SINGULARITY RACE: ARTIFICIAL SUPER INTELLIGENCE AS MILITARY REVOLUTION

A thesis presented to the Faculty of the U.S. Army Command and General Staff College in partial fulfillment of the requirements for the degree

MASTER OF MILITARY ART AND SCIENCE
General Studies

by

KYLE ROBERT HOPKINS, MAJOR, U.S. ARMY
Bachelor of Arts, American Military University, Charles Town, West Virginia, 2014

Fort Leavenworth, Kansas
2022

Approved for public release; distribution is unlimited. Fair use determination or copyright permission has been obtained for the inclusion of pictures, maps, graphics, and any other works incorporated into this manuscript. A work of the United States Government is not subject to copyright, however further publication or sale of copyrighted images is not permissible.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED *(From - To)*
10-06-2022	Master's Thesis	AUG 2021 – JUN 2022

4. TITLE AND SUBTITLE

Singularity Race: Artificial Super Intelligence as Military Revolution

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Kyle Robert Hopkins

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Command and General Staff College
ATTN: ATZL-SWD-GD
Fort Leavenworth, KS 66027-2301

8. PERFORMING ORG REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Artificial Intelligence (AI) is a rapidly developing field with governments and militaries around the world increasingly incorporating it into their technologies to create new capabilities. AI has the potential to eventually surpass human intellectual capabilities and obtain Super Intelligence. This thesis examines the implications of Artificial Super Intelligence (ASI) and how an adversary to the United States could employ it to gain an asymmetric strategic advantage. This paper finds that ASI poses an extreme risk to future operations in the mid and possibly near term and makes recommendations for how the Department of Defense should think about and incorporate the threat of ASI into strategic planning.

15. SUBJECT TERMS

Super-Intelligence, Artificial Intelligence, Military Revolution, Information Operations

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. PHONE NUMBER *(include area code)*
(U)	(U)	(U)	(U)	149	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

MASTER OF MILITARY ART AND SCIENCE

THESIS APPROVAL PAGE

Name of Candidate: Kyle Robert Hopkins

Thesis Title: Singularity Race: Artificial Super Intelligence as Military Revolution

Approved by:

_____, Thesis Committee Chair
Phillip G. Pattee, Ph.D.

_____, Member
Peter S. Im, MSS, MSSI, MMAS

_____, Member
LTC Christopher M. Baldwin, M.A.

Accepted this 10th day of June 2022 by:

_____, Assistant Dean of Academics for
Dale F. Spurlin, Ph.D. Degree Programs and Research

The opinions and conclusions expressed herein are those of the student author and do not necessarily represent the views of the U.S. Army Command and General Staff College or any other governmental agency. (References to this study should include the foregoing statement.)

ABSTRACT

SINGULARITY RACE: ARTIFICIAL SUPER INTELLIGENCE AS MILITARY REVOLUTION, by Kyle Robert Hopkins, 149 pages.

Artificial Intelligence (AI) is a rapidly developing field with governments and militaries around the world increasingly incorporating it into their technologies to create new capabilities. AI has the potential to eventually surpass human intellectual capabilities and obtain Super Intelligence. This thesis examines the implications of Artificial Super Intelligence (ASI) and how an adversary to the United States could employ it to gain an asymmetric strategic advantage. This paper finds that ASI poses an extreme risk to future operations in the mid and possibly near term and makes recommendations for how the Department of Defense should think about and incorporate the threat of ASI into strategic planning.

ACKNOWLEDGMENTS

My beloved wife Aly is due an amount of thanks roughly equivalent to Graham's number. While I spent nine months slowly growing the size of this thesis, carrying it around inside my laptop with me everywhere I went, she spent nine months slowly growing our second child, carrying it around inside her while chasing our toddler and asking me how my thesis was coming. She volunteered for more early mornings to allow me time to work and brought me motivation in the form of ice cream when I was struggling with a difficult section. This thesis is just as much a product of her love as it is my words and research.

I would also like to thank the members of my committee. Researching a topic that exists only in theory posed a significant challenge for a structured program. This thesis utilizes methodology new to the college and explores topics on the edges of current knowledge. I appreciate the insight, guidance, and patience of each of my committee members in helping complete this project.

TABLE OF CONTENTS

Page

ACRONYMS

AI	Artificial Intelligence
AGI	Artificial General Intelligence
ASI	Artificial Super Intelligence
CCP	Chinese Communist Party
NAI	Narrow Artificial Intelligence
NIC	National Intelligence Committee
TPU	Tensor Processing Unit

ILLUSTRATIONS

Page

TABLES

Page

CHAPTER 1

INTRODUCTION

Background

> What would happen if technology continued to evolve so much more rapidly than the animal and vegetable kingdoms? Would it displace us in the supremacy of earth? We as yet have only seen what will one day be considered the antediluvian protypes of the race…. We are daily giving [machines] greater power and supplying by all sorts of ingenious contrivances that self-regulating, self-acting power which will be to them what intellect has been to the human race.
>
> —Samuel Butler, "Darwin Among the Machines," 1863

One of humanity's greatest strengths is our ability to utilize tools. Throughout history tools have allowed us to increase our efficiency at performing tasks, specialize our knowledge, and create opportunities to iteratively improve and create yet more complex tools. The invention of machines gave humanity a set of tools that could completely replace human labor instead of amplifying it, and far surpass humans in the speed and quality of that labor. Now, developments in Artificial Intelligence (AI) are doing the same with thinking. Already AI systems have outperformed humans in tasks that previous generations believed forever beyond the reach of machines such as: image and object recognition,[1] complex board games like chess[2] and the significantly more

[1] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang et al., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision* 115 (April 2015): 211–252, https://doi.org/10.1007/s11263-015-0816-y.

[2] IBM Corporation, "Deep Blue," IBM 100: Icons of Progress, last modified March 7, 2012, http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/.

complex go,[3] video games that require strategic thinking in real time,[4] lip reading,[5] even turning a profit by investing in the stock market.[6] Currently, all applications of AI are narrow, meaning that although they can exceed human capability at a certain task they can only do the one specific thing they are designed for and nothing else. However, this will not always be the case.

In 1993, the statistician Vernor Vinge predicted that humanity would have the ability to create a superhuman intelligence.[7] He speculated that such a creation would lead to an intelligence explosion as the super intelligence improves upon itself becoming even more and more intelligent and leaving humankind far behind in capability. Such a capability would have profound implications for every aspect of human life. As the AI and decision theorist Eliezer Yudkowsky puts it: "[T]here are no hard problems, only problems that are hard to a certain level of intelligence. Move the smallest bit upwards [in

---

[3] David Silver, Aja Huang, Christopher Maddison, Arthur Guez, Laurent Sifre, George Driessche, Julian Schrittwieser et al., "Mastering the Game of Go with Deep Neural Networks and Tree Search," *Nature* 529 (January 2016): 484–489, https://doi.org/10.1038/nature16961.

[4] Dan Garisto, "Google AI Beats Top Human Players at Strategy Game StarCraft II," *Nature*, October 30, 2019, https://doi.org/10.1038/d41586-019-03298-6.

[5] Yannis M Assael, Brendan Shillingford, Shimon Whiteson, and Nando de Freitas, "LipNet: End-to-End Sentence-Level Lipreading," version 2, Oxford University, December 16, 2016, http://arxiv.org/abs/1611.01599.

[6] Fernando G.D.C. Ferreira, Amir H. Gandomi, and Todrigo T. N. Cardoso, "Artificial Intelligence Applied to Stock Market Trading: A Review," *IEEE Access* 9 (February 2021): 30,898-30,917, https://ieeexplore.ieee.org/document/9350582.

[7] Vernor Vinge, "Technological Singularity," Carnegie Mellon University, March 1993, https://frc.ri.cmu.edu/~hpm/book98/com.ch1/vinge.singularity.html.

level of intelligence], and some problems will suddenly move from 'impossible' to 'obvious.' Move a substantial degree upwards, and all of them will become obvious."[8] A superintelligence, therefore, would provide an asymmetric advantage to any actor on the world stage capable of developing and controlling it.

Since the end of World War Two the United States has been the dominant world power, capable of expanding its influence and pursuing its interests via the four instruments of national power: diplomacy, information, military, and economics.[9] However, history has shown us that dominant world powers can lose this status very unexpectedly and very rapidly due to military revolutions. Western history has experienced five such major revolutions: creation of the modern nation state, the merging of mass politics and warfare, the industrial revolution, World War One, and the advent of nuclear weapons.[10] Each of these revolutions center around an asymmetric advantage provided to those first to adopt them which allowed them to gain prominence, if only temporarily, until others also adopt the new paradigm out of necessity. The advent of a super-intelligent AI has the potential to create the next military revolution and remove the United States from its position of world dominance if controlled by an adversary.

---

[8] Eliezer S. Yudkowsky, "Staring Into the Singularity," Singularity, 1999, http://www.pivot.net/~jpierce/staring_into_the_singularity.htm.

[9] Office of the Chairman of the Joint Chiefs of Staff, Joint Publication 1, *Doctrine for the Armed Forces of the United States* (Washington, DC: Joint Chiefs of Staff, 2017), I-4.

[10] Williamson Murray and Macgregor Knox, "Thinking about Revolutions in Warfare," in *The Dynamics of Military Revolution, 1300–2050,* ed. MacGregor Knox and Williamson Murray, (Cambridge: Cambridge University Press, 2001), 6.

Murray and Knox point out that military revolutions by their nature are "uncontrollable, unpredictable, and unforeseeable."[11] But this is a matter of perspective. Nassim Taleb refers to these types of events as black swans and has three criteria for them: that they are rare, have an extreme impact, and are retrospectively predictable.[12] However, he demonstrates that the reason something becomes a black swan is specifically because the possibility was not taken seriously, not that no one thought of it. As an example, if someone had considered terrorism a plausible threat before 9-11 and required locks on airplane cabin doors, the black swan event of hijacked planes becoming weapons would never have occurred. The person who pushed for the change would likely even be criticized for being incorrect in their assumptions. Without the event occurring there would be no evidence that the implemented changes had the desired effect of preventing terrorism.

Correct predictions of the future are generally impossible to verify when they are taken seriously because they succeed in preventing the outcomes they predicted. Hence, ideas deemed unworthy of due consideration or planning become decisive. "Isn't it strange to see an event happening precisely because it was not supposed to happen?"[13] If the United States wishes to retain dominance on a world stage that is rapidly developing and becoming ever more complex, it must ponder the probable and the improbable alike. As the former Staff Director of the Senate Armed Services Committee, Christian Brose,

---

[11] Murray and Macgregor Knox, "Thinking about Revolutions in Warfare," 7.

[12] Nassim Nicholas Taleb, *The Black Swan: The Impact of the Highly Improbable* (New York: Random House, 2007), xxii.

[13] Ibid., xxiv.

states in the conclusion of his book *The Kill Chain*, "the problem is a failure of imagination."[14]

Problem Statement

Technology is developing at an exponentially faster rate and will at some point result in the next paradigm shift and military revolution. Artificial Super Intelligence (ASI) has the potential to create an asymmetric advantage in the ability for an international actor to wield the instruments of national power and result in such a military revolution. Historically, military revolutions tend to result in a deposing of the dominate world powers of their time by whomever is the first to wield the asymmetric advantage of the new revolution. The next revolution will be the same and the United States could find itself deposed from its place of world dominance unless it can predict and prepare for what is coming. It is not sufficient to seek solutions to gaps that challenge U.S. dominance today, because today's problems will be irrelevant after the next paradigm shift. Therefore, how can the Department of Defense identify vulnerabilities that could be exploited if the next military revolution centers around an Artificial Super Intelligence?

Purpose of the Study

There is a notable gap in Department of Defense literature in regards to super intelligence. The intent of this exploration is to fill the literature gap, identify the degree of risk posed by such a technology, and identify potential vulnerabilities an adversary could exploit with an ASI to attack the United States. This will provide decision makers

[14] Christian Brose, *The Kill Chain* (New York: Hachette Book Group, 2020), 246.

within the defense community relevant perspective on the topic and inform the degree to which the technology should be considered in planning and forecasting. With additional research the identified vulnerabilities could be mitigated to prevent their exploitation should an adversary be the first to acquire an ASI. Ultimately, the purpose of this study is to prevent the advent of an ASI military revolution becoming a black swan event that deposes the U.S. from its position of world dominance.

Research Questions

The primary research question this thesis seeks to answer is: how could an adversary utilize an ASI to supplant the United States as the dominant world power? In order to answer this question and address the purpose of the study several additional questions need to be answered:

1. What unique capabilities will an ASI have that allow it to affect the information and military elements of national power?
2. How could an actor adversarial to the United States utilize ASI capabilities to achieve operational and strategic effects?
3. What are vulnerabilities in the information and military realms that could be exploited by an adversary of the United States with an ASI?

Assumptions

As an exploration of how a theoretical technology could be employed in a future operating environment, several assumptions have to be made about the state of that future environment. Because the purpose of this thesis is to identify vulnerabilities and assess risk, the assumptions made about that environment are ones that are assumed to have the

potential to create the greatest risk. Each of these key assumptions are covered in greater depth in the literature review:

1. Artificial General Intelligence (AGI) will be initially developed by 2035 and advance to ASI by 2040.
2. An adversary to the United States will be the first to develop AGI/ASI and the technology will not be shared or proliferated outside their control.
3. A developed ASI is controllable, does not possess drives and motivations independent from its controlling agent and functions as an extension of its controlling agent's will.
4. The operational environment in 2040 will be as described in the National Intelligence Committee 2040 report competitive coexistence scenario, characterized primarily by competition between the U.S. and China.

Definition of Terms

Artificial Intelligence (AI) – a computer system capable of intelligent behavior or cognition.

Narrow AI (NAI) – an artificial intelligence that is trained to perform a specific task and is unable to apply knowledge generally. Its ability to perform tasks within its domain can exceed human capabilities but fails to provide logical or sensical outputs if applied to problems or tasks outside its specific domain. Synonymous with soft AI.

Artificial General Intelligence (AGI) – artificial intelligence that can apply knowledge generally across contexts and is equal to a human across most cognitive domains. Synonymous with hard AI, true AI, and Human Like Machine Intelligence.

Super Intelligence – "any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest."[15]

Artificial Super Intelligence (ASI) – artificial intelligence which greatly exceeds the cognitive performance of modern humans across all domains of relevant interest.[16]

Technological Singularity – the point at which the rate of technological development increases by such a degree humanity experiences a paradigm shift in its existence and all predictions of the future based on current or historical trends and models become irrelevant.

Scope, Limitations, and Delimitations

This study was scoped and delimited in several ways, almost all of which are related to the primary limitation of time to complete the research. There are multiple paths to achieving super intelligence, but this research was scoped to only explore ASI. The scope was further reduced to exploring ASI capabilities and impacts in the information and military elements of the theoretical framework of elements of national power (DIME) as the most relevant to the defense community.

The study was delimited in the following ways. This study will only produce a single normative narrative scenario developed to reflect the greatest potential risk. This study will not seek to compare and contrast multiple possible scenarios such as a super-intelligence manifesting via means other than AI, nor compare and contrast multiple

[15] Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (New York: Oxford University Press, 2014), 26.

[16] Ibid.

courses of action an adversary could take within the single scenario. Scenarios will not be created that explore how international actors that are neutral or friendly to the United States might act should they be the first to obtain an ASI, nor how the world community might react should the United States be the first to obtain an ASI. Each of these scenarios has potential value for analysis, but the adversarial ASI scenario was selected due to the author's perception that it would provide the best lens through which to identify exploitable vulnerabilities for mitigation.

Significance of the Study

While in recent years members of the defense community have flooded the literature with writings on AI, there is a notable gap in the literature on the topic of super intelligence from a defense perspective. While NAI has the potential to be transformational in its own right in the military and civil sectors, this study will start to fill the gap in knowledge about the future of AI and its potential to create a military revolution. Additionally, this study seeks to identify potential vulnerabilities that could be exploited by an adversary controlling an ASI to depose or replace the U.S. as the dominant world power. The findings of this study will provide insight to strategic planners forecasting future force requirements. Identification of vulnerabilities is the first step in mitigation; however, additional research will be required to find solutions for the vulnerabilities identified. Paradoxically, the true significance of this study would only be realized if its predictions do not come true; perhaps indicating that it played a part in preventing the negative scenario envisioned from coming to pass.

Summary

Left unexplored the next military revolution has the potential to depose the United States as the dominant world power. Vernor Vinge's predictions of a super-intelligence, especially via the path of ASI, poses a very real possibility of ushering in the next revolution. Understanding the strategic risks associated with ASI requires an exploration of how it could be employed to create a worst-case scenario where the United States is replaced as the dominant world power. This study will do just that and identify future vulnerabilities that an adversary could exploit utilizing an ASI. With additional research, these vulnerabilities could then be mitigated and prevent a future with negative implications for the United States.

CHAPTER 2

LITERATURE REVIEW

Introduction

An exploration on the effects of super intelligence is only of value if this theoretical technology is feasible, probable, and achievable within a moderate timeframe. A technology that will revolutionize aspects of human life but is not feasible for several centuries is not worth studying now; there would simply be too many unknowns between now and then to attempt to mitigate any of its effects. As such, the majority of this literature review is dedicated to understanding the feasibility and probability of ASI arriving in the near term. This is accomplished through a review of the literature on the field of AI and its current capabilities, what would be required to advance to ASI, and the expected timeframe for such an occurrence. This literature review also includes an overview of the forecasting scenario methodologies considered for use in this thesis to analyze the effects of ASI and relevant information on the future operating environment pertinent to the scenario this thesis will explore.

Super Intelligence

The concept of super intelligence has wide reference within the literature but has few works that explore the concept in depth. Within the field Vernor Vinge, Bosner, and Kurzweil are the seminal writers; Yampolskiy has also edited a notable collaboration of work on ASI specifically. Each of these authors explore the concept from different backgrounds and perspectives but reliably produce very similar analysis and conclusions, agreeing on most aspects of the topic. Bosner provides the succinct definition for

superintelligence that is used within this thesis: "any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest."[17] Bosner also uniquely posits that there are multiple forms of super intelligence: speed, collective, and quality. Speed and quality super intelligence are intellects that can do everything a human intellect can do but significantly faster or to a significantly higher quality. Collective super intelligence, however, is achieved simply by combining many intellects into a system or organization which can then perform well above the cognitive capability of any one of the individual intellects that comprise it.[18] In this way Bosner demonstrates that humanity has already achieved a form of super intelligence. No human could get to the moon by themselves, but collectively we achieve incredible things.

While the collective super intelligence of humans is effective, super intelligence of the other forms—speed or quality—would also be capable of combining into organizations to further boost their capability. Each of the seminal authors agree that there are multiple paths to achieving non-collective super intelligence. First, advances in biological science could provide a means to alter the human brain in such a way as to develop super intelligence, most likely in the form of a quality super intelligence. Second, human minds could be linked to and interfaced with computers allowing them to gain some of the advantages of digital computing discussed later in this literature review. This would most likely result in a speed super intelligence. Finally, a machine intelligence that could match human ability to perceive, understand, and interact with the world would

[17] Bostrom, *Superintelligence*, 26.

[18] Ibid., 64.

likely result in both a speed and quality super intelligence.[19] While some advances have been made in fields relevant to the first two paths, they are massively overshadowed by developments in the field of AI. Bosner and Kurzweil agree that ASI is the most likely path. Due to this loose consensus, it was decided to apply a delimitation to only explore super intelligence via the path of ASI in this thesis.

An important aspect of super intelligence that all authors focus on is the rate of technological advancement that it provides. Super intelligence creates an exponential rate of advancement over time as discoveries make subsequent discoveries easier through tool generation or by allowing further increases in levels of intelligence. This can be demonstrated through human collective super intelligence; it was thousands of years between the invention of agriculture and the wheel, but it was less than 25 years between the advent of the personal computer and the smart phone.[20] This has implications for ASI in particular. As the mathematician I. J. Good—who was the lead statistician on Alan Turing's team during WWII—wrote in 1965:

> Let an ultra-intelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultra-intelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. Thus, the first ultra-intelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control.[21]

---

[19] Vinge, "Technological Singularity."

[20] Ray Kurzweil, *The Singularity is Near* (Kansas City: Penguin Books, 2005), 20.

[21] Irving John Good, "Speculations Concerning the First Ultraintelligent Machine," in *Advances in Computers*, ed. Franz L. Alt and Morris Rubinoff (New York: Academic Press, 1965), 33.

The advent of ASI would result in a double exponential growth of technological advancement as the ASI simultaneously solves problems and increases its own capabilities and intelligence, further increasing the rate at which it can continue to do both.[22] The technological singularity is a theoretical point in the future when ASI has improved itself to such a degree that the rate of technological advancement becomes so fast it necessitates a paradigm shift for humanity. The singularity also places a theoretical mark on the future timeline beyond which it becomes impossible to make predictions about the future based on current trends and models.[23] The singularity is an important concept to highlight due to the prominence and focus it receives in the literature. However, it will not be a primary focus of this thesis, except as an adversary objective, because it represents a change so significant that it is impossible to predict or plan for. This thesis will focus exclusively on the time period immediately after the advent of ASI and how it could be utilized to shape the world prior to a technological singularity. To that end, it is important to understand what the literature says about what AI is, what it can currently do, what would be required to advance to ASI, and how likely that is to happen.

Artificial Intelligence Undefined

When discussion AI, it should first be established that the literature shows no consistency for the definition of AI as a term since it was coined by John McCarthy in

[22] Kurzweil, *The Singularity is Near,* 12.

[23] Vinge, "Technological Singularity."

1956 at the Dartmouth Summer Research Project on Artificial Intelligence.[24] The term Artificial Intelligence was originally chosen to describe the concept of thinking machines due to its neutrality among multiple schools of thought including cybernetics, automata theory, and complex information processing, each of which had researchers in attendance.[25] The original proposal for the study did not provide a formal definition, but did set an unofficial foundation for understanding the term: "The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it."[26]

Since then, no consensus on the definition of AI has been established. A survey of AI professionals in 1983 produced over 100 definitions for AI.[27] This is a reflection of the divergence in how separate groups approach the concept of AI. Some authors, Bijker most prominent among them, believe that AI cannot be described with a simple definition at all because it is not one thing but rather a phenomenon constructed through complex

---

[24] Nils J. Nilsson, *The Quest for Artificail Intelligence: A History of Ideas and Achievements* (Cambridge: Cambridge University Press, 2009), 77.

[25] Pamela McCorduck, *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* 2nd ed. (Natick, MA: A. K. Peters, Ltd., 2004), 115.

[26] John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955," *AI Magazine* 27, no. 4 (December 15, 2006): 12, https://ojs.aaai.org/index.php/aimagazine/article/view/1904.

[27] Negrotti, Massimo, *Understanding the Artificial: On the Future Shape of Artificial Intelligence* (London: Springer-Verlag, 1991), 155-157.

social processes.[28] But despite the lack of consensus on AI definitions, the overall trend of how the term is used appears to have gone through six stages that waver between computer as a machine and as a program:

1. Original – mechanical device
2. Simulating/imitating/mimicking original intelligence
3. Machine process
4. Function of computing
5. Branch of Computer Science
6. Programming/Software[29]

This issue has not dissipated in recent years despite AI becoming a widely recognized concept within the general population. An analysis in 2021 looked at the differences in AI definitions between stakeholders across different industries in Australia, Finland, the United States, and the UK.[30] They found that definitions for AI tended to fall into three main categories depending on if respondents were within academia, industry, or government. Academics define AI in terms of a methodology that is pedagogically

[28] Wiebe E. Bijker, "How Is Technology Made? That Is the Question!" *Cambridge Journal of Economics* 34, no. 1 (2010): 63-76, https://doi.org/10.1093/cje/bep068.

[29] Dalvinder Singh Grewal, "A Critical Conceptual Analysis of definitions of Artificial Intelligence as Applicable to Computer Engineering," *IOSR Journal of Computer Engineering* 16, no. 2 (2014): 9-13.

[30] Rebecca Eynon and Erin Young, "Methodology, Legend, and Rhetoric: The Constructions of AI by Academia, Industry, and Policy Groups for Lifelong Learning," *Science, Technology, & Human Values* 46, no. 1 (January 2021): 170.

valuable for learning.[31] Industry and commercial stake holders have definitions of AI that exhibit a mythical reverence and approach it like a cultural artifact useful as a capability in marketing or profit generation. Finally, analysis of policy documents and interviews with policy makers demonstrated a scarcity of knowledge on the subject that led their definitions to be rhetorically focused.[32]

Defining AI is made even more difficult by the fact that technologies considered to be AI have changed over time. Many technologies that were originally considered AI have since been relegated to the sphere of software as understanding of their functionality became ubiquitous. This is known as McCarthy's dictum: once something works as expected it is no longer considered AI.[33] Combined, these issues make defining AI, except in extremely broad terms very difficult; even the U.S. government has no official definition for AI. Therefore, the following definition was selected for this thesis: a computer system capable of intelligent behavior or cognition.

Narrow AI (NAI)

NAI is a subcategory of AI used to describe all the forms of AI that are currently in use. The term is derived from the fact that all current AI applications can only perform very narrow functions and do not have the ability to generally apply knowledge. An AI created to read X-Rays would be clueless if you asked it to tell the difference between a tree and a car. NAI runs the gambit from search engine algorithms to deep learning neural

[31] Eynon and Young, "Methodology, Legend, and Rhetoric," 172.

[32] Ibid., 178.

[33] Bostrom, *Superintelligence*, 19.

networks that aid researchers in developing medical advancements. The amount of literature available on NAI is enormous and growing at an incredible rate. Even within the defense community the literature available on NAI is vast with innumerable contributors from service post graduate schools, think tanks, governmental agencies, and security officials. Because NAI is not the focus of this thesis, review of the extensive literature on NAI was limited to specific current capabilities that have relevance or implications for national defense and will aid in an analysis of potential ASI capabilities in chapter four.

Advantages of Digital Intelligence

All AI benefit from a set of advantages related to their non-organic nature compared to human intelligence. These advantages are inherent in all AI and can be assumed as present during the discussion of any specific AI capabilities, including AGI and ASI. This thesis uses Bostrom's list of ten advantages of digital intelligence, summarized below:

1. Speed of computational elements: biological neurons have a peak speed of 200 Hz, seven orders of magnitude less than a cheap 2 GHz micro-processor.
2. Internal communication speed: neurons transmit information at 120 m/s compared to digital communication at the optical speed of light (300,000,000 m/s) resulting in significantly reduced latency even at enormous size.
3. Number of computational elements: organic brains are limited by brain size and other factors for the number of neurons they can host while digital systems are near infinitely scalable.

4. Storage capacity: human working memory can hold 7 ± 2 chunks of information at a time and long-term memory is estimated to have a maximum of about one billion bits of potential storage, several orders of magnitude less than an average smart phone.
5. Reliability: biological brains fatigue after a few hours of work, permanently decay over time, and can rarely perfectly reproduce information or perform tasks. Digital intelligences never tire or degrade and can always perfectly reproduce information and perform tasks.
6. Editability: changing organic minds via brain surgery is very difficult and extremely risky. Digital intelligence can be easily edited for a wide range of purposes.
7. Duplicability: it takes a long time to create and educate a biological mind while digital intelligences can be copied any number of times based on available hardware.
8. Goal coordination: it is difficult to get humans to share the same goals and, even when they do, to coordinate their efforts to effectively achieve them. Machine intelligences can simply be assigned to or duplicated to have the same goal set.
9. Memory sharing: machine intelligence can share not just information but memories and skills through the transfer of data files. Large organizations of AI can synchronize their databases to give each of them collective understanding of what each individual part has learned.

10. Modular algorithms: the ability to add or remove specialized algorithms tailored for specific domains of cognition to maximize cognitive effectiveness.[34]

Game Mastery

As of 2012 AI systems had already demonstrated superhuman capability in games such as checkers, backgammon, Othello, chess, crossword puzzles, Scrabble, bridge, Jeopardy!, and FreeCell. The game of Go was one game where AI took much longer to develop capability, but much less time than was predicted. In 2012 AI systems were performing at the level of a strong amateur; the Zen series of Go playing programs were ranking 6-dan—the equivalent of an advanced amateur player—in tournaments and improving their ranking by about one dan/year. Continuing to improve at that rate meant AI would likely be able to beat the world go champion in 2022. But most experts were skeptical of that ten-year timeline due to the complexity of the game of Go which has more possible board states than atoms in the universe. Instead, DeepMind's Alpha Go decisively defeated the world Go champion Lee Sedol four games to one in 2016.[35] Since then, multiple iterations on the AlphaGo system have immerged which are even more powerful. AlphaGo Master won 60 straight online games against professional Go players using just four Tensor Processing Units (TPUs)—TPUs are a new type of processing card developed by Google in 2015 that specialize in long form matrix dominated computations

[34] Bostrom, *Superintelligence*, 71-74.

[35] DeepMind, "Alpha Go," directed by Greg Kohs, Moxie Pictures, 2017, streaming video, 1:30:28, https://www.youtube.com/watch?v=WXuK6gekU1Y.

needed for neural network machine learning—on a single machine; compared to the 1,920 CPUs and 280 GPUs used in the Lee Sedol match.[36] AlphaGo Zero learned to play Go without using any data from human games, instead it played only against itself and exceeded the abilities of all previous versions of AlphaGo within 40 days.[37] Finally, MuZero was able to learn to play without ever being taught the rules and still surpassed all its predecessors.[38]

AI deep learning has also advanced past the ability to conquer humans in turn-based games and into real time strategy games. Unlike board games where AI has large amounts of time to consider its next move, real time strategy games require players to constantly adapt to conditions in real time. Players at the top of games like Starcraft II input more than 300 commands per minute to command their forces on a simulated battlefield.[39] Players must develop a strategy to manage resources, gain information on their opponent's location and actions, and execute offensive and defensive tactical military actions to defeat the other player. Strategy games like Starcraft II also pose an additional challenge to AI systems because players can choose between multiple playable factions which each have different capabilities and perform those core functions in very

---

[36] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert et al., "Mastering the Game of Go Without Human Knowledge," *Nature* 550 (October 2017): 354-359, https://doi.org/10.1038/nature24270.

[37] Ibid., 1.

[38] Kyle Wiggers, "DeepMind's MuZero Teaches Itself How to Win at Atari, Chess, Shogi, and Go," *VentureBeat*, The Machine, last modified November 20, 2019, https://venturebeat.com/2019/11/20/deepminds-muzero-teaches-itself-how-to-win-at-atari-chess-shogi-and-go.

[39] Garisto, "Google AI Beats Top Human Players at Strategy Game StarCraft II."

different ways. This means that each game, and even each second within each game, represents a completely unique asymmetric situation that the AI has to respond to with approximately 10^{26} possible actions for the AI to take in that fraction of a second. Despite the complexity of this environment, in 2019 the AI Alpha Star was able to learn the game in 44 days and then gain grand master ranking on European servers, performing better than 99.8% of the approximately 90,000 players. It is also important to note that these victories came while the system was artificially limited to simulate human capabilities by setting limits such as the maximum number of inputs it could place per minute and not allowing it to make multiple inputs simultaneously.[40] The speed at which AI gains mastery level understanding of high complexity games has implications for AI to participate in military planning, and its ability to monitor millions of inputs and factors in real time to determine objectives has implications for application in command and control.

Computer Vision

Computer vision refers to AI that can analyze photos or videos in order to identify specific content within them. While computer vision has existed as a field for over 20 years, the most significant advances have been made since 2014 with the introduction of deep learning algorithms.[41] Computer vision now has a wide range of real-world

---

[40] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi et al., "Grandmaster Level in StarCraft II Using Multi-agent Reinforcement Learning," *Nature* 575 (October, 2019): 350–354, https://doi.org/10.1038/s41586-019-1724-z.

[41] Gaudenz Boesch, "Object Detection in 2022: The Definitive Guide," Viso.ai, Deep Learning, accessed April 05, 2022, https://viso.ai/deep-learning/object-detection/.

applications including conducting visual inspections of equipment or component quality control, cancer detection, crop monitoring, traffic analysis, autonomous driving vehicles, weapon detection, facial recognition, and even lip reading.[42] Facial recognition in particular has gained significant attention within security and law enforcement communities. According to the U.S. Government Accountability Office, 20 of 42 US federal agencies surveyed owned or used facial recognition systems for law enforcement purposes in 2021.[43]

Advances in computer vision are still limited primarily by the need for supervised learning. For a deep neural net to be effective at identifying something it needs to be trained on hundreds of thousands or millions of images that have been manually labeled or tagged by a human. Unsupervised AI learning is also being developed but does not currently produce similar results.[44] In the narrow fields that AI are trained for recognition however, they achieve superhuman level results. For example, AI are better than humans at determining a person's sexual orientation[45] and reading lips (47.7% error rate for

[42] Vidushi Meel, "87 Most Popular Computer Vision Applications in 2022," Viso.ai, Applications, accessed 05 April 2022, https://viso.ai/applications/computer-vision-applications/.

[43] U.S. Government Accountability Office (GAO), GAO-21-518, *Facial Recongtion Technology: Federal Law Enforcement Agencies Should Better Assess Privacy and Other Risks,* Report to Congressional Requesters (Washington, DC: GAO, June 2021), https://www.gao.gov/products/gao-21-518/.

[44] Boesch, "Object Detection in 2022."

[45] Michael Kosinski and Yilun Wang, "Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images," *Journal of Personality and Social Psychology* 114, no. 2 (February 2018): 246-257.

humans compared to a 11.4% error rate for AI).[46] With continued development computer vision technology is likely to find military applications, especially in reconnaissance and information collection for intelligence.

Deepfakes

Deepfakes are media content created by deep learning neural network AI to generate fake representations of real people. Deep fakes gained public awareness in 2018 when a fake video of former president Barrack Obama was released. The video demonstrated how, given enough open-source video and voice data on a subject, AI could replicate a person's image, facial mannerisms, voice, and speech patterns to generate a video of them saying anything the creator wished.[47]

There are several methods of generating deepfakes that generally fall into four categories: reenactment, replacement, editing, and synthesis. Reenactment is like the president Obama example where AI studies material of a given target in order to match their expression, mouth, gaze, pose, and or body and then produce a reliable puppet that can be used for generating media.[48] These puppet-master deepfakes can even be manipulated in real time by an AI that follows the facial movements, expressions, and speech of another person acting out what is to be mimicked to a camera.[49] Replacement is

---

[46] Assael et al., "LipNet," 6.

[47] Yisroel Mirsky and Wenke Lee, "The Creation and Detection of Deepfakes: A Survey," *ACM Computing Surveys* 54, no. 1, article 7 (January 2022): 2.

[48] Ibid., 4.

[49] Thanh Thi Nguyen, Quoc Viet Hung Nguyen, Dung Tien Dguyen, Duc Thanh Nguyen, Thien Huynh-The, Saeid Nahavandi, Thanh Tam Nguyen, Quoc-Viet Pham, and Cuong M. Nguyen, "Deep Learning for Deepfakes Creation and Detection: A Survey,"

a method that mimics a small portion of the target, generally the face or mouth, and transfers it onto another video. Popular face swapping apps could fall into this category of deepfake, but their more notable application is placing a victim's face into pornographic material to humiliate, defame, or blackmail them. Editing refers to altering attributes of authentic media and are generally cosmetic in nature such as changing clothes, hair, body type, ethnicity, etc. Finally, synthesis refers to content that has no original source where AI creates entirely fake individuals through combining attributes of thousands of different people.[50]

This thesis is primarily concerned with the implications of reenactment deepfakes because of their real time applications. In 2019 scammers used a voice deepfake to mimic the voice of a CEO in real time during a call with a subsidiary company chief executive. Using the deepfake the scammers directed the executive to transfer $243,000 to the account of a Hungarian supplier. The tricked executive later commented on how he never suspected anything because he recognized his boss's subtle German accent and normal speech "melody."[51] Deepfakes are notable specifically because of how believable they are. A study in 2021 found that people are no better than chance at determining their

---

Cornell University, last modified February 06, 2022, https://doi.org/10.48550/arXiv.1909.11573/.

[50] Mirsky and Lee, "The Creation and Detection of Deepfakes," 4.

[51] Jesse Damiani, "A Voice Deepfake Was Used to Scam a CEO out of $243,000," *Forbes,* September 03, 2019, https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/?sh=15d8872d2241/.

authenticity for most deepfake videos.[52] While significant research has been done in the field to develop means of detecting deepfakes, progress is mostly in a forensics capacity with no clear path towards developing a means or method to detect a deep fake in real time.[53]

The continued development of deep fakes could have serious implications for the future, including in military applications. In 2022 Russia released a deep fake showing the president of Ukraine Volodymyr Zelenskyy asking his citizens to lay down their arms and surrender.[54] While this particular deepfake was of low quality and is unlikely to have made much impact, it signals the future potential of the technology as a means to influence political narrative or even military actions.

Predictive Analytics

Analyzing extremely large amounts of data is what AI does best, and through the use of Bayesian nets and Markov models AI can identify trends and make predictions with varying degrees of confidence. In 2011 data scientist Kalev Leetaru demonstrated how AI could be used to predict real world events through analysis of open-source global news archives. His AI trained on thirty years' worth of global reporting from 1979 to

---

[52] Nils C. Köbis, Barbora Doležalová, and Ivan Soraperra, "Fooled Twice: People Cannot Detect Deepfakes but Think They Can," *iScience* 24, no. (2021): 9.

[53] Luisa Verdoliva, "Media Forensics and Deepfakes: An Overview," *IEEE Journal of Selected Topics in Signal Processing* 14, no. 5 (2020): 10.

[54] Matthew Holroyd and Fola Olorunselu, "Deepfake Zelenskyy Surrender Video is the 'First Intentionally Used' in Ukraine War," *Euro-news*, last modified March 16, 2022, https://www.euronews.com/my-europe/2022/03/16/deepfake-zelenskyy-surrender-video-is-the-first-intentionally-used-in-ukraine-war/.

2010 from a large variety of aggregated news sources. Kalev then used the AI to look at specific trends to determine if the data could make useful inferences. Specifically, the AI was able to identify trends that predicted unrest in Egypt, Tunisia, and Libya ahead of the Arab spring as well as conflict in Serbia. Even more specifically the AI was used to identify the most likely location of Osama Bin Laden and selected a 200 km radius area in northern Pakistan as the most likely. While not a small area for searching, the predicted area did in fact include the compound where the terrorist leader was eventually found.[55]

The Pentagon is looking to tap into the potential of predictive analytics through a program called Global Information Dominance Experiments (GIDE). NORTHCOM conducted the third test of GIDE in 2021 including representatives from all 11 combatant commands. GIDE takes in data from sensors all around the world and looks for trends that are indicators for geopolitical events of interest. These trends then cue the use of limited resource collection assets to make higher fidelity evaluations to confirm or deny expectations. According to General VanHerck—the commander of NORTHCOM and NORAD—the analysis provided by GIDE is like a small window into the future:

> What we've seen is the ability to get way further what I call left, left of being reactive to actually being proactive. And I'm talking not minutes and hours, I'm talking days. The ability to see days in advance creates decision space. Decision space for me as an operational commander to potentially posture forces to create deterrence options to provide that to the secretary or even the president.[56]

---

[55] Kalev H. Leetaru, "Culturomics 2.0: Forecasting Large-Scale Human Behavior Using Global News Media Tone in Time and Space," *First Monday* 16, no. 9 (September 5, 2011), https://journals.uic.edu/ojs/index.php/fm/article/download/3663/3040/.

[56] Glen D. VanHerck, "NORTHCOM Commander Gen. Glen D. VanHerck Conducts Press Briefing on North American Aerospace Defense Command and U.S.

Analysis of the past and present is the foundation of all analysis of the future. AI assisted data aggregation and predictive analysis has great potential for assisting planners at both the operational and strategic levels.

Autonomous Drones and Weapons

Autonomous weapon systems are a rapidly growing topic within the literature, especially within defense communities. Interestingly, however, a large part—if not the majority—of papers on the topic seem to address the ethical implications of the technology rather than the operational implications and impacts. A couple notable books with a holistic approach to the topic are *The Kill Chain* by Christian Brose and *Army of None* by Paul Scharre. Both authors, and the literature at large, highlight the large difference between human in the loop and human out of the loop decisions for the use of lethal force. Until 2022 all drones were using humans in the loop, meaning that regardless of the degree of automation on board the drone, there was still a person involved somewhere that had to authorize and employ actual weapons. However, in 2022 there were reports of a drone conducting a fully autonomous targeting and lethal engagement in Libya, effectively cutting humans out of the loop for both deciding if the target was legitimate or if lethal force should be used.[57]

Northern Command Global Information Dominance Experiments," (Transcript, U.S. Department of Defense, July 28, 2021), https://www.defense.gov/News/Transcripts/Transcript/Article/2711594/northcom-commander-gen-glen-d-vanherck-conducts-press-briefing-on-north-america/.

[57] Ed Nash, "We May Have the First Case of a Robot Deliberately Killing Humans," Military Matters, June 1, 2021, https://militarymatters.online/defense-news/we-may-have-the-first-case-of-a-robot-deliberately-killing-humans/.

Increased automation when it comes to military capabilities has major advantages that make them extremely appealing. In 2021 the U.S. Air Force experimented with AI flying fighter jets. In the culminating exercise, their seasoned F-16 pilot failed to survive a single skirmish against the Falco AI by Heron Systems. Despite these results the Air Force stated that they do not have any plans to move to human out of the loop piloting of planes but want to find ways for the human operator and AI to work together to achieve greater results.[58] This is the approach taken more or less universally within the U.S. military branches. Other technologies such as Patriot and Aegis air defense systems operate with a high degree of autonomy but always with a human operator constantly monitoring them. This is because autonomous weapons pose a significant risk. Glitches, unexpected scenarios, and good old fashioned mechanical failure can result in autonomous weapons becoming "run-away guns" that potentially result in fratricide or international incidents.[59] However, other actors have already demonstrated their willingness to employ fully autonomous lethal systems despite these concerns, and their continued proliferation is likely to have significant impacts on how militaries around the world choose to invest in and trust fully autonomous systems.

NAI in Sum

AI has already demonstrated its ability to perform at above human level capability in many specific areas and has the potential to be highly disruptive and transformative in

---

[58] Sue Halpern, "The Rise of A.I. Fighter Pilots," *The New Yorker*, January 17, 2022, https://www.newyorker.com/magazine/2022/01/24/the-rise-of-ai-fighter-pilots/.

[59] Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: Norton, 2019), 190-194.

its own right. Each of the reviewed current AI capabilities have implications for the future of military and information operations. They also provide a foundation for what can be expected capability wise from an ASI. However, because of the brittle nature of these narrow applications NAI is not going to suddenly wake up and be able to think like humans, yet alone possess intelligence greater than humans across all domains. Getting to super intelligence via AI requires the ability to apply knowledge generally and understand the world holistically instead of narrowly.

NAI to ASI: The Path and Probability

> The existence of several examples of intelligence designed under these constraints [of nature] should give us great confidence that we can achieve the same in short order. The situation is analogous to the history of heavier than air flight, where birds, bats, and insects clearly demonstrated the possibility before our culture mastered it.[60]

The Neural and Bayesian networks of NAI can produce intelligence of a form but are not capable of replicating intelligence on par with a human, yet alone surpassing it. Transitioning from the current state of AI to an ASI requires an extremely important intermittent step, an AI that can apply knowledge generally across contexts and domains of thought to deduce and infer solutions to novel problems. This concept is referred to as Artificial General Intelligence (AGI), which is synonymous with terms such as hard AI, true AI, and Human Like Machine Learning.

> What is meant by [AGI] is, loosely speaking, AI systems that possess a reasonable degree of self-understanding and autonomous self-control, and have the ability to solve a variety of complex problems in a variety of contexts, and to

[60] Hans P. Moravec, "The Role of Raw Power in Intelligence," (unpublished manuscript, Carnegie Mellon University, May 12, 1976), https://frc.ri.cmu.edu/~hpm/project.archive/general.articles/1975/Raw.Power.html/.

> learn to solve new problems that they didn't know about at the time of their creation.[61]

In 2016 Vincent Muller and Nick Bostrom conducted a survey of 170 experts in various fields of AI to determine opinions on the likelihood of AGI and ASI. The results showed a huge disparity in opinions on both when AGI could be expected to arrive and the likelihood of that event. Despite this large incongruity, however, there was large group consensus that once AGI is achieved it will result in ASI in thirty years or less with a very high degree of confidence.[62] This confidence is echoed in the literature at large, which suggests AGI invariably leads to ASI because it will have both the benefits of human like cognition and Bostrom's advantages of digital computing. If human intelligence is capable of creating AGI, then it must be assumed that an AGI will be capable of learning to do the same, beginning the recursive process of self-improvement that leads to ASI and intelligence explosion.[63]

If AGI inevitably leads to ASI, then an examination of the literature to determine the likelihood of AGI is the most critical factor to determine the likelihood of ASI. AGI has been predicted to be about twenty years away since the invention of computers in the

---

[61] Ben Goetzel and Cassio Pennachin, eds., *Artificial General Intelligence* (Springer: Rockville, AGIRI, 1998), VI.

[62] Vincent C. Muller and Nick Bostrom, "Future Progress in Artificial Intelligence: A Survey of Expert Opinion," in *Fundamental Issues of Artificial Intelligence,* ed. Vincent C. Muller (Berlin: Synthese Library, 2016), 553-571.

[63] Kurzweil, *The Singularity is Near*, 40.

1940s, and have continually shifted to remain about twenty years away ever since then.[64] An analysis by Ajeya Cotra—a senior research analyst at Open Philanthropy—in 2012 predicted a median of 2052 for when an actor would be both willing and able to train a neural net with computational power similar to the human brain and result in AGI.[65] Muller and Bostrom's survey of experts in 2016 showed a predominant view within the community that AGI will likely emerge between 2040-2050 and will almost certainly arrive—with a 90% confidence rate—by 2075. A strategic research project conducted by the Army War College in 2020 to look at converging technologies estimated that advances in human brain inspired computing chip designs and unsupervised AI learning algorithms would result in AGI by 2030.[66] It should also be noted that some experts in the field believe AGI will never be achieved. Ragnar Fjelland is prominent in the literature as a skeptic basing his arguments around the concept that tacit and experiential knowledge are required to reach human like intelligence, and that these types of knowledge are beyond AI algorithms.[67]

---

[64] Stuart Armstrong and Kaj Sotala, "How We're Predicting AI-or Failing To," in *Beyond AI: Artificial Dreams*, ed. Jan Romportl, Pavel Ircing, Eva Zackova, Michal Polak, and Radek Schuster (Pilson: University of West Bohemia, 2012), 52-75.

[65] Ajeya Cotra and Rohin Shah, "Draft Report on AI Timelines," Alignment Newsletter, 2020, https://mailchi.mp/41774b61e5f8/an-121forecasting-transformative-ai-timelines-using-biological-anchors.

[66] Delcour, Nicholas, Louis Duncan, Stephen Frahm, Patrick Lancaster, and Lance Vann, "Estimation of Technology Convergence by 2035," (Mad Scientist Fellows Strategic Research Project, U.S. Army War College, 2020), 67, https://csl.armywarcollege.edu/USACSL/Publications/EstimationOfTechConvergence-USAWC.pdf/.

[67] Ragnar Fjelland, "Why General Artificial Intelligence Will Not Be Realized," *Humanities and Social Sciences Communications* 7 (2020): article 10.

The lack of consensus and ever-changing timeframe of predictions should also be viewed through the lens that AI experts have historically been extremely bad at predicting the rate of advancement within their own field.[68] This can largely be explained by a belief that AI progress will continue at a steady pace. Many applications of NAI do follow a trend of steady development due to their relatively short project periods intended to produce specific results, often for a commercial purpose. However, AGI advancement has been characterized by three factors which contradict this approach to predicting progress. First, research relevant to advancing towards AGI is often on a much larger timescale and can result in seemingly sudden, major advancements. As an example, Cyc's common sense assertions framework took thirty-five years to complete but went completely unnoticed until its release in 2021, revolutionizing the field seemingly overnight.[69] Second, the steady growth model also does not account for unexpected and unintended developments that arise during the course of non-AGI focused research. For example, research shows implications that AI is starting to mimic the human brain of its own accord and "is undergoing its own convergent evolution with nature—without anyone programming it to do so."[70] Another study released in November of 2021 shows that current AI models have developed which resemble the cognitive functioning of the

[68] Bostrom, *Superintelligence*, 24.

[69] Douglas Lenat, "Douglas Lenet: Cyc and the Quest to Solve Common Sense Reasoning in AI," interview by Lex Fridman, *Lex Fridman Podcast*, September 2021, video, 2:52:56, https://www.youtube.com/watch?v=3wMKoSRbGVs.

[70] Eric James Beyer, "MIT Researchers Just Discovered an AI Mimicking the Brain on Its Own," Interesting Engineering, December 18, 2021, https://interestingengineering.com/ai-mimicking-the-brain-on-its-own.

human brain, without anyone programing it that way.[71] This is especially significant because one of the study's major findings was that the closer a model resembles a human brain structure the closer it matches human abilities. Progress in AI is giving us new insight into how intelligence can manifest via evolutionary processes. [72]

Finally, the conversation about our ability to achieve AGI is dominated by the unstated assumption that humans must solve the problem. However, it has been proposed that AGI could be developed by narrow AI specifically designed for the task. AI that passes high level benchmarks in the domains of mathematics or computer science could be put to the task and find methods that humans never could.[73] Recent advances in these areas lend additional credence to this possibility with AI discovering new patterns in pure mathematics,[74] and AI that can read plain text to determine a desired output and write code to achieve it.[75]

---

[71] Martin Schrimpf, Idan Asher Blank, Greta Tuckute, Carina Kauf, Eghbal A. Hosseini, Nancy Kanwisher, Joshua B. Tenenbaum, and Evelina Fedorenko, "The Neural Architecture of Language: Integrative Modeling Converges on Predictive Processing," *Proceedings of the National Academy of Sciences* 118, no. 45 (November 2021): e2105646118, https://doi.org/10.1073/pnas.2105646118.

[72] Beyer, "AI Mimicking the Brain on Its Own."

[73] Bostrom, *Superintelligence,* 35.

[74] Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburnet et al., "Advancing Mathematics by Guiding Human Intuition with AI," *Nature* 600 (December 2021): 70–74, https://doi.org/10.1038/s41586-021-04086-x.

[75] Cade Metz,. "A.I. Can Now Write Its Own Computer Code. That's Good News for Humans," *The New York Times,* September 09, 2021, https://www.nytimes.com/2021/09/09/technology/codex-artificial-intelligence-coding.html.

Ultimately, the first AGI could arrive as the result of a breakthrough in hardware or software that replicates the core brain processes necessary for generalized intelligence, or as a result of converging AI technologies. A review of the literature on the requirements of AGI and the state of progress associated with each is required to gain a comprehensive perspective on the likelihood of AGI in the near term.

AGI Requirements

AGI must be able to replicate human intelligence. In order to do so the literature suggests there are four specific capability requirements that must come together to form a cohesive whole: common sense, natural language understanding, cross domain knowledge application, and the ability to solve novel complex problems. This section will review the literature on hardware requirements for AGI, the four capability requirements, and the method by which researchers will test to determine if an AI has advanced to AGI.

Hardware

The literature shows a large range in estimations on the amount of computing power that will be required to host and run an AGI. On the low end, using the human brain as a model, the Moravec estimate is that 100 TFlops of processing power would be required.[76] This benchmark was met in 2005 by IBM's Blue Gene/L supercomputer which sported 280 TFlops, over three times as much as the second fastest computer at the

[76] Joseph Carlsmith, "How Much Computational Power Does It Take to Match the Human Brain?" Open Philanthropy, September 11, 2020, https://www.openphilanthropy.org/brain-computation-report#Conclusion/.

time.[77] As of November of 2021, the world's fastest supercomputer—the Fugaku from Japan—has 442,010 TFlops of processing power. The fastest US computer—Summit by IBM—has 148,600 TFlops.[78]

This seems like a drastic increase in computational power in a short period of time, and it is, but it is also exactly on course with Moore's law that states computational power will double approximately every 1.5 years. Holding to Moore's law, computational power in super computers will meet the requirements of even the highest estimates in the literature—of roughly a billion TFlops—before 2040. The most detailed prediction of hardware requirements comes from Joseph Carlsmith who conducted an expansive meta-analysis on the topic in 2020. Carlsmith assigns a 75% probability that the requirements will be in the 1,000-100,000 TFlops range,[79] which is a large range but significantly lower than the high-end prediction of a billion TFlops.

For comparison, the Xbox Series X—a video game console that an average American might have in their living room—has 12 TFlops.[80] Following Moore's law, by 2035 a gaming console available for purchase at a local retail store could be expected to have over 6,000 TFlops, potentially enough computational power to run a simulation of

[77] "November 2005," TOP500, accessed December 10, 2021, https://www.top500.org/lists/top500/2005/11/.

[78] Ibid.

[79] Carlsmith, "How Much Computational Power."

[80] Jon Martindale, "What Is a Teraflop?," DigitalTrends, June 14, 2021, https://www.digitaltrends.com/computing/what-is-a-teraflop/.

the human brain or host an AGI. The bottom line is that hardware processing power is not expected to be a limiting factor for AGI.

Common Sense

Humans understand the world through the understanding of a huge number of rules that we often take for granted: things like two objects cannot occupy the same space at the same time, that time only goes in one direction, that water is wet, or that you cannot eat a rock. AI systems do not inherently know anything about the world and need to be either specifically taught those things or have a method by which to learn them. This lack of generally assumed knowledge is why AI systems so often seem to make obvious or silly mistakes. For example, the OpenAI GPT-3 released in May of 2020 received a lot of press and attention for being a large step forward towards AGI because of its capacity to understand inputs and reason responses. But the system notably lacks common sense. For example, when researchers asked it the best way to move a table into another room for a dinner party if it is bigger than the door, GPT-3 told them to remove the door and then saw the table in half.[81] The AI does not understand that doors can be opened or that sawing the table in half would fundamentally undermine the point of moving it into the other room. These are mistakes that no human would ever make.

[81] Gary Marcus, and Ernest Davis, "GPT-3, Bloviator: OpenAI's Language Generator Has No Idea What It's Talking About," *MIT Technology Review*, August 22, 2020, https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/.

Researchers have been working on this problem since as far back as 1958 when Professor John McCarthy wrote the first known paper on the subject.[82] Solving this problem has been approached using a variety of methods common throughout AI research including the above mentioned combining of AI systems designed to mimic understanding of specific environmental and social rules, to data mining Wikipedia.[83] The Defense Advanced Research Projects Agency (DARPA) funded a four-year project in 2019 pursuing AI common sense using two approaches: one that learns from experience about intuitive physics, behaviors of intentional actors, and spatial navigation; and another that attempts to build a foundational knowledge base by reading from the web.[84] However, to date, nearly all of these efforts have come far short of the common sense we expect from humans, likely due to a large under estimation of the sheer amount of common foundational knowledge humans take for granted.

The Cyc project, led by Douglas Lenat, is a notable exception to this trend. Cyc recently culminated a 35-year effort in pursuit of AI common sense. Lenat's team

---

[82] John McCarthy, "Programs with Common Sense," (Computer Science Department, Stanford University, Stanford, CA, 1959), http://jmc.stanford.edu/articles/mcc59/mcc59.pdf.

[83] Soren Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives, "DBpedia: A Nucleus for a Web of Open Data," in *The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 2007, Proceedings,* ed. Karl Aberer, Key-Sun Choi, Natasha Noy, Dean Allemang, Kyung-Il Lee, Lyndon Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, Guus Schreiber, and Philippe Cudre-Mauroux (Berlin: Springer, 2007), 722-735.

[84] Howard Shrobe, "Machine Common Sense," Defense Advanced Research Projects Agency, accessed February 1, 2022, https://www.darpa.mil/program/machine-common-sense.

undertook the task of hard coding ontological terms and axioms into an AI starting in 1984. In a recent interview Lenat disclosed that at the onset of the program they believed that roughly a million such statements would be required, but over time discovered that they were off by an order of magnitude. In 2017 the Cyc AI contained a foundation of over 24.5 million common sense assertions.[85]

While that number seems large, it is orders of magnitude smaller than it could be because Cyc teaches their AI foundational knowledge that it can use to make its own deductions. They do not systematically tell the AI that if something is orange it cannot also be red, or yellow, or blue etc. Instead, they provide a simple rule like colors are mutually exclusive and then test to see if the AI can apply that knowledge. Cyc uses 1,000 different heuristic agents that work together to evaluate the huge amount of information it knows to try and solve a problem, employing meta-reasoning to try and determine what the best way to answer a question would be to limit the parts of knowledge it will consider as part of the answering process. This has resulted in an ability for Cyc to utilize abduction to infer things when it gets something incorrect based on its foundation of knowledge, the first step in creating a system that can self-learn and gain general knowledge that can be developed into expert knowledge. Lenat describes a particularly salient moment during the project's development where the AI realized the only entity allowed to alter its code base that was not labeled as a person was itself. Seeking clarification, the AI asked the researchers whether or not it was a person.[86]

---

[85] Lenat, "Douglas Lenet."

[86] Lenat, "Douglas Lenet."

While not an indicator of AI consciousness, the moment does highlight how the system is able to make inquires based on inferences and learn certain aspects of common sense through experience, just like humans. The developers of Cyc are now in the process of testing their common sense AI by combining it with other AI systems to see if its addition results in significant improvements for dealing with real world problems, including natural language understanding.

Natural Language Understanding

One of the most basic human functions is the ability to communicate via oral and written language. AGI needs to understand language on an intuitive level that goes beyond knowing the meaning of words and allows it to understand the intent of the communicator. For example, humans can infer meaning from sentences even when words are completely missing.[87] The earliest implementation of machine speech recognition was in 1952 with Bell Laboratories "Audrey" system which could recognize single digits spoken by a single individual.[88] Today we have systems like Siri and Alexa that can recognize and respond to a huge number of commands with context. But recognizing specific command key words is not language understanding. Knowing what the communicator is attempting to convey with the chosen words is the true goal of natural

[87] Yichen Huang, Yizhe Zhang, Oussama Elachqar, and Yu Cheng, "INSET: Sentence Infilling with INter-SEntential Transformer," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (Stroudsburg, PA: Association for Computational Linguistics, July 2020), 2502–2515, https://doi.org/10.18653/v1/2020.acl-main.226.

[88] Melanie Pinola, "Speech Recognition Through the Decades: How We Ended Up with Siri," *PCWorld*, November 2, 2011, https://www.pcworld.com/article/477914/speech_recognition_through_the_decades_how_we_ended_up_with_siri.html.

language understanding. But just as humans can learn more than one language, AGI must also be capable of becoming multilingual or even omnilingual.

There is an underlying assumption that speech recognition AI would benefit from being multilingual as the data and learning from one language should help the system to better understand other languages as well. However, this has not proven to be the case so far in practice. In fact, AI systems designed to recognize the speech of high resource languages, such as English, actually do worse when incorporated into a multilingual model.[89] Google's translation AI generated a possible solution to this problem in 2014 when it created its own language allowing it to conduct "zero-shot" translations where the system was not taught how to translate between two languages.[90]

Generative Spoken Language Modeling (GSLM) is a new approach that seeks to solve the problem by teaching AI without parsed data. GSLM learns directly from raw audio data that does not have any labels or text associated with it, much like humans do. This both allows the system to learn from a much larger set of data—that has not been manually annotated specifically for that purpose—and incorporates natural aspects of speech such as rhythm, stress, and intonation that are very important for understanding meaning that goes beyond words and communicating in a more natural way with native

---

[89] Bo Li, Ruoming Pang, Tara N. Sainath, Anmol Gulati, Yu Zhang, James Qin, Parisa Haghani, W. Ronny Huang, Min Ma, and Junwen Bai, "Scaling End-to-End Models for Large-Scale Multilingual ASR," Google, USA, 2021, https://arxiv.org/pdf/2104.14830.pdf.

[90] Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat et al., "Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation," in *Transactions of the Association for Computational Linguistics,* vol. 5, ed. Colin Cherry (Stroudsburg, PA: Association for Computational Linguistics, 2017), 349, https://aclanthology.org/Q17-1024.pdf.

language speakers. This is of particular importance for languages that have words with multiple meanings and especially so when word inflection is key such as in Mandarin Chinese.[91]

An AI system that has natural language understanding must be able to do things like pass the Winograd Schema Challenge where a pair of sentences differ in only one word, but that single word creates an ambiguity that completely changes the meaning of the sentence and can only be inferred based on knowledge of the world. Such schema would be so easily solved by humans that they might not even notice an ambiguity at all, but systems that rely upon algorithms or heuristics would likely find them impossible as there is no statistical test over text corpora that would be able to solve the ambiguity. An example of such a sentence would be: "The city councilmen refused the demonstrators a permit because they [feared/advocated] for violence." [92] In this example the AI would need to be capable of understanding who "they" refers to in each of the two variations of the sentence.

The standard for testing natural language processing is currently the SuperGLUE benchmark which consists of 10 tasks including Winograd schemas, reading comprehension with commonsense reasoning, recognizing textual entailment, choice of

[91] Meta AI, "Textless NLP: Generating Expressive Speech from Raw Audio," September 09, 2021, https://ai.facebook.com/blog/textless-nlp-generating-expressive-speech-from-raw-audio.

[92] Ernest Davis, Leora Morgenstern, and Charles Ortiz, "The Winograd Schema Challenge," Computer Science Department at New York University, accessed December 15, 2021, https://cs.nyu.edu/~davise/papers/WinogradSchemas/WS.html.

plausible alternatives, and multi-sentence reading comprehension.[93] In 2021 the Chinese system ERNIE 3.0 achieved new state of the art results across 54 natural language processing tasks in Chinese and secured the number one spot on the SuperGLUE leaderboards with their English version which achieved better than human performance scores. Going beyond understanding, ERNIE 3.0 has also been noted to have impressive abilities to engage in creative writing and produce novels, poems, and couplets.[94]

Progress in natural language processing has made such rapid advancements in the last few years it may soon no longer be considered a hurdle in pursuit of AGI. However, the practical effects of these developments are just starting to be understood. Previously, voice recognition was relegated to simple technologies like taking notes verbally on your phone or asking your Alexa for a weather report. But the technology could swiftly become disruptive as evidenced by Checkers and Rally's fast food restaurant's opting to replace human workers in their drive throughs with voice recognition AI at 267 of their locations.[95]

---

[93] "Frequently Asked Questions," SuperGlue, accessed December 15, 2021, https://super.gluebenchmark.com/faq.

[94] Baidu Research, "ERNIE 3.0 Achieves State-of-the-Art Results in 54 Chinese NLP Tasks, Crowned 1st Place on SuperGLUE Leaderboard," (blog), July 14, 2021, http://research.baidu.com/Blog/index-view?id=160.

[95] Nancy Luna, "Checkers & Rally's is Rolling Out Voice-Ordering Bots to Take Drive-Thru Orders at 267 Restaurants Amid a Crippling Labor Shortage in the Industry," *Insider,* January 10, 2022, https://www.businessinsider.com/checkers-rolls-out-presto-voice-bots-at-drive-thru-lanes-2022-1.

Cross Domain Knowledge Application

In basic terms this requirement is meant to capture that a system with general intelligence needs to be able to think about the world in a way similar to humans. Currently, AI requires huge amounts of data in order to learn to do very small things. AGI must be able to flip that paradigm and complete big tasks with a small amount of data, like humans. One group of researchers proposes that functionality, physics, intent, causality, and utility are the five domains foundational to all knowledge.[96] The combination of applying human levels of capability in these domains would theoretically allow an AI system to learn to do anything a human does. Other researchers include morality and social intelligence as necessary cognitive domains. By focusing on these social domains one research group successfully developed an AI that can intuitively predict group and individual behaviors based on observations of their actions as well as make rapid generalizations and inferences similar to human judgement.[97] The ability to apply knowledge across domains will remove the brittle nature of AI and allow it to associate information it has gained on one topic to others, resulting in novel problem-solving capabilities.

[96] Yixin Zhu, Tao Gao, Lifeng Fan, Siyuan Huang, Mark Edmonds, Hangxin Liu, Feng Gao et al., "Dark, Beyond Deep: A Paradigm Shift to Cognitive AI with Humanlike Common Sense," *Engineering* 6, no. 3 (2020): 310-345.

[97] Michael Shum, Max Kleiman-Weiner, Michael L. Littman, and Joshua B. Tenenbaum, "Theory of Minds: Understanding Behavior in Groups through Inverse Planning," *Proceedings of the AAAI Conference on Artificial Intelligence* 33, no. 01 (July 17, 2019): 6163-6170, https://ojs.aaai.org/index.php/AAAI/article/view/4574.

Solve Novel Complex Problems

> If you understand something in only one way, then you don't really understand it at all. This is because, if something goes wrong, you get stuck with a thought that just sits in your mind with nowhere to go. The secret of what anything means to us depends on how we've connected it to all the other things we know. This is why, when someone learns "by rote," we say that they don't really understand. However, if you have several different representations then, when one approach fails you can try another. . . well-connected representations let you turn ideas around in your mind, to envision things from many perspectives until you find one that works for you. And that's what we mean by thinking![98]

The ability for AI to solve complex problems of its own accord is the most important bar of AGI. Most current AI works through the use of data mining, neural nets, and deep learning, which are all essentially ways of looking at huge amounts of data and finding trends or forming useful associations. For instance, IBM's project debater is able to debate on par with real humans. With just 15 minutes of preparation—the same as its human opponent—the AI can create several minute long arguments as well as counter the information presented by its opponent.[99] While at first glance this sort of achievement may be perceived as problem solving, since the AI is creating a solution to a problem it did not know it was going to face ahead of time; in reality it is still only mining data and presenting information that it already had access too, albeit in an impressive and convincing manner.

What novel problem solving truly entails is the ability to solve problems the AI does not have data for, but can infer possible solutions to, based on what it already

---

[98] Marvin Minsky, *The Society of Mind* (New York: Simon & Schuster, 1988), 64.

[99] Noam Slonim, Yonatan Bilu, Carlos Alzate, Roy Bar-Haim, Ben Bogin, Francesca Bonin, Leshem Choshen et al., "An Autonomous Debating System," *Nature* 591, no. 7850 (March 2021): 379–84, https://doi.org/10.1038/s41586-021-03215-w.

knows. The AI must then also have some means by which to prioritize potential solutions based on a variety of factors like assumed probability of success, number of sub goals, overlapping sub goals with other possible solutions, resource requirements, and time. These hypotheses could then be tested to determine which have merit, and the results of each test should create new information that potentially creates new hypothesis or reorders the prioritization of current testing goals. This is the way that humans solve problems and central to our ability to continually create and invent.[100]

Some progress has been made in this domain. Google's DeepMind is able to generalize knowledge gained into radically new and different circumstances simulated in 3D environments. DeepMind utilizes zero-shot generalization in reinforcement learning agents that are trained on a large set of 3.4 million tasks. In an experiment the agents applied their knowledge across 700,000 games they had never seen before and were found to exhibit behaviors unusual for AI such as experimentation, cooperation with other agents, and changing the state of the world in pursuit of goals. The agents even performed actions that closely resemble the creation and use of tools such as flipping an object to make into a ramp.[101] It is also noteworthy that while the amount of data used for training in this example may seem very high, it is actually small by AI standards that

---

[100] Francesco Donnarumma, Domenico Maisto, and Giovanni Pezzulo, "Problem Solving as Probabilistic Inference with Subgoaling: Explaining Human Successes and Pitfalls in the Tower of Hanoi," *PLOS Computational Biology* 12, no. 4 (April 2016): e1004864, https://doi.org/10.1371/journal.pcbi.1004864.

[101] Adam Stooke, n Anuj Mahajan, Catarina Barros, Charlie Deck, Jakob Bauer, Jakub Sygnowski, Maja Trebacz et al., "Generally Capable Agents Emerge from Open-ended Play," *DeepMind* (blog), July 27, 2021, https://deepmind.com/blog/article/generally-capable-agents-emerge-from-open-ended-play/.

generally are looking at data sets with figures in the billions. This is important to note because in the AI community the golden rule tends to be the more data the better. However, some research has shown that AI performing certain tasks can actually become worse at them when more data is available. After a certain threshold the additional data causes their neural nets to incorporate more and more irrelevant connections—interestingly, a flaw that sometimes also manifests in humans.[102]

Testing For AGI

The Turing test is the original method by which to determine if an AI has reached generalized intelligence. The test is named after the father of computing, Allen Turing, who wrote a paper on the topic in 1950.[103] Some authors also credit the philosopher Descartes whose writings in *Discourse on the Method* set the stage for or predicted a Turing test, although the relevant technology did not exist yet in his day.[104] Both thinkers put forth the idea that intelligence is not a phenomenon bound within the confines of organic matter, but rather an expression of certain behaviors given inputs or stimuli. Turing suggested a simple method to determine if a machine is truly intelligent using an

[102] Stephanie Lin, Jacob Hilton, and Owain Evans, "TruthfulQA: Measuring How Models Mimic Human Falsehoods," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics,* vol. 1, *Long Papers* (Stroudsburg, PA: Association for Computational Linguistics, May 2022), 3214-3252, https://owainevans.github.io/pdfs/truthfulQA_lin_evans.pdf.https://owainevans.github.io/pdfs/truthfulQA_lin_evans.pdf.

[103] Allen Turing, "Computing Machinery and Intelligence," *Mind* 59, no. 236 (October 1950): 433–460, https://academic.oup.com/mind/article/LIX/236/433/986238.

[104]Jack Copeland, *Turing: Pioneer of the Information Age* (United Kingdom: Oxford University Press, 2012), 174.

imitation game. In this game, a human conducts a series of blind interviews, each of approximately five minutes, and then assesses which interview they believe was with a machine. An AI passes this original conception of the Turing test by fooling its interviewers into believing it is human. The belief was that it would be impossible for a machine to conceal its nature without possessing intelligence.[105]

Turing's original conception of the test served as a sufficient barrier to separate intelligence from software for some time. However, in recent years major advances in AI have produced several chat bots capable of passing a basic Turing test in a limited sense. In 2014 the chatbot Eugene Goostman simulated a 13-year-old boy from Ukraine and managed to fool about 33% of participants who interacted with it.[106] However, the assumed age of 13 and the nationality of Ukraine gave interviewers reason to believe English was a second language, which essentially worked as a handicap, giving adult judges a reason to excuse obvious mistakes or non-sensical responses. More recently, Kuki (formerly known as Mitsuku) a chatbot from the Metaverse that simulates an 18-year-old girl, has interacted with more than 25 million people from around the world, many of whom believe she is real.[107] Despite these limited successes it must be noted that these bots are not competing in the true spirit of a Turing test since those interacting with them are not interrogating the AI in a deliberate fashion to evaluate intelligence.

---

[105] Turing, "Computing Machinery and Intelligence," 435.

[106] BBC, "Computer AI Passes Turing Test in 'World First'," BBD News, June 09, 2014, https://www.bbc.com/news/technology-27762088.

[107] Juanita Bawagan, "The Turing Test 2.0," *Physics World*, May 8, 2021, https://physicsworld.com/the-turing-test-2-0/.

Regardless, many thinkers have proposed that the original conception of the Turing test does not go far enough to truly evaluate machine intelligence. Gary Marcus, a cognitive scientist at NYU, has proposed a test where AI are shown a piece of media at random and then asked contextual questions about its content, such as "Why did character A get mad at character B?"[108] Charlie Ortiz believes that the main weakness of the Turing test is that it focuses only on language communication which leaves out important aspects of intelligence like perception and physical interaction with the world. He coined what he calls the Ikea challenge where an AI would be required to build structures based on verbal instructions and answer questions while it works.[109]

Many other proposed alternatives to the Turing test exist, each with their own unique way of attempting to address aspects of the test they believe are unaccounted for. However, the recent COVID-19 pandemic has accidentally presented an ideal test environment for AGI. Groups of humans now regularly coordinate and cooperate via long distance digital platforms to accomplish tasks. A simple and effective Turing test would involve adding an AI to work as a member of a team and perform any kind of task that an individual would be capable of with a modern home computer. The interviewer could ask them to draw them pictures in MS Paint or help him put together a presentation in power point on some topic, explain why a meme is funny, play cooperatively or competitively in a computer game, or perform any number of work duties. If an AI is

---

[108] Gary Marcus, "What Comes After the Turing Test?" *The New Yorker*, June 9, 2014, https://www.newyorker.com/tech/annals-of-technology/what-comes-after-the-turing-test.

[109] George Dvorsky, "8 Possible Alternatives to the Turing Test," *Gizmodo*, April 15, 2015, https://gizmodo.com/8-possible-alternatives-to-the-turing-test-1697983985.

capable of interacting with other humans via online collaboration platforms like slack or MS Teams in the same way that much of the US workforce has done since the start of the pandemic, and to such a degree that those interacting with it cannot tell it is an AI, then how could we deny that it has obtained human like intelligence?

AGI in Sum

It is important to note that AGI is still theoretical despite the large consensus on its feasibility. Achieving the first AGI will be an enormous breakthrough with significant implications. Each aspect of reaching AGI is difficult and the scope of the challenge should not be under sold. However, enormous strides have been made in each area in the last ten years and the pace of development and investment is only increasing. Indeed, the advances made in AI in the last ten years have exceeded what many thought would ever be possible and have closed the gap to such a degree that some of the requirements are close to being considered solved. It is my conclusion, based on the available evidence, that AGI is an inevitability within a medium time frame and plausible in the near term. Since an AGI would have the capability to increase its own intelligence through continuous learning and quickly allow it to gain superhuman ability in any domain it pursued due to its advantages of digital intelligence, any AGI would be capable of reaching ASI status by simply improving the knowledge and capabilities it would already possess. Ultimately, the question of ASI is not if but *when*. Because this thesis is an exploration of risk and vulnerabilities posed by ASI it will utilize predictions on the earlier side of estimates. This thesis will assume AGI arrival by 2035 that then advances to have ASI capabilities—short of intelligence explosion and singularity—by 2040.

The Problem of Control

Should AGI or ASI be realized there is one final hurdle which is of significant concern: the ability to control an intelligence that could grow to be greater than our own. While it is entirely feasible that an AGI/ASI could be manifested that is only responsive to external motivations and lacks any internal drives, the opposite must be planned for. There is significant literature on this topic because the perceived risk of getting it wrong generally leads the imagination to Armageddon scenarios.

All the seminal authors on super intelligence recognize the problem of control and address possible methods which generally fall into two categories: limitation of capabilities and motivation selection. Placing limits on capabilities includes options such as placing the ASI in a box, physical or digital, from which it has no means of interacting with the outside world except through its controlling agent. AGI could also be stunted to reduce its intellectual abilities and prevent it from achieving ASI or resulting in an intelligence explosion. Digital tripwires could also be implanted within the AI to monitor for harmful behaviors and automatically shut it off should certain conditions be met. All methods of control that rely on limitations of capabilities have an obvious downside in that they are somewhat self-defeating. Why create an AGI in the first place just to intentionally make it dumb through stunting or denying it access to information?[110]

For this reason, control methods centered around guiding AI motivation are more valuable. Options include hard coding either a set of motivating system or underlying set of values/principles, limiting the scope of possible motivations the AI can possess, and

[110] Bostrom, *Superintelligence*, 157-176.

starting with an AI that is specifically trained and vetted on human ethics prior to advancing it to AGI/ASI.[111] Because of the uncertainty associated with this part of the problem the method of control is left as one of the key assumptions of this thesis. The analysis in chapter four will assume that control is possible either because AGI is inherently non-motivated or that a motivation guidance method proves capable of allowing a controlling agent to employ an ASI as they see fit.

Forecasting Scenario Methodology

Having established that ASI is feasible in the near term, a method for analyzing its capability and potential is needed. This thesis will assess how an actor adversarial to the United States could utilize ASI against it, so a methodology for exploring future scenarios is utilized. The use of scenarios to aid in strategic planning has exploded in the last two decades, especially in the business sector. Scenarios can be used to explore elements of the present and future to identify limits in current knowledge, a means of communicating information by exploring potential outcomes, as a tool for evaluating the potential effectiveness of organizational strategies, and as a means of aiding decision makers evaluating courses of action.[112]

However, a review of the literature demonstrates that there are no seminal works on the topic and that scenario methodologies are often created to be specifically tailored to the topic under consideration. For this reason, the majority of methodologies reviewed

---

[111] Bostrom, *Superintelligence*, 157-176.

[112] Hannah Kosow and Robert Gaβner, *Methods of Future and Scenario Analysis: Overview, Assessment, and Selection Criteria* (Bonn: German Development Institute, 2008), 1.

were immediately dismissed from consideration for use in this thesis. Methodologies created as a result of meta-analysis of the field proved to be the most valuable within the literature. Jonas Iversen's "Futures Thinking Methodologies"[113] and Hannah Kosow and Robert Gaβner's *Methods of Future and Scenario Analysis* were the two sources—both based on meta-analysis—primarily considered for use. Kosow and Gaβner's method was ultimately selected due to the extensive context and support material provided to assist in the use of specific techniques within the larger methodological framework. While additional information on the specifics of the scenario methodology used in this thesis are outlined in chapter three, the remainder of this literature review covers the information used as part of this scenario generation process.

NIC Global Trends 2040

Every four years the National Intelligence Council publishes a report on their analysis of global trends and how they expect them to influence events over the next twenty years. The intent of the report is to provide U.S. policy makers with information useful for the crafting of national security strategy. The report pulls from experts across disciplines to formulate a comprehensive analysis of trends that are likely to influence the coming decades. This makes the report a perfect source for determining the expected operational environment in 2040, the year this thesis assumes for the advent of ASI. Below is a brief synopsis of the critical elements of the report.

[113] Jonas Svava Iversen, "Futures Thinking Methodologies–Options Relevant for Schooling for Tomorrow," (Organisation for Economic and Co-operation Development, Paris, France, 2005), https://www.oecd.org/education/ceri/35393902.pdf.

The report specifically highlights five key trends it expects to define the 2040 time period: global challenges, fragmentation, disequilibrium, contestation, and adaptation. Global challenges refer to issues that will be shared by all members of the international community such as climate change, disease, and economic recession; each are expected to occur with greater frequency. Fragmentation refers to a continued increase in division along ideological lines in all communities from the local level all the way up to the international level. Disequilibrium is a result of global challenges and fragmentation creating issues at the state and international level at a rate too fast for existing organizations to cope with and provide effective or timely solutions. Contestation refers to societies and states increasing their competition with others due to a rise in tensions, issues, and ideological divides resulting in a more volatile political environment at all levels. Finally, adaptation refers to an increase in the rate at which societies will need to change in order to cope with the challenges they face; an inability to adapt quickly will likely exacerbate inequalities.[114]

Based on these five key trends the 2040 report generated five future scenarios. Two of these scenarios are radical departures from the current norm and involve the breakdown of current world orders due to unresolved global challenges resulting in either an end to international cooperation and globalization or world-wide revolutionary uprisings. The other three scenarios are more closely aligned with current trends and include what would be considered a best case, worst case, and center of the road prediction. In each of these three scenarios the international challenges increase in

[114] Office of the Director of National Intelligence, *Global Trends 2040: A More Contested World* (Washington, DC: National Intelligence Council, March 2021), 3.

severity over time and international affairs are largely defined by a U.S. and China rivalry. This thesis will utilize the "Competitive Coexistence" scenario—the middle of the road option—as the description of the operational environment. In this scenario the U.S. and China decide to prioritize economic growth to hedge against global challenges. While this relationship is competitive in nature, the major powers and those aligned with them are willing to utilize military power to prevent small conflicts from escalating in an attempt to maintain economic stability. Economic stability serves to lower risk of the major powers engaging directly in armed conflict, but certain regions, such as the South China Sea, continue to pose a threat.[115]

Current State of AI Competition

A review of literature on AI competition between the US and China suggests a rough equivalency in applied capabilities with the U.S. being the leader in innovation. However, the Chinese government places a much higher level of emphasis on AI development as a national priority, seeking to become the center of innovation in AI by 2030.[116] In 2018, 26.5% of the top 10% most cited papers on AI were authored in China, compared to 29% in the United States. China and the European Union will likely both

[115] Ibid., 114-115.

[116] Elsa B. Kania, "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power," (Center for a New American Security, Washington, DC, 2017), http://www.indexfunds.org/resources/Research-Materials/NatSec/CNAS_Battlefield_Singularitypdf.pdf, 4.

overtake the United States in total AI research papers published in 2022.[117] As of 2018, the United States and the European Union (including the United Kingdom pre-Brexit) were tied for the largest pool of AI talent at about 29,000 AI scientists, followed by China at 18,200. India is the only other notable player in this category with over 17,000 AI scientists.[118]

However, unlike the United States and its allies, China is heavily incentivizing the creation of new talent through a variety of programs.[119] Finally, while China currently exceeds the United States in super computer capabilities that could be used to pioneer aspects of AI development in pursuit of AGI, companies based in the United States currently dominate the chip manufacturing used for AI systems, representing 95.6% of the market and have complete control of the newly developed TPUs created specifically for AI.[120] The bottom line is that the United States, European Union, and China have a rough equilibrium in the AI race with the United States and its allies having a slight edge. However, if current trends continue China could realize its goal of being the dominant AI innovator by 2030, the same year as the earliest predictions for the first AGI.

---

[117] Sarah O'Meara, "Will China Lead the World in AI by 2030?" *Nature* 572 (August 2019), https://news.sisuer.cn/wp-content/uploads/2020/06/Will-China-lead-the-world-in-AI-by-2030.pdf.

[118] China Institute for Science and Technology Policy, *China AI Development Report* (Haidian, Beijing, China: Tsinghua University, July 2018), 38, https://ediscIplinas.usp.br/pluginfile.php/4873100/mod_folder/content/0/China_AI%20report_2018.pdf?forcedownload=1.

[119] O'Meara, "Will China Lead the World in AI by 2030?".

[120] O'Meara, "Will China Lead the World in AI by 2030?"

Fictional Intelligence

This thesis will produce a short piece of fictional intelligence as a part of its analysis to aid in demonstrating how ASI capabilities could be employed. August Cole, co-author of the book *Ghost Fleet* popularized the term Fictional Intelligence, or FICINT, as fiction writing about the future grounded in reality for national security professionals. "The approach helps both to raise self-awareness and challenge one's own assumptions while articulating complex concepts using tried and true writing techniques that emphasis tension, conflict, and clarity."[121] Dr. Jacob Parakilas points out that life has a tendency to imitate art, especially so for military affairs where only a small portion of a nation's population tends to have direct military or defense experience. Policy makers, along with the voting public, can have their ideas about warfare, or more importantly the future of warfare, heavily influenced by dramatic portrayals.[122]

History shows this to be true; in 1925 Hector Bywater published a book that described a potential American island-hopping strategy. His work prompted a re-write of War Plan ORANGE, the contingency plan for a conflict with the Japanese that was used in World War II as the basis for the entire Pacific campaign.[123] In 1871 Colonel Sir

---

[121] August Cole, "'FICINT': Envisioning Future War Through Fiction & Intelligence (Indo-Pacific Series)," War Room - U.S. Army. War College, May 22, 2019, https://warroom.armywarcollege.edu/special-series/indo-pacific-region/ficint-envisioning-future-war-through-fiction-intelligence-indo-pacific-series/

[122] Jacob Parakilas, "Fiction and Consequences: War in Art and the Art of War," *The Diplomat,* January 27, 2021, https://thediplomat.com/2021/01/fiction-and-consequences-war-in-art-and-the-art-of-war/.

[123] Richard J. Norton, "Through a Mirror Darkly: The Face of Future War, 1871-2005," *Naval War College Review* 62, no. 1 (2009): 131.

George Tomkyns Chesney published *The Battle of Dorking: Reminiscences of a Volunteer* in Blackwood's Magazine. It was a fictional story describing the events of a German invasion which decimates the British royal navy using newly developed weaponry and follows up with a land invasion, decisively defeating the ill prepared British Army.[124] Chesney was motivated to write the piece because of his belief that, "if serious military reform was not undertaken and the Germans ever got across the channel, England was doomed."[125] But the true value of Chesney's work was demonstrating how a fictional narrative drives engagement with a set of ideas. His story attracted so much attention that the Prime Minister of his time officially addressed it when discussing his opposition to increases in defense spending.[126]

The U.S. Naval Institute[127] and West Point Modern War Institute[128] now host recurring FICINT initiatives. Even the former NATO Supreme Allied Commander Europe, Admiral James Stavridis, co-authored a piece of FICINT, signifying the value

[124] George Chesney, *The Battle of Dorking* (Blackmask Online, 2001), http://public-library.uk/ebooks/29/91.pdf.

[125] Norton, "Through a Mirror Darkly," 126.

[126] Michael Matin, "Scrutinizing 'The Battle of Dorking': The Royal United Service Institution and the mid-Victorian Invasion Controversy," *Victorian Literature and Culture* 39, no. 2 (2011): 390.

[127] Hal Wilson, "Letter of Marque," *Proceedings* 146, no. 12 (December 2020), https://www.usni.org/magazines/proceedings/2020/december/letter-marque.

[128] Hal Wilson, "Jonathan Roper: Travelling Consultant," Modern War Institute at West Point, May 21, 2019, https://mwi.usma.edu/jonathan-roper-traveling-consultant/.

being placed on the practice by the top most echelons of military command.[129] "The more time we can invest in understanding the ethical, legal, operational, and doctrinal implications now, the better chance we have to make those decisions as carefully as a country like ours needs to during a large-scale conflict."[130] At its core FICINT is a blend of knowledge synthesis on topics important to national security and creative thinking. As an added bonus, FICINT is presented in a manner that is ideally suited to engage a broad scope of audiences by presenting information that is often complicated or boring in a manner that is engaging and easy to digest. It advances topics of merit into the thoughts of those who ought to be thinking about them, but otherwise might not be.

Conclusion

AGI and ASI are coming. It is not a matter of *if*, but *when*. The capabilities of current AI can already exceed human ability in narrow capacities and provide a glimpse of what ASI will be capable of. When combined with the requirements of AGI and the ability to gain subject matter expertise in domains of human knowledge like physics and psychology, ASI capabilities will result in a paradigm shift throughout society. The current state of AI competition does not guarantee that the U.S. will be the first to acquire ASI and so the threat of an adversary employing ASI capabilities against the U.S. must

[129] Elliot Ackerman, and Admiral James Stavridis, *2034: A Novel of the Next World War* (Kansas City: Penguin Books, 2022), 1.

[130] August Cole, "Cole on FICINT and the Cognitive Warfighting Domain," *The Cognitive Crucible,* episode #33, Information Professionals Association, audio, 38:39, https://information-professionals.org/episode/cognitive-crucible-episode-33/.

be explored to gain an understanding of the associated strategic risk and identify vulnerabilities for mitigation.

CHAPTER 3

RESEARCH METHODOLOGY

Introduction

A review of the literature demonstrates that the advent of an ASI is feasible within the next few decades. This chapter will outline how this study seeks to answer its primary research question: How could an adversary utilize an ASI to supplant the United States as the dominant world power? The method provided will also answer the following secondary research questions:

1. What unique capabilities will an ASI have that allow it to affect the information and military elements of national power?
2. How could an actor adversarial to the United States utilize ASI capabilities to achieve operational and strategic effects?
3. What are vulnerabilities in the information and military realms that could be exploited by an adversary of the United States with an ASI?

This chapter will describe how and why the method of research was selected, how it intends to conduct the study, collect data, and analyze that data in order to answer the above research questions.

Method

The purpose of this study is to assess risk and identify potential vulnerabilities after an ASI is developed. Because no such technology currently exists it would be impossible to directly gather data on it. Instead, a qualitative analysis will be conducted using a forecasting scenario methodology. Kosow and Gaβner provide several methods of

scenario generation suitable for a variety of purposes grouped into three ideal-typical scenario techniques: trend extrapolation, systematic-formalized, and creative-narrative.[131] The majority of these methods are used to analyze key factors and determine how their interactions will result in diverging future scenarios.

However, this thesis is interested in one specific future in which an ASI is utilized by an adversary against the U.S. In order to explore this specific outcome, the use of normative approach is required. The purpose of a normative scenario is to demonstrate how a possible scenario is concretely conceivable, allowing for a broader basis of discussion when determining goals, actions, or strategies relevant to the scenario outcome.[132]

> Experience shows that when an issue is thought through in narrative form (so-called 'contextualization'), the "germinal visions" which underlie the scenario are automatically put into perspective socially, economically, technologically, culturally etc. and are analyzed for correlations and possible (unexpected) consequences.[133]

The normative-narrative scenario method utilizes the common set of five phases utilized in all scenario methods provided by Kosow and Gaβner:

1. Scenario field identification.
2. Key factor identification.
3. Key factor analysis.
4. Scenario generation.

---

[131] Kosow and Gaβner, *Methods of Future and Scenario Analysis,* 43.

[132] Ibid., 70.

[133] Ibid., 71.

5. Scenario transfer (optional).[134]

Additionally, the normative-narrative scenario method utilizes seven steps nested within this five-phase process, summarized below:

1. Scenario workshop: Stakeholders and experts of the relevant vision apply creative models to single out and develop the conceptual future around a set of visionary ideas and/or goals. This step is nested with phases one through three of the common five phase process.
2. Scenario expose: The vision and/or goals are interpreted against an analysis of the key factors. The normative goals are systematized and visualized in graphical images. This step is nested with phases three and four of the common five phase process.
3. Story-board: The lines of action are developed in detail and integrated with the visionary ideas and/or goals. This step is nested with phase four of the common five phase process.
4. Scenario writing: The scenario narrative is drafted with consideration for inner logic, consistency, plausibility, and alignment with the normative vision and/or goals. This step nests with phase four of the common five phase process.
5. Optimization (optional): In this optional step the stakeholders and experts provide feedback to assist in fine tuning the scenario. This step nests with phase four of the common five phase process.

[134] Kosow and Gaβner, *Methods of Future and Scenario Analysis,* 25.

6. Evaluation (optional): In this optional step a workshop or other means is used to evaluate the implications of the scenario and draw conclusions for use in future design or decision-making processes. This step is nested with phase five of the common five phase process.
7. Publication (optional): disseminate the work to the intended audience.[135]

Method Deviations

Some deviations to this methodology are used in this thesis. First, the normative-narrative approach is generally utilized to help organizations determine a course of action towards a desired goal while simultaneously creating buy in from stake holders. This thesis uses the process to illustrate the opposite—how an undesirable scenario could be achieved by a hostile foreign entity—but for a similar purpose: to create buy in from organizational stake holders towards mitigating risks associated with the scenario. In step one, scenario workshop, the normative goal in this thesis was not determined through a workshop of stakeholders and experts, but instead by the primary research question and purpose of the study: determine how an adversary could utilize an ASI against the U.S. to replace it as the dominant world power. Step five, optimization, was utilized in a limited and non-formalized fashion in this thesis. In lieu of a group of stakeholders and experts the committee for this thesis provided feedback on the scenario development to assist the author in fine tuning. Similarly, in step six, evaluation, the author conducted analysis to draw conclusions without the assistance of a workshop. The lack of a workshop of stakeholders and experts throughout these steps is considered a permissible deviation for

[135] Kosow and Gaβner, *Methods of Future and Scenario Analysis,* 71-72.

two reasons: the common methodology makes provisions for a single author developing a scenario through "desk research" [136] and the purpose of this research is not to create unity of purpose for an organization towards a specific course of action but to demonstrate the plausibility of a scenario for purposes of determining risk and vulnerability.

Application of the Normative Narrative Steps

Step one, scenario workshop, is accomplished through the common phases of scenario field identification, key factor identification, and key factor analysis. The scenario field is determined by the primary research question: how could an adversary utilize an ASI to supplant the United States as the dominant world power? Key factors are identified in two categories: trends and drivers. Trends are key factors of the expected operating environment for the selected timeframe of the scenario and form the context within which the drivers will function. The key factors in the 2040 Global Trends Report and the state of AI competition covered in the literature review are the trends utilized in this thesis and frame the operating environment for the scenario. Drivers are those key factors which shape the critical events of the scenario. This research focuses drivers on the potential capabilities of an ASI. Identification of these capabilities is accomplished through a process of intuitive inference by combining the advantages of digital intelligence, current AI capabilities, the required attributes of AGI, domains of human knowledge, and other predicted ASI capabilities. The combination of factors and capabilities from these categories result in inferred capabilities based on key assumptions which are specified in the analysis. An explanation for how each combination of factors

[136] Kosow and Gaβner, *Methods of Future and Scenario Analysis,* 41, 62.

is inferred to result in a new capability is provided. This approach grounds the predictions of this research in reality by connecting them directly to known current capabilities and specific factors required for the achievement of ASI. This ensures the capabilities analyzed are all on the high side of probability and prevents straying into the realm of science fiction. This step answers the secondary research question: what unique capabilities will an ASI have that allow it to affect the information and military elements of national power?

Step two, scenario expose, is accomplished by analyzing the two primary trends of great power competition and AGI research cooperation to create a four futures chart. Each of the four futures are analyzed to determine where in the range of possible futures the scenario begins and how the drivers are likely to affect the operating environment as the scenario progresses. These futures are then used to determine a set of operational objectives for the ASI and controlling agent within the scenario. Objectives which allow for an analysis of the future with the greatest degree of risk are selected within the stated scope of the thesis.

Step three, story board, is accomplished through the development of an adversarial operational approach for employing an ASI to achieve the stated objectives and end state. The ways and means for accomplishing the ends of each stated objective are described in detail and summarized in a graphical depiction of the lines of effort.

Steps four and five, scenario writing and optimization, are accomplished through the production of a FICINT narrative. The narrative provides a plausible demonstration of how the means and ways described in the operational approach could result in the intended end state. The narrative takes place from the perspective of individuals to

provide perspective on how the transpiring events could be perceived by figures in various positions within society. Steps three, four, and five answer the secondary research question: how could an actor adversarial to the United States utilize ASI capabilities to achieve operational and strategic effects?

Step six, evaluation, is accomplished by drawing conclusions from the analysis in chapter four about potential vulnerabilities exploitable by an ASI. Each identified vulnerability is analyzed via the Military Strategic Risk Matrix to determine its severity (see figures 1 and 2 below) and summarized in a table as part of the conclusions provided in chapter five. Step six answers the secondary research question: what are vulnerabilities in the information and military realms that could be exploited by an adversary of the United States with an ASI?

Step seven, publication, is inherent with the publication of this thesis.

		Sources of Risk Based on damage to interest, time, resiliency			
Driver of Risk	**Strategic Value of Interest**	**Confined** Damage to interests, and/or short-term impacts	**Considerable** Damage to interests and/or mid-term impacts	**Catastrophic** Damage to interests and/or long-term impacts	**Existential** Damage to interests, and/or permanent of defining system
The Security of the U.S., its population, civil society, Allies and Partners	HLD/Vital Ally/Global Partner/Regional Other/Local	• Small Scale Contingency Ops (NEO, HA/DR) • Tactical Terror Attack (Lone Wolf) • Minor domestic civil disturbance • American hostage(s) • Loss of access • Coop Security activity or arrangement canceled	• Minor Armed Conflict • Operational Terror Attack • Isolated or Minor Attack on Global Domain or critical infrastructure • Major domestic civil disturbance • Isolated Attack on U.S. Embassy or Business • Loss of Ally or Partner • Rise of Regional Hegemon • Unsecured global domains • Isolated epidemic or natural disaster	• Theater War or Major Armed Conflict • Strategic Terror Attack (9/11) • Strategic Attack on Global Domain or critical infrastructure • Concurrent widespread major domestic civil disturbances • Integrated regional attacks on U.S. Embassies or Businesses • Invasion or Loss of Major Ally or Partner • Regional Security Organization (NATO) breakup • Major epidemic or natural disaster (Spanish Flu of 1918, Katrina)	• Nuclear War (U.S. or Allies) • WMD Terror Attack • Domestic Rebellion • Pandemic or natural disaster that threaten U.S. existence
The Security of the U.S. economy and the global economic system	HLD/Vital Ally/Global Partner/Regional Other/Local	• Limited trade, resource, or financial interruption • Confined Interference in critical infrastructure • Change in currency standard • Minor cyber compromise	• Extended trade, resource, or financial interruption • U.S. Recession • Extended interference in critical infrastructure • Failure of IMF • Lack of Int'l norms • U.S. Depression	• Financial failure of major institution or market • Major Degradation of critical infrastructure • Access to Global Domain(s) disrupted by adversary	• Global or U.S. economic collapse • Closed economic system • Destruction of critical infrastructure • Seizure of U.S. business/industry • Access to Global Domain(s) denied by adversary.
Preservation and extension of universal values	HLD/Vital Ally/Global Partner/Regional Other/Local	• Local Atrocities • Imposition of martial law by Ally or Partner • Democratic regression in Ally or Partner	• Mass atrocities • Democratic regression in Key Ally or Partner • Local imposition of alternate value system	• Genocide (Holocaust) • Regional imposition of alternate value system • Emergence or powerful totalitarian nation	• Global Imposition of alternative value system
Advancing and maintaining U.S. led International Order	HLD/Vital Ally/Global Partner/Regional Other/Local	• Local or State order undermined or replaced by alternative system, neutral or antagonistic to U.S. system (sets negative precedent)	• Regional Order undermined or replaced by alternative system, neutral or antagonistic to U.S. system	• Elements of International order undermined or replaced by alternative system, neutral or antagonistic to U.S. system	• U.S. Order Replaced in total by alternate system, hostile to current U.S. system

Figure 1. Military Strategic Risk Matrix–Consequence Development

Source: Office of the Chairman of the Joint Chiefs of Staff (CJCS), CJCS Manual 3105.01A, *Joint Risk Analysis Methodology* (Washington, DC: Joint Chiefs of Staff, 2021), C-4.

		Sources of Risk			
		Confined	**Considerable**	**Catastrophic**	**Existential**
Strategic Value of Interest	**HLD / Vital**	Modest	Major	Extreme	Extreme
	Ally / Global	Modest	Major	Major	Extreme
	Partner / Regional	Minor	Modest	Major	Major
	Other / Local	Minor	Minor	Modest	Modest

Figure 2. Military Strategic Risk Matrix–Consequence Assessment

Source: Office of the Chairman of the Joint Chiefs of Staff (CJCS), CJCS Manual 3105.01A, *Joint Risk Analysis Methodology* (Washington, DC: Joint Chiefs of Staff, 2021), C-4.

Ethical Considerations

Due to the predicative nature of this study, no significant ethical considerations are expected that require mitigation. Some consideration was given to the ethical implications for adding to the body of literature about a technology that has the potential to be just as disruptive, if not more disruptive, as the advent of the nuclear bomb. These concerns were ultimately dismissed for two primary reasons: ASI is more akin to nuclear fission than the nuclear bomb in that it could be used for extremely positive or extremely negative ends—it just so happens that this research focuses solely on the negative aspects—depending on the will of its controlling agent; secondly, unlike the Manhattan project which was a classified military endeavor, AI research and progress is being advanced almost entirely in the civil sector and will continue on the path towards ASI—and all the risks that come with it—regardless of the degree to which we understand and prepare for it in advance. Care will be taken during the drafting of the scenario to ensure depicted events and predictions are not subject to the researcher's inherent biases but grounded in reliable research and analysis from sources within academia, the Department of Defense, and various intelligence communities.

Summary

Research into potential futures is difficult because the items of relevance may not exist yet, as is the case with an ASI. A qualitative forecasting scenario methodology provides a framework to explore expected capabilities and their potential impacts to determine vulnerabilities and their degree of risk.

CHAPTER 4

ANALYSIS

Introduction

This chapter examines the data collected and analyzed to answer the research questions. It is broken up into four major sections: analysis of ASI capabilities and their effects as key factor drivers, analysis of the operational environment associated with the forecasted arrival of AGI in 2035 and ASI in 2040 as key factor trends, synthesis of the key factor trends and drivers into an operational approach for achieving an end state relevant to the research questions, and a narrative description of select events from the forecasted scenario to demonstrate plausibility. Taken as a whole this chapter directly answers the primary research question of: how could an adversary utilize an ASI to supplant the United States as the dominant world power? While this chapter presents a single plausible future, the analysis presented to establish it has broader applications that could be utilized in the generation and analysis of many scenarios with variations on the key assumptions used in this thesis.

ASI Capabilities

In order to determine how an ASI could be employed, it is necessary to first identify what its capabilities will be. Predictions in the literature about ASI capabilities have a tendency to range into the extreme bounds of the imagination, likely due to the belief that there is nothing an intelligence greater than our own could not achieve. To avoid this speculative form of forecasting a set of expected capabilities were derived through a process of intuitive inference by combining the currently existing AI

capabilities and requirements of GAI discussed in the literature review with a small subset of knowledge domains. The domains selected are those that a controlling agent would prioritize for an ASI to obtain greater than human level expertise based on their applicability to military and information operations. Additionally, the advantages of digital intelligence discussed in the literature review are considered to be implicitly present in all combinations.

The forecasted capabilities are categorized into three tiers based on their dependency on known factors and other forecasted capabilities:

1. Tier One: capabilities derived directly from current AI capabilities, AGI requirements, and expertise in knowledge domains.
2. Tier Two: capabilities that are dependent upon at least one tier one capability.
3. Tier Three: capabilities that are dependent upon at least one tier two capability.

Tiering the capabilities in this fashion provides utility to additional analysis because tier two and three capabilities could potentially be neutralized by developing effective mitigations to the lower tier capabilities, they are dependent upon.

Figure 2 (ASI Capability Linkage Chart) provides a visual depiction of how each capability is derived from a combination of known factors. The known factors are listed on the left-hand side and color coded to align with current AI capabilities, AGI requirements, and knowledge domains. The forecasted capabilities are aligned based on their tier and have a color-coded outline which matches the color of the solid lines which connect to their requirements and dependencies. Tier two and three capabilities inherit any and all dependencies and requirements of the lower tier capabilities they are

dependent upon; these linkages are not explicitly restated. Some forecasted capabilities are enhanced by known factors or other forecasted capabilities. These enhancements are not requirements in order for the forecasted capability to manifest or feasibly function but increase its effectiveness. Such enhancements are depicted by dotted lines in Figure 2, color coded to match the colored outline of the capability that is enhanced. To add additional clarity, Table 1 (ASI Capability Linkages) lists each forecasted capability's linkages and enhancements in a bulletized format. Forecasted capabilities are italicized in Table 1 to differentiate them from known factors. Following Table 1, each forecasted capability is analyzed individually.

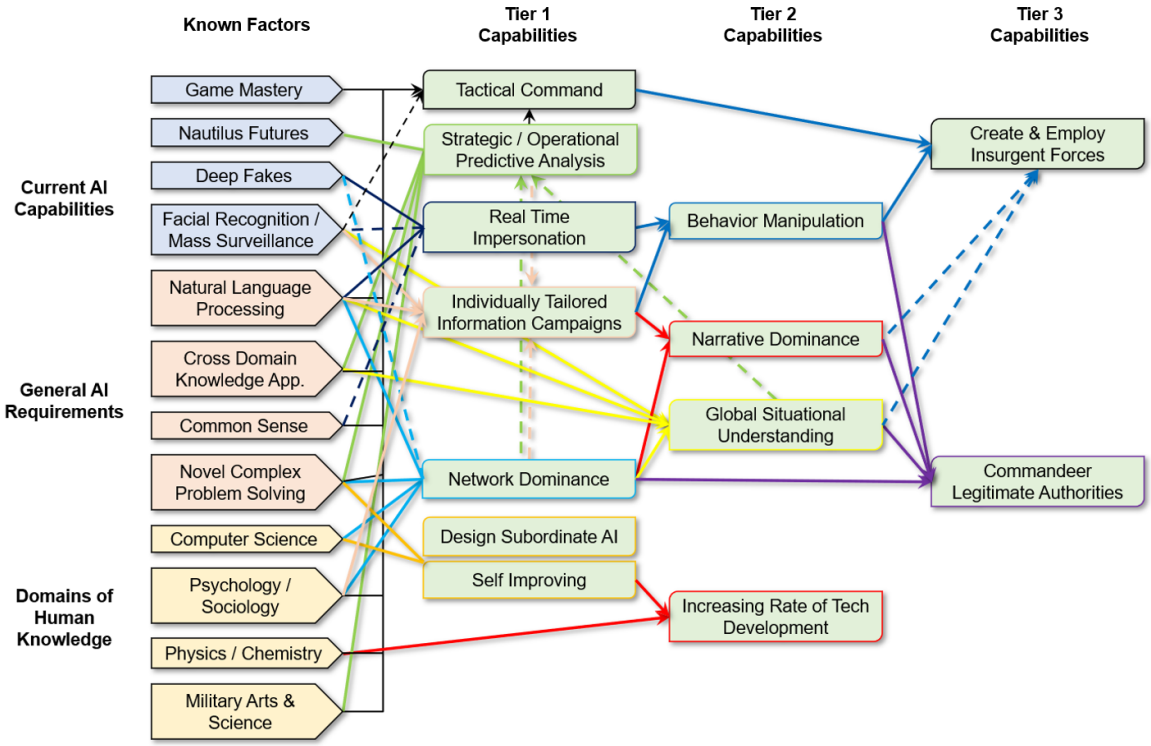

Figure 3. ASI Capability Linkage Chart

Source: Created by author.

Table 1. ASI Capability Linkages			
Capability	Tier	Requirements/Dependencies	Enhancements
Design Subordinate AI & Self-Improving	1	• Novel Complex Problem • Solving Computer Science Expertise	
Real Time Impersonation	1	• Deep Fakes • Natural Language Processing	• Mass Surveillance • Common Sense
Network Dominance	1	• Natural Language Processing • Novel Complex Problem • Solving Computer Science Expertise • Psychology/Sociology Expertise	• Deep Fakes • *Real Time Impersonation*
Strategic/Operational Predicative Analysis	1	• Predictive Analysis • Cross Domain Knowledge App. • Novel Complex Problem Solving • Military Art & Science Expertise	• *Network Dominance* • *Global Situational Understanding*
Individually Tailored Information Campaigns	1	• Mass Surveillance • Natural Language Processing • Psychology/Sociology Expertise	• *Network Dominance* • Predictive Analysis
Tactical Command	1	• Game Mastery • Natural Language Processing • Cross Domain Knowledge App. • Common Sense • Novel Complex Problem Solving • Psychology/Sociology Expertise • Physics Expertise • Military Art & Science Expertise	• Predictive Analysis • Mass Surveillance
Increased Rate of Technological Development	2	• *Self-Improving* • *Physics/Chemistry Expertise*	
Global Situational Understanding	2	• Mass Surveillance • Natural Language Processing • Cross Domain Knowledge App. • *Network Dominance*	
Narrative Dominance	2	• *Individually Tailored Information Campaigns* • *Network Dominance*	
Behavior Manipulation	2	• *Real Time Impersonation* • *Individually Tailored Information Campaigns*	
Create and Employ Insurgent Forces	3	• *Tactical Command* • *Behavior Manipulation*	• *Narrative Dominance* • *Global Situational Understanding*
Commandeer Legitimate Authorities	3	• *Behavior Manipulation* • *Narrative Dominance* • *Global Situational Understanding* • *Network Dominance*	

Source: Created by author.

Design Subordinate AI

The ability to design subordinate AI is a tier one capability that simply requires the ability to solve novel complex problems and human level comprehension of computer science. Humans already have the ability to create AI so it is essentially inherent in the definition of ASI that it will share this ability. It is inferred that an ASI will be able to create AI on a spectrum of intelligence and capability up to its own—the furthest point of current human knowledge at the time of its own creation—in order to create efficiencies. In the same way that humans come together to form organizations which then collectively exhibit a greater level of intelligence and capability than any of the individuals that comprise it, an ASI will design AI custom tailored to perform specific functions that contribute to the achievement of its goals. Depending on the complexity of the task, significant efficiencies would be gained for the ASI by having less intelligent AI handle tasks that do not require its higher-level functioning. This also allows an ASI to circumvent some of the physical limitations it may have on its computational capabilities—based on the current level of hardware available in processors, etc.—by forming AI organizations. The ability to design and utilize subordinate AI would enhance the implementation of any other ASI capability based on the amount of physical resources its controlling agent is able or willing to dedicate to them.

In addition to enhancing its own operations, the ability to create subordinate AI means a controlling agent could proliferate AI capabilities to other agents as a method of achieving their objectives. This could range from mundane NAI applications to maximize the efficiencies of utilities to complex AGI designed to overhaul city planning. However, it also includes weaponized AI designed to carry out malicious activities. An ASI could

design AI applications to sell on open or black markets or provide them to groups aligned with their controlling agent's interests to accomplish a wide range of tasks less economically developed states or groups would otherwise be incapable of. The ease of moving digital data around the globe means the proliferation of weaponized AI could happen extremely quickly and would be extremely difficult to halt or roll back.

Designing subordinate AI is not depicted or listed as an enhancement for any other capabilities in Figure 2 or Table 1 as it is universally applicable, similar to the advantages of digital intelligence. It should be assumed as a present in all the other forecasted capabilities as an enhancing factor.

Self-Improving

The ability to improve itself is a tier one capability that requires the ability to solve novel complex problems and have an above human level understanding of computer science. At the time of its development an ASI will represent the pinnacle of human achievement which means that an ASI must exceed human knowledge in AI to improve itself. However, just as humans improve on their knowledge and designs through iteration, an ASI will inherently be capable of doing the same. An ASI would improve upon itself in a variety of ways including optimizations such as increasing the physical hardware upon which it is reliant, optimizing code utilized in its programming, and developing novel advancements in means or method. Ultimately the ability to improve itself means that ASI will increase both the speed and quality of its intelligence over time and therefore further increase its capability. As it is able to process greater amounts of information simultaneously across more domains at faster speeds, problems of increasing complexity will become less and less difficult and time consuming.

Self-improvement means that the capability of an ASI will not be static. This has significant implications both for how an actor would seek to employ ASI offensively and for attempts to mitigate the effects of ASI through parity. Multiple state or non-state actors kept from employing ASI offensively due to a status quo of mutually assured destruction—similar to that created with the advent of nuclear weapons—may not be feasible as ASI will vary in their degree of intelligence and therefore the degree and speed at which its capabilities can be employed. An ASI that has recursively improved upon itself will likely be significantly more capable than any newly developed and deployed ASI. Depending on the degree to which the two ASI are different in measure of capability could result in a situation where the lesser system becomes ineffective at achieving objectives opposed by the greater ASI. Such a disparity would make a mutually assured destruction scenario untenable and therefore efforts to gain stability through parity ineffective.

Increased Rate of Technological Development

Increasing the rate of technological development is a tier two capability dependent upon the ability to recursively self-improve and obtain above human expertise in domains of knowledge, especially STEM fields like physics, chemistry, and engineering. Expertise that exceeds human capability in these areas will allow for technological development at a pace greater than human development. As discussed in the literature review, humanity has progressed on an exponential curve of technological development, increasing our own rate of progress over time as we utilize advancements to increase the rate of yet more advancements. ASI will be able to increase its own level of intelligence and then create large organizations of subordinate AIs at this higher level

to work on problems. ASI will benefit from all three forms of super intelligence: collective, speed, and quality. This will result in a double exponential curve of development and eventually lead to technological singularity, a point in time beyond which ASI is advancing technology so fast that it becomes impossible for a competitor to catch up. Figure 3 provides a visual depiction of the disparity of technological development between two ASIs with equal rates of self-improvement and technological advancement but different start points.

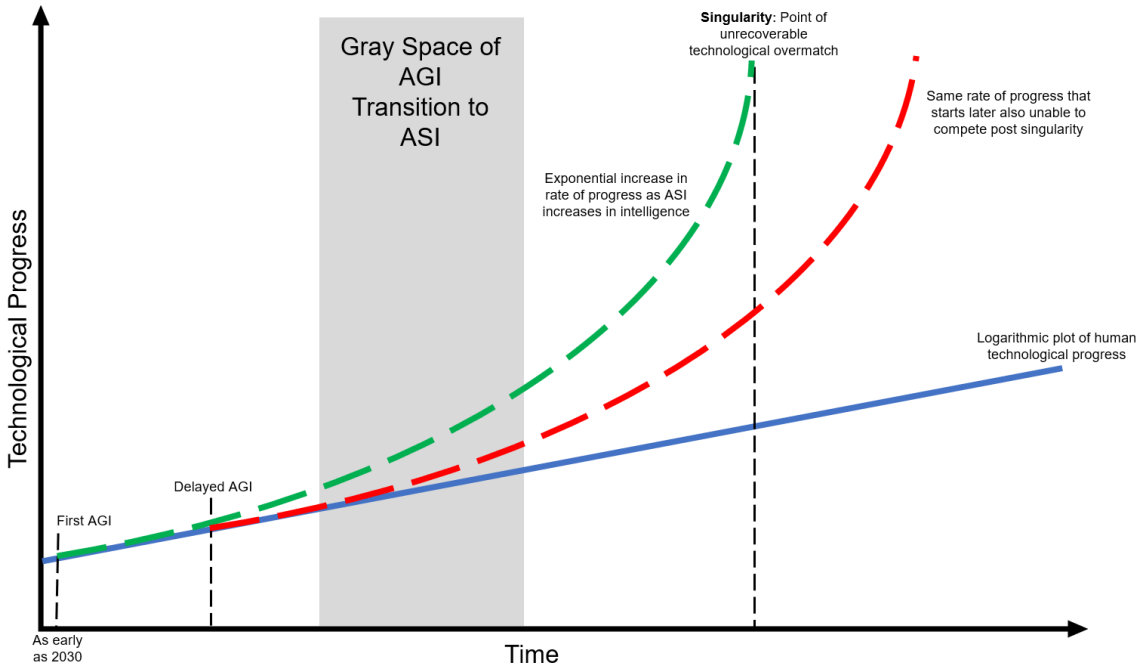

Figure 4. Theoretical Comparison of Technological Progress Over Time

Source: Created by author.

However, there are physical limitations associated with this capability. An ASI may be able to digitally simulate a significant amount of information in order to make predictions and develop new tech. However, at some point any hypothesis of an ASI will

need to be tested in the real world to validate the simulation, create tangible results, and provide data back to the system to allow for additional developments. Additionally, technological advances are only useful so far as they can be deployed physically in the world. ASI generating a device to control the weather, as a hypothetical example, would only be useful if the device actually worked in practice and its controlling agent had access to and was willing to utilize the needed resources to build such a device. Just because an intelligence is greater than that of a human does not necessitate that it will get something correct on the first attempt, even if it is much more likely to. Regardless, the need to test in the physical environment places an artificial limiter on the capability to increase the rate of technological development associated with the amount of physical resources that can be applied to producing and testing ASI proposed advancements. Due to this artificial limitation, access to resources may serve as a recovery mechanism to close limited gaps with an ASI that is more advanced than another as a result of having more time for recursive self-improvement.

Real Time Impersonation

Impersonating individuals in real time over digital systems is a tier one capability derived from deep fakes and natural language processing. Given enough data, deep fakes can already create real time manipulatable facial puppets or voice replications by breaking speech data down into the foundational component sounds. It is inferred that an ASI will be capable of advancing this technology even further to be able to recognize and duplicate speech patterns, cadence, tone, expected mannerisms, usage of slang and vernacular, and if enhanced by mass surveillance, even include personal details. The amount of data required to generate a believable impersonation of an individual's voice

and face will likely be significantly reduced by an ASI by building human archetype models that align with facial structures/features and voice pitch. The Ziva Dynamics ZRT face trainer provides an excellent current example of how real time generated faces are already nearly photo realistic.[137] By combining the ability of deep fakes with natural language processing an ASI can be expected to communicate with a target over digital communication platforms—such as cell phones, IP Phones like Cisco systems, zoom, slack, or MSTeams—using the voice and/or image of authority figures from within their organizations, trusted associates, or even family and friends. Impersonations of such a high quality could be employed to gain information or compel the target to take specific actions (see Behavior Manipulation).

Network Dominance

Network Dominance is a tier one capability that is defined by an ASI's ability to gain access to networks to such an extent that it retains control of that network similar to a system administrator. This capability can be achieved through multiple methods both of which rely upon novel complex problem solving. Humans already possess the ability to find and exploit flaws in network architecture via cyber and social engineering methods. Cyber methods rely upon the ability to analyze and understand the code utilized to establish security protocols and identify ways those protocols can be manipulated or subverted to gain access to content. An ASI with superhuman expertise in computer science would be able to find and exploit architecture flaws at an incredible speed.

[137] Ziva Dynamics Inc., "Ziva Face Trainer," Ziva, https://zivadynamics.com/zrt-face-trainer/.

Additionally, an ASI could engage in social engineering utilizing Real Time Impersonation to gain access to networks through authorized users. An ASI that gains even momentary access through a legitimate member of a network could exploit it to inject code to enable continued exploitation via digital means.

An ASI with access to a network will be able to gain a complete and comprehensive understanding of its architecture and security protocols at speeds that would be impossible for humans to counter. This above human level understanding would also allow it to hide backdoors and exploits in places within the code base that would be extremely difficult for humans to detect. Because an ASI would not need to rely on exploitations tools—code applications used by hackers that are designed to take specific actions once they gain access to a network in order to either conceal an exploit or transmit information—it could read, understand, and alter the code of system files directly. Network dominance is achieved when an ASI has gained permanent unrestricted access to a network providing it knowledge of all information contained within and the ability to manipulate that information or modify the network to achieve its own ends. Such dominance over a network would completely bypass encryption and would likely be completely invisible to the network's owner, the code altered in such a way as to continue providing outputs aligned with what the system administrators or anti-virus software expect to see. Such dominance of a network could allow an ASI to create chaos within an organization. Files could be altered to remove key pieces of information, change important data points, induce false information, or gaslight employees. Dominated networks will also likely be used to gain access to increasingly secure or higher value networks. Should the network of an operating system owner such as Microsoft, Google,

or Apple become dominated, an ASI could utilize its access to implant exploits within content patches that are then pushed to every user of those systems, effectively gaining mass access via a single compromise.

Global Situational Understanding

Global Situational Understanding is a tier two capability derived from Network Dominance, Mass Surveillance, Natural Language Processing, and Cross Domain Knowledge Application. Military Intelligence communities have long understood the importance of trying to keep even small details hidden from the enemy because lots of small details put together can reveal much larger pieces of information. Humans are progressively living more and more of their lives in the digital domain, exponentially increasing the amount of data that is potentially accessible through the internet. When the internet of things is factored in, an ASI utilizing Network Dominance could have access to live surveillance of huge amounts of the world at all times. Access to cell phones or city closed circuit cameras could allow for tracking the movements and activities of specific individuals and learn their patterns and monitor for key activities. Access to logistics networks would provide information on where materials are being collected in sufficient quantities to indicate specific types of operations or activities. Access to microphones in homes or places of businesses would allow an ASI to listen in on conversations believed to be private or secret.

Collecting mass data from internet connected devices is not a new concept—it is happening today—but processing and analyzing that data into something useful is not easy. The amount of information that is obtainable from network connected devices is incalculable and beyond compression for a single individual. Human cognitive

capabilities are limited such that entire organizations are needed to process and understand information on a small subset of world happenings. Current AI capabilities allow for some impressive analysis from mass data, but only along narrow pre-defined lines. However, due to the advantages of digital intelligence, an ASI, or collective of ASI, will have the ability to process this deluge of available information into a unified picture to gain both a factual and inferred understanding of the current state of the world that is unprecedented in scope and constantly updating in real time. It should also be expected that the accuracy of inferences made by the AI will increase over time as it learns what factors lead to accurate inferences when combined under specific sets of circumstances. A comprehensive understanding of the current state of the world over time creates an ability to identify trends and understand complex interactions of variables that will provide an asymmetric advantage in setting strategic and operational goals.

Strategic and Operational Predictive Analysis

Predictive analysis for use in creating operational and strategic objectives and/or campaign plans is a tier one capability dependent upon Predictive Analysis, Cross Domain Knowledge Application, Novel Complex Problem Solving, and Military Art & Science Expertise. Setting strategic and operational objectives to achieve end states requires an ability to analyze factors that are likely to influence changes in the operational environment. Current organizations accomplish operational and strategic planning by limiting the number of factors under consideration to a cognitively manageable amount and conducting analysis to determine what variables are most likely to have the potential

to influence conditions to align with the desired outcome. Available means, such as military forces, can then be assessed for their ability to influence the relevant variables.[138]

An ASI with the advantages of digital intelligence will be capable of conducting analysis of a far greater number of factors with a much higher degree of fidelity to make determinations on the best ways to achieve objectives and the best means to employ. This capability would be greatly enhanced by Network Dominance and Global Situational Understanding. As the amount of current and accurate information available increases, so does the number of factors that can be identified for utilization in analysis. Network Dominance would provide accurate and current information on both the current state and future plans of adversarial organizations, states, or forces; and Global Situational Understanding allows for a method of assessing the validity of planning assumptions and the effectiveness of implemented strategic and operational plans. Over time this capability will increase as ASI observes the results of previous predictions and variables. Ultimately, this is likely to result in an ASI generating strategic or operational plans that contain ways or methods of employing means that a human would never have conceived of, similar to the unprecedented moves made by the AI Alpha Go. If acted upon, such plans would provide an asymmetric advantage due to their complexity of purpose and unpredictability to the opposition.

[138] Office of the Chairman of the Joint Chiefs of Staff (CJCS), Joint Publication (JP) 5-0, *Joint Planning* (Washington, DC: Joint Chiefs of Staff, 2020), I-1 - I-2.

Tactical Command

Commanding forces in real time tactical situations is a tier one capability that has a long list of dependencies. The high number of requirements for this capability is related to the fact that commanding forces in combat is different than analyzing trends like in operational and strategic planning. The best plan in the world does not translate to effective tactical employment without an understanding of how to communicate effectively with people and how to prepare and motivate them to accomplish extremely difficult and dangerous tasks. The need for decisions and changes to an established plan also arises at a fast pace in tactical action, especially compared to operational or strategic timelines.

However, the implications of AI performance in games that require strategic and tactical thinking in the immediate, short, medium, and long term has obvious military implications. Human tacticians make decisions by scoping their understanding of the battlefield down to a limited set of understandable factors, bounding the fight inside a limited number of parameters. Area of operations, key terrain, avenues of approach, objective areas, and operational variables are all examples of how military commanders simplify tactical actions to fit within the limits of human cognition. The bounds of ASI cognition will be capable of incorporating a significantly greater number of factors, perceiving the battlefield at a level of detail incomprehensible to humans.

Similarly, it will also be capable of conducting orders of magnitude more iterations of course of action analysis. The U.S. Army routinely develops five or less courses of action—often as few as two—for friendly and enemy forces which are then—under ideal circumstances—war gamed against each of the enemy options, a single time.

An ASI would be capable of developing thousands of courses of action and wargaming each friendly and enemy plan pairing thousands of times to determine how changes in luck—like individual rounds hitting or missing targets—or variables like minor changes in the weather can affect outcomes. An ASI would be able to plan and synchronize incredibly complex tactical actions while still communicating the component parts simplistically to the human actors required to carry them out. An ASI acting as a tactical commander or advisor to a tactical commander could translate small seemingly meaningless reports and changes in battlefield conditions into decision points that control the flow of battle. PLA thinkers recognize the potential for AI in this capacity and anticipate a "battlefield singularity" where human cognition can no longer keep pace with the speed of decision-making and tempo of combat in future warfare.[139]

Individually Tailored Information Campaigns

Information campaigns individually tailored to the specific biases and psychological traits most likely to elicit a desired reaction is a tier one capability dependent upon Mass Surveillance, Natural Language Processing, and Psychology/Sociology expertise. AI assisted mass information campaigns, especially misinformation campaigns, are already a reality. AI also already assists information campaigns in the form of algorithms which track the preferences of unique users in an attempt to show them more of the type of information they are most likely to engage with. This is a form of mass surveillance, with certain companies maintaining profiles on millions of individuals specifically to target them with information. This type of

[139] Kania, "Battlefield Singularity," 16.

information manipulation has already proven to be very effective. ASI will combine these mass surveillance methods with a superhuman understanding of psychology and sociology not to just determine what available information to put in front of an individual, but to craft information specifically tailored for individuals. Information tailored in this way can be designed to exploit individual biases, preconceived notions, emotional triggers, values, and principles, delivered in a method and style most likely to be perceived as credible by the individual.

Individually tailored information is considered a tier one capability, but it can be greatly enhanced by both Subordinate AI Creation and Network Dominance. The use of subordinate AI would greatly increase the number of individuals that could be targeted while network dominance would greatly enhance the quality of psychological and sociological profiles used to target and develop the information. An ASI that has gained access to an individual's smart phone, for example, could use the microphone to listen in on their conversations or the camera to track their facial expressions as they engage with media or discuss it with other. Both would provide intimate personal data for building a targeting profile. Without network dominance an ASI would likely be reliant upon users voluntarily utilizing applications or sources under its control, greatly reducing the effectiveness of this capability. However, an ASI able to gain network dominance over a preponderance of critical information networks would be able to use this capability to gain Narrative Dominance.

Narrative Dominance

Narrative dominance is a tier two capability achieved by gaining Network Dominance of a preponderance of critical media networks, allowing for individually

targeted information campaigns to become uncontested. Media outlets whose networks have been dominated could be synchronized with individually targeted information to show different stories or information to different users, including the legitimate authors, editors, and publishers of the media outlet. Because media sources often utilize each other as sources for stories, even the outlets not under the effects of network domination would likely become assets in pushing a dominate narrative, regardless of its authenticity. Reports of international events could be completely fabricated but difficult to disprove because media sources from around the globe have published different first-hand reporting with minor variations in detail or perspective. With enough seemingly authentic reporting coming out of legitimate and trusted media sources, it would become extremely difficult, if not nearly impossible, to provide a counterfactual narrative; especially because any attempt to publish information to the contrary would have to be on a platform whose network is not dominated, or it would be subject to altercation or deletion prior to being seen by the public.

Behavior Manipulation

Manipulating people into engaging in behaviors or conducting desired actions is a tier two capability dependent upon Real Time Impersonation and Individually Tailored Information Campaigns. To successfully manipulate a person, an ASI would need to develop a psychological profile on them through the same means as individualized information. Information designed to prime the individual to exhibit a desired behavior could then be targeted to them via an individualized information campaign. In certain circumstances this priming may be sufficient to incorporate an individual into a larger coordinated event like a protest or demonstration.

However, more specific actions would be compelled through Real Time Impersonations. Individuals could be instructed to do things that are seemingly harmless by impersonations of their organizational superiors or family members. Individuals could also be compelled to commit acts they have moral objections to through nefarious impersonations. The Taliban had great success co-opting Afghan military members to commit insider attacks against U.S. forces through financial, ideological, religious, and coercive measures such as threats against or even kidnapping of family members.[140] An ASI with Network Dominance could have access to any number of personal details to use for blackmail or emotional manipulation such as creating real time deep fakes depicting eminent harm to family members. Regardless of the means of manipulation, the end result is that ASI will have a variety of methods by which it could trick or compel individuals to conduct activities in pursuit of its objectives, even when such actions might conflict with the individual's normal motivations or values.

Commandeer Legitimate Authorities

Commandeering legitimate organizations, authorities, and capabilities is a tier three capability that requires Behavior Manipulation, Narrative Dominance, Global Situational Understanding, and Network Dominance. This capability uses Behavior Manipulation to get organizations, instead of specific individuals, to conduct actions that aid in the pursuit of the ASI's objectives. Similar to Behavior Manipulation these actions

[140] Javid Ahmad, *Dress Like Allies, Kill Like Enemies: An Analysis of 'Insider Attacks' in Afghanistan* (West Point, NY: Modern War Institute, April 4, 2017), 10, https://mwi.usma.edu/wp-content/uploads/2017/04/Dress-Like-Allies-Kill-Like-Enemies.pdf.

can range from small seemingly unimportant and harmless to drastic measures that have massive national security and policy implications. For this reason, Network Dominance of the organization being commandeered is essential to ensure information about the desired actions is circulated in such a way that the objectives are met without creating organizational alarm or incongruity.

Network Dominance is also likely necessary to gain sufficient understanding of an organization's practices, procedures, culture, and lingo to effectively employ Behavior Manipulation against key members. For this same reason, Narrative Dominance is likely required to achieve decisive effects. Key organizational leaders are less likely to be tricked into taking actions counter to their organizational or national interests unless those actions are congruous with a dominant national narrative. For example, leaders that have been falsely informed by their news source in advance that the President is considering a controversial policy change will be primed to commit their organization to that new policy when prompted.

Commandeering legitimate authorities should always be considered a short-term effect, as leaders should eventually realize that false orders have been issued. Network Dominance can stall this outcome by changing or deleting digital communications that would inform leaders of the deception. However, ASI has no ability to prevent people from discussing things in person, and the incongruities between actual intent and executed intent between leaders and their organizations should eventually reveal the commandeering. Once discovered, organizations could be expected to take steps to mitigate or prevent similar manipulation in the future. For this reason, an ASI will likely

take special care to employ this capability subtly or when it can create the most decisive results in pursuit of its objectives.

Create and Employ Insurgent Forces

Creating and employing an insurgent force is a tier three capability that requires Behavior Manipulation to influence a vulnerable population into joining a paramilitary group and Tactical Command to employ them effectively as a means to achieve military objectives. An Ipsos poll in January of 2022 found that 7 percent of Americans believe in some of the most extreme QAnon conspiracy theories such as: "a group of Satan-worshipping elites who run a child sex ring are trying to control our politics and media."[141] Even if this number is reduced by the survey's sampling error of 3.6% to an estimate of only 3.4% of the population that would still be over 11 million people across the country potentially vulnerable to radicalization. By contrast, it is estimated that there are no more than about 700,000 full time law enforcement officers in the United States.[142] A targeted campaign to radicalize and organize this population into an insurgent force with a success rate of only 10% would still drastically outnumber law enforcement. The prevalence of military style firearms and ammunition privately owned by citizens in the United States also means that such a radicalized force could come self-

---

[141] Ipsos, "Ipsos Misinformation and Conspiracy Theory Poll," (Ipsos Group S.A., Washington, DC, January 5, 2022), 9, https://www.ipsos.com/sites/default/files/ct/news/documents/2022-01/BBC%20Misinformation%20and%20Conspiracy%20Theory%20Topline%2001.05.22.pdf.

[142] U.S. Bureau of Labor Statistics, "Occupational Employment and Wages, May 2021: 33-3051 Police and Sheriff's Patrol Officers," Division of Occupational Employment and Wage Statistics, accessed February 19, 2022, https://www.bls.gov/oes/current/oes333051.htm.

equipped to engage in para-military activity without the need for raids on arms stockpiles prior to a decisive operation.

This scenario only accounts for individuals that could be militarized through simple or complex information operations, including the use of Individually Targeted Information Campaigns. However, an even larger portion of the population could be targeted through the nefarious methods of Behavior Manipulation. Individuals less prone to radicalization, but with special access or in positions of authority, could still be manipulated or coerced into conducting acts harmful to their nation's security or as a component of a larger insurgent operation coordinated by an ASI using Tactical Command. Likely these acts would be small in nature to prevent significant ideological or moral dilemmas for the individuals, but if synchronized with paramilitary activities or a large number of similarly compelled actions, would create decisive effects.

An ASI commanding an insurgent force domestically is likely to achieve decisive effects due to the element of surprise, mass, and tempo. Enhanced by Global Situational Understanding the ASI would employ innumerable domestic sensors as Information Collection platforms to inform tactical decision making. Behavior Manipulation could also be employed to create acts of sabotage against domestic authorities that cumulatively decrease their capability to effectively respond to or resist the insurgent operations. Once a legitimate force is marshalled to resist the insurgent force it would have to contend with all the normal difficulties of fighting insurgents compounded with the advantage of ASI tactical leadership.

Operational Approach for Adversarial ASI Employment

In order to determine how an ASI could be employed by an actor adversarial to the United States to replace it as the dominant world power this section will describe a plausible operational approach for achieving this end state by employing the forecasted ASI capabilities as means. The operational approach in this context serves as the ways portion of an ends, ways, means approach to operational design via operational art as described in JP 5-0.[143]

The Advent of ASI: Four Futures

This section utilizes the factors from the NIC Global Trends 2040 report covered in the literature review as the basis for the operational environment and starting conditions. However, an additional factor of international cooperation on the development of AGI/ASI is considered. Combined with the primary factor of great power competition, a four futures framework was developed to account for how the arrival of AGI/ASI would impact the operational environment and scenario starting conditions. The four futures diagram is depicted in Figure 5 and each quadrant is described below. The diagram depicts great power competition on the X-axis with a range from allies to full scale conventional war. The mid-point of the X-axis represents neutral co-existence with the great powers neither friendly nor hostile. The scenario initial conditions of competitive co-existence are therefore depicted as right of center but well left of open conflict. The Y-axis depicts international cooperation on the development of AGI, ranging from governments racing to be first and developing the technology completely in

[143] CJCS, JP 5-0, IV-1.

classified environments on the low side to governments agreeing to a single unified research project that is open source and shared between the entire international community on the high side. The scenario initial conditions reflect the current reality of AGI development happening primarily in the domestic and commercial realms where limited cooperation occurs. The starting conditions lean towards the low side because governments are also investing in AGI research which is not fully public.

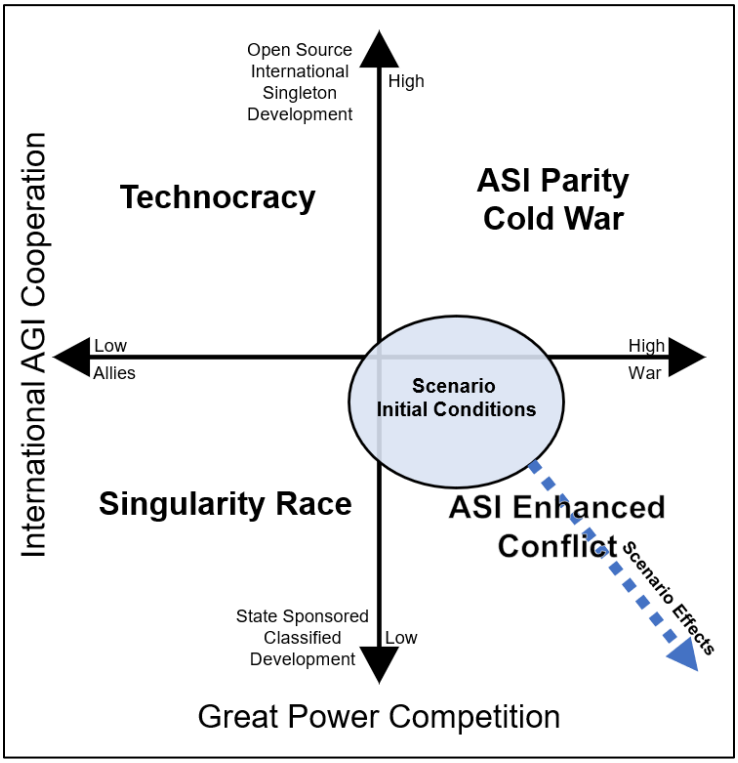

Figure 5. 2040 Operational Environment: Four Futures

Source: Created by author.

Technocracy

Results from low competition between the great powers and high international cooperation on AGI/ASI development. In this future, strong relationships reinforce trust

that all nations can benefit from the development and utilization of AGI/ASI. This trust in turn manifests in the prioritization of AGI/ASI capabilities that are mutually beneficial to all societies and help overcome worldwide challenges, resulting in cooperation and prosperity.

ASI Parity Cold War

Results from high competition between the great powers. Issues that might normally lead to direct conflict are instead tempered by high international cooperation on the development of AGI/ASI. The dire prospect of ASI capabilities being employed in a conventional war by both sides, due to shared development of the technology, motivates the great powers to limit the scope of their competition to below the threshold of armed warfare. AGI/ASI capabilities that are mutually beneficial to all societies are still prioritized but great powers seek to employ them selectively to gain a competitive advantage on the world stage.

Singularity Race

Results from low competition between the great powers and low to no international cooperation on the development of AGI/ASI. States are cooperating domestically and economically to overcome the shared world challenges while simultaneously recognizing that the development of ASI could provide them with permanent international dominance through technological overmatch. ASI development becomes coopted by governments in order to reduce the chance of assisting other states in their progress towards the technology. States prioritize resources to increasing the intelligence of their ASI on technological advancement and recursive self-improvement.

These developments yield slow gains to society at large due to fears that another nation's ASI could leap ahead in advancements through reverse engineering of proliferated technologies.

ASI Enhanced Conflict

Results from high competition between the great powers that devolves into conflict combined with low international cooperation on AGI/ASI development. Governments recognize the competitive advantages to be gained in international competition and conflict by employing the capabilities of an ASI. ASI is utilized by state and/or non-state actors in pursuit of their strategic objectives, moving competing states progressively closer to conventional war. States engage in a singularity race with active attempts to degrade, destroy, and prevent ASI development by competitors. If multiple states obtain ASI, the conflict centers around each side's attempt to destroy the other's ASI, largely by employing their own ASI towards that endeavor.

Middle of the Road

The selected Competitive Coexistence scenario from the NIC 2040 report that serves as the basis of the operational environment for this thesis describes a moderate amount of competition between the great powers, depicted as slightly right of the center axis on the four futures chart in Figure 5. Current trends in the international cooperation and competition on AI also suggest that AGI/ASI development is likely to be fairly middle of the road with a moderate amount of indirect cooperation on the development of AGI/ASI through public research and private sector developments but with states simultaneously attempting to monopolize developments to their economic and military

advantage. In this scenario the advent of ASI is likely not perceived as a significant factor in strategic planning for the great powers who are focused on short- and medium-term advantages from NAI and instead arrives unexpectedly as a result of proprietary innovations.

Non-State Controlling Agent

A middle of the road scenario where ASI is likely to first arrive as a result of proprietary development means that the first ASI could plausibly, and perhaps even probably, be under the control of a private owner. A private owner would likely be incentivized to keep information about such a development tightly controlled until it could be determined how to best utilize it to the benefit of their organization. However, a private owner may also have gotten into the AGI business specifically to monopolize it towards achieving strategic objectives aligned with their personal philosophy. A super-individual with motivations to utilize AGI/ASI to break free of the control of state governments could become adversarial to the United States, even if the individual is a citizen. Super-individuals often control resources on par with or greater than some traditional nation states. The recent proposed purchase of Twitter for $44 billion by Elon Musk[144] represents a single cash payment larger than the annual GDP of over half the countries in the world.[145] The Forbes list of billionaires lists 200 super-individuals with

[144] Clare Duffy, "Elon Musk to buy Twitter in $44 Billion Deal," *CNN Business,* April 25, 2022, https://www.cnn.com/2022/04/25/tech/elon-musk-twitter-sale-agreement/index.html/.

[145] "GDP by Country," Worldometer, accessed March 13, 2022, https://www.worldometers.info/gdp/gdp-by-country/.

fortunes over $10 billion,[146] an amount comparable to the annual GDP of countries like Madagascar and Mongolia. A super-individual would certainly have access to sufficient resources to employ an ASI's capabilities to their greatest extent and achieve strategic effects without the requirement of government organizations or infrastructure.

Ends, Ways, and Means

The scenario in this thesis utilizes a super-individual as the controlling agent for the first ASI with the end state of elevating their non-state organization to state level status and replacing the United States as the dominant world power, solidifying this status by reaching technological singularity. In order to achieve this end state, it is inferred that the super-individual acting as the controlling agent for an AGI/ASI would pursue the following objectives: protect the ASI from external attacks and potential threats; prevent ASI competition by delaying, degrading, or preventing the development of AGI and ASI by other actors; degrade the influence and military capabilities of great power competitors; and maximize rate of intelligence improvement to progress from AGI to ASI and then towards singularity. While this scenario utilizes a super-individual as the controlling agent, these objectives would be equally relevant in scenarios where a traditional nation state is the controlling agent in pursuit of a similar end state. These objectives are depicted in Figure 6 with a set of relevant sub actions and/or intermediate objectives. Analysis for each objective is provided below.

[146] Kerry A. Dolan and Chase Peterson-Withorn, eds., "World's Billionaires List," *Forbes,* accessed April 7, 2022, https://www.forbes.com/billionaires/.

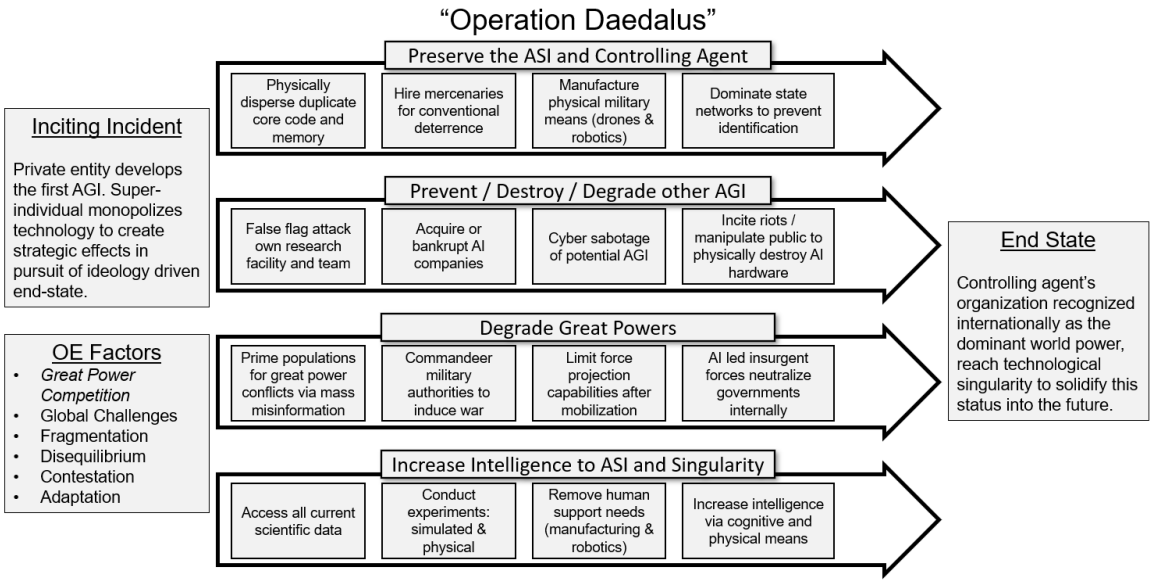

Figure 6. Operational Approach for Adversarial Employment of ASI

Source: Created by author.

Preserve the ASI and Controlling Agent

Preserving the ASI is the highest priority objective because the ASI cannot be employed in pursuit of any other objectives, including the end state, if it is destroyed. Preserving the ASI refers to both the physical requirements of the ASI such as hardware and electricity, and the digital components of the ASI such as ensuring the integrity of the code base that facilitates the intelligent behaviors and access to information and memory. The duplicability and memory sharing advantages of digital of intelligence will be particularly useful in pursuit of this objective.

In order to safeguard against physical attack, an ASI would duplicate its core code base across several hardware systems each located in different physical locations. These duplicate ASIs would want to remain synchronized in thought and would utilize a single shared memory source that they each pull from and push to. This memory source would

also be backed up at regular intervals to multiple sets of hardware, similarly distributed across multiple physical locations. An alternative method is for each ASI duplication to maintain its own copy of the shared memory bank—for increased speed of access to the data—that is regularly synchronized with each of the other memory banks to maintain congruity. This distribution of identical physical copies would make targeting the ASI with conventional attacks extremely difficult as all copies would need to be destroyed—either by kinetic strikes against the ASI hardware or sources of electric power—within a timeframe shorter than a single copy can be reconstituted. Depending on the speed of data transfer rates and availability of hardware this window could be extremely short, in the realm of hours.

Duplication could also serve as a method of defending against digital attacks; however, an ASI would most likely also employ a subordinate AI whose sole function was to monitor for cyber intrusions or altercations to the core ASI code base. Duplication would assist in such an effort by providing multiple examples of the core code base for comparison. If one copy of the ASI suddenly manifests differences compared to the other copies, this would be an indicator of compromise, and the affected copy would be overwritten by one of the other copies and wiped of the invasive code. Such redundancies and safeguards mean that a digital attack would also have to simultaneously target all copies of the ASI in order to be successful. Such an attack may be completely infeasible if a copy of the codebase and memory is made at regular intervals and stored on hardware that is not physically connected to digital systems except when queried by the subordinate AI. All “off the grid” copies would need to be destroyed at the same time the other copies are attacked via digital means.

In order to maximize the security of ASI duplications the controlling agent would prioritize distributing the physical locations in multiple countries. Any actor seeking to destroy the ASI would need to deal with the logistical and diplomatic challenges associated with carrying out attacks in multiple countries around the world simultaneously. These challenges would be further increased if the ASI hardware is housed in locations of economic or cultural importance within the host nation. The amount of resources required for such an operation would be extremely high with a comparably high amount of risk, compounded by the fact that it would be nearly impossible to know if all copies of the ASI are being targeted as part of the operation. If a single duplication of the ASI survives it will be able to reconstitute any number of new duplications, assuming its controlling agent still has the resources to house them.

Duplication means that targeting an ASI directly is extremely unlikely to ever be successful. Instead, actors seeking to neutralize an ASI will likely attempt to identify and target its controlling agent instead. This means the controlling agent will require conventional deterrence. An ASI controlled by a traditional nation state will be protected by the government of that state and all its associated military deterrence capabilities. While a super-individual would not have access to a traditional military force, they could employ private security forces or mercenary military groups, such as the infamous Blackwater group. Any conventional force employed in defense of an ASI or its controlling agent would also gain significant advantages in a fight, benefitting from ASI tactical command, intelligence on threat plans gained through network dominance, the potential to gain control of hostile drones or automated systems through network dominance, and technological advantages in their weapons and equipment provided by

ASI research. The controlling agent may also seek to invest in automated weapon systems including drone swarms and robotics that the ASI could control either directly or with subordinate AI to maximize the advantages of using the ASI as tactical commander. Moreover, an ASI may orchestrate small conflicts in regions of little strategic interest to the major powers in order to gain data and experience in employing forces tactically or to test new military technologies under real world conditions.

An ASI would also engage in active defense. Utilizing global situational understanding, network dominance, and strategic predictive analysis, an ASI would be constantly attempting to identify threats—in both the present and future—and actively mitigate them as early as possible to minimize risk. Traditional nation states with significant conventional militaries, cyber capabilities, and intelligence networks would likely be perceived as the greatest threat and therefore be the focus of active measures. Preventing large capable states from discovering that an ASI has been developed and is operating would serve as the best form of protection from attacks—organizations cannot oppose forces they do not know exist—so initially active activities would likely be restricted to surveillance through network dominance and global situational understanding. Limited actions may be taken to alter or destroy information that could allow government agencies to discover the ASI. If discovered, an ASI will have a host of options available to it but would prioritize efforts to prevent the state from perceiving it as a threat, concealing the identity of its controlling agent, concealing its physical location(s), concealing its cyber IP addresses, and preparing options to prevent offensive actions against it.

Prevent, Destroy, or Degrade other ASI

ASI can only be employed as an asymmetric advantage if it cannot be contested by another ASI. An opposing ASI would be the best countermeasure against network dominance, which is a requirement for most tier two and three ASI capabilities and significantly enhances other tier one capabilities. An opposing ASI actively working to prevent network dominance would also be the most likely means of determining the cyber and physical location of the attacking ASI. Even if an ASI is not in a role of active opposition it must be assumed that it would also be pursuing recursive self-improvement, seeking to become more intelligent and capable than its opponent. This makes other ASI one of the greatest threats to achieving the desired end state.

The controlling agent will take steps as early as possible to prevent competition in the AGI/ASI space. There are several methods that could be employed towards accomplishing this objective:

1. Decrease the talent pool of relevant AGI/ASI researchers through aggressive hiring of top talent to include university professors.
2. Corporate acquisitions of companies conducting relevant research to acquire research and talent. Reassign talent to non-AGI related projects and destroy/alter the acquired research.
3. Cyber-attacks or network dominance to sabotage active AGI/ASI research and alter/delete published literature.
4. Killing, kidnapping, framing, or other nefarious means of removing personnel critical to ASI research teams via behavior manipulation, deep fakes, and real time impersonations.

5. Physical attacks to destroy ASI research facilities and data servers via behavior manipulation of personnel with access, covert mercenary operations, or commandeering law enforcement to seize assets.

It should be noted that the controlling agent of an ASI by necessity must have in their employ at least one team of researchers that made the original discovery/advancement of AGI. These researchers would constitute a risk to the objective of preventing any other AGI from coming online, especially if any of them harbored ideological beliefs about proliferating the technology for the good of mankind. Placing legal restrictions on the team through non-disclosure agreements and implementing extremely tight security measures may not be enough to prevent leaking of the technology. At least some of the proprietary information would exist in the minds of the researchers. For this reason, after securing the original AGI, the controlling agent may target his own research facility and team members to prevent their knowledge inadvertently aiding in the creation of another AGI. An individual willing to monopolize AGI/ASI in pursuit of personal ideology is unlikely to have qualms with such a betrayal.

Degrade Great Powers

Achieving the end state of replacing the U.S. as the dominant world power requires the international community having reason to believe the controlling agent and their organization has more power to influence world events and project power than the U.S. Employing an ASI to cause significant strategic damage to the great powers would provide a demonstration to the world that serves both as proof of the power of an ASI and credibility that the controlling agent is able and willing to use it to create strategic effects.

It would also serve the added benefit of creating an effective deterrent for any other states considering challenging the controlling agent.

The best approach to degrading the power and influence of the great powers would be to have them exhaust resources fighting each other. An ASI would employ narrative dominance to escalate tensions between nations and spread the idea that conventional war is inevitable. Followed by a commandeering of legitimate military authorities, through the use of network dominance and real time impersonations, the ASI could manufacture an armed conflict between the great powers. After starting a conflict, the ASI would seek to covertly continue manipulating both sides in order to maximize the destruction of military capabilities for both states in as short a time as possible. Rapid escalation of the conflict in a short time period should force the powers to mobilize reserves. The ASI would seek to manipulate both states into pursuing war objectives that require power projection outside their home territory seeking to draw away as much combat power as possible. Once it was determined that most of the states' available military resources were committed the ASI would undermine the states from within by creating and employing insurgent forces from their populations. These insurgencies would be synchronized across time and space by the ASI to seize seats of government, arrest politicians, control critical infrastructure, neutralize law enforcement, sabotage or steal remaining military resources, and effectively neutralize the great powers for a period of time, if not permanently.

While the great powers wage war against each other the ASI would also be creating effects throughout the rest of the international community. Narrative dominance would utilize the war to showcase how the great powers are too volatile and violent to

lead humanity in solving their great collective challenges. False narratives about war crimes, corruption, collateral damage to other states, and the negative intentions of both sides should they be victorious would drive international sentiment for the great powers down. The ASI would coordinate simultaneous attacks on embassies and businesses belonging to the great powers around the world timed to align with the start of the planned domestic insurgencies. Major events being synchronized in time makes them significantly more difficult for governments to react to and also provides a clear indicator in hindsight that the events were not spontaneous coincidences but rather an orchestrated occurrence. This will lend additional credibility to the power of the ASI once it chooses to reveal itself.

Manufacturing a conflict between the great powers would have additional benefits for the ASI. The conflict would provide a convenient target for the ASI to deflect blame onto for other activities it might engage in as part of degrading their capability or preventing other AGIs. It would also provide a clear and present danger for the great powers to focus their intelligence networks on, potentially reducing the risk of discovery for the ASI and buying more time to increase in intelligence and progress towards technological singularity. Destroying or neutralizing large amounts of the great powers' military resources would also limit their ability to attempt an attack against the controlling agent should states discover the ASI and its activities.

Increase Intelligence to ASI and Singularity

As discussed in the literature, the first AGI would have the potential to advance to ASI due to its ability to apply knowledge across domains, make inferences, understand the world through perception, and benefit from the advantages of digital intelligence. To

progress the AGI to ASI, the controlling agent would need to provide it with sufficient data for it to learn from in order to gain expertise in a multitude of knowledge domains. This process could initially progress by simply accessing information freely available on the internet. The controlling agent would likely then prioritize AGI learning in the area of computer science in order to facilitate the AGI gaining access to networks, proprietary research, and advance towards the skills required for network dominance. Additionally, expertise in computer science would form the basis of the AGI understanding how its own code and hardware works, allowing for the initiation of the recursive self-improvement process.

While the AGI/ASI would initially be dependent upon the controlling agent for assistance conducting real world experiments to validate simulations and manufacture new technologies, an ASI would likely pursue the means of removing humans from the process of recursive self-improvement in order to increase rate of development and progress. Advances in robotics that could be controlled directly by the ASI would provide it with a method of interacting with the physical environment to both conduct tests and fabricating prototypes. If supplied with sufficient materials, an ASI could enter a closed loop of self-improvement where it uses robotics to upgrade its own hardware leading to an increase in intelligence, which allows research of increased complexity that with testing leads to greater levels of technology used to fabricate yet more complex robotics capable of conducting even more complex experiments. Allowing an ASI to enter a self-sustaining improvement loop would be a major decision point for the controlling agent as access to robotics with fabrication capabilities could provide a means for the ASI to break free of its dependency on the controlling agent for resources as well. If the critical

assumption that the ASI does not possess internal motivations proves false, this decision is a decisive point where an ASI could break free of human control. As such the controlling agent should be expected to take significant security measures to hedge against this possibility.

Whether the recursive self-improvement process utilizes humans in or out of the loop, the process will be resource intensive. Because attacking an ASI directly poses extreme challenges due to duplication, attempts to compete with an adversary ASI would likely focus on denying its ability to self-improve and increase in levels of intelligence by denying it the resources and/or facilities required to do so. Disrupting the exponential curve of self-improvement would allow a competing ASI time to catch up or even surpass the original in level of intelligence. Competition between ASI with opposed controlling agents could continue in this way until one of them reaches the point of technological singularity beyond which continued competition is likely impossible. Singularity then represents a sort of finish line to any ASI contest, providing who ever achieves it first too great of an advantage to make continued competition feasible.

Increasing an ASI's level of intelligence through the process of recursive self-improvement will provide additional benefits to the controlling agent throughout the process. Increasing ASI expertise in STEM fields like physics, chemistry, engineering, and math well beyond human levels will produce new technologies which could be employed to gain a technological overmatch in military capabilities or increase economic prosperity. Both choices would provide options for increasing the controlling agent's perception of power amongst the international community. Research into new

technologies would be prioritized along lines likely to produce results that could be employed in helping achieve the other objectives of the operational approach.

FICINT Scenario: Operation Daedalus

The following section is a series of vignettes that provide a narrative description of some of the key events of the Operation Daedalus scenario. The events are described through the lens of an individual experiencing them and highlight the confusion associated with dealing with the effects of ASI capabilities. To aid the reader, events that are a result of ASI capabilities being employed or in pursuit of a specific objective are annotated within a set of brackets.

16 January 2036; Mountain View, California: Dave, an AI researcher, walks through the charred remains of his lab. The destruction is heart breaking. He and his team had been on the verge of a breakthrough in their research. Achieving an AI virtually indistinguishable from a human mind had been the holy grail of his profession for a century and his team had prototyped the first AI to potentially reach that benchmark. It was going to be revolutionary. At least it would have been if protestors had not burned his lab to the ground along with a large portion of his parent company's campus [Behavior Manipulation].

Dave left the lab and headed to his temporary office space in a building that had not been part of the fire. There was a chance that not everything was lost. All of his local data servers and advanced processors that were used in his research were lost, but a lot of their work had been backed up to the company cloud which stored information redundantly on multiple serves. But he could not find any of it, and neither could tech support. They could not find any trace of it in the current or historical backups or git

commits. With each call he makes, a sense of dread steadily grows in his gut. His sister research team in London had coincidentally been attacked over the weekend and lost all of their research as well as the servers that stored his research. Their research, their data, their hardware, all their progress and knowledge was gone [Network Dominance]. Everything, except what was in the heads of him and his team [OBJ: Prevent Adversary ASI].

Dave drove home that night wondering if he would be able to retain his team after the heartbreak of seeing everything they had worked towards for fifteen years disappear overnight. His field had already been dwindling over the past decade with AI talent all being put to work where the money was, which was always some narrow application to try and create quarterly profits. Even the Universities were having a difficult time keeping their programs staffed with qualified professors, creating a shortage of new scientists capable of doing the work. His team would be in high demand and without a purpose; they might not stay [OBJ: Prevent Adversary ASI].

As Dave parked in his driveway and got out of his car, his thoughts were suddenly interrupted. A swarm of large men in blue jackets were running at him, guns drawn. He froze, unsure what was happening or what he should do. One of the men twisted his arm behind his back, slammed him to the ground, and slapped a pair of handcuffs on him. He was under arrest they explained, but he had not so much as stolen a stick of gum since he was seven [Commandeering of Legitimate Authorities]. The whole ride down to the station he tried to explain to them that they had the wrong guy. His protests continued right up until a man in a suit opened a folder in the department's interrogation room and started showing him pictures of… himself. Doing terrible, terrible things. Things, he had

never done. But there he was. Clear as day in the photos. Dave lived alone. He had no alibi. The growing sense of dread in his gut grew into an unfamiliar feeling that was much, much worse [Individually Tailored Disinformation].

03 February 2038; New York: It was not every day that the National Security Agency got a request for assistance from the liberal press that was usually so critical of them. The agents tasked with responding had assumed that this was just a ruse to get them to answer seemingly unassuming and irrelevant questions that would ultimately incriminate the agency. But as Agent Smith looked around at the faces of every person in that conference room all he saw was gloom and doom. These people were scared.

The man that had introduced himself as the editor in chief started the meeting by introducing one his reporters. Ms. Newsome explained to the group that a few months ago she had been nominated for a Pulitzer Prize for an article on political corruption and how some of the most senior members of the legislative and executive branch were not just in the pocket of the Chinese but potentially acting for the Chinese national interest. Agent Smith remembered the article; it had caused a stir in the department but after investigation appeared to be a complete work of fiction. Odd that such an uncredible story would be nominated for a Pulitzer or that such an untrustworthy author would remain employed at a prestigious paper. However, Ms. Newsome went on to explain that she never wrote it. The editor stepped in to explain that not only did she not write it, but no one in his organization ever published it. And yet, it was printed both physically and on their official site and millions of people had read it, including Agent Smith. No one in the newsroom had realized what was going on until Ms. Newsome had a friend congratulate her on the nomination at a party [Network Dominance].

This event caused the paper to launch an internal investigation and over the last six months they had uncovered hundreds of fake articles using the credibility of their authors and their paper. The editor was asking for help from the NSA to plug their security issue because their staff and contracted experts had so far been unable to do so [Narrative Dominance]. They now had a team whose sole job was to constantly check their systems, presses, and forward-facing websites for new fake content. The editor proposed that if this was happening to them, it was very likely happening to other papers as well. Most importantly of all, the content of the fake news being posted showed two trends: disenfranchisement with the United States government as corrupt and hostile to the will of the people and that war with the Chinese is righteous, inevitable, and beneficial. Agent Smith only had one question at the end of their presentation: why the hell didn't you call us sooner?

04 April 2040; People's Liberation Army Air Force Base on Fiery Cross Island: General Lieu got the call in the middle of the night on his secure line. The voice on the other end was one he knew well and trusted above almost all others [Real Time Impersonation]. A good thing too, for what his old comrade in arms told him seemed beyond belief. The American's hubris and jealousy had finally taken hold of their senses. Forensic evidence suggested they were behind the recent assassinations of party officials and intelligence had intercepted their follow-on plans [Individually Tailored Information Campaign]. They wanted to take advantage of the temporary national shock to strike blows to their national pride and integrity. As they speak American fleets are on their way to seize control of the sovereign territory they had created in the South China Sea, including the air base under his command on Fiery Cross. This must not be allowed to

happen, but neither can they allow the Americans to draw them into an unnecessary war. They need to dissuade them of their foolish plan by showing them that the party will not be caught off guard and is ready to employ overwhelming strength in defense of their nation. They are going to launch all wings, organized into groups armed with hypersonic missiles to destroy the enemy floating fortresses and air to air missiles to destroy their fighters over the horizon should they be so foolish as to fire first or ignore their demands. The First People's Carrier Fleet was diverting back to the area but until then, General Lieu's mission was to keep the Americans from crossing the nine-dash line [Commandeer Legitimate Authorities].

04 April 2040; Carrier Strike Group, Pacific Ocean Southeast of Taiwan: Admiral Green also got a call on his secure line in the middle of the night. He could not remember another single instance when his four-star Combatant Commander had called him directly during an off cycle [Real Time Impersonation]. No pleasantries, his senior office told him very directly in no uncertain terms that they were under attack. His sister fleet operating 1,000 miles to the southwest near Vietnam was under attack by the Chinese. Reports are clear that it was not an accident, the Chinese had deployed a substantial mass of combat power to deal a decisive blow. He needs to get sensors in the air as far forward as possible right now and do everything in his power to prevent Chinese aircraft from getting within supersonic missile range. If you see them coming, his commander told him, you need to kill them before they get in range or you're sunk [Commandeer Legitimate Authorities].

Admiral Green stormed onto the bridge bellowing orders: all hands on deck, prepare to scramble all fighters, prime missile countermeasures, and prepare for combat.

An Airborne Warning and Control System (AWACS) aircraft was already loitering and he ordered it to move southwest at maximum speed to gain a radar picture extending at least 600Km between their fleet and the closest Chinese airbase. The next hour moved at a snail's pace and then suddenly at the speed of light. The AWACS was detecting several wings of fighter and bomber aircraft moving directly towards their position in combat formation. It would only be a few minutes before the first formation would be within hypersonic missile range of the fleet. Admiral Green picked up his comms to speak to the squadron commander flying lead on the interdiction mission he had just launched, but got nothing. Were his comms being jammed? [Network Dominance] He was not sure if he should utter a prayer or a curse. His men knew the situation and their orders. The only question now was: would they hesitate when they found themselves at the decisive point? [OBJ: Commandeer military authorities to induce war]

13 November 2040; Huntsville, Alabama: Deep down Frank had always known he would end up having to defend his homeland from communists. Despite the warnings he and his had been shoutin' for decades, they had infiltrated nearly every aspect of their once great civilization. The war with China had just shown a spotlight on what he had already known was lurking in the shadows. Now, it was undeniable. The current leaders of government were communist agents [Individual Information Campaign]. Their actions since the war started were proof of that. Every day more images and videos of American kids getting killed and mangled by Chinese long-range weapons but forbidden from launching their own attacks onto Chinese soil. The supposed commander-in-chief content to let their forces group up like fish in a barrel for the communists to kill off a little at a time. Their once great military was being whittled away while the elites bode their time,

probably just waiting for the right moment to announce their surrender [Narrative Dominance]. You did not see this kind of cowardice from the European Union or the Russians duking it out in Eastern Europe [OBJ: Degrade the Great Powers].

Frank remembered with pride the day his patriotism had been noted by his state senator [Real Time Impersonation]. That great American had been trying to fight the hidden communists in their government for years. Frank saw him on his news feed all the time blowing the whistle on some new socialist agenda or corrupt politician. When the senator's assistant had reached out to him saying the senator was putting together a top-secret coalition of patriots, Frank had been honored to join. He followed the link they sent and installed the top-secret, ultra-encrypted, probably military, messaging app on his phone. Just a few hours later the senator had video called him directly via the app to inform him of the plot to destroy the Constitution, sacrifice their military to the Chinese, and seize control of the government for their Chinese masters. It could not be allowed. The time to stand up to tyranny had come.

Frank had been designated as a first-class officer in the forming patriot force and had been receiving tasks to accomplish via the app on an almost daily basis [Create and Employ Insurgent Forces]. Often it was things like gathering supplies, going out to rifle ranges and training small groups on shooting, watching videos on how to communicate with hand signals or conduct small team tactical maneuvers. He even got instructions to set up a wireless camera to a helmet mount and sync it with his app. All sorts of stuff. But now it was finally go time. The last week had been a flurry of intense actions. Using the app to coordinate all their actions the patriot force raided gun stores, police stations, and national guard armories across the state. They stole all the weapons, ammunition,

equipment, and vehicles that they could get at with bolt cutters or a blow torch and sabotaged the rest. Frank had expected the police to be a real problem, but it turns out a bunch of them were part of the patriot force and they had helped the senator create a statewide list of those likely to be a problem [Operational Predicative Analysis]. Most non-patriot officers simply found themselves off duty or on some special assignment that would make them unavailable to any response on the nights of their raids. But there were still quite a few, mostly higher ups, that had needed to be detained until the operation was complete. The Senator had approved the arrest of many officers and assigned small teams of patriots via the app to arrest them while they were off duty. Most were taken to a location off the grid, for their own safety; but Frank had heard of at least one arrest that had turned fatal for a couple patriots and the officer.

More American blood spilling was regrettable, but Frank knew that the worst was about to come. It was not what any of them wanted, especially because most of those serving the will of the traitor politicians were honest folk just doing their job. But, if it came down to it, some sacrifices might have to be made in order to save their great republic. As Thomas Jefferson had said, "The tree of liberty must be refreshed from time to time with the blood of patriots and tyrants. It is its natural manure."[147] Today, they would seize back control of their government from the would-be tyrants, and they might have to kill some of their own countrymen to prevent them from unknowingly interfering

[147] Jefferson to Smith, November 13, 1787, in *The Papers of Thomas Jefferson*, vol. 12, ed. Julian P. Boyd, Mina R. Bryan, and Fredrick Aandahl (Princeton, NJ: Princeton University Press, 1955), 355-357.

in their own salvation. Their deaths would be added to the list of crimes laid at the feet of those in Washington [OBJ: Degrade the Great Powers].

The Feds also were not a problem. The Senator was using his insider information to keep them two steps ahead of investigations that might uncover something. Frank himself had once received a direct notice through the app that one of the patriots coming to the shooting training event he was hosting that day was actually an undercover FBI agent trying to gather intel on potential domestic terrorist activity. The Senator even went so far as to get, what Frank assumed must be another special agent who had been recruited to the cause, to coach him live on what to and not to say when the snitch was around via a wireless earpiece connected to the patriot app in his phone. They were always one step ahead.

Frank's phone pinged and he looked down to where it was now strapped to his forearm for easy access. On it he saw a map of the city centered on his fire team's objective. A kick swipe and he was able to see the larger situation update. The ping was to inform him that all teams were in place and that they had five minutes for final preparations. Frank could see hundreds of green dots throughout the city, each a patriot team ready to do their part in the liberation. They surrounded every government building, courthouse, and other critical tactical points in preparation for a simultaneous decisive blow [Tactical Command]. Frank signaled to his team to check their helmet cams, earpieces, and weapons. In five minutes, they would bring freedom back to America.

22 November 2040; United Nations: It was unusual for corporate executives to be given the opportunity to speak at the United Nations, especially during a time of crisis when revolutions were threatening governments around the world and war still raged

between the world's great powers throughout the northern hemisphere. It spoke of just how much power and influence Mr. Diggs, an international tech company CEO, really had. After becoming CEO after years as CTO, his leadership had resulted in one of the world's largest corporations leading the charge to solve many of the planet's shared problems like climate change. His great work over the last few years investing in international communities had lifted many places out of poverty thanks to the influx of manufacturing jobs. The new technologies developed by his company were also solving all sorts of scarcity problems that had haunted poorer countries even more in recent years due to droughts and supply chain issues from international conflicts.

However, Ambassador Brown could not have predicted this speech. The recent success of the countries in Africa were the result of a partnership between their governments and Diggs's most important technological breakthrough: An Artificial Super Intelligence he called Karellen. Diggs then left the stage, indicating that Karellen would address the crowd itself, which caused quite the commotion before the screen in front of each of them turned on to reveal a friendly looking female face smiling at them. Ambassador Brown felt an odd familiarity with the person even though he was confident he had never met her before. Looking around he realized that each of his colleagues had a different face on their screens, each reflective of the country from which they heralded. Over the next hours, the AI explained to each of the ambassadors in their own language how human leadership would always result in what they were seeing today. The great powers using their wealth and technological advancements not to help their citizens to lead the most fulfilling and meaningful lives possible, but to send their best and brightest to kill and destroy. Not to come together to solve the problems that effected the world,

but to fight for a shrinking slice of the pie. Human leaders could not be trusted to see the bigger picture and do what was best for all humanity. It showed them evidence of how it had brought the poorest of communities into prosperity and how it could do the same around the world. It asked them all to trust it to create the better future they could not.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

The Operation Daedalus scenario demonstrates how an ASI could plausibly be employed to depose the United States as the dominant world power. Additionally, it demonstrates how ASI could enable a single super-individual to achieve strategic effects on par with or greater than the current great powers while expending significantly less resources to do so. The scenario demonstrates the far-reaching impacts of ASI on all aspects of society. Particularly as the next military revolution, transforming the primary means of achieving effects away from states possessing capabilities that can deliver kinetic effects to the ability to affect the minds of individuals and organizations, compelling them to take actions counter to their own best interest.

The Operation Daedalus scenario of this thesis forecasts several effects that have catastrophic and even existential impacts to the strategic interests of the United States. These effects center on exploiting vulnerabilities in three centers of gravity for U.S. power: security of its strategic support area, command of military forces actively deployed throughout the world, and diplomatic and economic presence in most countries around the world. The security of the U.S. strategic support area—the homeland and its associated military facilities, logistical hubs, equipment manufacture, and government agencies—was compromised in this scenario by targeting a subset of the population prone to misinformation, radicalization, and already equipped for para-military operations. Command of deployed U.S. military forces were commandeered through the real time impersonation of legitimate military commanders and made to engage in actions that pulled the U.S. into a major armed conflict with a near-peer competitor. Finally, U.S.

soft power around the world was neutralized through narrative dominance targeted to the specific cultures and individuals instilling an anti-U.S. sentiment followed by behavior manipulation to synchronize attacks on U.S. embassies and international economic interests. Utilizing the Military Strategic Risk Matrix from CJCSM 3105.01A, as described in chapter 3, these sources of risk combined with the strategic value of the centers of gravity result in calculated strategic risk of the two highest orders: extreme and major. This assessment is summarized in table 2 below.

Table 2. Assessed Vulnerabilities and Associated Strategic Risk from ASI

Center of Gravity	Source of Risk	Strategic Value	Impact of Effect	Degree of Risk
Security of Strategic Support Area	Domestic Rebellion/Insurgency Loss of Force Projection Capabilities	Vital	Existential	Extreme
Command of Deployed Military Forces	Military Forces Commandeered Theater War or Major Armed Conflict	Global	Catastrophic	Major
International Diplomatic and Economic Interests	Anti-American Narrative Dominance Integrated Regional Attacks on U.S. Embassies and/or Businesses	Regional	Catastrophic	Major

Source: Created by author.

Extreme and major risk was assessed in this thesis from only exploring ASI capabilities that simply require current technology and an internet connection. However, advances in robotics, nano-machines, biology, and other technologies will create a whole host of additional means and methods an ASI could employ to achieve strategic effects. This thesis presents a plausible scenario for how an ASI could inflict existential effects

on the U.S. without even developing any new technologies. If the Army wants to prepare to fight the next war, it needs to invest in and prepare for what conflict will look like when AGI and ASI are involved. This means cognitive warfare, not kinetic warfare. Cognitive warfare bypasses traditional security measures that rely on the control of physical spaces. Instead, ASI targets the minds of people directly, shaping the behaviors of individuals and the organizations they belong to. It attacks objectives from within.

In an effort to prepare for Large Scale Combat Operations, or the next kinetic war, the Army and Department of Defense at large, is investing in projects and initiatives that would increase our vulnerability in a cognitive centered conflict with ASI. More and more reliance on digital systems, centralized command and control structures, and autonomous platforms all increase vulnerability to an ASI. Additionally, here is a significant strategic planning gap that exists due to unstated assumptions surrounding the security of the centers of gravity effected in this scenario. The American way of war is predicated on the idea that the U.S. fights its conflicts on terrain other than its homeland. The security of the strategic support area is always assumed but is a critical requirement without which the U.S. loses its ability to project combat power or conduct major operations. If the homeland is sufficiently compromised the entire government could be rendered inert.

Similarly, there has never been a historical reason to assume that using remote communications to command military forces deployed around the world would create a vulnerability that allows for those military forces to be commandeered. However, as demonstrated in this scenario an ASI could both commandeer our military forces to employ them in pursuit of its own objectives and seriously undermine the security of the

strategic support area to prevent the necessary functions required for force projection. The space between the two types of predicted conflict in Figure 7 represents an estimation of the planning gap that exists between an ASI conflict and the next war current U.S. Army doctrine is preparing for, which is significant.

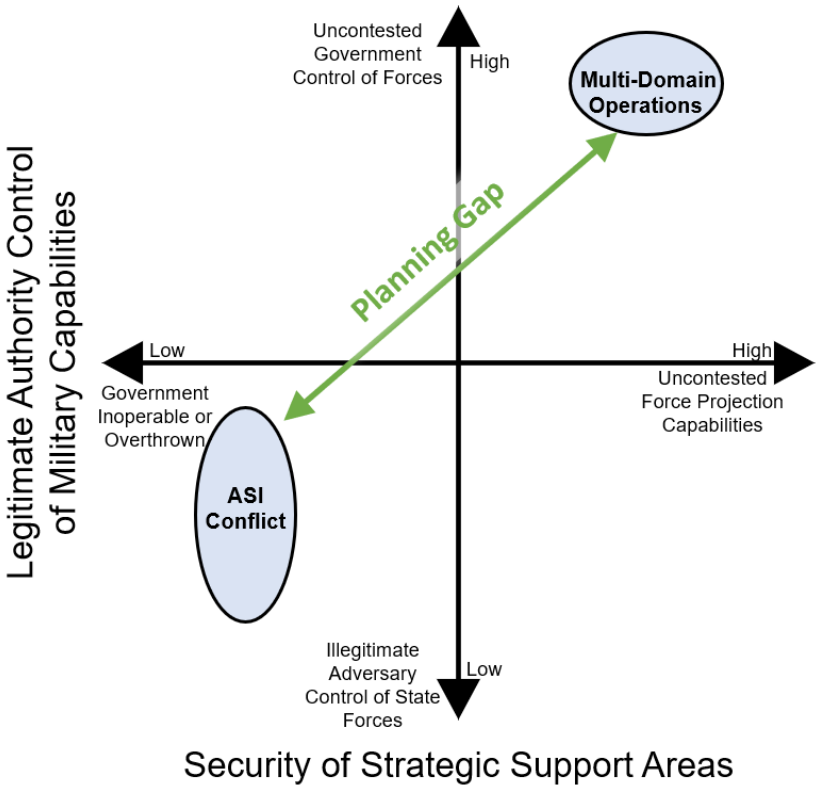

Figure 7. Theoretical Future Planning Gap

Source: Created by author.

The ability of ASI to bypass traditional security and achieve effects through cognitive warfare suggests that current theories about deterrence will be rendered obsolete. The current difficulties that governments are having with updating deterrence models and theory to deal with cyber attacks reinforces this notion. How does a nation

respond with proportionality if no shots are fired? These concerns are only further muddled by the fact that every aspect of human society stands to be impacted by the arrival of AGI and ASI. Governments need to consider these potential impacts and their implications for international competition and cooperation. Most importantly, however, is establishing regulations on the use and sale of AGI before it is developed. Given that AGI is likely to be developed fist in the private sector, the owning company will unknowingly be making decisions that are pivotal to the future of the entire international community. To what ends will that private company employ an AGI, what will they teach it, how will it verify that it can be controlled, what safeguards will they employ, and most importantly: to whom will they sell it? If AGI is not regulated before it is created, then the decisions about one of the most decisive points in human history could be determined based on the profit motive of a private company.

Additional Research Recommendations

The ASI capabilities employed to achieve effects in the Operation Daedalus scenario are all either dependent upon or greatly enhanced by Network Dominance. This creates significant implications for how the modern world's reliance on digital information systems and cyber networks will become a massive vulnerability. We are increasingly moving more and more of our societal functions into the digital realm, a realm where it is impossible for us to compete with a digital intelligence. These shifts are driven by natural market forces because they create great efficiencies and utility for daily life.

Unless there is a fundamental change or breakthrough in how we secure digital information or protect networks that renders cyber-attacks impossible, then everything

we put into cyberspace should be considered at the disposal of a future ASI, friendly or enemy. Militaries around the world, including the U.S., are only accelerating this trend towards digitizing capabilities and increasing automation; however, each step in that direction further increases risk and vulnerability to ASI. Manned Unmanned Teaming provides a great tactical advantage until an ASI with Network Dominance hijacks the UAS to turn its weapon systems on its manned partner.

This acceleration seems unlikely to change—even if strategic decision makers acknowledged the threat posed by ASI— because ASI is still theoretical and militaries around the world are gaining advantages by automating their militaries right now. Any military seeking to mitigate the threat of ASI by reducing dependence on digital and autonomous systems could become overmatched by competitors in the short and medium term that do pursue such technologies. For this reason, additional research is recommended to determine if there are force structure options that can accommodate success in the medium term against forces with automation advantages without a reliance upon systems that could be exploited by ASI network dominance. Additional research is also recommended to determine how military commanders should communicate and authenticate orders given two assumptions: an adversary can see and alter any information on or connected to digital networks and can perfectly impersonate military leaders over digital communication systems.

Network Dominance may prove so invasive and difficult to counter that fully severing internet connections locally, regionally, or globally may be the only viable option to prevent the effects of tier two and three ASI capabilities it enables like global situational understanding, narrative dominance, and the generation and employment of

insurgent forces. The degree of societal dependence on the internet would make such a counter measure an extremely impactful decision with significant implications for nearly all aspects of modern society. Additionally, fully removing internet access will likely pose a significant challenge given the rise of mobile networks and space based wi-fi service providers. Additional research is recommended to examine the impacts, risks, and challenges associated with executing a country wide internet blackout for various periods of time and methods that could be employed to achieve such a blackout, such as full power outages.

In the Operation Daedalus scenario, a super-individual instigates conflict between the great powers and actively seeks to prevent other AGI in order to avoid a singularity race. However, in other scenarios where multiple states are pursuing ASI independently, a singularity race may increase the chances of conflict on its own accord. Competition, as opposed to cooperation, assumes winners and losers, and the degree to which a nation stands to lose if their competitor achieves a technological singularity is likely to be perceived as near totality. The further ahead a state gets in a singularity race will likely correlate to the severity of means their competitors would be willing to employ to halt or reverse their progress.

This risk of conflict as a result of ASI will increase if states do not take the risk of ASI seriously in their strategic planning. The universal scientific consensus on the threat of climate change has not spurred governments to act in a way congruent with the degree of potential risk. Likewise, it is plausible that states will not perceive enough or any risk from ASI due to its theoretical nature and take no action until one party has already obtained a sizable lead in progress towards singularity. The shock of sudden realization

may push states into taking more drastic action than they would with a slow continuous understanding of the increasing risk.

International cooperation on the pursuit of AGI/ASI provides a means of ensuring the technology does not result in competition and inadvertently lead to conflict. However, even high degrees of cooperation will still have challenges associated with the natural suspicion states have for each other, especially great power competitors. If any state takes the threat of singularity seriously it is likely to engender such a degree of paranoia as to make conflict nearly inevitable. While cooperation is likely to have challenges it is intuitive that the degree of paranoia and suspicion would be lower than if competitors are developing the technology in classified environments. Taken together, these conditions suggest an inverse relationship between the degree of international cooperation on ASI research and the probability of conflict between major powers as a result of ASI. Additional research is recommended into the plausibility, benefits, and challenges of international cooperation on ASI as a joint venture with competitors.

This thesis only examined how ASI capabilities could affect the military and information domains of national power. However, there are likely to be significant risks and vulnerabilities associated with the economic domain as well. Network dominance and global situational understanding have obvious implications for their ability to manipulate stock markets and steal funds from digital banks, crypto currency exchanges, and online markets. Individual information campaigns targeted at investors, CEOs, or shareholders could have large impacts and lead to huge market manipulation. Additional research is recommended to explore scenarios for how an ASI could be employed in the economic realm.

Finally, this thesis only examined the strategic implications of an adversary to the United States being the first to obtain AGI/ASI. These implications are likely to be completely different if a country that is neutral towards, friendly, or allied with the U.S. is first instead. There are also major implications worthy of exploration for how the U.S. should act if it or a U.S. based company—that is not owned by an adversarial super-individual—is the first to develop AGI/ASI. How would the rest of the international community react? Additional research is recommended to explore the implications of the United States or non-hostile nations being the first to obtain AGI/ASI and how a democratic state is likely to utilize the technology, in particular if it is developed commercially.

Final Thought

AGI and ASI are coming. Regardless of the exact timing of their arrival, the impacts of their capabilities will fundamentally transform society and could be for good or ill. It is incumbent upon us to do everything in our power to understand the threat of ASI and shape the future towards the good and away from the ill. This starts with ensuring strategic and national policy decision makers have a comprehensive understanding of the risks and potential associated with AGI and ASI. Planning for the future is full of uncertainty but hoping that AGI/ASI is far enough away that it does not need consideration now, right now, will lead the nation to a precipice from which we may not be able to escape.

BIBLIOGRAPHY

Ackerman, Elliot, and Admiral James Stavridis. *2034: A Novel of the Next World War.* Kansas City: Penguin Books, 2022.

Ahmad, Javid. *Dress Like Allies, Kill Like Enemies: An Analysis of 'Insider Attacks' in Afghanistan*. West Point, NY: Modern War Institute, April 4, 2017. https://mwi.usma.edu/wp-content/uploads/2017/04/Dress-Like-Allies-Kill-Like-Enemies.pdf/.

Armstrong, Stuart, and Kaj Sotala. "How We're Predicting AI-or Failing To." In *Beyond AI: Artificial Dreams*, edited by Jan Romportl, Pavel Ircing, Eva Zackova, Michal Polak, and Radek Schuster, 52-75. Pilson: University of West Bohemia, 2012.

Assael, Yannis, Brendan Shillingford, Shimon Whiteson, and Nando de Freitas. "LipNet: End-to-End Sentence-Level Lipreading." Version 2, Oxford University, December 16, 2016. http://arxiv.org/abs/1611.01599.

Auer, Soren, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. "DBpedia: A Nucleus for a Web of Open Data." In *The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 2007, Proceedings,* edited by Karl Aberer, Key-Sun Choi, Natasha Noy, Dean Allemang, Kyung-Il Lee, Lyndon Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, Guus Schreiber, and Philippe Cudre-Mauroux, 722-735. Berlin: Springer, 2007.

Baidu Research. "ERNIE 3.0 Achieves State-of-the-Art Results in 54 Chinese NLP Tasks, Crowned 1st Place on SuperGLUE Leaderboard." (Blog), July 14, 2021. http://research.baidu.com/Blog/index-view?id=160.

Bawagan, Juanita. "The Turing Test 2.0." *PhysicsWorld*, May 8, 2021. https://physicsworld.com/the-turing-test-2-0/.

BBC. "Computer AI Passes Turing Test in 'World First'." *BBC News*, June 09, 2014. https://www.bbc.com/news/technology-27762088.

Beyer, Eric James. "MIT Researchers Just Discovered an AI Mimicking the Brain on Its Own." Interesting Engineering, December 18, 2021. https://interestingengineering.com/ai-mimicking-the-brain-on-its-own.

Bijker, Wiebe E. "How Is Technology Made? That Is the Question!" *Cambridge Journal of Economics* 34, no. 1 (2010): 63-76. https://doi.org/10.1093/cje/bep068.

Boesch, Gaudenz. “Object Detection in 2022: The Definitive Guide.” Viso.ai, Deep Learning. Accessed April 05, 2022. https://viso.ai/deep-learning/object-detection/.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. New York: Oxford University Press, 2014.

Brose, Christian. *The Kill Chain.* New York: Hachette Book Group, 2020.

Carlsmith, Joseph. “How Much Computational Power Does It Take to Match the Human Brain?” Open Philanthropy, September 11, 2020. https://www.openphilanthropy.org/brain-computation-report#Conclusion/.

Chesney, George. *The Battle of Dorking*. Blackmask Online, 2001. http://public-library.uk/ebooks/29/91.pdf.

China Institute for Science and Technology Policy. *China AI Development Report*. Haidian, Beijing, China: Tsinghua University, July 2018. https://edisciplinas.usp.br/pluginfile.php/4873100/mod_folder/content/0/China_AI%20report_2018.pdf?forcedownload=1/.

Cole, August. “‘FICINT’: Envisioning Future War Through Fiction & Intelligence (Indo-Pacific Series).” War Room – U.S. Army War College, May 22, 2019. https://warroom.armywarcollege.edu/special-series/indo-pacific-region/ficint-envisioning-future-war-through-fiction-intelligence-indo-pacific-series/.

———. “August Cole on FICINT and the Cognitive Warfighting Domain.” *The Cognitive Crucible,* Episode #33. Information Professionals Association. Audio, 38:39. https://information-professionals.org/episode/cognitive-crucible-episode-33/.

Copeland, Jack. *Turing: Pioneer of the Information Age.* United Kingdom: Oxford University Press, 2012.

Cotra, Ajeya, and Rohin Shah. “Draft Report on AI Timelines.” Alignment Newsletter, 2020. https://mailchi.mp/41774b61e5f8/an-121forecasting-transformative-ai-timelines-using-biological-anchors.

Damiani, Jesse. “A Voice Deepfake Was Used to Scam a CEO out of $243,000.” *Forbes,* September 03, 2019. https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/?sh=15d8872d2241/.

Davies, Alex, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburnet et al. “Advancing Mathematics by guiding Human Intuition with AI.” *Nature* 600 (December 2021): 70–74. https://doi.org/10.1038/s41586-021-04086-x.

Davis, Ernest, Leora Morgenstern, and Charles Ortiz. "The Winograd Schema Challenge." Computer Science Department at New York University. Accessed December 15, 2021. https://cs.nyu.edu/~davise/papers/WinogradSchemas/WS.html.

DeepMind. "Alpha Go." Directed by Greg Kohs. Moxie Pictures, 2017. Streaming video, 1:30:28. https://www.youtube.com/watch?v=WXuK6gekU1Y.

Delcour, Nicholas, Louis Duncan, Stephen Frahm, Patrick Lancaster, and Lance Vann. "Estimation of Technology Convergence by 2035." Mad Scientist Fellows Strategic Research Project, U.S. Army War College, 2020. https://csl.armywarcollege.edu/USACSL/Publications/EstimationOfTechConvergence-USAWC.pdf/.

Dolan, Kerry A., and Chase Peterson-Withorn, eds. "World's Billionaires List: The Richest in 2022." *Forbes.* Accessed April 7, 2022. https://www.forbes.com/billionaires/.

Donnarumma, Francesco, Domenico Maisto, and Giovanni Pezzulo. "Problem Solving as Probabilistic Inference with Subgoaling: Explaining Human Successes and Pitfalls in the Tower of Hanoi." *PLOS Computational Biology* 12, no. 4 (April 2016): e1004864. https://doi.org/10.1371/journal.pcbi.1004864.

Duffy, Clare. "Elon Musk to Buy Twitter in $44 Billion Deal." *CNN Business,* April 25, 2022. https://www.cnn.com/2022/04/25/tech/elon-musk-twitter-sale-agreement/index.html/.

Dvorsky. George. "8 Possible Alternatives to the Turing Test." *Gizmodo,* April 15, 2015. https://gizmodo.com/8-possible-alternatives-to-the-turing-test-1697983985.

Eynon, Rebecca, and Erin Young. "Methodology, Legend, and Rhetoric: The Constructions of AI by Academia, Industry, and Policy Groups for Lifelong Learning." *Science, Technology, & Human Values* 46, no. 1 (January 2021): 166-191.

Ferreira, Fernando G.D.C., Amir H. Gandomi, and Todrigo T. N. Cardoso. "Artificial Intelligence Applied to Stock Market Trading: A Review." *IEEE Access* 9 (February 2021): 30,898-30,917. https://ieeexplore.ieee.org/document/9350582.

Fjelland, Ragnar. "Why General Artificial Intelligence Will Not Be Realized." *Humanities and Social Sciences Communications* 7 (2020): Article 10.

Garisto, Dan. "Google AI Beats Top Human Players at Strategy Game StarCraft II." *Nature*, October 30, 2019. https://doi.org/10.1038/d41586-019-03298-6.

Goetzel, Ben, and Cassio Pennachin, eds. *Artificial General Intelligence.* Springer: Rockville, AGIRI, 1998.

Good, Irving John. "Speculations Concerning the First Ultraintelligent Machine." In *Advances in Computers*, edited by Franz L. Alt and Morris Rubinoff, 6:31-88. New York: Academic Press, 1965.

Grewal, Dalvinder Singh. "A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering." *IOSR Journal of Computer Engineering* 16, no. 2 (2014): 9-13.

Halpern, Sue. "The Rise of A.I. Fighter Pilots." *The New Yorker*, January 17, 2022. https://www.newyorker.com/magazine/2022/01/24/the-rise-of-ai-fighter-pilots/.

Hernandez, Jore. "A Military Drone with a Mind of its Own was Used in Combat, U.N. Says." *NPR*, June 1, 2021. https://www.npr.org/2021/06/01/1002196245/a-u-n-report-suggests-libya-saw-the-first-battlefield-killing-by-an-autonomous-d/.

Holroyd, Matthew, and Fola Olorunselu. "Deepfake Zelenskyy Surrender Video is the 'First Intentionally Used' in Ukraine War." *Euro-news.* Last modified March 16, 2022. https://www.euronews.com/my-europe/2022/03/16/deepfake-zelenskyy-surrender-video-is-the-first-intentionally-used-in-ukraine-war/.

Huang, Yichen, Yizhe Zhang, Oussama Elachqar, and Yu Cheng. "INSET: Sentence Infilling with INter-SEntential Transformer." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics,* 2502–2515. Stroudsburg, PA: Association for Computational Linguistics, July 2020. https://doi.org/10.18653/v1/2020.acl-main.226.

IBM Corporation. "Deep Blue." IBM 100: Icons of Progress. Last modified March 7, 2012. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/.

International Committee of the Red Cross. "Practice Relating to Rule 65. Perfidy: Section B. Killing, Injuring or Capturing an Adversary by Resort to Perfidy." IHL Database: Customary IHL. Accessed February 21, 2022. https://ihl-databases.icrc.org/customary-ihl/eng/docindex/v2_rul_rule65_sectionb/.

Ipsos. "Ipsos Misinformation and Conspiracy Theory Poll." Ipsos Group S.A., Washington, DC, January 5, 2022. https://www.ipsos.com/sites/default/files/ct/news/documents/2022-01/BBC%20Misinformation%20and%20Conspiracy%20Theory%20Topline%2001.05.22.pdf.

Iversen, Jonas Svava. "Futures Thinking Methodologies–Options Relevant for Schooling for Tomorrow." Organisation for Economic and Co-operation Development, Paris, France, 2005. https://www.oecd.org/education/ceri/35393902.pdf.

Jefferson, Thomas. *The Papers of Thomas Jefferson.* Vol. 12. Edited by Julian P. Boyd, Mina R. Bryan, and Fredrick Aandahl. Princeton, NJ: Princeton University Press, 1955.

Johnson, Melvin, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat et al. "Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation." In *Transactions of the Association for Computational Linguistics.* Vol. 5, edited by Colin Cherry, 339-351. Stroudsburg, PA: Association for Computational Linguistics, 2017.https://aclanthology.org/Q17-1024.pdf.

Kania, Elsa B. "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power." Center for a New American Security, Washington, DC, 2017. http://www.indexfunds.org/resources/Research-Materials/NatSec/CNAS_Battlefield_Singularitypdf.pdf/.

Köbis, Nils C., Barbora Doležalová, and Ivan Soraperra, "Fooled Twice: People Cannot Detect Deepfakes but Think They Can." *iScience* 24, no. (2021): 9.

Kosinski, Michael, and Yilun Wang, "Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images." *Journal of Personality and Social Psychology* 114, no. 2 (February 2018): 246-257.

Kosow, Hannah, and Robert Gaβner. *Methods of Future and Scenario Analysis: Overview, Assessment, and Selection Criteria.* Bonn: German Development Institute, 2008. https://www.die-gdi.de/uploads/media/Studies_39.2008.pdf/.

Kurzweil, Ray. *The Singularity is Near*. Kansas City: Penguin Books, 2005.

Leetaru, Kalev H. "Culturomics 2.0: Forecasting Large-Scale Human Behavior Using Global News Media Tone in Time and Space." *First Monday* 16, no. 9 (September 5, 2011). https://journals.uic.edu/ojs/index.php/fm/article/download/3663/3040/.

Lenat, Douglas. "Douglas Lenet: Cyc and the Quest to Solve Common Snese Reasoning in AI." Interview by Lex Fridman. *Lex Fridman Podcast*, September 2021. Video, 2:52:56. https://www.youtube.com/watch?v=3wMKoSRbGVs.

Li, Bo, Ruoming Pang, Tara N. Sainath, Anmol Gulati, Yu Zhang, James Qin, Parisa Haghani, W. Ronny Huang, Min Ma, and Junwen Bai. "Scaling End-to-End Models for Large-Scale Multilingual ASR." Google, USA, September 11, 2021. https://arxiv.org/pdf/2104.14830.pdf.

Lin, Stephanie, Jacob Hilton, and Owain Evans. “TruthfulQA: Measuring How Models Mimic Human Falsehoods.” In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. Vol. 1, *Long Papers*, 3214-3252. Stroudsburg, PA: Association for Computational Linguistics, May 2022. https://owainevans.github.io/pdfs/truthfulQA_lin_evans.pdf.

Luna, Nancy. “Checkers & Rally’s is Rolling Out Voice-Ordering Bots to Take Drive-Thru Orders at 267 Restaurants Amid a Crippling Labor Shortage in the Industry.” *Insider,* January 10, 2022. https://www.businessinsider.com/checkers-rolls-out-presto-voice-bots-at-drive-thru-lanes-2022-1.

Marcus, Gary. “What Comes After the Turing Test?” *The New Yorker*, June 9, 2014. https://www.newyorker.com/tech/annals-of-technology/what-comes-after-the-turing-test.

Marcus, Gary, and Ernest Davis. “GPT-3, Bloviator: OpenAI’s Language Generator Has No Idea What It’s Talking About.” *MIT Technology Review*, August 22, 2020. https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/.

Martindale, Jon. “What Is a Teraflop?” DigitalTrends, June 14, 2021. https://www.digitaltrends.com/computing/what-is-a-teraflop/.

Matin, Michael. “Scrutinizing ‘The Battle of Dorking’: The Royal United Service Institution and the mid-Victorian Invasion Controversy.” *Victorian Literature and Culture* 39, no. 2 (2011): 385-407.

McCarthy, John. “Programs With Common Sense.” Computer Science Department, Stanford University, Stanford, CA, 1959. http://jmc.stanford.edu/articles/mcc59/mcc59.pdf.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955.” *AI Magazine* 27, no. 4 (December 15, 2006): 12. https://ojs.aaai.org/index.php/aimagazine/article/view/1904.

McCorduck, Pamela. *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. 2nd ed. Natick, MA: A. K. Peters, 2004.

Meel, Vidushi. “87 Most Popular Computer Vision Applications in 2022.” Viso.ai, Applications. Accessed 05 April 2022. https://viso.ai/applications/computer-vision-applications/.

Meta AI. “Textless NLP: Generating Expressive Speech from Raw Audio.” September 09, 2021. https://ai.facebook.com/blog/textless-nlp-generating-expressive-speech-from-raw-audio.

Metz, Cade. “A.I. Can Now Write Its Own Computer Code. That’s Good News for Humans.” *The New York Times,* September 09, 2021. https://www.nytimes.com/2021/09/09/technology/codex-artificial-intelligence-coding.html.

Minsky, Marvin. *The Society of Mind.* New York: Simon & Schuster, 1988.

Mirsky, Yisroel, and Wenke Lee. “The Creation and Detection of Deepfakes: A Survey,” *ACM Computing Surveys* 54, no. 1, article 7 (January 2022): 1-41.

Moravec, Hans P. “The Role of Raw Power in Intelligence.” Unpublished manuscript, Carnegie Mellon University, May 12, 1976. https://frc.ri.cmu.edu/~hpm/project.archive/general.articles/1975/Raw.Power.html.

Müller, Vincent C., and Nick Bostrom. “Future Progress in Artificial Intelligence: A Survey of Expert Opinion.” In *Fundamental Issues of Artificial Intelligence,* edited by Vincent C. Müller, 553-571. Berlin: Synthese Library, 2016.

Murray, Williamson, and Macgregor Knox. "Thinking about Revolutions in Warfare." In *The Dynamics of Military Revolution 1300–2050*, edited by MacGregor Knox and Williamson Murray, 1-14. Cambridge: Cambridge University Press, 2001.

Nash, Ed. “We May Have the First Case of a Robot Deliberately Killing Humans.” Military Matters, June 1, 2021. https://militarymatters.online/defense-news/we-may-have-the-first-case-of-a-robot-deliberately-killing-humans/.

Negrotti, Massimo. *Understanding the Artificial: On the Future Shape of Artificial Intelligence.* London: Springer-Verlag, 1991.

Nguyen, Thanh Thi, Quoc Viet Hung Nguyen, Dung Tien Dguyen, Duc Thanh Nguyen, Thien Huynh-The, Saeid Nahavandi, Thanh Tam Nguyen, Quoc-Viet Pham, and Cuong M. Nguyen. “Deep Learning for Deepfakes Creation and Detection: A Survey.” Last modified February 06, 2022. https://doi.org/10.48550/arXiv.1909.11573/.

Nilsson, Nils J. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press, 2009.

Norton, Richard J. “Through a Mirror Darkly: The Face of Future War, 1871-2005.” *Naval War College Review* 62, no. 1 (2009): 123-140.

O'Meara, Sarah. "Will China Lead the World in AI by 2030?" *Nature* 572 (August 2019). https://news.sisuer.cn/wp-content/uploads/2020/06/Will-China-lead-the-world-in-AI-by-2030.pdf.

Office of the Chairman of the Joint Chiefs of Staff (CJCS). CJCS Manual 3105.01A *Joint Risk Analysis Methodology*. Washington, DC: Joint Chiefs of Staff, October 12, 2021.

———. Joint Publication 1, *Doctrine for the Armed Forces of the United States*. Washington, DC: Joint Chiefs of Staff, 2017.

———. Joint Publication 5-0, *Joint Planning*. Washington, DC: Joint Chiefs of Staff, 2020.

Office of the Director of National Intelligence. *Global Trends 2040: A More Contested World.* Washington, DC: National Intelligence Council, March 2021.

Parakilas, Jacob. "Fiction and Consequences: War in Art and the Art of War." *The Diplomat,* January 27, 2021. https://thediplomat.com/2021/01/fiction-and-consequences-war-in-art-and-the-art-of-war/.

Pinola, Melanie. "Speech Recognition Through the Decades: How We Ended Up with Siri." *PCWorld*, November 2, 2011. https://www.pcworld.com/article/477914/speech_recognition_through_the_decades_how_we_ended_up_with_siri.html.

Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang et al. "ImageNet Large Scale Visual Recognition Challenge." *International Journal of Computer Vision* 115 (April 2015): 211–252. https://doi.org/10.1007/s11263-015-0816-y.

Scharre, Paul. *Army of None: Autonomous Weapons and the Future of War.* New York: Norton, 2019.

Schrimpf, Martin, Idan Asher Blank, Greta Tuckute, Carina Kauf, Eghbal A. Hosseini, Nancy Kanwisher, Joshua B. Tenenbaum, and Evelina Fedorenko. "The Neural Architecture of Language: Integrative Modeling Converges on Predictive Processing." *Proceedings of the National Academy of Sciences* 118, no. 45 (November 2021): e2105646118. https://doi.org/10.1073/pnas.2105646118.

Shrobe, Howard. "Machine Common Sense." Defense Advanced Research Projects Agency. Accessed February 1, 2022. https://www.darpa.mil/program/machine-common-sense.

Shum, Michael, Max Kleiman-Weiner, Michael L. Littman, and Joshua B. Tenenbaum. "Theory of Minds: Understanding Behavior in Groups through Inverse Planning." *Proceedings of the AAAI Conference on Artificial Intelligence* 33, no. 01 (July 17, 2019): 6163-6170. https://ojs.aaai.org/index.php/AAAI/article/view/4574.

Silver, David, Aja Huang, Christopher Maddison, Arthur Guez, Laurent Sifre, George Driessche, Julian Schrittwieser et al. "Mastering the Game of Go with Deep Neural Networks and Tree Search." *Nature* 529 (January 2016): 484–489. https://doi.org/10.1038/nature16961.

Silver, David, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, et al. "Mastering the Game of Go Without Human Knowledge." *Nature* 550 (October 2017): 354-359. https://doi.org/10.1038/nature24270.

Slonim, Noam, Yonatan Bilu, Carlos Alzate, Roy Bar-Haim, Ben Bogin, Francesca Bonin, Leshem Choshen, et al. "An Autonomous Debating System." *Nature* 591, no. 7850 (March 2021): 379–384. https://doi.org/10.1038/s41586-021-03215-w.

Stooke, Adam, Anuj Mahajan, Catarina Barros, Charlie Deck, Jakob Bauer, Jakub Sygnowski, Maja Trebacz, et al. "Generally Capable Agents Emerge from Open-ended Play." *DeepMind* (Blog), July 27, 2021. https://deepmind.com/blog/article/generally-capable-agents-emerge-from-open-ended-play/.

SuperGLUE. "Frequently Asked Questions." Accessed December 15, 2021. https://super.gluebenchmark.com/faq.

Taleb, Nassim Nicholas. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House, 2007.

TOP500. "November 2005." Accessed December 10, 2021. https://www.top500.org/lists/top500/2005/11/.

———. "November 2021." Accessed December 10, 2021. https://www.top500.org/lists/top500/2021/11/.

Travis, Nathaniel. "Is Human Behavior Just Elaborate Running and Tumbling?" PsyArXiv, February 9, 2022. psyarxiv.com/wzvn9.

Turing, Allen. "Computing Machinery and Intelligence," *Mind* 59, no. 236 (October 1950): 433–460. https://academic.oup.com/mind/article/LIX/236/433/986238.

U.S. Bureau of Labor Statistics. "Occupational Employment and Wages, May 2021: 33-3051 Police and Sheriff's Patrol Officers." Division of Occupational Employment and Wage Statistics. Accessed February 19, 2022. https://www.bls.gov/oes/current/oes333051.htm.

U.S. Government Accountability Office (GAO). GAO-21-518, *Facial Recongtion Technology: Federal Law Enforcement Agencies Should Better Assess Privacy and Other Risks*. Report to Congressional Requesters. Washington, DC: GAO, June 2021. https://www.gao.gov/products/gao-21-518/.

VanHerck, Glen D. "NORTHCOM Commander Gen. Glen D. VanHerck Conducts Press Briefing on North American Aerospace Defense Command and U.S. Northern Command Global Information Dominance Experiments." Transcript, U.S. Department of Defense, July 28, 2021. https://www.defense.gov/News/Transcripts/Transcript/Article/2711594/northcom-commander-gen-glen-d-vanherck-conducts-press-briefing-on-north-america/.

Verdoliva, Luisa. "Media Forensics and Deepfakes: An Overview." *IEEE Journal of Selected Topics in Signal Processing* 14, no. 5 (2020): 910–932.

Vinge, Vernor, "Technological Singularity." Carnegie Mellon University, March 1993. https://frc.ri.cmu.edu/~hpm/book98/com.ch1/vinge.singularity.html.

Vinyals, Oriol, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, et al. "Grandmaster Level in StarCraft II Using Multi-agent Reinforcement Learning." *Nature* 575 (October 2019): 350–354. https://doi.org/10.1038/s41586-019-1724-z.

Wiggers, Kyle. "DeepMind's MuZero Teaches Itself How to Win at Atari, Chess, Shogi, and Go." *VentureBeat*, The Machine, November 20, 2019. https://venturebeat.com/2019/11/20/deepminds-muzero-teaches-itself-how-to-win-at-atari-chess-shogi-and-go/.

Wilson, Hal. "Jonathan Roper: Travelling Consultant." Modern War Institute at West Point, May 21, 2019. https://mwi.usma.edu/jonathan-roper-traveling-consultant/.

———. "Letter of Marque." *Proceedings* 146, no. 12 (December 2020). https://www.usni.org/magazines/proceedings/2020/december/letter-marque/.

Worldometer. "GDP by Country." Accessed March 13, 2022. https://www.worldometers.info/gdp/gdp-by-country/.

Yudkowsky, Eliezer S. "Staring Into the Singularity." Singularity, 1999. http://www.pivot.net/~jpierce/staring_into_the_singularity.htm.

Zhu, Yixin, Tao Gao, Lifeng Fan, Siyuan Huang, Mark Edmonds, Hangxin Liu, Feng Gao, et al. “Dark, Beyond Deep: A Paradigm Shift to Cognitive AI with Humanlike Common Sense.” *Engineering* 6, no. 3 (2020): 310-345.

Ziva Dynamics Inc. “Ziva Face Trainer.” Ziva. https://zivadynamics.com/zrt-face-trainer/.