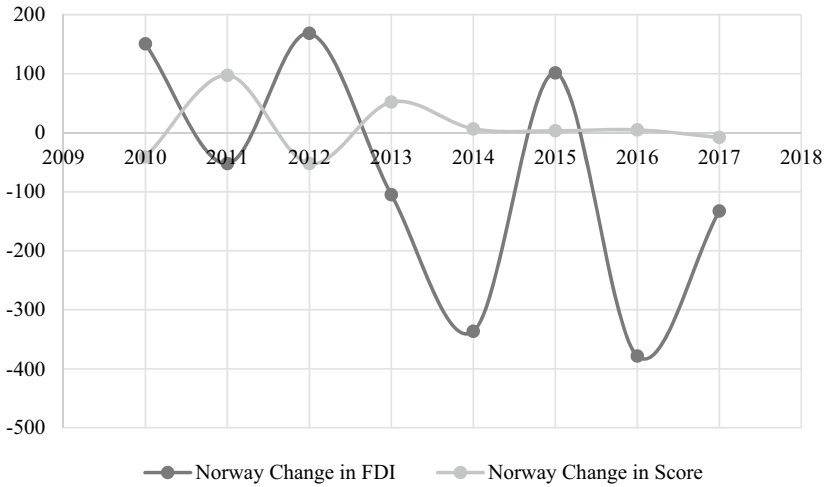


Fig. 7.6 Norway (Source Authors)

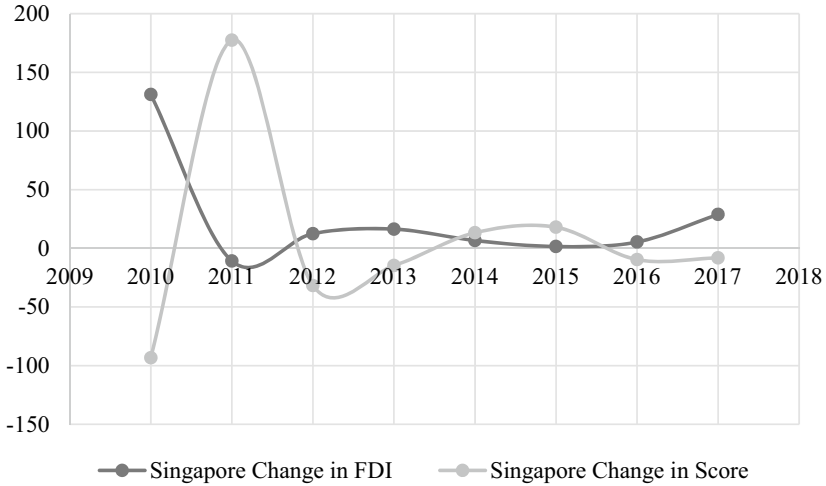


Fig. 7.7 Singapore (*Source* Authors)

In the eight-year period under study, Sweden observed the highest positive correlation between the two variables among the high scorers at 0.6798. The country observed growths in technological readiness and new FDI influxes in the years 2010–2011; it also saw a correlation in decline in its growth of technological readiness in 2012; however, the country subsequently saw no changes in technological readiness between 2013 and 2017, but saw declines in new FDI in 2013–2015, as well as increases in new FDI in 2016–2017 (Fig. 7.8).

In the 2009–2017 period, Switzerland observed a Pearson r correlation score of 0.1926, demonstrating a weak positive correlation. Thus, weak positive correlations were noted for Switzerland, alongside Denmark, Finland, Hong Kong, Luxembourg, and the Netherlands (Fig. 7.9).

In the period under study, the United Kingdom observed a negative Pearson r correlation score of -0.3634 , which indicates a weak negative correlation. The country observed co-variations in 2013 alone. On the other hand, there were inverse correlations between 2010–2012 and between 2014–2017. The country improved its technological readiness score and saw decreased FDI in the following years: 2011 and

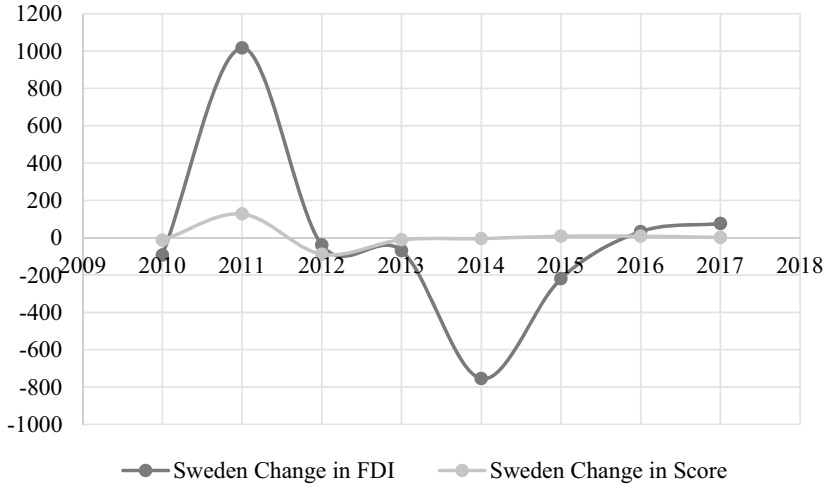


Fig. 7.8 Sweden (*Source* Authors)

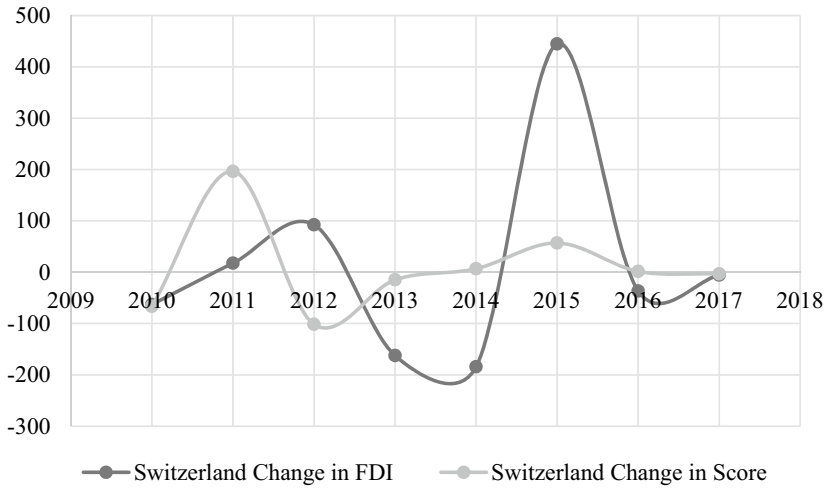


Fig. 7.9 Switzerland (*Source* Authors)

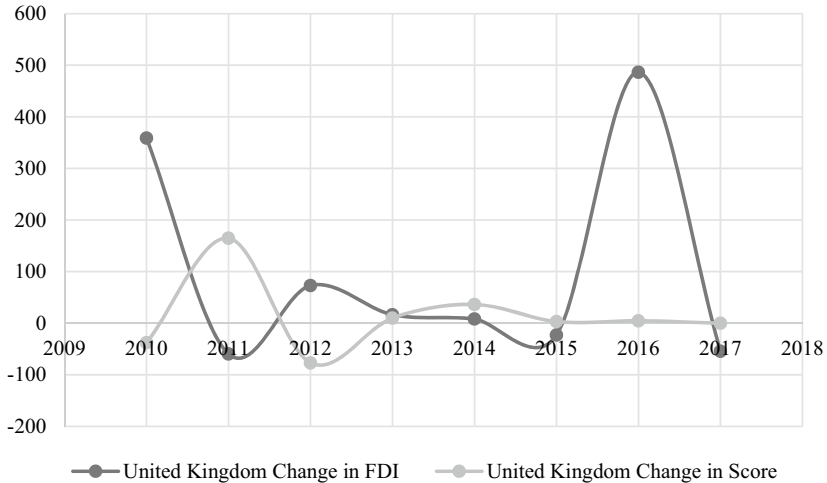


Fig. 7.10 United Kingdom (*Source* Authors)

2014–2015. On the other hand, the country failed to improve its technological readiness score and still saw new FDI in the years 2012, and 2016–2017 (Fig. 7.10).

7.6.2 *Periphery Countries*

For the period under study, Bolivia observed a Pearson r score of 0.0473, indicating a weak positive correlation between the independent variable (technological readiness) and the dependent variable (new FDI). Years of improvement in technological readiness and growth in new FDI were noted in 2014–2015, whereas improvements in technological readiness correlated with a decline in FDI in 2010–2012. Finally, declines in technological readiness are still correlated with new FDI in 2017 (Fig. 7.11).

In the 2009–2017 period, Burundi observed the strongest correlation between the two variables among the low scorers, with a moderate positive correlation score of 0.608. Co-variations between changes in either variable were noted from 2010–2017. Burundi thus saw all instances of improved scores result in new FDI influxes, and likewise saw decreases

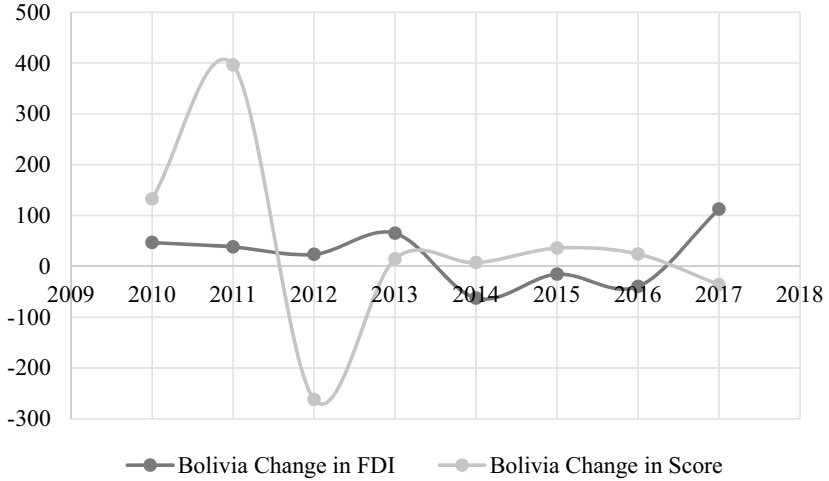


Fig. 7.11 Bolivia (*Source* Authors)

in technological readiness correlate with declines in new FDI (most noticeably in 2012 and 2017) (Fig. 7.12).

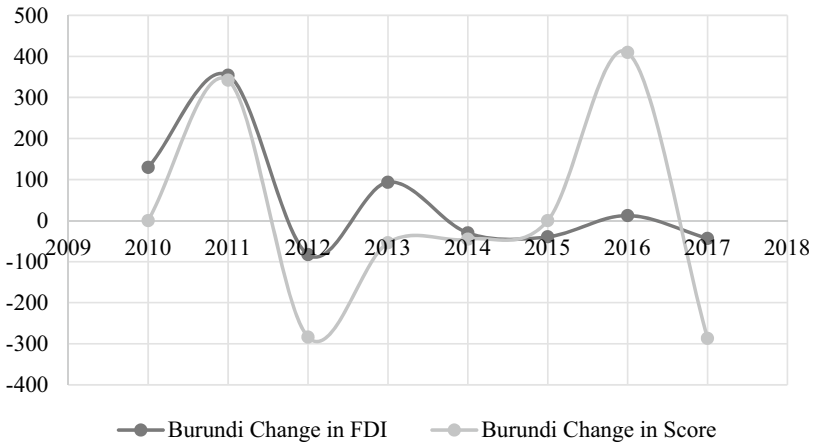


Fig. 7.12 Burundi (*Source* Authors)

In the period under study, Chad noted a weak negative correlation score of -0.2037 . We can observe positive correlations in 2011, and 2016–2017, whereas we can note inverse correlations between the two variables in the years 2012–2015. In 2012 and 2013, we note Chad not improving its score, but still saw growth in new FDI, whereas 2014 and 2015 saw the country improve its technological readiness but observe decreasing new FDI (Fig. 7.13).

In the period studied, Ethiopia observed a Pearson r score of -0.3022 , indicating a weak negative correlation between the two variables. Covariation was noted in 2010, as well as 2015. On the other hand, inverse correlations were seen in 2011–2017 (with the 2015 exception) (Fig. 7.14).

For Lesotho, in the period under study, the Pearson r score is -0.137 , indicating a weak negative correlation. Inverse correlations, with improvements in the score correlating with decreases in FDI, noted in 2011 and 2015, while declines in technological readiness score still saw new FDI in 2012–2014. Improvements in the technological readiness score led to new FDI in 2016 and 2017 (Fig. 7.15).

Between 2009 and 2017, Mauritania saw a Pearson r score of 0.3983 , indicating a weak positive correlation between new FDI and

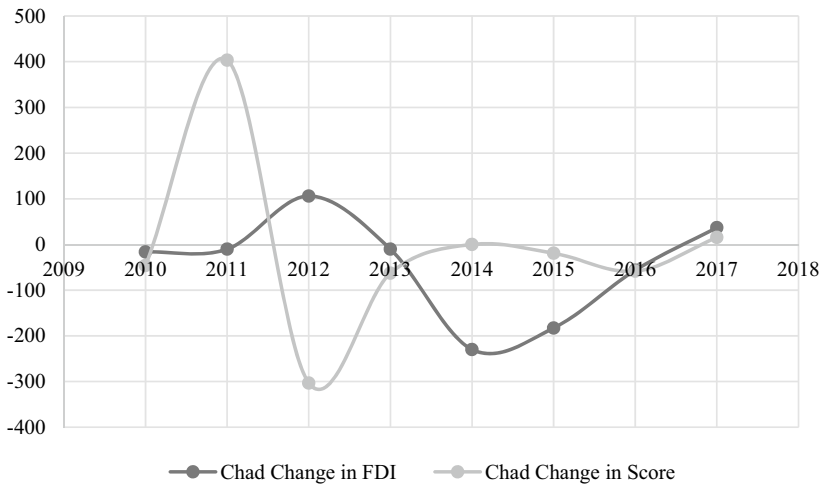


Fig. 7.13 Chad (Source Authors)

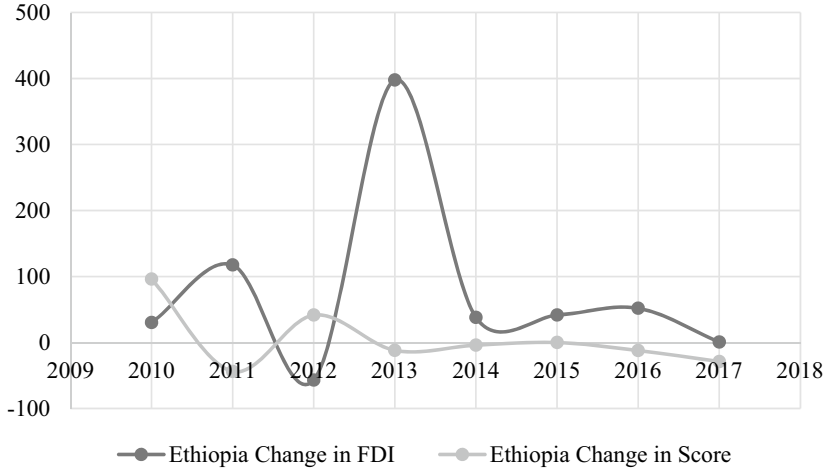


Fig. 7.14 Ethiopia (*Source* Authors)

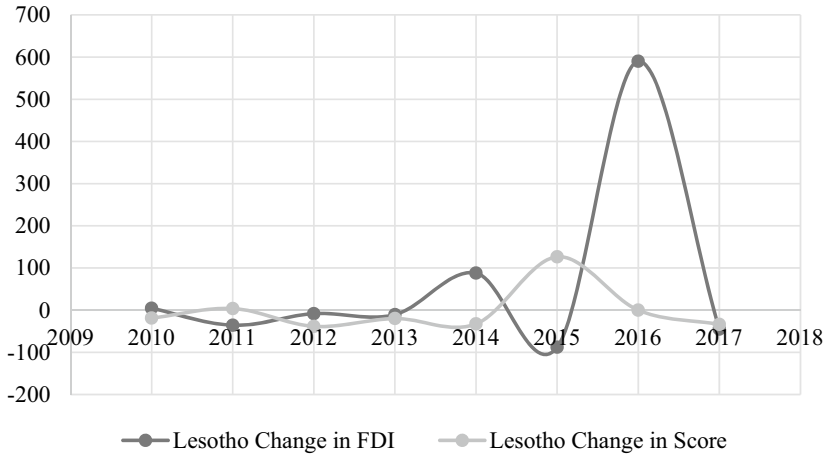


Fig. 7.15 Lesotho (*Source* Authors)

growths/declines in the country's technological readiness score. Co-variations are noted in 2011, 2013, and 2016 with mutual increments and declines. In 2012, the country's increment in its technological readiness yielded FDI, though not concomitant with the magnitude of the increment (i.e., though the year still saw an FDI influx, this was a comparative decline since its technological readiness score had been higher than the preceding year) (Fig. 7.16).

In the 2009–2017 period, Nepal observed a weak positive correlation score of 0.3666 between the two variables. Co-variations are noted in 2010–2013, and 2016–2017, whereas inverse correlations are noted in 2014–2015, with, respectively, an improvement in technological readiness correlating with a decline in FDI and decline in technological readiness still correlating with new FDI influxes (Fig. 7.17).

In the period under review, Rwanda's Pearson r score is -0.0365 , indicating a weak negative correlation between the two variables. The country saw co-variations in 2011, 2013–2016. On the other hand, inverse correlations are noted in 2012 and 2017 (Fig. 7.18).

For the period under study, Tanzania observed a Pearson r score of -0.3284 , indicating a weak negative correlation. The country's improvements in technological readiness correlated with new FDI influxes in 2012, whereas declines in technological readiness resulted in FDI declines

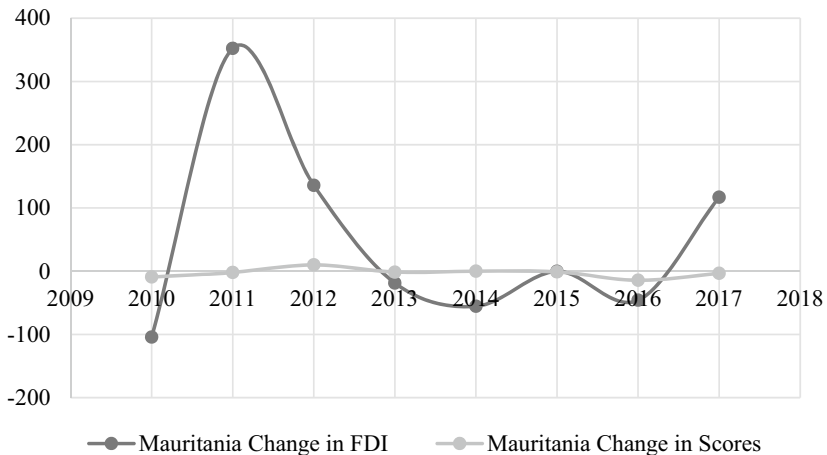


Fig. 7.16 Mauritania (Source Authors)

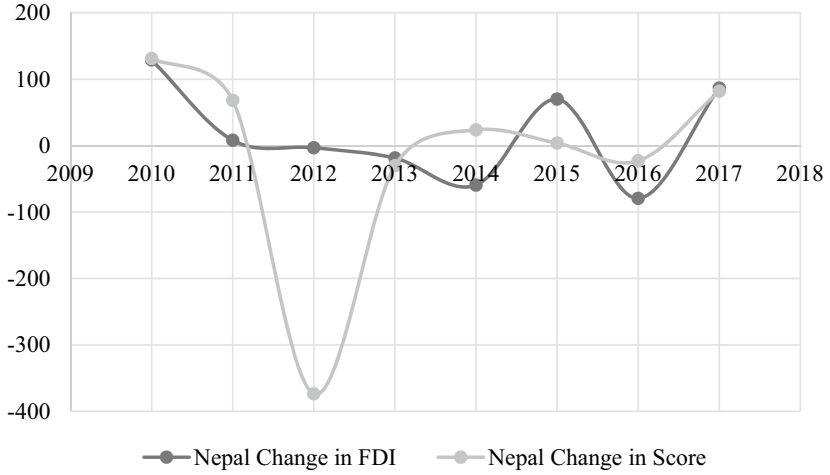


Fig. 7.17 Nepal (*Source* Authors)

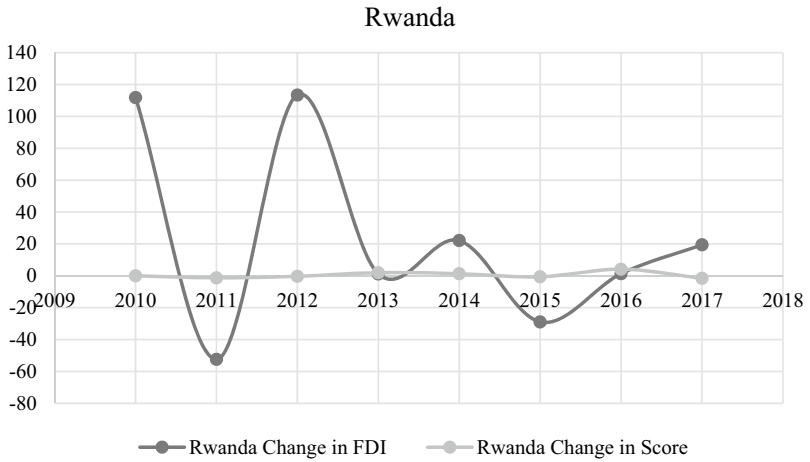


Fig. 7.18 Rwanda (*Source* Authors)

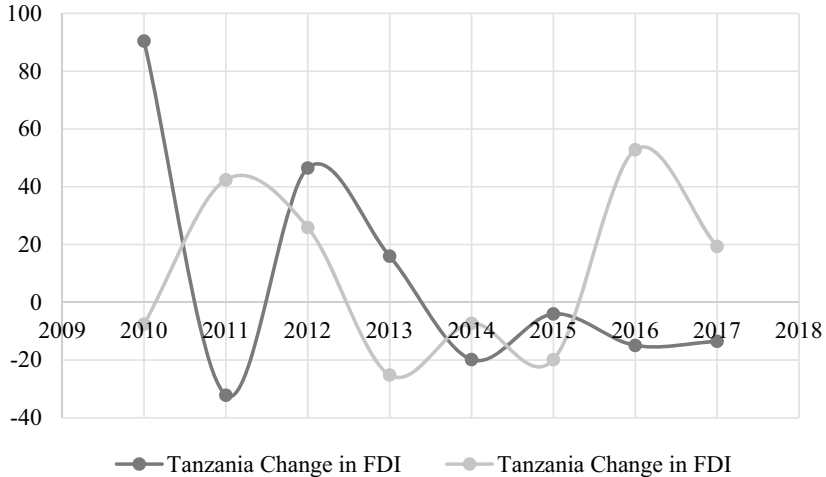


Fig. 7.19 Tanzania (*Source* Authors)

in 2013 and 2015. Inverse correlations are therefore noted in the years 2010–2011, and 2014, and 2016–2017 (Fig. 7.19).

In the 2009–2017 period, Zimbabwe observed a positive correlation score of 0.6061. Co-variations are noted in 2010–2016, while an inverse correlation is noted in 2017. Improvements in the technological readiness score correlated with new FDI in the following years: 2010–2011 and 2013. Declines in the technological readiness score correlated with declines in FDI influges in the years 2012 and 2014–2015 (Fig. 7.20).

7.7 DISCUSSION

Proposition 1 stated that increases in the Technological Readiness indicator in a given year directly correlate with an increase in FDI influx. In this regard, we found that among the top tier scores, six countries saw a correlation between improvements in the 4IR readiness proxy and new FDI; on the other hand, five countries among the lower tier scorers saw a positive correlation between the 4IR readiness proxy and new FDI (see Table 7.3). Weak positive correlations between improvements in technological readiness and new FDI were noted for Denmark, Finland, Hong Kong, Luxembourg, the Netherlands, and Switzerland among the high

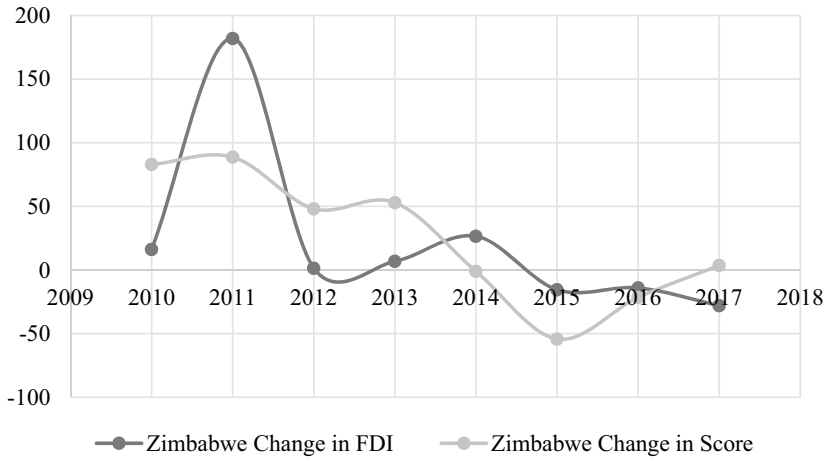


Fig. 7.20 Zimbabwe (*Source* Authors)

scorers. On the other hand, Burundi, Zimbabwe, Mauritania, Nepal, and Bolivia saw improvements in the 4IR proxy correlate with increased new FDI.

Proposition 2 hypothesised that overall, there would be a difference in FDI influxes for low and high-scoring countries in 4IR readiness indicators per year. Dissecting the table in two demonstrates that overall, the countries' results present a mixed picture; 4 of the countries in the top correlation score ranking of countries whose inward FDI influx was linked to their improved technological readiness scores in the WEF index, were among the countries in the group of low scorers, whereas three of the countries which had their inward FDI being the least linked to their score on the 4IR readiness indicator were Norway, the United Kingdom, and Singapore, respectively.

In terms of Proposition 3, we hypothesised that regardless of improvement in 4IR readiness, low WEF index scorers experience no new FDI influxes. This is to test whether global financiers were indifferent to improvements in the initial low scorers, whilst potentially investing in the high scorers regardless of their performance. Below in Fig. 7.21, we present a radar chart with 2 groupings (i.e., aggregated outcomes for the high and low scorers) and 7 possible outcomes per group (WEF score growth and increased FDI; WEF score decline and decreased FDI; WEF

Table 7.3 Ranking of correlations scores

<i>Rank</i>	<i>Country</i>	<i>Pearson r score</i>	<i>P value (<0.05)</i>	<i>Country classification</i>
1	Sweden	0.6798	0.63626	High scorer
2	Burundi	0.608	0.109788	Low scorer
3	Zimbabwe	0.6061	0.111209	Low scorer
4	Luxembourg	0.4454	0.26875	High scorer
5	Mauritania	0.3983	0.328413	Low scorer
6	Hong Kong	0.3966	0.330673	High scorer
7	Nepal	0.3666	0.371729	Low scorer
8	Netherlands	0.2849	0.494015	High scorer
9	Switzerland	0.1926	0.647706	High scorer
10	Denmark	0.0539	0.899133	High scorer
11	Bolivia	0.0473	0.911445	Low scorer
12	Finland	0.018	0.966257	High scorer
13	Rwanda	-0.0365	0.943784	Low scorer
14	Lesotho	-0.137	0.746321	Low scorer
15	Chad	-0.2037	0.629703	Low scorer
16	Ethiopia	-0.3022	0.467237	Low scorer
17	Tanzania	-0.3284	0.427686	Low scorer
18	Norway	-0.3355	0.417287	High scorer
19	United Kingdom	-0.3634	0.381053	High scorer
20	Singapore	-0.6677	0.070795	High scorer

score growth and decreased FDI; WEF score decline and increased FDI; stagnant WEF score and new FDI; stagnant WEF score and decreased FDI; stagnant FDI and no new FDI).²

From the generalised incidence summary chart above, we note that both groupings actually had an equal share in terms of incidences of decreasing technological readiness and receiving less FDI year-on-year. This occurred a total of 21 times for both groups out of the possible 77 (see incident type 2). Given this coalescence and apparent similarity in sensitivity of incoming FDI to technological readiness decline, this may perhaps indicate that the WEF's typology of "innovation" and "factor-based economies" is not rigid. In this way, these "factor-based

² None of the country types saw zero movement in technological scores coincide with no movement in FDI. Thus it appears that global financial flows into all the countries, regardless of type to which they fall into and regardless of the lack of change in score, are likely to persist. This indicates a potential spuriousness and necessitates further isolation-based studies.

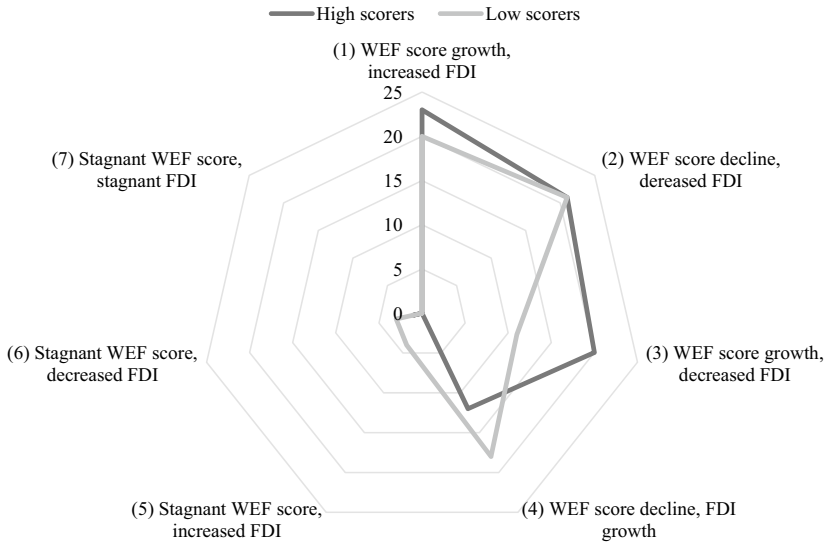


Fig. 7.21 Generalised incidence summary, 2009–2017 (Data sourced from WEF. Chart and calculations by authors)

economies” can attract new FDI in both factor *and* innovation related terms.

In terms of the third proposition is to be further noted that the decline in their technological readiness score for the low scorers on the index still saw much less FDI being “withheld” compared to their higher scoring counterparts (11 incidences, compared to 20 for their high scoring counterparts; see incident type 4 in Fig. 7.21). Indeed, on incident type 5 of the radar chart, we note that lack of improvement on the WEF score still yielded new FDI to the low-scoring countries on a higher and more consistent basis than their higher scoring counterparts (18 incidents, compared to 12 for their higher scoring counterparts). However, we can still note that a lack of improvement on the technological readiness score among the low scorers also results in a loss of FDI for the low scorers compared to the high scorers. This could be hypothetically explained by the fact that the higher scoring countries have a smaller gap to close than their lower scoring counterparts since they have the technology in higher comparative terms. Moreover, this only occurred in 3 incidences among

the low scorers, and in 1 incident among the high scorers, leading to a difference of only 2.

Overall, based on the generalised findings among the case studies (consisting of 20 countries), it appears as though there is no permanent “global Matthew effect,” in which some states have an exclusive inflow of FDI regardless of their performance. Indeed it would appear as though the countries presently ranked as low scoring in this chapter are nearly just as likely to attract new FDI (see incident type 1) if they attain a consistent improvement in their technological readiness score. Indeed, based on incident type 3, these states have a higher chance of attracting more FDI, even when their technological readiness indicator score declines. The issue, however, is the degree to which these countries can start and keep up increments in their technological readiness scores; herein lies the insight of the WEF’s typology. For if these states do not rapidly close the gap, possibly within the next decade, their low placement may become harder to reverse and the 4IR era may be rendered further unequal in similar fashion as the first three industrial eras.

7.8 CONCLUSION

The present chapter determined whether measures of 4IR readiness are associated with new FDI inflows. Secondly, the findings indicated the degree to which laggards in 2009 have since improved in terms of their 4IR readiness. Finally, the chapter determined the degree to which the 4IR era is set to be characterised by the same structure and players which defined the third industrial era and how that may be due not to a global Matthew effect, but instead the internal channelling of the FDI, of which the lower scoring countries have a disproportionate likelihood of receiving regardless of improvements on technological readiness, towards better technological performance in the long-term. Further, within the lower scoring countries, there also appear to be outliers; countries such as Ethiopia (whose FDI grew by 620% in the eight-year period) and Rwanda (whose FDI grew by 188% over the period under study), whose growth in FDI outweighs the other lower scorers (such as Chad, whose FDI declined by some 45.34% in the same period). Thus, among the present low scorers in the technological readiness indicators, some have better positioned themselves to take advantage of global FDI for sustained periods and, in turn, to be positioned to take advantage of the fourth industrial revolution. However, questions still linger about their ability to

fully do so, and the timeframes within which they may achieve this. In this regard, further econometric studies could shed light on how much, if at all, these countries' inward FDIs are channelled towards technological catch-up, and how in turn that contributes to their incremental gain in the long-term with respect to technological readiness. Dependency theory may yet account for the division of labour and patterns of dependency in the wake of AI's ubiquity in industry and productive sectors. All other things being equal, however, we anticipate the number of periphery-classified countries to expand and the core to shrink as "principal–client" relations modulate themselves into the 4IR.

BIBLIOGRAPHY

- Cardoso, F. H. 1979. *Development under Fire*. Mexico D.F.: Instituto Latinoamericano de Estudios Transnacionales, DEE/D/24 i, Mayo (Mexico 20 D.F., Apartado 85–025).
- Conway, D., and N. Heynen. 2006. *Globalization's Contradictions: Geographies of Discipline, Destruction and Transformation*. New York: Routledge.
- Conway, D., and N. Heynen. 2014. Dependency Theories: From ECLA to Andre Gunder Frank and Beyond. In *The Companion to Development Studies*, eds. Vandana Desai, and Rob Potter, 111–115. New York: Routledge.
- Groizard, J. L. 2009. Technology Trade. *The Journal of Development Studies* 45 (9): 1526–1544.
- Inikori, Joseph E. 2002. *Africans and the Industrial Revolution in England: A Study in International Trade and Economic Development*. Cambridge: Cambridge University Press.
- Ndzendze, B., and T. Marwala. 2021. *Artificial Intelligence and Emerging Technologies in International Relations*. Singapore: World Scientific Press.
- Piketty, T. 2014. *Capital in the Twenty-First Century*. Cambridge, MA: The Belknap of Harvard University Press.
- Schwab, Klaus. 2018. *The Global Competitiveness Report 2017–2018*. Geneva: World Economic Forum.
- Shafaeddin, Mehdi. 1998. How Did Developed Countries Industrialize? The History of Trade and Industrial Policy: The Cases of Great Britain and the USA. UNCTAD Discussion Paper No. 139.
- Tausch, Arno and Heshmati, Almas. 2009. *Re-Orient? MNC Penetration and Contemporary Shifts in the Global Political Economy*. IZA DP No. 4393. Bonn: Institute for the Study of Labor.
- WEF. 2009. The Global Competitiveness Report 2009–2010. *World Economic Forum*. Available at: https://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2009-10.pdf.

Yokota, Kazuhiko, and Akinori Tomohara. 2010. Modeling FDI-Induced Technology Spillovers. *The International Trade Journal* 24 (1): 5–34.



The English School and Artificial Intelligence

8.1 INTRODUCTION

The English School of IR is the first concerted attempt at theorising international relations processes outside the American mainstream, as represented by structural realism and liberalism. The School claims to present an account of international relations that is a combination of theory and history, as well as morality and power (Dunne 2011). Combining insights from both realism and liberalism, the school provides a basis for studying and understanding contemporary international relations and world history in terms of the “social structures of international orders” (Stivachtis 2017: 29). That is, unlike other theoretical systems, the School presents an attempt at a holistic approach to IR (Stivachtis 2017: 29). The theory makes distinctions between three key concepts. These are the “international system,” “international society” and “world society” (Stivachtis 2017: 29). Accordingly, this makes the school take a middle position between the two main theories (Stivachtis 2017: 29).

The theory has considerable conceptual power and can provide clarity on several recent developments in the intersection of AI and international relations. In its distinction between international society and the international system, it provides a unique and nuanced approach to the understanding of anarchy. Specifically, it allows us to conceptualise of anarchy in an exclusionary context. That is, there is no universal anarchy—the phenomenon is circumscribed among groups of states that

are allied and otherwise share a semblance of intertwined interests and even identity. For our purposes, the school helps us understand the pre-existing pathways that shape the lenses through which states approach and perceive AI, an unsettling technology in many ways. As such, recent policies by the United States have sought to extract exclusionary cooperation among its allies, while the PRC and Russia have sought to establish their own AI realms, at times in collaboration with one another. In the middle are hundreds of states reminiscent of the non-aligned during the Cold War. Competition has also intensified to influence the International Telecommunications Union (ITU) as the “standards keeper” of the world’s digital realm. Thus, whereas realism might only see states pursuing their self-interest and operating as the main actors in the international system (to the exclusion of coalitions and supranational entities on one end or corporations and INGOs on the other), and where liberalism might mainly see commercial interests bonding states together, the school can explain both phenomena, while also identifying gaps in between. On the one hand, it can account for the existence of “communities of states” within the system and why they take the form they do. For our world of AI, this means the English school should be able to offer explanations (or at the very least hypotheses) for the manner in which “AI communities” (as we may term them) are emerging. Regime-type explanations do not suffice; as seen in Chapter 5, the onset of geopolitics over AI has taken on a domestic turn, resulting in marginally fewer liberties even in democratic states through bans of certain apps and devices due to their Chinese origin. In the early age of AI (and perhaps as a result of its newness), states have become beholden to their fears over their higher ideals. The consequence, we argue, is a fractured world along “identitarian” fault lines among states.

Section 8.2 is a breakdown of the English School in terms of its origins, the assumptions it makes, and the contributions it has made by way of its arguments. Section 8.3 examines the relevance of the concept of international society in the advent of AI and its growing ubiquity by looking at the role of technology in the forming and maintenance of these. Section 8.4 examines the prospect of a rules-based AI world system and the contributions to be drawn from the English School, and the limits inherent within it.

8.2 THE ENGLISH SCHOOL: ASSUMPTIONS, ARGUMENTS, AND CONTRIBUTIONS

As a specialisation, the history of the English School began with the work of the “British Committee on the Theory of International Politics” (though even earlier origins are to be traced to the lectures of Charles Manning and Martin Wight). This Committee began in the late 1950s, with financial assistance from the Rockefeller Foundation, focusing on the international community as a central theme in their work. Interdisciplinary in scope, the British Committee convened a series of meetings composed of diverse groups of scholars and diplomatic practitioners (Buzan 2001: 472). In its current iteration, it is generally understood that the English School was ushered in with the 1992 special issue of the journal *Millennium*, titled “Beyond International Society”, and by the ECPR workshop in Limerick in the same year. According to Dunne (2011), the major distinguishing factor of the School is the synthesis it provides with regard to how the world system operates. This is particularly seen in its weaving together of the three pillars: the international system, the inter-state society, and world society.

In *The Anarchical Society*, Hedley Bull (1977: 9) defined an international system as a context in which “two or more states have sufficient contact between them, and have sufficient impact on one another’s decisions, to cause them to behave—at least in some measure—as parts of a whole.” And, he argued that “where states are in regular contact with one another,” these interact in such a way as to inform one another’s strategic thinking through their own behaviour(s) (Bull, 1977: 10). The international system is a power-conditioned form of interaction in an anarchic context. On the other hand, an international society comes about when a group of “like-minded” consciously see themselves and one another as being part of a commonly subscribed-to set of rules in their international relations. This is also underlined by common institutions (Bull, 1977: 13). Later, Bull and Watson (1984: 1) provided this further explication of international society:

We mean a group of states (or, more generally, a group of independent political communities) which not merely form a system, in the sense that the behaviour of each is a necessary factor in the calculations of the others,

but also have established by dialogue and consent common rules and institutions for the conduct of their relations, and recognize their common interest in maintaining these arrangements.

International society is not naturally occurring but has expanded through the centuries. Accordingly, the changing boundaries of inclusion and exclusion are central to the history of the broadening of international society (Dunne 2011). For example, China was not included in decisions and was in fact even denied sovereign statehood following loss in the Opium Wars of the nineteenth century and effectively becoming colonised along its eastern coast through a series of unequal treaties. This state of affairs only changed in January of 1942 when western states renounced their unequal treaties towards China. Questioning why this was so, Dunne (2011) uses the English school's framework and comes to the conclusion that membership is defined by a so-called standard of civilisation that sets up value- and belief-laden conditions that corresponded to the Europe of the time (particularly that of the nineteenth century). Imperial and republican China was deemed to not meet this standard (Kissinger 2011). (Incidentally, during its imperial phases, China's own standards of civilisation had effectively rendered its neighbours in East Asia as tributaries instead of equal sovereign states in their own right [Kerr 2013: 86]). This highlights the important role of cultural difference, as China was deemed to differ too fundamentally from the Europeans' experience of international society. Though necessary, however, mutual recognition (on the basis of cultural identity among others) is not sufficient for the formation of international society within the English school. As such,

The actors must have some minimal common interests, such as trade, freedom of travel, or simply the need for stability...The higher the levels of economic interdependence, the more likely it is that states will develop institutions for realizing common interests and purposes. (Dunne 2011)

Thus, contrary to liberalism, overlapping economic interests do not by themselves suffice to realise true interdependence but must be accompanied by fundamental mutual regard. This is also seen in the English School's conceptualisation of institutions. For the English school, international "institutions" are long-term practices such as diplomacy and even war instead of mere organisations such as the United Nations. The success

of the latter relies on the existence of the former. In keeping with this standard, international organisations are defined as “pseudo-institutions” or “secondary institutions” within the English school.

Finally, Bull posited the normative concept of world society to advance the notion that “the ultimate units of the great society of all mankind are not states ... but individual human beings” (Bull 1977: 21). Thus, the concept of world society transcends the state system and sees individuals, non-state actors, and ultimately the global population as units that should be the focus of international societal identities and arrangements (Stivachtis 2017).

8.3 INTERNATIONAL SOCIETY AND AI

Examining and observing IR through the lenses of the international system also shines a light on the differential levels of access to technology. Understanding this, and what it means for the distribution of material power enables us to be aware of the differences in the “interaction capacity of the units” (Dunne 2011). Taken altogether, “levels of technology can be thought of as attributes of the units; an obvious case in point is whether a state has nuclear weapons technology or not” (Dunne 2011). Accordingly, adherents of the English school argue for the utility of thinking through technology in systematic terms in such dimensions as “communication, transportation, and levels of destructive capacity” (Dunne 2011). One of the scenarios pointed to by the English school is incongruence in technologies, such as between two or more systems in which modes of transportation are mismatched—horse-drawn wagons in one, and supersonic jets in another. Such stark difference characterises our world. As technologies spread out, they can transform the quality as well as the character of what can be described as an “interaction capacity” of the entire system (Buzan et al. 1993: 70; Dunne 2011). But can such incongruences be deliberate? In other words, is the assumption of an inevitable flow of technology an unmerited assumption bucked by trends, historically and in the contemporary world? This has been the case with AI today.

The English school’s deliberate distinction between the international system and international society helps demarcate the nature and pattern of relations among certain states. As such, relations among the European Union member states represent the existence of European international society, while relations between the European Union and the world

outside of it (with exceptions such as the United States, Canada and Switzerland) reflects participation in an international system. For most of history, there was no single international society or even system. Rather, there were numerous regional international societies that were characterised by their own distinct rules and sets of institutions. These were reliant upon civilisational and commonly religious conceptions of the world. These informed governance and legal systems (Stivachtis 2017: 29). Consequently, relations among polities that were situated in varying regional international societies could not be carried out on the basis of similar legal and moral basis as could be expected among those of the same system, which encompasses organic cultural conditions (Stivachtis 2017: 29). Historically, there has never been a single body or rules and institutions universally agreed upon among these regional international societies (Stivachtis 2017: 29). The end of WWII brought about the possibility of realising this. It was a moment that saw the consolidation and codification of a universal international humanitarian law. The UN Charter and the 1948 Universal Declaration of Human Rights gave explicit prominence to universality of justice, rights, and fundamental freedoms (Dunne 2011: 741). The end of the Cold War is perhaps the closest the world has come to such a truly universal international society. The opportunity seems to have been missed through a number of steps, including growing mismanagement of the US–Russia and EU–Russia relationships at every turn, US–China incongruence, and growing authoritarianism. Coming into a compartmentalised international system characterised by international societies, AI technology is distributed along predetermined pathways, and so are disputes over it. Within AI, there are the specialised areas of robotics, NLP, and ML. The last of these three is the leading area of growth, with some 60% of funding in AI being in ML in the United States. Briefly, the ML value chain consists of the following steps: (1) Data collection; (2) Data storage; (3) Data preparation; (4) Algorithm training; and (5) Application development. These stages involve (1) obtaining raw data; (2) placing these raw data in data centers; (3) conversion, formatting, and labelling of these raw data; (4) configuring an algorithm that can make predictions based on the data; and (5) converting these algorithmic predictions into commercial products and applications such as software with installations in mobile apps, medical diagnostic tools, military equipment, geospatial observation and self-driving automobiles. Along any of these stages persists the prospect of politicisation and power politics. For example, US concerns over the

CPC infiltrating US information (deemed baseless by China's government) have led to a technological dimension in their trade war (thereby politicising stages 1 and 5 of the value chain). At the heart of these lie two concerns, which will find increasingly sharp relief: security and commercial competition.

An emerging trend is growing governmental scrutiny towards technology companies, many of which are central to the AI landscape and ecosystem. This has, in some instances, led to legal cases being brought against these companies. Beginning in earnest in the mid-2000s, competition authorities, primarily in the EU, have initiated fine-carrying probes into the market tendencies of Amazon, Apple, Facebook, and Google (Teleanu 2020). To what extent this can be done on a global scale is questionable, precisely because of differences in legal systems. Additionally, there has been a growing effort towards global taxation of these companies, dubbed as "Big Tech." This is indicative of a backlash against corporate power and exploitation to be sure, but it is also to be read against the backdrop of states reasserting themselves. However, these entities will continue to gain leverage due to two main factors: they have made themselves indispensable, but they are also able to play off states against one another by moving across different international societies. This shows how much the English school needs to accommodate twenty-first-century realities, and the technology-producing MNC is one of the era's most significant species. How it goes about navigating the kaleidoscope of international societies and what it represents for the realisation of world society is the School's major challenge that should be front of the agenda.

8.4 CONCLUSION

One of the key issues of debate within the English school is on the two poles of pluralism and "solidarism." Pluralism advances the notion of multiple (plural) international *societies* that have a scarcity or absence of shared norms, rules, and institutions. On the other hand, solidarism refers to a type of international society in which there is a high degree of convergence. In the wake of different AI systems, we are witnessing differing approaches to AI ethics. The dichotomy between pluralism and solidarism is essentially one of how to bring about world society (i.e., relations among people). According to the English school, the cause of disagreement relates to tensions between the needs of states and those

of humankind (Stivachtis 2017: 31). This debate assumes solidarity is a possibility. As seen, intangible phenomena such as culture and history have a role to play in the pathways of enmity and cooperation in the era of AI. They foster group-mindedness and othering. In turn, the expansion of institutions has been one often characterised by colonialism and misogyny aided by technology. The English School downplays aspects of this history. Critical theories are cognisant of it. We turn to them in the next chapter.

REFERENCES

- Bull, H. 1977. *The Anarchical Society: A Study of Order in World Politics*. London: Macmillan.
- Bull, Hedley, and Adam Watson. 1984. *The Expansion of International Society*. Oxford: Clarendon Press.
- Buzan, Barry. 2001. The English School: An Underexploited Resource in IR. *Review of International Studies* 27 (1): 471–488.
- Buzan, Barry, C.A. Jones, and R. Little. 1993. *The Logic of Anarchy: Neorealism to Structural Realism*. New York: Columbia University Press.
- Dunne, Tim. 2011. The English School. In *The Oxford Handbook of Political Science*, ed. Robert E. Goodin. Oxford: Oxford University Press.
- Kerr, Gordon. 2013. *A Short History of China: From Ancient Dynasties to Economic Powerhouse*. Harpenden: Oldcastle Books.
- Kissinger, A. 2011. *On China*. New York: Penguin.
- Stivachtis, Y. 2017. *Europe after Enlargement*. London: Routledge.
- Teleanu, Sorina. 2020. *Top Digital Policy Developments in 2019: A Year in Review*. DiploFoundation. <https://www.diplomacy.edu/blog/top-digital-policy-developments-2019-year-review#9>.



Critical IR Theories and Artificial Intelligence: Constructivism, Postcolonialism, Feminism, and Green Theory

9.1 INTRODUCTION

Among some of the most impactful emerging paradigms in IR are constructivism, postcolonial theory, feminism, and green theory. Whereas realism and liberalism have historical evolution on their side, these theories (termed critical theories) are advantaged by their ability to speak to the experiences and unique conditions which are the reality of states and peoples on the margins of power and influence in the IR world (indeed which are deliberately overlooked by the exclusive focus on states by realism and on international organisations, powerful corporations, and individuals by liberalism). And while they have their own differences, they are joined by their implicit and explicit argument that there are patterns of behaviour that should, upon accurate diagnosis, be *changed*.

Constructivism takes a critical stance on language and its influence on IR processes by questioning concepts that are taken for granted, including anarchy and sovereignty. In our analysis, constructivism showcases how technologies which heighten information acquisition and communication create informational fogs which can result in miscommunication and even conflict. However, at the core, these are driven by human ideas and intentions. The postcolonial theory highlights the skewed relationship between former colonial metropolises and former colonies even after the nominal

end of colonial administration. In the era of AI, these technologies stand to perpetuate these skewed power relations through data and commercial exploitation. New forms of violence stand to be inflicted. At the same time, postcolonial perspectives bring to the fore the coloniality and Eurocentrism of dystopian views of AI.

Feminism highlights the patriarchal values which underpin IR and the disproportionate effects on IR processes. Meta-analytically, it showcases the role of mainstream (or malestream) IR theories in perpetuating this through its ignorance of gender as a unit of analysis in IR despite its visible role. In the era of AI, feminism further contributes to approaching a critical stance on the lifecycles of AI technologies from design and creation, to deployment. Green theory, on the other hand, provides lenses for understanding the long term and wider environmental consequences of industrial productivity, growth, and change. AI is at the forefront of the 4IR and is characterised by considerable energy demands. Given the diversity of AI technologies, a green theoretical perspective must help differentiate among various forms of AI to identify those which can have utility in combatting environmental change, and those which exacerbate it.

Section 9.2 of this chapter unpacks constructivist theory, in terms of its origination and advancement by scholars in North America and Europe and their respective traditions. The section also presents the contributions the theory has made and the contemporary relevance it carries and continues to have. The section moves on to examine the relevance of constructivism for AI. Similarly, we conceptualise postcolonial theory and explore its contributions in the era of AI in Sect. 9.3. In Sect. 9.4 we turn to feminism, and to green theory in Sect. 9.5. We conclude in Sect. 9.6.

9.2 CONSTRUCTIVISM

Constructivism, which encompasses a variety of approaches and methods has, since the 1980s and 90s (at least explicitly in IR), had a major impact on the direction of the study of IR. It has provided original and insightful perspectives for understanding social as well as international reality through a focus on the fundamental role played by ideas, identity, and norms in formulating the preferences of states, and the making of world politics (Jung 2019: 1). In its IR variant, constructivism is focused on the very nature of social science and, by extension, the discipline of IR (Adler 2005: 88). It came to the fore during the “Third Great

Debate” in IR, taking place in the 1980s between the rationalists and early critical international theorists. Robert Keohane noted its emergence and its validity as a new approach in his 1988 address at the ISA Annual Conference. He termed it “reflectivist” at the time (Keohane 1988). The theory would be given its current name by Nicholas Greenwood Onuf the following year.

Adler (1997: 332) articulates that constructivism is the notion that the material world is conditioned by human action and is dependent upon epistemic interpretations of it, which are dynamic and normative. Later, Guzzini (2000) put forth that constructivism is a “metatheoretical commitment” that takes a critical stance on knowledge and social reality. Accordingly, he argues that “as an epistemological claim, knowledge is socially constructed; as an ontological claim, social reality is constructed; finally, as a reflexive claim, knowledge and reality are mutually constitutive” (Jung 2019: 1). Onuf asserted that states, like individuals, “live” in a “world of our making.” In the book *World of Our Making* (1989), Onuf argued that there were phenomena, termed “social facts,” which are made by human action. This is juxtaposed with the concept of independent realities, which he termed “brute facts,” that have no reliance on human actions and minds for their ontological existence. Famously, concepts such as anarchy, the axiom, and mainstay of realist thinking and theorising, were put under the microscope by Alexander Wendt and his colleagues. Similarly, liberal claims have been put under scrutiny, including globalisation, human rights, and “collective security” (Behrvesh 2011).

Like all the theories covered in this book, constructivism has been placed under scrutiny. At the root of this were criticisms of its testability. In particular, critics argued that the theory gave us little to nothing in the way of substantive knowledge or testable hypotheses with regard to states or the international system they make up (Behrvesh 2011: 1). This has been substantially responded to by newer scholarship, including through the use of positivist research. In the mid-1990s, in the wake of the end of the Cold War, the global scholarly environment was conducive to the rise of the constructivist school in light of the diminishing status of realism and other rationalist explanations of the international system (Behrvesh 2011).

The constructivist school has two broad varieties at a high level, one North American and another European. These mainly differ in the questions they pose about IR and the process of foreign policy-making and

in the methods they utilise. North American scholars tended to emphasise the importance of “social norms” and “identities” in the construction of international politics and the production of foreign policy outcomes. These scholars make more use of positivist methods. On the other hand, European constructivist scholars focus their research on the role of “language,” and “social discourses” in constructing social reality, including internationally. This variant is dominated by “post-positivist” or “interpretivist” methods. This means less emphasis on the effects of identity, and more on interrogating their origins (Behravesch 2011).

We briefly explore the intersection of language, technology and international relations outcomes using the unification of Germany as a brief case study. The unification was achieved following the Franco-Prussian War of 1870 and 1871. The war was a decisive Prussian triumph and, embarrassingly, the new German Emperor was crowned in the Hall of Mirrors in Versailles, and not in Berlin. The war a unifying moment for the greater Prussian state (whose king and descendants would be the kings of the emerging Empire of Germany) and the minor German-speaking kingdoms and territories in the south. Following the victory, many of these smaller states accepted the inevitable prospect of Prussian annexation, while others took more persuasion and force. Few, particularly the Empire of Austria (with its multilingual empire to the east of Europe), were excluded. As a result, modern Germany seems like what is known as “Kleindeutschland”, or little Germany, as it failed—to unite all German-speaking communities in central and western Europe in a “Grossdeutschland” or greater Germany.

The war with France was the result of a clash over language and was set off by the now infamous Ems Telegram. This refers to the telegram sent by the King of Prussia to his Chief Minister Otto von Bismarck, minuting the outcome of his meeting with the ambassador of the Emperor of the Second French Empire, Napoleon III. It contained a poorly translated phrase that the latter would perceive as a slight and therefore cause for rapid French mobilisation against Prussia. Why was the telegram written, and what did it contain? In short, it stemmed from France’s protest against the offer of the Spanish crown to a member of the Prussian royal family. Napoleon III demanded that the Prussian candidate decline the throne since this would upset the balance of power in Europe. When the King of Prussia successfully convinced the candidate to withdraw his name, the French emperor was not satisfied. He insisted that the Prussian king should renounce for his family all future candidacies for the throne

of Spain. Summarising his response, that he could make no such promises to his French counterpart, he sent the minutes through the recently introduced technology of electronic telegraphy called the telegram. He added that he informed the French representative that his communications would now be directed to the French ambassador's "adjutant." The telegram was afterwards leaked to the press by Bismarck, who tactically deleted several passages to portray the event as insulting to the French.

This last part of the telegram—of direction of all communications through the adjutant—set France on a warpath in two days of the news reaching Paris. What was so controversial about the term? In French, "adjutant" can be interpreted as a junior non-commissioned officer instead of a high-ranking secretary as meant in German. It was as if the French ambassador, and thus the French monarch and nation, had been deliberately mistreated and not accorded proper protocol. The consequent war has been analysed from a variety of perspectives. For some historians, the then-new technology increased the pace of human communication such that random confusion was bound to occur. For others, it removed the human touch in diplomacy and made statesmen prisoners of affairs. And still for some, it was the machinations of the gifted but self-admittedly deceitful statesmanship of Bismarck, who saw war as a chance to humiliate the French, get territorial concessions and unify the smaller states under Prussia [see his memoirs, Bismarck, 1898], and the credulity and recklessness of Napoleon III, ready to live up to the military adventures of his more famous uncle, Napoleon I, that put the two countries on a war course. Whatever the case may be, this incident shows that language is political and that emerging technologies (as the telegraph was in the 1870s) play a catalytic role in how states communicate and miscommunicate. Governing over a newly united and influential German Empire, Bismarck hosted the notorious Berlin Conference in 1884 and 1885, which, fatefully, carved up the African continent without any regard to social and linguistic dynamics on the ground.

9.3 POSTCOLONIAL THEORY

The onset of European colonialism in Africa occurred over centuries, starting in the 1600s but was consolidated in the nineteenth century at the Berlin Conference. Following independence, some leaders on the African continent quickly recognised the significance of language from the 1950s to the 1970s. Such recognition has been institutionalised into the

AU, and the OAU before it. The significance of language and culture, was established early on by these nation-builders on the continent. In Tanzania, this was apparent in Julius Nyerere's policy of Swahili as a national language. To support this, Nyerere translated Shakespeare into Swahili. In Senegal, it is evident in Leopold Senghor and his contributions to the philosophy of Negritude and his works were translated into Serer his native language for the first time in recent years. Kenneth Kaunda's African humanism used ethics to create a nation of human rights and dignity, whereas Milton Obote in his 1969 pamphlet *Common Man's Charter*, wanted to entrench a cultural transformation by freeing Uganda of tribalism. These African leaders, and others across the continent, recognised the positive role of language. The extent to which they succeeded in the policies they wanted to enact is debatable, but their idea is a correct one. Contrarily, when Emperor Haile Selassie of Ethiopia ignored the political, cultural, and linguistic independence of Eritrea by making it a province in 1962, he intensified secessionist sentiment in Eritrea and set Ethiopia on a path to civil war that culminated in Eritrean independence in 1993. The annexation of Ethiopia broke the unification provisions set out by the United Nations in 1950 via UN Resolution 390(A). Today, the function of language as a bridge-building resource has been accepted by China through its Confucius Institutes and Germany itself through the Goethe Institutes. Both these countries have many language and cultural institutes set up across the world, including in the African countries.

The African continent's guiding framework is Agenda 2063, which recognises the role of language and culture. Aspiration 5 of the document envisages "An Africa with a strong cultural identity, common heritage, shared values and ethics." Having been envisioned in the early 2010s and published in 2013, the plan runs in a world that has since swiftly changed. One of these areas of significant change is in the technological space, particularly the emergence of the 4IR. This involves changes in the world of work, as well as in the social and cultural domains as robust new intelligent technologies such as artificial intelligence, the internet of things and blockchain become widely used. This needs some changes in social institutions, especially in promoting lifelong learning and computational thinking in our schools. A joint study by the South African Department of Arts and Culture, the Cultural Observatory, and the Nelson Mandela Metropolitan University published in 2018 describes the threats and opportunities for the cultural and creative industries and

the terms for success in the 4IR. Key among the conditions for success are a whole-of-society tactic and multi-sector partnerships.

Elsewhere we have also observed the opportunities presented for African languages by AI. The first of these is in preserving those languages that are threatened to go extinct. We have also observed the publication of apps utilising AI to translate major and less well-known African languages. One such app is the OBTranslate, based in Nigeria, which has a library of 2000 African languages. Another app is Reah, the South African Sign Language (SASL) translation system offered in Android and scheduled for publication for Apple devices.

The onset of the 4IR in real-time presents an opportunity for its localisation into the African context. Language is indispensable in this regard. Conceiving the 4IR technologies and thinking of them in African languages gives a sufficient opportunity for innovation. Put differently, the ability to describe the 4IR in local languages offers a chance to invent them from—and for—local settings. To be certain, the continent has made advances in the convergence of culture and technology. Synthesis and comprehensive tactics are required, however. The AU's Centre for Linguistic and Historical Studies by Oral Tradition (CELHTO) can adopt such a role. It has already taken significant steps to exploit the possibilities presented by technologies for its mandate, including the digitisation of oral histories. The AU declared the theme for 2021 to be "Arts, Culture And Heritage: Levers for Building the Africa We Want." This presents an opportunity to explore and consolidate the role of 4IR technologies in safeguarding existing arts, the creation of new ones, and recovering and repatriating those wrongly taken during the colonial era. On all fronts, blockchain offers promise, given its applications in asset management.

One major issue with language and AI is bias. For the global South and the postcolonial world, this requires advocating for bias-eliminating global standards in AI and fairer language practices within the continent itself. Currently, the two countries are mired in ethnolinguistic conflicts. There is a civil conflict between the Francophone-dominated government and the disaffected Anglophone minority population (roughly 20%) after decades of language-based oppression in Cameroon. In Ethiopia, the ethnic-based federal system has resulted in a war between the federal government and the Tigray-speaking region north of the country. It has also been observed that even among the stabler democracies, governments tend to favour certain ethnicities over others due to their voting relevance. This is perhaps to be expected, as, in most African countries, an

average of 70% of the same language speakers tend to be in the contiguous territory. Such has been the case in Kenya, where voting preferences are strongly correlated with language and ethnicity. This makes the diverse capital's places of intense post-election violence. A coherent language policy for the 4IR can only follow from social cohesion and democracy.

Christof Heyns, a renowned South African professor of law at the University of Pretoria and a giant in the human rights world, brilliantly wrote about Africa's philosophical and ethical contribution to the problem of drone warfare vis-à-vis human rights in 2016. The ethical foundation of many African cultures, Ubuntu, demands that even when we attain a state of maximum efficiency for AI-powered drones, human supervision over life and death decisions should not be surrendered to machines. As he put it, "The current turn to algorithms may in some respects be a step forward but it also presents the risk of a return to the anonymous exercise of power which does not have to meet the requirements of justification and accountability." Working within the pan-African paradigm, Lamola (2021) makes the observation that innovation in AI promises to present "the re-conceptualization of human existentiality along the paradigm of posthumanism." Making use of a critical culturo-historical methodology, his work pushes back against what he terms "the Eurocentric genitive basis of the philosophical anthropology that underpins this technological posthumanism, as well as its dystopian possibilities." Making use of the pan-Africanist ideas of Pixley ka Isaka Seme, the paper presents an alternative paradigm for understanding the place of technology in the future of humanity (Lamola 2021: 2). Lamola's paper thus cautions against the assumption of European fears around AI and locates these dystopian visions as rooted in—and thus their uncritical incorporation in non-Western mindsets as indicative of—a form of neo-colonialism. These are some of the many powerful ways postcolonial theory can contribute to the era of AI.

9.4 FEMINISM

Feminism seeks to bring a gender lens into the study of IR, which is widely seen as missing and in dire need (Ruiz 2005). The feminist theory begins with the observation that power flows along lines of gender. Similarly, significant IR phenomena, including conflict and commerce, are disproportional in their effects. As such, feminism observes that the majority of the actors in the field of IR, including diplomats, heads of

government, and other policymakers and academic professionals, have historically been mostly male, a trend that continues in most countries. Feminism asserts that this is not a coincidence but is an outcome of the nature of society from which these actors emerge and preside. Crucial to this are the patriarchal pathways of most societies in states, which in turn embed such practices and values in the international system. In turn, IR as a field of study operates under this patriarchal social order. As such, debates within IR are mainly restrained by a lack of cognisance of women's roles in world politics due to scholars not being trained or grounded in the experiences and perspectives of women (Ruiz 2005). Youngs (2004) further observes that feminist theorising has identified mainstream IR theory as "one of the discourses that help perpetuate a distorted and partial world view that reflects the disproportionate power of control and influence that men hold, rather than the full social reality of the lives of women, children and men." Feminism is more reflective and seeks to be comprehensive in its description of the world of IR; this enables it to make informed prescriptions about facets that need to be changed, and how such changes may be carried out. This includes the problematisation of the hierarchies, which differently place people and groups as differing in "capacity, control, influence and freedom" (Youngs 2004).

In their deconstruction of IR theory, feminist theorists posit that the theoretical foundations of the field are male-centred, resulting in juxtapositions of male and female. In this dichotomy, the former is rendered the normal (with feminists thus supplanting "mainstream" with "malestream" [Youngs 2004]), and the latter is the abnormal or undesirable "other". Implicitly and explicitly, this excludes women from taking up roles in the public sphere in which IR is necessarily carried out. In an overview of the theory's evolution, Youngs (2004: 76) observes that:

Feminist International Relations has expanded, and built on, the work of feminist political and economic theory to examine the masculinist framing of politics and economics and associated institutions, including notably the state and its key military and governmental components, as well as the discourses through which these institutions operate and are reproduced over time.

Feminism is itself a layered and multifaceted theoretical framework, with many internal debates, including critiques of Western feminism's shortcomings in Southern and other non-Western settings. As an IR school of thought, feminism draws from Western works, which is a significant point of criticism, as it arguably ignores or misdiagnoses issues in non-Western settings.

Enloe's work focuses on central IR issues such as war, militarism, and security. In her work, she highlights "the dependence of these concepts on gender structures" (Enloe 1989). She explores these in her book *Bananas, Beaches and Bases: Making Feminist Sense of International Politics* (1989) and includes a critique of the masculine idea of war-making as virtuous since it is linked to the concept of men as warriors, protectors, conquerors, and exploiters of those that are deemed "feminised," or otherwise objectified and othered, through defeat. In her significant work *Gender in International Relations: Feminist Perspectives on Achieving Global Security*, J. Ann Tickner (1992) poses questions that are seen as having been elemental to the consolidation of feminist thought in IR. These include the following: "Why is the subject matter of my discipline [IR] so distant from women's lived experiences? Why have women been conspicuous only by their absence in the worlds of diplomacy and military and foreign policy-making?" (Tickner 1992; see also Youngs 2004: 79).

How might feminism look in practice? Tickner, writing in 2001 (48–49), when considering security, wrote that:

Whereas conventional security studies has tended to look at causes and consequences of wars from a top-down, or structural, perspective, feminists have generally taken a bottom-up approach, analyzing the impact of war at the microlevel. By so doing, as well as adopting gender as a category of analysis, feminists believe they can tell us something new about the causes of war that is missing from both conventional and critical perspectives. By crossing what many feminists believe to be mutually constitutive levels of analysis, we get a better understanding of the interrelationship between all forms of violence and the extent to which unjust social relations, including gender hierarchies, contribute to insecurity, broadly defined.

One of the major policy contributions of feminist lenses was the passing of the UN Security Council Resolution 1325, which acknowledged the disproportionate extent to which women and girls were affected by conflict in 2000. The resolution, having been developed by a multiplicity of stakeholders across civil society, including feminist actors and scholars,

committed the international community to deepen understanding of how conflict affects women and girls,¹ as well as to put in place corrective measures. One of these measures was increasing the inclusion, representation, and participation of women in UN structures and in individual states: “[the UNSC] Urges Member States to ensure increased representation of women at all decision-making levels in national, regional and international institutions and mechanisms for the prevention, management, and resolution of conflict.” In July 2010, the United Nations General Assembly created UN Women (more formally known as the United Nations Entity for Gender Equality and the Empowerment of Women), initially headed by former Chilean president Michelle Bachelet and then former South African deputy president Phumzile Mlambo-Ngcuka, to carry forward and institutionalise these efforts.

In the course of the theory’s evolution, feminist IR has highlighted three major related phenomena. Firstly, it is observed that both the state and market, both in the philosophies which underpin them and their patterns of practice, are “gendered by masculinist assumptions and structures” (Youngs 2004: 77). Secondly, echoing Tickner, Youngs observes that “the dominant conceptualization of political and economic agency in male-dominated terms ignores both women’s realities and their active contributions to political and economic life.” Finally, Youngs notes the overlooking of gender as a category. This is a significant omission given the overwhelming impact of gender in the social construction of genders and their roles in IR activities, which are portrayed as necessarily different and permanent (Youngs 2004: 77).

Feminist scholarship in IR, therefore, shows that we live in a world that has largely been shaped by one gender, and reflects a specific gender’s outlook (Hooper 2001). AI is notoriously male-dominated, with around 90% of executives in US technology companies being male.² While we do not know the exact figures in China, the churning out of largely male engineers (indeed the country is disproportionately male by some 70 million people), it is reasonable to deduce that the gender imbalance there largely skews male as well. The current age of AI seems to have primarily cohered with feminism’s predictions. Needed, then, are

¹ The resolution notes “the need to consolidate data on the impact of armed conflict on women and girls.”

² This statistic is drawn from a study which examined 6000 companies, representing 180,000 employees and 15,000 founders.

more gendered lenses on AI throughout its technological lifecycle. This includes interrogation of the fundamental causes of the gender bias in the AI-making process; conceptualisation and operationalisation of equitable technology transfers along gender-cognisant lines, gender-inclusive processes, and ethical foundations for AI.

9.5 GREEN THEORY

As we have seen, there is a myriad of critical theoretical traditions. But these have a shortcoming. They exclusively address relations between humans and note patterns among human communities and political entities. This is to the exclusion of analysis of their relations with the nonhuman environment around them. Green theory points out, for example, liberalism's emphasis on individual liberties of choice and consumption as being ignorant of the environmental consequences of such consumption (Dyer 2017: 85). In the 1960s, public recognition of the environmental crisis came to the fore. This was dubbed a "tragedy of the commons." This concept refers to a situation wherein self-interested actors excessively use a limited shared resource such as land or water.

Beginning in earnest in the 1970s, the UN organised the first conference on the environment and pollution, while the 1980s saw the emergence of "green" political parties pushing for environmentally responsible alternatives (Dyer 2017: 84). Responding to a need for a green theory of politics, the 1990s saw IR take recognition of the natural environment as an originator of the research agenda of the discipline, and which required theoretical and policy attention in its own right. This was given special impetus by the growing heaps of evidence that human industrial activity was affecting the global climate, leading to security problems alongside ecological devastation (Dyer 2017: 84). This saw the emergence of the green theory. Some early works in this regard showed the centrality of environmental scarcity to the emergence of conflicts within states and between states. The green theory seeks to move beyond "environmentalism." In its critique of mainstream approaches to environmentalism, it argues that:

Most forms of environmentalism seek to establish theoretical positions and practical solutions through existing structures, or in line with existing critiques of such structures. If less critical in orientation, then these views

are likely to be compatible with a liberal position in IR (viewing international cooperation as being of general benefit to states). (Dyer 2017: 85)

In advancing analysis beyond both environmentalism and political ecology, the theory is more radical in its posture towards existing structures of a political, social, and economic nature (Dyer 2017: 85). Specifically, it is suggested that a defining feature of green theory vis-à-vis mainstream environmentalism is its conceptualisation of “a coherent moral vision,” or a “green theory of value” which stands independently of a theory of practices or political agency that is human-centric. As an example, “green morality” suggests that human development may need to be curbed if so doing will be to the broader benefit of non-human nature (Dyer 2017: 86). Green theory’s message in this regard has become ever more salient as the climate continues to change, to some extent beyond reversal. Dyer (2017: 86) observes that climate change is the central challenge and is a direct consequence of fossil fuels. Green theory, therefore, allows us to understand this phenomenon through a long-term perspective: “From the perspective of green theory, this technical impasse requires a change in human values and behaviour and therefore presents an opportunity for political innovation or even a transformative shift in global politics” (Dyer 2017: 86).

This philosophy has filtered into the realm of policy, most recently in the form of the Paris Climate Agreement of 2016 and the Green New Deal advocated by the Democratic Party in the United States since 2018. Both recognise that industrial activity is the primary reason behind climate change, and should therefore be curtailed and replaced with more sustainable alternatives. The former commits major states to provide economic support to the developing world as they seek such alternatives. Fundamentally, then, a green theoretical perspective of climate change places collective choices at the centre. More recent works, incorporating post-colonial and feminist lenses, have sought to ensure that the theory also appreciates the racial and gendered impacts climate change has and the disproportionate contributions these different groupings make.

Green theory highlights the need to be alert to the environmental and ecological impacts of AI. As a critical theory, green lenses remind us that climate change is an injustice, with real consequences for real communities and other lifeforms on the planet. It brings home the need for

solutions, the pursuit of which requires us to abandon perceived short-termism in mainstream paradigms and generate “an ecocentric theory of value and a more ethical than instrumental attitude to human relations in our common future” (Dyer 2017: 86). AI is energy- and infrastructurally intensive, requiring hyperscale data centres and investments in electric power, which is currently predominantly non-renewable. At the same time, the onset of AI brings promise, especially for a more precise approach to agriculture and environmental restoration. As such, green approaches to AI will need to differentiate among different types of AI and their applications. Indeed, some AI technologies may be catalysts and tools in combatting the uses of other, environmentally destructive AI tools. Moreover, environmental effects should inform the current debates on AI ethics and standards and assume a more prominent place.

9.6 CONCLUSION

Having reviewed the impact and implications of AI for critical theories, we note that these theories perform quite accurately in terms of predicting the manner and pathways through which power functions, including in the era of AI. Importantly, this era is only at its beginning. As such, these theories will need to be cognisant of new power pathways by heeding emerging trends and conducting their analyses in cross-pollination and intersectionality. Indeed such crosspollination would benefit many of the other theoretical paradigms examined in this book. Additionally, greater uptake of quantitative methods alongside qualitative work would bolster parsimony and replicability of the findings of the critical schools. We turn to the possibilities which can be explored in these efforts in the conclusion of the book in the next chapter.

BIBLIOGRAPHY

- Adler, Emmanuel. 1997. Seizing the Middle Ground: Constructivism in World Politics. *European Journal of International Relations* 3 (1): 319–363.
- Adler, Emanuel. 2005. *Communitarian International Relations: The Epistemic Foundations of International Relations*. London and New York: Routledge.
- Behraves, Maysam. 2011. *Constructivism: An Introduction*. E-IR. https://www.e-ir.info/2011/02/03/constructivism-an-introduction/#_edn3.

- Dyer, HC. 2017. Green Theory. In *International Relations Theory*, eds. S. McGlinchey, R. Walters, and C. Scheinflug, 84–90. Bristol: E-International Relations.
- Enloe, Cynthia. 1989. *Bananas, Beaches and Bases: Making Feminist Sense of International Politics*. London: Pandora.
- Guzzini, S. 2000. A Reconstruction of Constructivism in International Relations. *European Journal of International Relations* 6 (1): 147–182.
- Hooper, Charlotte. 2001. *Manly States: Masculinities, International Relations and Gender Politics*. New York: Columbia University Press.
- Jung, Hoyoon. 2019. The Evolution of Social Constructivism in Political Science: Past to Present. *SAGE Open* 9 (1): 1–10.
- Keohane, R. O. 1988. International Institutions: Two Approaches. *International Studies Quarterly* 32 (4): 379–396.
- Lamola, Malesela John. 2021. The Future of Artificial Intelligence, Posthumanism and the Inflection of Pixley Isaka Seme’s African Humanism. *AI & Society*. <https://doi.org/10.1007/s00146-021-01191-3>.
- Ruiz, Tricia. 2005. Feminist Theory and International Relations: The Feminist Challenge to Realism and Liberalism. *Soundings Journal* 88 (1): 1–7.
- Tickner, J. Ann. 1992. *Gender in International Relations: Feminist Perspectives on Achieving Global Security*. New York: Columbia University Press.
- Tickner, J. Ann. 2001. *Gendering World Politics: Issues and Approaches in the Post-Cold War Era*. New York: Columbia University Press.
- Youngs, Gillian. 2004. Feminist International Relations: A Contradiction in Terms? Or: Why Women and Gender are Essential to Understanding the World ‘We’ Live In. *International Affairs* 80 (1): 75–87.



Conclusion

10.1 INTRODUCTION

This conclusion assesses whether the field of IR is headed towards increased debate or, alternatively, typological theory-building with AI as a common factor. We anticipate more cross-theoretical debates as being more likely to be the case given the different sets of axioms each theory brings and thus the analytical methodology and foci they encompass; some on AI's inequitable distribution and others on possible positive-sum game aspects, others on domestic determinants (and thus inside-out processes) over transnational ones (and thus outside-in ones), and still others on the anarchical components as opposed to emergent regulatory frameworks.

Section 10.2 discusses some opportunities stemming from the rise of AI and how the technology may be used to mitigate conflict, while Sect. 10.3 highlights the critical insights made by the book in the different theoretical examinations individually and in relation to each other and lays out areas for further research based on the insights made and some of the shortcomings. The final Sect. 10.4 is an invitation for greater use of AI in IR research as well as teaching and learning.

10.2 AI, INFORMATION ASYMMETRY, AND IR PROCESSES

On the face of it, the rise and implementation of AI in state-making are net neutral; after all, machines can only execute what their human operators instruct and program them to. It is crucial to note in this regard that advances in technologies are both symptomatic and catalytic of progress (see Chapter 7 in this book). Machines are made by their human engineers, and in turn, reconfigure social and political realities and policy options. Conflict is no exception among these. It was advances registered in this sphere during the Second Industrial Revolution that resulted in the Second World War being the most devastating up to that point. But it was also technology that mitigated that conflict. It is true that new innovations in weapons and systems enhanced state capacity for destruction, but through the same processes machine-based cryptanalysis, most prominently the Ultra decoder invented by Alan Turing, were used by the Allies to infiltrate and decode German war communications so that they could execute them.

This centres on a phenomenon that is at the core of existing models of conflict, namely information symmetry. Attempts at closing the gap between what is known and not known about potential belligerents can be done by assessing the potential opponent's war-related industries in areas such as arms manufacturing and nuclear stockpiles where relevant. Short of espionage, however, which can itself not be fully effective in the face of counterespionage, states are unable to fully know the state of readiness enjoyed by the other state for a military clash. Locked in such a predicament, states are sometimes driven by overconfidence in their own capabilities to initiate conflict. With the growing capabilities and usage of AI, this problem can be mitigated and thus reduce the probability of conflict because reducing the information asymmetry can act as a deterrent. Ndzendze (2018) argues that through artificial intelligence wars that come about as a result of bluffing can be expected to be a rarity: since artificial intelligence can allow states to calculate as precisely as possible the military capabilities of another state, would-be initiators of conflict that are in a lesser state of readiness will be incentivised to avoid military confrontation. Moreover, AI can be a useful tool for ensuring attacks only on strategic military targets rather than populated areas and thus avoid civilian casualties. While this is not set to take place, the emergence of this option places a greater ethical responsibility on states to act in line with the Geneva Convention. In a similar vein, AI can be enlisted to

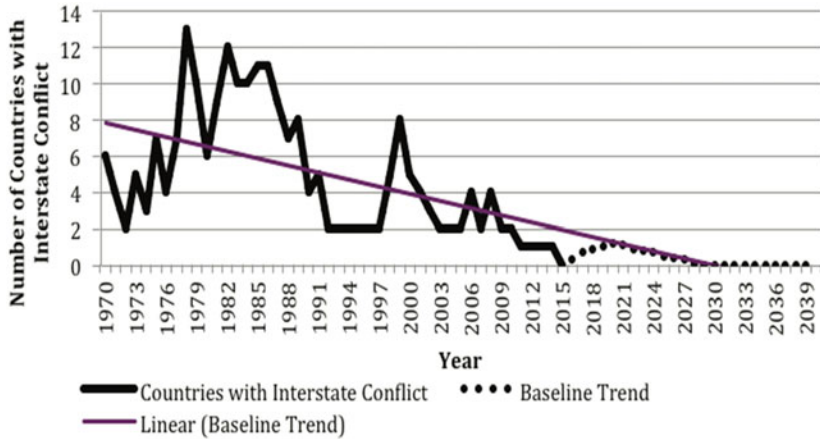


Fig. 10.1 Countries with ongoing interstate violence past and future, 1970–2040 (*Source* RAND ‘Conflict Trends and Conflict Drivers: An Empirical Assessment of Historical Conflict Patterns and Future Conflict Projections’ [2015] https://www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1063/RAND_RR1063.pdf)

mitigate threats from non-state actors, such as terrorist organisations by predicting probable sites for future attacks.

These early years in the history of AI present an opportunity to attain transnational “AI integration” in order to circumvent some of the scenarios described in Chapter 5 on the threatened future of interdependence. Open source and global collaboration will be crucial in this regard. But, unavoidably, state power will continue to matter and set the trajectory of international politics. In the book *Winning the War on War*, Goldstein (2012) observes that:

Around the world, almost all nations maintain capable military forces. These forces contain some 20 million soldiers worldwide, several million of them in designated combat zones. They are armed to the teeth with guns, planes, ships, missiles, submarines, helicopters, drones, flamethrowers, bayonets, nuclear weapons, and every other conceivably useful means of destruction...But nowhere in the world are those military forces fighting each other. (Goldstein 2012: 276)

Figure 10.1 shows the declining trend since the 1980s, occasional disturbances notwithstanding. With the advent of AI, more warfare is likelier to take place in the digital realm and only pour over into the conventional and proxy warfare of past and contemporary times when objectives cannot be obtained as discussed in Chapter 5.

10.3 THEORETICAL INSIGHTS

Our analysis distils seven key findings. Firstly, *all* theories will be impacted by AI and will have to think its implications through for their theoretical traditions. Secondly, traditional theories have more predictive power for relations among states, whereas non-traditional ones have more explanatory value, and can more meaningfully interject in terms of policy interventions at the sub-state and international organisational levels. Thirdly, and specific to certain theories, we have identified the need to incorporate AI into the realist arsenal through the concept of latent power. A model was presented in this regard in Chapter 4. Fourth, towards liberalism, we show that competition over AI has led to geopolitical tensions, which have resulted in democracies paradoxically curbing liberties in the name of preventing access for authoritarian regimes into their states. We also note the need to revise long-standing liberal assumptions, namely, the monadic argument within the democratic peace and the commercial interdependence thesis. In the case of the former, we note that the introduction of AI into information systems that are crucial for domestic audiences (including information warfare, fake news, and Deepfakes) skews the assumed mechanism through which civilians hold their leaders accountable. In the case of the latter thesis (commercial interdependence), we note that the rise of automation likely means a disruption towards the standing patterns of interdependence which have characterised globalisation since the 1990s.

Fifth, through critical lenses, we note bias and disproportionate cause and effect along the lines of race, gender and industrialisation stages. Sixth, we observe that conventional conflict is less likely, and will be less sharpened but more protracted due to the blurred nature of battlelines with the rise of cyber-attacks, which will be made more accurate by AI. Finally, the analysis showcases opportunities for combining insights from both traditional and non-traditional theoretical paradigms. Opportune is the mutual focus on power by realism and critical theories. The former explains its imbalance (and the importance thereof) at the state-to-state

level, whereas the latter does so at the sub-state level, including along lines of race and gender. These emerge as areas for further research and typological development modelling.

10.4 METHODOLOGICAL AND PEDAGOGICAL OPPORTUNITIES

It is useful to conclude by returning to some of the fundamental questions of how IR is studied, and new theoretical knowledge discovered. As seen in Chapter 2, scholars in the field make use of a variety of methods for generating, supplementing, and testing paradigms. Both qualitative and quantitative scholars have much to gain from tapping into AI, for at its core it is a way of acquiring data and putting it to use. AI can be useful for modelling, predicting, and replicating events and thus corroborate many of the assumptions which inform our analysis. Thus, with time, AI-generated results in IR should soon become the norm. The key will be in interdisciplinary work, in addition to the attainment of knowledge of AI and algorithms by IR scholars. In relation to this, there also arises opportunities for using AI in the IR classroom. Many events covered in today's IR syllabus are decades and even centuries old. This renders them too abstract for many of our undergraduate students. For some, the events of the Vietnam War are as archaic as those of the Second World War. AI presents an opportunity for immersive, experiment-based teaching and learning. Early research and development in virtual and augmented reality (most recently receiving momentum following Facebook's announced focus on accelerating the creation of a "metaverse") can be given application in the education sector as well. Through this research-as-education model, students will be able to take up roles as simulated agents in such critical historical events, while also using AI to model contemporary and even future developments on key transnational issues with a theoretical angle. In so doing we will put theories to the test and require them to live up to their predictive component in addition to their explanatory function.

BIBLIOGRAPHY

- Goldstein, Joshua. 2012. *Winning the War on War: The Decline of Armed Conflict Worldwide*. New York: Plume.
- Ndzendze, B. 2018. In Era of Artificial Intelligence, Optimists for Peace are Warranted. *Independent Online*. Available at: <https://www.iol.co.za/lifestyle/opinion/in-era-of-artificial-intelligence-optimists-for-peace-are-warranted-ce2e424dd829-5d09-bf07-ff499d3cc358>.
- RAND. 2015. Conflict Trends and Conflict Drivers: An Empirical Assessment of Historical Conflict Patterns and Future Conflict Projections. RAND. https://www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1063/RAND_RR1063.pdf.

INDEX

A

Africa, 7, 8, 25, 28, 45, 111,
145–147
African Union (AU), 146, 147
Agriculture, 42, 154
Apple Computer, 147
Artificial intelligence (AI), 1–12, 16,
27, 29, 33–44, 46–51, 56, 57,
60–64, 66, 69, 70, 74, 76, 77,
80–85, 89, 93–101, 107, 134,
138–140, 142, 147, 148, 151,
153, 154, 157, 158, 160, 161
Automobiles, 35, 138

B

Bell Labs, 38
Belt and Road Initiative (BRI), 97
Bias, 9, 152, 160
Blockchain, 3, 146, 147
Brazil, 6

C

Canada, 50, 99, 138

China, 4, 8, 12, 21, 35, 41–47, 50,
61–65, 68, 75, 87, 89, 93,
97–100, 136, 139, 146, 151
Climate change, 2, 25, 63, 153
Competitiveness, 62, 109, 110
COVID-19, 3, 63, 77
Culture, 58, 83, 84, 140, 146–148

D

Democracy, 5, 11, 18, 20, 58, 63, 73,
74, 76–80, 83, 84, 90, 147, 148,
160
Democratic peace thesis, 11, 18, 74,
77, 80, 81, 101
Development, 7–11, 20, 28, 41–44,
47, 62, 64, 66, 70, 74, 76, 84,
89, 93, 95–97, 99–101, 108,
109, 133, 138, 153, 161
Diplomacy, 6, 8, 10, 37, 50, 94, 136,
145, 150
Drones, 4, 6, 49, 79, 94, 100, 148,
159

E

Economic development, 18, 25, 28
 Elections, 65, 78, 79
 European Union (EU), 6, 50, 137, 139
 Evolution, 4, 10, 11, 16, 25, 35, 69, 141, 149, 151

F

5G (fifth generation connectivity), 3, 45
 Foreign direct investment (FDI), 2, 9, 10, 48, 50, 105–107, 110–113, 115, 117, 119, 121, 123, 125–130
 4IR (Fourth Industrial Revolution), 3, 12, 29, 36, 45, 49, 105–107, 109–113, 125, 126, 129, 130, 142, 146–148
 Functionalism, 8, 99

G

Gender, 9, 12, 27, 29, 142, 148, 150–152, 160, 161
 Germany, 65, 144, 146
 Globalization, 18, 76, 82
 Global value chains, 82

I

Industrial Revolution, 12, 55, 105, 107
 Information warfare (IW), 56, 67–69, 79, 97, 160
 Infrastructure, 3, 5, 29, 64, 84, 93, 99, 108, 109
 Innovation, 44, 45, 47, 62, 66, 74, 76, 93, 101, 105, 107, 109, 127, 128, 147, 148, 153, 158
 International Telecommunications Union (ITU), 134

Internet of Things (IoT), 3, 35, 36, 49, 146

Israel, 68, 93, 100

J

Japan, 39, 46, 50, 65, 67, 75, 97

L

Language, 6, 46, 51, 141, 144–148
 Latin America, 6, 7, 45, 108

M

Machine learning (ML), 33, 34, 38, 40, 46, 87, 94, 95, 101, 138
 Mexico, 50
 Microsoft, 101
 Middle East, 7, 88
 Migration, 25
 MIT, 38

N

National strategies (artificial intelligence), 47
 Nuclear weapons, 25, 65, 67, 68, 88, 137, 159

O

Open source software (OSS), 159

P

Patents, 45, 46, 49, 66, 99

R

Regional integration, 2

S

Schwab, Klaus, 3, 109
 Second Industrial Revolution, 158

Silicon Valley, 8
Software, 35, 50, 64, 108, 138
South Africa, 3, 50, 68
South Korea, 50, 93
Sputnik, 39, 87, 97
Start-ups, 97

T

Trade, 2, 3, 6, 9, 10, 16–19, 48, 50,
75, 81–83, 90, 97, 107, 136,
139

U

United Nations (UN), 75, 136, 138,
146, 151, 152

Universities, 3, 19, 42, 44, 146, 148

V

Von Clausewitz, Karl, 69

W

War, 5, 8, 10, 16, 18–20, 22, 28, 43,
50, 51, 56, 67–69, 77–81, 84,
87, 89, 91–93, 96–98, 100, 136,
139, 144–147, 150, 158

World Economic Forum (WEF), 5,
66, 105, 106, 109–111, 126–129