NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DEFENSE ANALYSIS CAPSTONE REPORT

TIP OF THE SPEAR: CAN SPECIAL FORCES LEAD THE WAY FOR MILITARY APPLICATIONS OF AI?

by

Joakim Riberg and Lars J. Selin

June 2021

Thesis Advisor:	Leo J. Blanken
Second Reader:	Vinnie Monaco

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

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 instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the 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 Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE June 2021	3. REPORT TYPE AND DATES COVERED Defense Analysis Capstone Report

4. TITLE AND SUBTITLE TIP OF THE SPEAR: CAN SPECIAL FORCES LEAD THE WAY FOR MILITARY APPLICATIONS OF AI?	5. FUNDING NUMBERS N/A
6. AUTHOR(S) Joakim Riberg and Lars J. Selin	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Swedish Armed Forces, STOCKHOLM, N/A N/A	10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	12b. DISTRIBUTION CODE A

13. ABSTRACT (maximum 200 words)

As artificial intelligence (AI) continues to evolve with ever-increasing speed, especially in the commercial sector, the military community runs the risk of being left behind. This is especially problematic for Special Operations Forces (SOF) worldwide, as they are often the first to encounter new methods and technology employed by irregular adversaries. By studying contemporary research and interviewing experts within the field of AI, the authors employ a mixed-methods approach to explore what factors the SOF community must consider when acquiring and evaluating new AI capabilities. Furthermore, the authors attempt to clarify what qualifies as AI and what factors affect its adoption and user acceptance. The results of this study suggest that to avoid or minimize the traps associated with AI, the military must partner with other organizations, focusing on the right products and managing the education and expectations of the users within their organization. Rather than recommend specific products or solutions, this capstone report proposes a tentative model focused on risk mitigation and fast procurement chains to allow SOF to maintain its edge on the current battlefield.

14. SUBJECT TERMS artificial intelligence, AI, procurement, innovation, Special Operations Forces, SOF	15. NUMBER OF PAGES 129
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

TIP OF THE SPEAR: CAN SPECIAL FORCES LEAD THE WAY FOR MILITARY APPLICATIONS OF AI?

Joakim Riberg
Major, Swedish Army
B, Swedish Defence University, 2019

Lars J. Selin
Major, Swedish Army
B, Swedish Defence University, 2019

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN DEFENSE ANALYSIS
(IRREGULAR WARFARE)

from the

NAVAL POSTGRADUATE SCHOOL
June 2021

Approved by: Leo J. Blanken
Advisor

Vinnie Monaco
Second Reader

Douglas A. Borer
Chair, Department of Defense Analysis

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

As artificial intelligence (AI) continues to evolve with ever-increasing speed, especially in the commercial sector, the military community runs the risk of being left behind. This is especially problematic for Special Operations Forces (SOF) worldwide, as they are often the first to encounter new methods and technology employed by irregular adversaries. By studying contemporary research and interviewing experts within the field of AI, the authors employ a mixed-methods approach to explore what factors the SOF community must consider when acquiring and evaluating new AI capabilities. Furthermore, the authors attempt to clarify what qualifies as AI and what factors affect its adoption and user acceptance. The results of this study suggest that to avoid or minimize the traps associated with AI, the military must partner with other organizations, focusing on the right products and managing the education and expectations of the users within their organization. Rather than recommend specific products or solutions, this capstone report proposes a tentative model focused on risk mitigation and fast procurement chains to allow SOF to maintain its edge on the current battlefield.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1
- A. STRUCTURE 1
- B. DEFINITIONS 2

II. THE PROMISE OF ARTIFICIAL INTELLIGENCE 5
- A. AI AND THE MILITARY 7
- B. SUMMARY 12

III. RESEARCH APPROACH 13
- A. TENTATIVE MODEL FOR CONSIDERING AI PROCUREMENT 13
- B. METHODOLOGY 15
- C. DATA COLLECTION 16
  - 1. Secondary Sources 16
  - 2. Interviews 17
  - 3. Seminars and Product Fairs 17
- D. VALIDITY OF SOURCES 18
- E. EXPECTED RESULTS 19

IV. ANALYSIS 21
- A. TRANSPARENCY 21
  - 1. Transparency and Trust 22
  - 2. Transparency in AI 24
  - 3. Trust in AI 27
  - 4. Implications of Transparency for Trust in AI 28
  - 5. Conclusion Analysis of Transparency 31
- B. SECURITY 31
  - 1. Risks Related to ML 32
  - 2. Ensuring Secure Military ML 37
  - 3. Best Practices 39
  - 4. Conclusion Analysis of Security 45
- C. DATA 46
  - 1. Types of Data 48
  - 2. Data and AI 49
  - 3. Conclusion Analysis of Data 54
- D. INTERVIEWS 55
  - 1. Findings 55

E. CONCLUSION OF ANALYSIS.....61

V. RESULTS.....65
A. PROPOSED MODEL.....67
B. SUMMARY OF RESULTS.....70

VI. CONCLUSION.....71

APPENDIX A. HISTORY OF AI.....75
A. THE EMERGENCE OF AI.....75
B. ARTIFICIAL INTELLIGENCE AND ITS SUBFIELDS.....78
1. Machine Learning.....79
2. Deep Learning.....79
3. Robotic Process Automation.....80
4. Natural Language Processing.....81
5. Physical Robots.....81
C. DATA EXPLAINED.....82

APPENDIX B. INTERVIEWS.....85
A. INTERVIEWEES.....86
B. QUESTIONS FOR INTERVIEWS.....87

SUPPLEMENTALS: QUICK REVIEW AND ANIMATION.....95
A. CAPSTONE QUICK REVIEW.....95
B. SHORT ANIMATION AND INTRODUCTION.....95

LIST OF REFERENCES.....97

INITIAL DISTRIBUTION LIST.....111

LIST OF FIGURES

Figure 1. Proposed Model for Analysis.....68

Figure 2. Graphical Overview of AI. 78

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1. Human and Artificial Intelligence. 76

Table 2. Data Volumes Explained 83

Table 3. Interview Questions 87

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	Artificial Intelligence
COTS	Commercial of the Shelf
CRAAP	Currency, Relevance, Authority, Accuracy, and Purpose
CIA	Confidentialty, Integrity, Authenticity
CSC	Cyberspace Solarium Commission
DARPA	Defense Advanced Research Projects Agency
DL	Deep Learning
DOD	Department of Defense
JAIC	Joint Artificial Intelligence Center
ML	Machine Learning
NATO	North Atlantic Treaty Organization
NPS	Naval Postgraduate School
RPA	Robotic Process Automation
SDRA	Swedish Defense Research Agency
SIPRI	Stockholm International Peace Research Institute
SL	Supervised Learning
SOF	Special operations forces
UL	Unsupervised Learning

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First of all, we would like to acknowledge the great support offered by our advisor, Professor Leo Blanken, and our second reader, Professor Vinnie Monaco. Thank you for giving us ideas, constructive criticism and advice on improvements and clarification, which enriched the results of our research. We would also like to thank the Naval Postgraduate School for this unique opportunity to attend this program as foreign officers through the International Graduates Program Office and for the flexibility provided by the Defense Analysis Department to pursue this capstone. A special thank you also goes out to the interviewees; without your contributions, this capstone would not have been possible. Finally, for supporting the move abroad from family and friends in Sweden and making the COVID-19 pandemic tolerable, our families deserve our deepest gratitude: Thank you Karin, Emilia, and Viktor, and Annelie, Nathalie, and Emmylie.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

As current developments in artificial intelligence (AI) rapidly move forward, several implementations are emerging that promise enhancements to both the conduct of military operations as well as the functions within organizations. Nevertheless, since AI research is both expensive and often driven by the civilian sector, and thereby general public demand, the edge in research and development is mainly confined to large and economically strong nations such as China and the United States.[1] As special operations forces (SOF) often are at the forefront of implementing new technology within nations' militaries, the issue of identifying trends, as well as adapting the organization to accommodate new technologies, is highly relevant within the SOF community. To achieve and maintain an acceptable level of implementation and utilization of emerging technologies, SOF must understand the problems connected to implementing AI with its organization in order to determine which technologies and capabilities to pursue. By examining the different approaches other organizations have taken to implement AI, as well as trying to identify future capabilities predicted by both the civilian sector and academia, this capstone explores a way forward for the SOF community for acquiring AI. Hence, the question to be answered is: *What factors must the SOF community consider when acquiring and evaluating new AI capabilities?*

A. STRUCTURE

Initially, a context for the capstone is presented by looking at what AI is, and what AI promises to be, followed by an explanation of the methodology used throughout the research. The subsequent analysis is based on previous research regarding military acquisition of AI and identified problem sets. As the results of the analysis are presented, this capstone produces guidelines for the SOF community and suggests a SOF-specific model on how to evaluate and assess the utility of AI implementation. The conclusion is

[1] Lee, Kai-Fu, *AI Superpowers: China, Silicon Valley, and the New World Order* (Boston: Houghton Mifflin Harcourt, 2018), 168.

that the SOF community is well suited to gain the edge in the AI race provided the limitations of being a small organization are accepted and AI-specific pitfalls are mitigated.

B. DEFINITIONS

Within the framework of this capstone, the following definitions are used throughout the research. Further detail on the history and a literature review of AI and foundational knowledge regarding the subject of data is presented in Appendix A.

Artificial Intelligence is, as further discussed in Appendix A – History of AI, a highly contested subject. While some scholars argue that everything that can produce a calculation and reach a result void of human interaction constitutes as AI, others reject this notion and argue that only a machine that can adapt to context, learn new tasks, and function as a human should be considered intelligent.[2] Therefore, the following analysis focuses on Machine Learning (ML), something that most scholars agree to be AI.[3] Jason Bell, the author of the book *Machine Learning: Hands-On for Developers and Technical Professionals*, defines ML as:

> Machine learning is a branch of artificial intelligence. Using computing, we design systems that can learn from data in a manner of being trained. The systems might learn and improve with experience, and with time, refine a model that can be used to predict outcomes of questions based on the previous learning.[4]

While ML consists of several sub-divisions, such as assisted learning, unassisted learning, and deep learning, all ML is based on the use of data to learn new traits.[5]

---

[2] Kelley M. Sayler, *Artificial Intelligence and National Security*, CRS Report R45178 (Washington, DC: Congressional Research Service, 2019), 1–2, https://fas.org/sgp/crs/natsec/R45178.pdf.

[3] Rafay Chaudhary, "Artificial Intelligence: More than Machine Learning," Towards Data Science, September 2019, https://towardsdatascience.com/artificial-intelligence-more-than-machine-learning-fc95a1f8c2f5.

[4] Jason Bell, *Machine Learning: Hands-On for Developers and Technical Professionals* (Indianapolis, IN: John Wiley & Sons, 2014), 2, http://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=1818248.

[5] Yang Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access* 6 (May 2018): 35367, https://doi.org/10.1109/ACCESS.2018.2836950.

Confidentiality, Integrity, and Accessibility (CIA) is a well-known triad of cybersecurity.[6] The CIA triad, which describes the need for protection of data while still allowing access for authorized users, emerged as a measure for cybersecurity in the early 1980s.[7] Since then, the abbreviation has developed into the main parameters for measuring the security of a network, especially within the military.[8]

Strategy is often described as the effort to balance ways (how) and means (what) to achieve reachable ends (goals) while considering risks (threats).[9] In his article "Defining Military Strategy," Arthur Lykke described the relationship of ends, ways, and means as a three-legged stool used to counter the risks of a nation.[10] Using the analogy of the stool, Lykke argued that disproportion among the ends, means, and ways results in an unstable strategy to counter the risks.[11] Although the main purpose of this capstone is not to produce a strategy for AI, the ends, ways, and means construct provides the structure to suggest best practices to deflect the risks of implementing AI in a military setting.

---

[6] Mehul S. Raval, Ratnik Gandhi, and Sanjay Chaudhary, "Insider Threat Detection: Machine Learning Way," in *Versatile Cybersecurity*, ed. Mauro Conti, Gaurav Somani, and Radha Poovendran, Advances in Information Security (Cham, Switzerland: Springer International Publishing, 2018), 204, https://doi.org/10.1007/978-3-319-97643-3_2.

[7] Spyridon Samonas and David Coss, "The CIA Strikes Back: Redefining Confidentiality, Integrity And Availability In Security," *Journal of Information System Security* 10, no. 3 (2014): 21–22, http://www.proso.com/dl/Samonas.pdf.

[8] Samonas and Coss, 25.

[9] Mackubin Thomas Owens, "Strategy and the Strategic Way of Thinking," *Naval War College Review* 60, no. 4, Article 10. (2007): 112, https://digital-commons.usnwc.edu/nwc-review/vol60/iss4/10.

[10] Arthur F. Lykke, "Defining Military Strategy," *Military Review* 77, no. 1 (February 1997): 183–85, http://search.proquest.com/docview/225319163/abstract/B68261794C094016PQ/1.

[11] Lykke, 184.

THIS PAGE INTENTIONALLY LEFT BLANK

II. THE PROMISE OF ARTIFICIAL INTELLIGENCE

As new technologies emerge, there is often a sense that they will alter the world entirely, as was the case with the harnessing of electricity, the innovation of flight, and the completion of the nuclear bomb. Although these technologies have certainly changed the world as we know it in one way or another, however, the rate of change has seemingly always depended on society's ability to absorb and distribute change. Only the atom bomb can be said to have created a sudden and instant change in society.[12] Furthermore, all of these technologies have yet to achieve some of their purported impact; electricity has not yet eradicated fossil fuels as the primary source of propulsion, flight has yet to allow for mass travel around the world at supersonic speed, and much less did the advent of the atomic bomb provide the means to end to all wars.[13] Still, most scholars agree that AI promises change within several areas, such as world economics, transportation, and military capabilities in the near future.[14] Perhaps AI will prove the odd one out by radically and forever upsetting society within a matter of years in what some call an "AI revolution."[15] Nonetheless, even though AI promises a future where machines not only replace humans performing menial jobs, but also threaten even the fabric underpinning corporate culture and the world economy, there are reasons to be skeptical.[16] AI research

---

[12] Robert A. Buchanan, "History of Technology, WWII, The Japanese Surrender," *Britannica*, Technology, November 2020, https://www.britannica.com/event/World-War-II/Hiroshima-and-Nagasaki.

[13] George Orwell, "You and the Atom Bomb," The Orwell Foundation, April 2,2020, https://www.orwellfoundation.com/the-orwell-foundation/orwell/essays-and-other-works/you-and-the-atom-bomb/.

[14] Michael C. Horowitz et al., "Strategic Competition in an Era of Artificial Intelligence," *Artificial Intelligence and International Security* (Washington, DC: Center for a New American Security, July 2018), 7; Katja Grace et al., "Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts," *Journal of Artificial Intelligence Research* 62 (July 31, 2018): 729, https://doi.org/10.1613/jair.1.11222; Joseph Byrum, "Taking Advantage of the AI Revolution," *ISE; Industrial and Systems Engineering at Work* 50, no. 6 (June 2018): 29, http://search.proquest.com/docview/2054133919/abstract/27A128515D48480DPQ/1; Lee, *AI Superpowers*, 144.

[15] Byrum, "Taking Advantage of the AI Revolution," 29.

[16] Byrum, 32.

and implementation has in fact already affected society during the last seventy years.[17] Furthermore, similar forecasts of displacement of labor were made as early as the 1980s by several scholars looking into the possibilities of AI.[18] It, therefore, seems hard to say that the "revolution" suddenly will upset the current way of life.

Just as most innovations plateau when they are at their most promising height, AI has also suffered from "AI Winters," as several scholars point out.[19] Furthermore, the widespread adoption of almost all inventions is reliant on customer demand and profitability, rather than on pure possibility and actual benefits to society.[20] With the construction of the Concorde, for example, the engineers were looking into accessing supersonic speeds for commercial travel. Uniting visions and promising a revolution in commercial flight, the prospect of traversing the globe in mere hours led countries and companies to pour significant funding into the project. In the end, this turned out to be a non-profitable market, resulting in the abandonment of the Concord in 2003.[21] At the moment, there is a sense of the same evolution facing AI; the mere promise of possibilities seems to have sparked a self-revolving revolution in AI. In fear of missing out and lagging behind the competition, government procurement organizations seem interested in almost anything labeled "AI."

---

[17] John Haugeland, *Artificial Intelligence: The Very Idea*, 1st paperback ed. (Cambridge, MA: MIT Press, 1989), 176, https://ieeexplore-ieee-org.libproxy.nps.edu/xpl/ebooks/bookPdfWithBanner.jsp?fileName=6302870.pdf&bkn=6276821&pdfType=chapter.

[18] Robert Trappl, ed., *Impacts of Artificial Intelligence: Scientific, Technological, Military, Economic, Societal, Cultural, and Political*, 1st ed. (New York, NY: Elsevier Science Publishers B.V., 1986), 39–42.

[19] James Hendler, "Avoiding Another AI Winter," *IEEE Intelligent Systems* 23, no. 2 (March 2008): 2, https://doi.org/10.1109/MIS.2008.20 ; Andreas Holzinger et al., "Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction: From Machine Learning to Explainable AI," in *Machine Learning and Knowledge Extraction*, ed. Andreas Holzinger et al., vol. 11015 (Cham, Switzerland: Springer International Publishing, 2018), 2, https://doi.org/10.1007/978-3-319-99740-7_1; Lee, *AI Superpowers*, 6–8.

[20] Bhaskar Chakravorti, "The New Rules for Bringing Innovations to Market," *Harvard Business Review*, 2004, 5.

[21] Florian Ion Tiberiu Petrescu and Relly Victoria Petrescu, *The Aviation History.* (Norderstedt, Germany: Books on Demand GmbH, 2012), 40–41, https://doi.org/10.13140/RG.2.1.1273.9285.

As Kai-Fu Lee, an acknowledged expert within the field of AI points out, society will continue to dilute the effect of inventions, which require new infrastructure, regulations, and norms that are time consuming and expensive to establish.[22] Driverless cars are one current example, promising safer transportation by eliminating the human factor.[23] As humans are reluctant to put their lives in the hands of pure automation, however, there is little evidence that the public would allow the total automation of all cars within the next few years.[24] This does not mean that driverless vehicles are utopic (they already exist in small numbers); instead, it suggests that it takes time for humans to accept and get accustomed to new inventions.

Considering the inherent inertia in society to swiftly adopt changes, militaries, as well as society, should therefore be cautious of products promising swift solutions for everything. Nevertheless, this does not mean that the advancements in AI should, or can, be neglected.[25] Instead, a more sensible approach for assessing new capabilities in AI might be one that considers the inertia of adaptation to new technology produced by legislation, culture, and resistance to change within organizations.[26] After all, AI will undoubtedly have some effect on both economic growth and military capabilities, as well as how society functions as a whole, as agreed by most scholars.

A. AI AND THE MILITARY

From a military perspective, the utility of AI promises not only smarter and better decisions at a lower cost but also new possibilities of facing the enemy at a lesser risk to

---

[22] Lee, *AI Superpowers*, 164.

[23] Lee, 101.

[24] Taulli, *Artificial Intelligence Basics*, 164–65.

[25] Byrum, "Taking Advantage of the AI Revolution," 31.

[26] Lee, *AI Superpowers*, 142–142.

human life.[27] In a society highly concerned with price and human welfare, this has sparked the race to be at the forefront of AI technology. This race has provided ample opportunities for AI researchers around the globe to present their products to the military.[28] In response to Chinese advances in the field, the U.S. Army has commissioned several Defense Advanced Research Projects Agency (DARPA) projects worth billions in revenues to exploit the new gains AI offers in processing big data.[29] The United States, supported by well-functioning academia, has so far maintained its advantage in the race by adopting a careful approach, where big companies such as Apple, Google, and Microsoft have been allowed to dominate the market.[30] Yet, while the U.S. approach from the government has focused on restricting the dissemination and use of data, the Chinese method is quite the opposite.

China's miraculous advances in the field, on the other hand, have relied more on mass, that is to say, free enterprising, unchained competition, blatant plagiarism, and a continuously growing number of scientists focused on AI, all supported and subsidized by the government.[31] This approach has allowed the Chinese government to kick-start development and leverage their control of information by facilitating unrestricted use of

---

[27] Elsa B. Kania, "Chinese Military Innovation in the AI Revolution," *The RUSI Journal* 164, no. 5–6 (September 19, 2019): 26–34, https://doi.org/10.1080/03071847.2019.1693803 ; Zhu Feng et al., "Inspiration for Battlefield Situation Cognition from AI Military Programs Launched by DARPA of USA and Development of AI Technology," in *Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems*, ed. Lin Zhang, Xiao Song, and Yunjie Wu, vol. 644 (Singapore: Springer Singapore, 2016), 566–77, https://doi.org/10.1007/978-981-10-2666-9_57 ; R. Rasch, A. Kott, and K.D. Forbus, "Incorporating AI into Military Decision Making: An Experiment," *IEEE Intelligent Systems* 18, no. 4 (July 2003): 18–26, https://doi.org/10.1109/MIS.2003.1217624.

[28] Edward Moore Geist, "It's Already Too Late to Stop the AI Arms Race—We Must Manage It Instead," *Bulletin of the Atomic Scientists* 72, no. 5 (September 2, 2016): 319, https://doi.org/10.1080/00963402.2016.1216672.

[29] Feng et al., "Inspiration for Battlefield Situation Cognition from AI Military Programs Launched by DARPA of USA and Development of AI Technology," 567.

[30] Lee, *AI Superpowers*, 136.

[31] Gregory C. Allen, *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security* (Washington, DC: Center for New American Security, 2019), 8–9, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

information within their own country.[32] So far, this unregulated approach by the Chinese has proven to be an effective method, according to Lee.[33] Currently, several researchers agree that the state-backed Chinese development process and the huge amount of unregulated data, allows China to compete at the same level as American academia and company-led research.[34]

By contrast, as the Swedish Defence Research Institute (SDRA) paper by Peter Svenmarck et al. "Possibilities and Challenges for Artificial Intelligence in Military Applications," prepared for the North Atlantic Treaty Organization (NATO), points out, regardless of the chosen approach, the three main challenges of transparency, security, and data must be considered when implementing AI into Military Activities.[35] First, whether the product relies on Machine Learning (ML) or Deep Learning (DL), the actual algorithms producing the result need to be known and evaluated.[36] Revealing the algorithms seldom lies in the interest of the seller of the product, however, since these are the actual source of the product, and revealing them makes the company more vulnerable to exploitation and loss of information.[37] Furthermore, even if the algorithms are known, the need for an educated procurement chain arises, entailing close cooperation with academia or constant recruitment and education of personnel, as described by Elsa B. Kania in her article "Chinese Military Innovation in the AI Revolution."[38] Further, in their recent article published in *War On The Rocks:* "Can Warfighters Remain the Masters of AI?," Harrison

---

[32] Horowitz et al., "Strategic Competition in an Era of Artificial Intelligence," 9.

[33] Lee, *AI Superpowers*, 82–84.

[34] Lee, 91–92; Taulli, *Artificial Intelligence Basics*, 167; Horowitz et al., "Strategic Competition in an Era of Artificial Intelligence," 12.

[35] Peter Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," *NATO Big Data and Artificial Intelligence for Military Decision Making,* (Stockholm: Swedish Defence Research Agency, May 2018), 4, https://www.researchgate.net/publication/326774966_Possibilities_and_Challenges_for_Artificial_Intelligence_in_Military_Applications.

[36] Svenmarck et al., 2.

[37] Lee, *AI Superpowers*, 91.

[38] Kania, "Chinese Military Innovation in the AI Revolution," 34.

Schramm and Jeff Kline highlight the problem of obtaining sufficiently educated personnel to utilize AI.[39]

Second, implementing AI offers new security considerations for organizations; how will commercial companies guarantee a military organization that the enemy does not have insight into and the ability to exploit weaknesses in these developments through manipulation of the data or the model?[40] Albeit companies and universities responsible for developing the algorithms might have a strong incentive for keeping them a secret, it has hardly prevented copying and exploitation in the past.[41] There seems to be a great risk of buying into a system at great cost, only to have the system become a risk rather than an asset for the organization. Furthermore, the possible manipulation of data is a problem that was noted as early as 1985 by Robert Trappl in his book *Impacts of Artificial Intelligence*, and emphasized in 2018 by Daniel S. Hoadley and Nathan J. Lucas in their Congressional Research Service paper "Artificial Intelligence and National Security."[42] As the systems themselves are getting more and more complicated, the possibilities to identify and correct a faulty system becomes harder and harder, if not impossible in a time-sensitive situation such as war.[43] This statement certainly applies today, as it is getting hard even for the designers to back-track the process within several layers of DL.[44]

Lastly, the available data quantities and quality are crucial in determining what kind of systems to consider and to what degree it will be effective enough to motivate the

---

[39] Harrison Schramm and Jeff Kline, "Can Warfighters Remain the Masters of AI?" *War on the Rocks*, February 6, 2020, https://warontherocks.com/2020/02/can-warfighters-remain-the-masters-of-ai/.

[40] Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," 7.

[41] Lee, *AI Superpowers*, 91.

[42] Taulli, *Artificial Intelligence Basics*, 37; Daniel S. Hoadley and Nathan J. Lucas, *Artificial Intelligence and National Security*, CRS Report R45178 (Washington, DC: Congressional Research Service, 2018), 28–30, https://digital.library.unt.edu/ark:/67531/metadc1157028/m1/1/.

[43] Trappl, *Impacts of Artificial Intelligence*, 37.

[44] Hoadley and Lucas, "Artificial Intelligence and National Security," 31.

investment.[45] As Lee points out, an organization might not reap the anticipated benefits of acquiring a costly AI system if the organization lacks access to considerable amounts of the "right" data to analyze.[46] For a relatively small community, such as SOF, this could cause a problem.[47] Therefore, it seems wise to consider working together with other countries and organizations to maximize the benefits of AI systems. However, this also implies that interagency regulations and international law, as well as national laws, must be taken into consideration, furthering the need for a complete judicial analysis before acquiring certain systems.[48]

Even though the considerations mentioned previously apply to military procurement of AI in general, they might provide a foundation for further development of a smaller set of guidelines more specifically applicable to the needs of SOF. Given that SOF units and the special circumstances surrounding their missions often require bespoke equipment, employed under special conditions, the Joint Publication 3-05 describes Special Operations as:

> Operations requiring unique modes of employment, tactical techniques, equipment and training often conducted in hostile, denied, or politically sensitive environments and characterized by one or more of the following: time sensitive, clandestine, low visibility, conducted with and/or through indigenous forces, requiring regional expertise, and/or a high degree of risk.[49]

As noticed by Michael A. Pfarrer in his 2000 Naval Postgraduate School (NPS) master's thesis "Optimizing Procurement of Special Operations Weapons," the usual procurement process of a nation's army, air force, or navy seldom meets the requirements

---

[45] Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," 8.

[46] Lee, *AI Superpowers*, 56.

[47] Elliot A. Cohen, *Commandos and Politicians: Elite Military Units in Modern Democracies*, Harvard Studies of International Affairs 40 (Cambridge, MA: Center for International Affairs at Harvard University, 1978), 54.

[48] Richard Kemp, *Legal Aspects of Artificial Intelligence* (London, UK: Kemp It Law, 2016), 1.

[49] Joint Chiefs of Staff, *Special Operations*, JP 3-05 (Washington, DC: Joint Chiefs of Staff, 2014), GL-11, http://edocs.nps.edu/2014/July/jp3_05.pdf.

of the SOF community.[50] Therefore, it seems likely that separate factors might apply for a smaller organization, such as SOF, when considering implementation of AI into the organization.

B. SUMMARY

Even though AI may continue to change several aspects of society, many questions still need addressing before this technology turns life upside down. History provides many examples when innovation has altered the world, often for the better in retrospect. Not many innovations have revolutionized society within a few years, but rather most have led to incremental changes to which the world has adapted. As the world continues to change, it seems likely that the integration and advances in AI will continue to gradually affect the military as well. Nonetheless, considering humans' reluctance to adapt to changes, SOF and their unique requirements might be fortuitously positioned at the tip of the spear in this technological race, if they can adapt a sound procurement process for the assessment and acquisition of AI.

[50] Michael A. Pfarrer, "Optimizing Procurement of Special Operations Weapons and Equipment" (master's thesis, Monterey, CA, Naval Postgraduate School, 2000), xii, https://calhoun.nps.edu/bitstream/handle/10945/7758/optimizingprocur00pfar.pdf?sequence=1&isAllowed=y.

III. RESEARCH APPROACH

Aiming to produce a SOF-specific model to assist in the assessment of whether a particular type of AI is viable for implementation, this capstone builds on the foundational considerations of transparency, security, and data identified in the previously mentioned cooperative paper from NATO and the SDRA, "Possibilities and Challenges for Artificial Intelligence in Military Applications." To develop a model, this research uses an iterative effort to identify SOF-specific factors through a combination of secondary sources and interviews. In the next section, the applicability of the chosen model in a SOF-specific context is explained, followed by a description of the proposed conduct and selected method of the research. Finally, a summary of the expected results of the study clarifies the aim of the chosen approach.

A. TENTATIVE MODEL FOR CONSIDERING AI PROCUREMENT

Representing the three areas of primary concern for military acquisition of AI products, the SDRA paper's aim is to present NATO forces with guidelines for what to consider before investing in AI. These areas are further elaborated on in the previous chapter "Military and AI" and consist of:

1. Transparency
2. Security
3. Data

Since these factors are presented as essential considerations when investigating the potential benefits and risks of AI, they will provide the foundation for analysis in this capstone.[51] As an integral part of a nation's military forces, SOF often provide the testbed for new tactics, equipment, and rely on different procurement chains, budget, and operational requirements.[52] Therefore, based on the assumption that SOF has specific

[51] Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," 1.

[52] Cohen, *Commandos and Politicians: Elite Military Units in Modern Democracies*, 31–32.

needs and structural differences from the military in general, the developed model should reflect these needs in other or new areas of consideration for SOF, as they relate to AI. Hence, the overarching question of this capstone is: *What factors must the SOF community consider when acquiring and evaluating new AI capabilities?*

To identify these factors and determine what the SOF community can leverage in order to maintain the edge within AI implementation, the following questions provide the framework for the analysis:

- *What is currently considered to be AI?*
  - As there is no consensus on what constitutes AI, does this debate provide opportunities, or does it only present hurdles to overcome?
- *What are the current ethical considerations regarding AI?*
  - What is perceived as the most fundamental ethical problem with AI and does this make some areas of AI more troublesome for implementation?
- *What are the demands for infrastructure?*
  - Do the requirements for infrastructure, maintenance, educated personnel, or need for data make some of the current AI inventions unsuitable for SOF?
- *How can AI be evaluated for effectiveness?*
  - How can the actual efficacy of implementation be measured and what does this imply for organizations acquiring products?
- *What is the cost of implementing AI?*

- Does the cost of certain AI-based systems prevent a smaller organization from acquiring equipment, regardless of perceived benefit?

- *How can SOF leverage its shorter procurement chains?*
  - Are there any benefits to acquiring existing AI "Commercial of the Shelf" (COTS) products compared to bespoke military products?

B. METHODOLOGY

Aiming is to discover something new rather than testing an existing theory, this capstone utilizes a mixed-methods approach, incorporating both qualitative and quantitative data. Since the research to a large extent will be iterative, based on reading and interpreting text, a hermeneutic approach fits the study well, as explained by Sebastian K. Boell and Dubravka Cecez-Kecmanovic in their article "A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches":

> The hermeneutic route to understanding is always iterative: an understanding of a text (a part) draws from the reader's preunderstanding of a context (a whole); and vice versa, the understanding of a context (a whole) develops from understanding individual texts or text equivalents (parts).[53]

Furthermore, with the intent to explore how AI fits into the greater context of military applications, a holistic, qualitative approach is chosen, rather than exploring the individual metrics of some particular products within the AI field.[54] On the other hand, as qualitative studies do not rely on quantifiable data, the validity of the research is often questioned; therefore, several sources of data are used to enhance the credibility of the

---

[53] Sebastian K. Boell and Dubravka Cecez-Kecmanovic, "A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches," *Communications of the Association for Information Systems* 34 (2014): 263, https://doi.org/10.17705/1CAIS.03412.

[54] Barbara M. Wildemuth, *Applications of Social Research Methods to Questions in Information and Library Science*, 2nd ed. (Santa Barbara, CA: ABC-CLIO, 2017), 319.

study through the use of data triangulation.[55] If there are existing quantitative data available to strengthen the findings of the research further, these are incorporated, as suggested by David A. Lake, in his *-ism* critical article: "Why "isms" Are Evil: Theory, Epistemology, and Academic Sects as Impediments to Understanding and Progress."[56]

C. DATA COLLECTION

A mix of textual analysis, interviews, and survey of the current product developments in the field of AI is used to provide the necessary information for the research. Although most information can be acquired through secondary sources like books, the internet, and papers written on the subject, the field of AI is continually changing. Therefore, the information gathered through secondary sources is complemented by interviews conducted with persons connected to relevant AI competencies within the commercial sector, military, and academia. Furthermore, throughout the research, recent developments in AI use in the military reported through news outlets and seminars are incorporated to maintain relevancy. Following is a brief clarification of the sources.

1. Secondary Sources

Current articles, books, and theses provide the foundation for the research. Both AI-oriented literature and works regarding military procurement are utilized to evaluate and produce a working model. Sources were scrutinized for relevance, credibility, and potential biases. Journals used for the research consist of military publications, such as *War on the Rocks*, conference proceedings, and RAND, as well as technology-centric publications like *MIT Technology Review* and other publications on AI development. However, as the AI field is continually changing, there is a risk that new material emerges during the research, which is why it becomes imperative to repeat and update the analysis constantly.

---

[55] Lisa A. Guion, David C. Diehl, and Debra McDonald, "View of Triangulation: Establishing the Validity of Qualitative Studies," *EDIS* 2011, no. 8 (August 2011): 1–2, https://doi.org/10.32473/edis-fy394-2011.

[56] David A. Lake, "Why 'Isms' Are Evil: Theory, Epistemology, and Academic Sects as Impediments to Understanding and Progress," *International Studies Quarterly* 55, no. 2 (2011): 475–76, http://www.jstor.org/stable/23019696.

2. Interviews

Through the Naval Postgraduate School's (NPS) collaboration with several local AI corporations in the San Francisco area, the authors gained access to commercial companies producing military applications, which allows the research to explore a non-military perspective on military procurement. Since this capstone uses an inductive method and qualitative research strategy with a focused group as its design, data is gathered through oral interviews using a pre-approved bank of questions in compliance with the NPS Human Research Protection Program Office & Institutional Review Board (IRB) process.[57] Although there are disadvantages to face-to-face meetings as they are time-consuming, and require coding and transcription, the interviews are conducted face-to-face through media platforms, allowing for the assessment body language, the ability to ask to follow-up questions, and the reduction in delay between questions and answers.[58] Furthermore, even though the interviews are semi-structured, meaning there is a set of questions to follow, this approach still allows for the interviewer and the interviewee to elaborate on answers and follow up on issues identified in during the interview.[59] At the same time, since commercial companies are driven by the need for continued profit, this was considered when evaluating the information collected from company representatives. The interviews were recorded, coded, and transcribed to allow for back-tracing.

3. Seminars and Product Fairs

Seminars oriented towards military application can provide information outside the previously mentioned sources, thereby adding a new perspective based on the actual market available compared to the more objective views of singular enterprises. However, to identify a possible divide between the needs of the SOF and the regular military, the

---

[57] "Human Research Protection Program Office & Institutional Review Board (IRB)," Naval Postgraduate School, accessed May 4, 2021, https://nps.edu/web/research/irb-home.

[58] Raymond Opdenakker, "Advantages and Disadvantages of Four Interview Techniques in Qualitative Research," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 7, no. 4 (September 2006): 1–14, https://doi.org/10.17169/fqs-7.4.175.

[59] Neil Stephens, "Collecting Data from Elites and Ultra Elites: Telephone and Face-to-Face Interviews with Macroeconomists," *Qualitative Research* 7, no. 2 (May 2007): 203–16, https://doi.org/10.1177/1468794107076020.

research on current capabilities offered helps identify potential areas targeted by the commercial sector when selling to specific customers.

Due to the current effect of Covid-19, most seminars and product fairs were suspended during the information-gathering process for this capstone. Although further direct access, apart from the seminars and company briefs attended, would have been preferable; articles, product leaflets, and news outlets have supplemented as sources for data gathering.

D. VALIDITY OF SOURCES

Information tends to be subjective, whether intentionally or not. Therefore, all sources used in the research are scrutinized according to the following factors based on the Currency, Relevance, Authority, Accuracy, and Purpose (CRAAP) test criteria:[60]

- *Type:* Relates to the actual nature of the information, such as peer-reviewed articles, interviews, seminars, product leaflets, or product spokespersons.
- *Possible dependence:* Does the source of information have connections or demands that might produce a biased result?
- *First or secondhand information:* Is the source providing its unique knowledge, or is the information given based on previous or other sources?
- *Tendencies:* Is the source likely to be favorable of a specific idea or concept?

Keeping this simple framework in mind, while conducting the research, the authors could reduce the risk of promoting a single point of view. Furthermore, to mitigate the

[60] Dawn Emsellem Wichowski and Laura E. Kohl. "Establishing Credibility in the Information Jungle: Blogs, Microblogs, and the CRAAP Test." In *Online Credibility and Digital Ethos: Evaluating Computer-Mediated Communication,* ed. Moe Folk and Shawn Apostel (Hershey, PA: IGI Global, 2013), 231. http://doi:10.4018/978-1-4666-2663-8.ch013.

apparent potential for bias, several types of sources and methods of data collection (triangulation of information) are used.[61] However, if a source provides valuable information but can be questioned regarding credibility or potential bias, the authors mention that in the text.

E. EXPECTED RESULTS

Since the commercial sector currently is at the forefront of development in AI, there is a myriad of innovations appearing on the market. The short lifespan of these inventions does not justify significant investment since there will soon be a new, better, faster, and more potent products on the market. Provided the opportunities a fast-moving field such as AI offers, readily available COTS products, combined with the quicker SOF procurement chain procurement, should present SOF with the opportunity to maintain the edge on the battlefield. On the other hand, since there is, and will continue to exist, a need to develop highly specialized systems for managing complex tasks that require large infrastructures for data management, maintenance, and know-how as well as special considerations regarding their effects on existing systems, these more demanding, bespoke systems present a problem for the smaller SOF communities. Yet, in recognizing the need for both rapid procurement and robust infrastructure, it should be possible to establish a division of responsibilities regarding the implementation of AI between SOF and the regular armed forces. Building on the traditional use of SOF as a platform for experimentation, with a focus on the implementation of COTS products, while the regular armed forces focus on the demands of more complex AI-systems, can form a division of responsibilities within AI-implementation capable of creating a mutually beneficial synergy between SOF and conventional forces.

---

[61] Sharan B. Merriam and Elizabeth J. Tisdell, *Qualitative Research: A Guide to Design and Implementation* (Newark, NJ: John Wiley & Sons, Incorporated, 2015), 259, http://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=2089475.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSIS

In this chapter, the three main pillars of transparency, security, and data proposed for the tentative model are analyzed and their actual meaning discussed. Using both historical and contemporary research, this chapter scrutinizes the implications and possibilities of the three pillars for their relevance when answering the overarching research question of what SOF must consider when acquiring AI. Each pillar is first presented and the connection to AI established, followed by an examination. Its historical meaning is contrasted against contemporary research and its relevance to AI is discussed. Finally, a summary of the analysis presents the identified specific implications for SOF procurement of AI.

A. TRANSPARENCY

In the case of AI, transparency often refers to whether it is possible to explain, understand, and interpret the actions and results of the AI-model, either focusing on how an AI-model reaches its decision or examining the results.[62] Based on the notion that AI will ultimately reach a point where it surpasses human understanding and develop a mind of its own, one prevailing argument has been that AI needs to be transparent in order to be trusted.[63] With the apocalyptic versions of society depicted in films like James Cameron's *The Terminator* and Stanley Kubrick's classic *2001: A Space Odyssey* where AI takes on a life of its own and refuses to cooperate with humans, as well as recent failures of self-driving cars in mind, trust in AI continues to be an issue that is highly debated.[64] Aiming to further explore the argument that transparency is necessary, the actual meaning of transparency and its relation to trust is explored, followed by its implications for

---

[62] Zachary C. Lipton, "In Machine Learning, the Concept of Interpretability Is Both Important and Slippery," *Queue* 16, no. 3 (May 2018): 6, https://dl.acm.org/doi/pdf/10.1145/3236386.3241340.

[63] Andrew Burt, "The AI Transparency Paradox," *Harvard Business Review,* summary. December 2019, https://hbr.org/2019/12/the-ai-transparency-paradox.

[64] George Hurlburt, "How Much to Trust Artificial Intelligence?" *IT Professional* 19, no. 4 (2017): 7, https://doi.org/10.1109/MITP.2017.3051326.

transparency in AI. Finally, building on the arguments made in the subsequent discussion, three suggestions are proposed to enhance the effects of transparency to promote trust in AI.

This analysis concludes that while transparency might be a good way to achieve quality assurance, institutions and organizations can only leverage transparency if they are capable of accurately evaluating the products. To best benefit from transparency, procurement of AI should rely on historically trustworthy partners and focus on increasing the knowledge within organizations and institutions to manage the user's expectations on AI. Trust in AI, then, is more reliant on human perception, interaction, and the trustworthiness of the companies developing future solutions than the actual transparency of AI itself.

1. Transparency and Trust

Transparency as a term, although widely used and hard to define, is often related to the ability to create trust and has been frequently used in the field of economics.[65] One example is Brad L. Rawlins' paper from 2008 "Measuring the Relationship between Organizational Transparency and Employee Trust," in which Rawlins suggests that an individual's perception of an organization's transparency is closely related to the level of trust.[66] Rawlins adheres to Elinor Ostrom and James Walker's definition of trust in the book *Whom Can We Trust?*, "the willingness to take some risk in relation to other individuals on the expectation that the others will reciprocate."[67] The idea that presenting vulnerability is necessary to build trust, not only correlates with transparency as a concept of being accountable and sharing information that is potentially harmful, but also builds on

---

[65] Stefan Larsson and Fredrik Heintz, "Transparency in Artificial Intelligence," *Internet Policy Review* 9, no. 2 (May 5, 2020): 4–5, https://doi.org/10.14763/2020.2.1469.

[66] Brad R. Rawlins, "Measuring the Relationship between Organizational Transparency and Employee Trust," *Public Relations Journal* 2, no. 2 (2008): 15, https://prjournal.instituteforpr.org/wp-content/uploads/Measuring-the-relationship.pdf.

[67] James Walker and Elinor Ostrom, "Trust and Reciprocity as Foundations for Cooperation," in *Whom Can We Trust? How Groups, Networks, and Institutions Make Trust Possible*, ed. Karen S. Cook, Margaret Levi, and Russell Hardin (New York, NY: Russell Sage Foundation, 2009), 91.

Ostrom's belief that reciprocity in order to be trusted, you need to show trust — is needed for trust to be beneficial. However, research more focused on the correlation between expectations and trust, conducted by René F. Kizilicec from Stanford University, suggests that although there is a correlation between transparency and trust, the perception of violated expectations greatly affects the benefits of transparency.[68]

Trust is also a topic that has been thoroughly researched in several different areas, such as sociology, economy, and political science.[69] With this in mind, trust is often explored through different types of social exchanges of goods between people. In his article "The Market for Lemons," Nobel Prize winning author George Akerlof posits that there is a clear link between information asymmetry, uncertainty, and the level of trust in a market. Akerlof argues the need for institutions to control and maintain a level of certainty for a market to exist.[70] Although these ideas do not directly assess the effects of transparency, the idea of limiting the information asymmetry and uncertainty for trust to exist can be related to the level of transparency of a product.

Contrary to Akerlof´s belief in institutional trust, another Nobel Prize winning scientist, Elinor Ostrom, argues in her article "Solving the Problem of the Commons," that the need for institutions is overemphasized since trust is based on reciprocity rather than sanctions alone.[71] Ostrom's conclusions are further supported by Karen S. Cook, Russel Hardin, and Margaret Levi, who conclude that "A trust relation emerges out of mutual interdependence and the knowledge developed over time of reciprocal trustworthiness."[72]

---

[68] René F. Kizilcec, "How Much Information?: Effects of Transparency on Trust in an Algorithmic Interface," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, CA: ACM, 2016), 2393, https://doi.org/10.1145/2858036.2858402.

[69] Russell Hardin, *Trust and Trustworthiness*, 4th ed. (New York, NY: Russell Sage Foundation, 2002), xviii.

[70] George A. Akerlof, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics* 84, no. 3 (August 1970): 488, https://doi.org/10.2307/1879431.

[71] Elinor Ostrom, "Coping with Tragedies of The Commons," *Annual Review of Political Science* 2, no. 1 (June 1999): 507–8, https://doi.org/10.1146/annurev.polisci.2.1.493.

[72] Karen S. Cook, Russell Hardin, and Margaret Levi, *Cooperation Without Trust*, vol. IX (New York, NY: Russell Sage Foundation, 2005), 2.

Furthermore, Cook, Hardin, and Levi dismiss the notion that trust is provided by regulations and institutions. Instead, they argue that institutions and regulations are necessary when trust does not exist, since trust is an encapsulated interest —our expectancy that our interest and the other party's interest coincide to be mutually beneficial— where the incentive of being trustworthy is incumbent upon the parties.[73]

In terms of trust in AI, one question is whether trust should be directed at the actual AI model or how the model is perceived by its users.[74] Marco Ribeiro, Sameer Singh, and Carlos Guestrin propose that for humans to trust AI, interpretability and explainability is crucial.[75] Based on their argument that "if the users do not trust a model or a prediction, they will not use it," they conclude that individuals need a level of transparency to trust AI.[76] On the other hand, Andrea Ferrario, Michele Loi, and Eleonora Viganò of the Mobiliar Lab for Analytics in Zürich contest this idea and argue that trustworthiness rather relies on the perception and education of the user rather than the actual AI model itself.[77] Furthermore, they refute the idea that lack of trust will prohibit the use of AI; instead, they relate the trust in AI to necessity. The AI model will be considered *relatively* trustworthy if the user is limited in options, but an AI is *absolutely* trustworthy if it is the best option available.[78]

2. Transparency in AI

Transparency in AI can simply be explained as "for a model to be fully understood, a human should be able to take the input data together with the parameters of the model

---

[73] Cook, Hardin, and Levi, *Cooperation Without Trust*, 5.

[74] Larsson and Heintz, "Transparency in Artificial Intelligence," 7.

[75] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," *ArXiv:1602.04938* (San Jose, CA: Association for Computational Linguistics August 9, 2016): 10, http://arxiv.org/abs/1602.04938.

[76] Ribeiro, Singh, and Guestrin, 1.

[77] Andrea Ferrario, Michele Loi, and Eleonora Viganò, "In AI We Trust Incrementally: A Multi-Layer Model of Trust to Analyze Human-Artificial Intelligence Interactions," *Philosophy & Technology* 33 (October 2019): 537, https://doi.org/10.1007/s13347-019-00378-3.

[78] Ferrario, Loi, and Viganò, 536.

and in reasonable time step through every calculation required to produce a prediction."[79] Simply put, a human has to be able to understand how and why an AI system has reached its conclusion in order to be able to trust the outcome.[80] Transparency, then, allows buyers to trust sellers and their product by evaluating the quality, as argued by Akerlof in "The Market for Lemons."[81] Following Akerlof's argument, institutions and other mechanisms can therefore be used to ensure the quality and limit the asymmetry of information in the market, making buyers trust the sellers.[82] Nonetheless, as research has shown quality assurance and product transparency can affect trade in the opposite way, making looser bonds between buyers possible as the risk of investment goes down.[83] This, then, would indicate that if transparency equates to quality assurance, it might also negate the need for trust under certain circumstances.

Similar to the arguments just described, Akerlof posits that uncertainty, and the information asymmetry, must be reduced between the buyer and seller of goods to create a functioning market. Therefore, if there is a way for institutions to certify the quality of a product, there will be a reason for the buyer to trust the seller, preventing the market from being filled with 'lemons.'[84] In this aspect, the need for transparency seems to hold true: If the asymmetry in information between the seller and buyer is minimized through the interpretability and explainability of the product, this should provide the basis for building trust.[85] Furthermore, in providing insight into the workings of the products, a manufacturer fulfills the parameter of exposing vulnerability, which aligns with Ostrom and Walker's

---

[79] Lipton, "In Machine Learning, the Concept of Interpretability Is Both Important and Slippery," 13.

[80] Keng Siau and Weiyu Wang, "Building Trust in Artificial Intelligence, Machine Learning, and Robotics," *Cutter Business Technology Journal* 31, no. 2 (2018): 53, https://www.cutter.com/sites/default/files/itjournal/2018/cbtj1802.pdf.

[81] Akerlof, "The Market for 'Lemons'," 490.

[82] Akerlof, 499–500.

[83] Peter Kollock, "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust," *American Journal of Sociology* 100, no. 2 (September 1994): 316, 318, 342, https://doi.org/10.1086/230539.

[84] Akerlof, "The Market for 'Lemons.'"

[85] Akerlof, 500.

concept of trust.[86] There are some indications, however, that transparency might be counterproductive to trust as well.

First, Brad Rawlings' somewhat brash definition of transparency put forward in his paper "Measuring the Relationship between Organizational Transparency and Employee Trust," where he states that "Simply put, transparency is the opposite of secrecy," raises questions about transparency's applicability in the realm of AI.[87] As several studies point out, attacks against AI can be significantly more effective if information about the target is known.[88] Therefore, it seems that although transparency can act as a factor in building trust, the gains of transparency must be weighed against the security implications for AI.[89] Hence, the vulnerability aspect seems to limit the utility of transparency under certain conditions.

Secondly, as René F. Kizilcec's paper "Much Information? Effects of Transparency on Trust in an Algorithmic Interface" suggests: "Transparency may promote or erode users' trust in a system by changing beliefs about its trustworthiness."[90] Kizilcec's study suggests that only when the expectations of the user have not been violated can increased transparency also affect trust positively. Conversely, too much information seemed to negate the positive effect when expectations were violated.[91] This would imply that in an

---

[86] Walker and Ostrom, "Trust and Reciprocity as Foundations for Cooperation," 91.

[87] Rawlins, "Measuring the Relationship between Organizational Transparency and Employee Trust," 6.

[88] Marcus Comiter, *Attacking Artificial Intelligence: AI's Security Vulnerability and What Policymakers Can Do About It*, Belfer Center Paper (Cambridge, MA: Belfer Center for Science and International Affairs, 2019), 63, https://www.belfercenter.org/sites/default/files/2019-08/AttackingAI/AttackingAI.pdf; Tong Chen et al., "Adversarial Attack and Defense in Reinforcement Learning-from AI Security View," *Cybersecurity* 2, no. 1 (December 2019): 20, https://doi.org/10.1186/s42400-019-0027-x; Ilja Moisejevs, "Poisoning Attacks on Machine Learning: A 15-Year-Old Security Problem That's Making a Comeback," Towards Data Science, July 2019, Attacker capabilities and Poisoning Attacks, https://towardsdatascience.com/poisoning-attacks-on-machine-learning-1ff247c254db.

[89] Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," 16.

[90] Kizilcec, "How Much Information?" 2390.

[91] Kizilcec, 2392.

AI context, transparency to a certain degree can maintain trust, but ultimately, trust is based on products meeting expectations.

3. Trust in AI

Trust and trustworthiness are complex issues that are hard to define. However, consistent with both Elinor Ostrom's argument about the need for communication and repeated interactions, as well as Akerlof's point about the need for quality assurance, Russel Hardin sums it up pretty well in his book *Trust and Trustworthiness*:

> If we consider all the trust relations we experience, we find that a large fraction of them fall into three categories: relationships or interactions that are iterated, those that are backed up by institutions, and those that are mediated by other (noninstitutional) third parties.[92]

In terms of AI, trust seems to fall into these three categories: people have re-occurring interactions with AI, they trust institutions and certain brands to validate the performance of AI being sold in the market, and they expect academics and inventors to strive for the greater good of the public. Furthermore, in order to clarify the relationship between trust and public relations, in their 1999 paper "Guidelines for Measuring Relationships in Public Relations," Linda Childers Hon and James Grunig point out that perceived integrity, dependability, and competence make up the three factors of trust in a relationship.[93] With this in mind, trust is the expectation for another party to consistently fulfill expectations based on perceived capabilities.[94] Transferred to AI, this would imply that it is our knowledge and expectations on the performance of AI models, based on the manufacturer's promise, that makes up the foundation for our trust.

In their article "In AI We Trust Incrementally: A Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions" Andrea Ferrario, Michele Loi, and Eleonora Viganò provide support for the previous argument. First, they posit that "In AI,

---

[92] Hardin, *Trust and Trustworthiness*, 25.

[93] Linda Childers Hon et al., *Guidelines Measuring Relationships*, Commission on Public Relations Measurement & Evaluation (Gainesville, FL: Institute for Public Relations, 1999), 3, http://painepublishing.com/wp-content/uploads/2013/10/Guidelines_Measuring_Relationships.pdf.

[94] Hon et al., 24–25.

the strength of evidence of trustworthiness regarding properties of AI systems decreases with the decreasing expertise of the agent interacting with the systems."[95] This implies that the level of education of the operator is relevant for the perceived trust in a certain system or model. This correlates with the argument previously mentioned by Kizilcec where information is only necessary to a certain point. Second, Ferrario et al. suggest that the trustworthiness of AI is directly related to the options available as an alternative solution.[96] In short, if there are no other solutions available, then users tend to subjectively trust the chosen solution. As previously mentioned, since AI has already proliferated in society, and lacking alternative solutions, we already subjectively trust AI.

4. Implications of Transparency for Trust in AI

Considering the arguments made previously, transparency seems to be a double-edged sword. On the one hand, transparency can promote trust by removing uncertainty and information asymmetry. On the other hand, there also seems to be a limit to how much transparency is beneficial for trust, and instead decreases trust by threatening to make the systems vulnerable. Crucial for transparency, then, is to acknowledge that trust in AI seems to depend at least as much on the user's perception of the product as the product itself. Considering that the effect of transparency is closely related to the education level of the personnel responsible for utilizing the AI, as both Kizilcec's and the Ferrario, Loi, and Viganò's papers argue, their logic seems reasonable: Building trust in AI is more about expectation management than the actual transparency or results of the products themselves.[97] Yet, given the proper education of personnel evaluating the AI product, transparency can both increase trust and serve as a quality assurance.

In order to clarify the argument just presented, a car might provide a good analogy. We trust our car to work, and when it does not, we usually hand it over to a mechanic for repairs. Even though all information about the inner workings of a car might be readily

---

[95] Ferrario, Loi, and Viganò, "In AI We Trust Incrementally," 536.

[96] Ferrario, Loi, and Viganò, 536.

[97] Kizilcec, "How Much Information?" 2392; Ferrario, Loi, and Viganò, "In AI We Trust Incrementally," 535.

available, few can say that they know everything about how their vehicle works. If problems keep occurring, the dependability of the car is in question, and we tend to lose trust in our car. While the mechanic can perceive the problems as minor, this does not mitigate the fact that we cannot use the car during the time it takes to repair it. Consequently, if we have similar experiences with several cars from the same manufacturer, the chances are that next time, we will choose a different brand. That said, we rarely lose trust in all vehicles and resort to walking, especially if there are no viable alternatives available.[98] Instead, we might end up with a new car, which we trust, until it too starts to break down.

In the context of AI, trust does not seem to be all that different. Although some argue that full transparency is required and persons operating AI systems should be experts, this hardly seems like a realistic solution. Instead, we should be practical and focus on maintaining and increasing the existing trust by providing a foundation where we can benefit from transparency. Following are three suggestions on how transparency can increase trust in AI:

a. Trustworthy Partners

In the field of AI, there are certain companies that invoke more trust than others, such as Google, Microsoft, or Apple.[99] As these companies stake their reputation on their products, they also make themselves vulnerable, providing a foundation for trust and a guarantee for quality.[100] Although the existing trust might not lie in specific AI models, their liability remains the same. The ability to build on this existing trust (brand naming) and providing sufficient transparency for the products to be evaluated by qualified personnel (institutions) should increase public trust in products along with the need for quality assurance, as argued by Akerlof.[101] Hence, in this case, transparency also relates

---

[98] Ferrario, Loi, and Viganò, 536.

[99] Gry Hasselbalch, "Why Trust in AI Is Not Enough," *Dataethics* (blog), September 3, 2019, https://dataethics.eu/why-trust-in-ai-is-not-enough/.

[100] Walker and Ostrom, "Trust and Reciprocity as Foundations for Cooperation," 91.

[101] Akerlof, "The Market for 'Lemons'," 500.

to the reciprocal part of trust, where companies are willing to provide transparency into their corporations to remain trusted by the users.

b. Education

As highlighted in the 2020 Cyberspace Solarium Commission (CSC) report "A Warning from Tomorrow," there is a deficit in the number of skilled personnel to operate and evaluate the systems, even in an AI superpower such as the United States.[102] With this shortcoming in mind, organizations and institutions should prioritize the education of a workforce responsible for the evaluation of AI. As argued by Ferrario, Loi, and Viganò, to best benefit from transparency, manufacturers need an educated counterpart to be able to evaluate their product. However, building on the example of trust in cars, perhaps we should not aim for a nation of mechanics, but rather a workforce of informed users with AI driver's licenses, backed up by mechanics.

c. Expectation Management

Although transparency might give an insight into the expected functionality of a new product, it will not be able to create trust if the product is perceived as insecure, undependable, or unable to produce the expected results. Therefore, Kizilcec's argument about expectation management seems crucial. However, since proper evaluation of products is necessary to produce reasonable expectations, Akerlof 's argument regarding the need for quality assurance and institutions also holds true. Hence, a certain level of transparency is necessary to allow institutions and organizations to evaluate products in order to produce reasonable expectations in users. Through providing sufficient transparency, producers can help manage the expectations of users to maintain and increase trust, as argued by Kizilcec.

---

[102] Peter Singer et al., *A Warning from Tomorrow* (Arlington, VA: Cyberspace Solarium Commission, March 2020), 43, https://drive.google.com/file/d/1ryMCIL_dZ30QyjFqFkkf10MxIXJGT4yv/view.

5. Conclusion Analysis of Transparency

AI is, and will continue to be, an issue of debate for the foreseeable future. Merely the fact that there is no clear definition of what actually constitutes AI is enough for several books to be written. However, in the debate about the need for transparency as a foundation for trust, three points can be made: First, although transparency is an essential part of quality assurance and allows for trust to be maintained, human perception and historical interaction seem to play a more important part in trusting AI. Furthermore, considering the security implications related to transparency in AI, maintaining trustworthy partners remains crucial. Second, the effect of transparency in building trust is closely related to the potential user's level of understanding and ability to evaluate the products. To benefit fully from transparency, organizations must focus efforts on educating operators supported by a core of highly skilled personnel, who are able to evaluate the potential benefits of incorporation of AI. Third, transparency can provide the foundation for organizations and institutions to manage the expectations for AI. Allowing for products to be scrutinized can help organizations align expectations with actual results delivered, thereby increasing and maintaining trust in AI when the products are consistently reliable and dependable at performing a given task.

Transparency, then, does not automatically equate to increased trust in AI. Transparency relies on the institutions, organizations, and the user's ability to evaluate the products and correctly assess the benefits and expected results of the products. In a world where AI is already incorporated throughout society, people have already begun to trust AI by proxy. Given this reality, transparency should be recognized as the means to enhance the already existing trust, instead of a requirement to build trust.

B. SECURITY

As most nations have already implemented different applications using AI within their societies, military establishments around the world find themselves faced with the fact that AI solutions are a part of both the offensive and defensive cyber capabilities.[103]

[103] Geist, "It's Already Too Late to Stop the AI Arms Race—We Must Manage It Instead," 318.

While the United States has adopted a gradual and restrictive approach, China opts for an unrestricted implementation, possibly outpacing the rest of the world in the near future.[104] In times of crisis, such as the Covid-19 outbreak, however, even the Western world has acknowledged that the benefits of speeding up the use of AI sometimes outweighs the risks.[105] At first glance, the gains of unrestricted and quick implementation of AI can seem beneficial for the military. Still, since AI and the sub-field of ML to a high degree depend on learning based on large quantities of data, this creates novel weaknesses in a nation's cyber defense.[106] Even though most states have cyber strategies in place, these seldom address the specific threats related to AI. Therefore, the risks associated with learning from data provides the foundation for the following AI specific security assessment.

Focusing on the potential dangers of implementing AI into existing structures, this analysis suggests that secrecy through regulations and quality assurance through institutions can add a much-needed layer of protection. To achieve this, the military needs to collaborate with trusted partners in the civilian sector, implement institutions for quality assurance, and set realistic goals.

1. Risks Related to ML

The use of AI-augmented hacking and counter-hacking has been going on for some time.[107] As the use of ML-based systems increases in this domain, the need for subsequent ML implementations to counter these incursions increases. In the foreseeable future, any system not implementing some sort of ML-driven defense will inevitably succumb to

---

[104] Allen, *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*, 9, 22.

[105] Karen Hao, "AI Is Helping Triage Coronavirus Patients. The Tools May Be Here to Stay," *MIT Technology Review*, April 23, 2020, https://www.technologyreview.com/2020/04/23/1000410/ai-triage-covid-19-patients-health-care/.

[106] Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications," 1.

[107] Stephanie Donahole, "A New Tool For Hackers – AI in Cybersecurity," *GlobalSign Blog* (blog), August 28, 2019, https://www.globalsign.com/en/blog/New-Tool-For-Hackers-AI-Cybersecurity.

attacks.[108] Nevertheless, as ML is incorporated into the defense of important infrastructure, it also allows for attacks directed at the heart of ML. Here, two main efforts of attack stand out and are explored further: Poisoning attacks and evasion attacks.[109] Poisoning attacks target the ML algorithms and learning data (the ML model) by introducing new associations into the model to corrupt the model.[110] Considering the "Black box" dilemma caused by the use of deep neural networks (DNN) in some parts of ML, where humans have a hard time analyzing how the system has arrived at a certain conclusion, the risk of having faulty models operating undetected increases.[111] Evasion attacks are instead aimed at misguiding an otherwise working model into faulty classification by presenting an adversarial example.[112] In the following paragraphs, poisoning, evasion attacks, and common ways to defend against these attacks are presented.

a. Poisoning Attacks

In order to train a machine, a data set is determined by the model's intended aim, be it analysis of pictures of objects for classification, data streams to determine flows in a network, or words for a linguistic model, and used for initial learning.[113] As further training of the model is needed, this will occur either through the relearning of specific data sets or directly during the use of the model. Regardless of the method, it is the data used

---

[108] Maria Korolov, "Why It's Time to Prepare for AI Wielding Hackers," Data Center Knowledge, June 3, 2019, https://www.datacenterknowledge.com/machine-learning/why-its-time-prepare-ai-wielding-hackers.

[109] Comiter, *Attacking Artificial Intelligence: AI's Security Vulnerability and What Policymakers Can Do About It*, 10.

[110] Ion Stoica et al., *A Berkeley View of Systems Challenges for AI*, Report No. UCB/EECS-2017-159 (Berkley, CA: Electrical Engineering and Computer Sciences, University of California, 2017), 5. https://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-159.pdf.

[111] Martin Ciupa and Keith Abney, "Conceptualizing AI Risk," in *Computer Science & Information Technology* 8, no.3 (February 2018): 73, https://doi.org/10.5121/csit.2018.80307.

[112] Ilja Moisejevs, "Evasion Attacks on Machine Learning (or "Adversarial Examples")," Towards Data Science, July 2019, Evading Machine Learning, https://towardsdatascience.com/evasion-attacks-on-machine-learning-or-adversarial-examples-12f2283e06a1.

[113] Jonathan M. Spring et al., *Machine Learning in Cybersecurity: A Guide*, Report No. CMU/SEI-2019-TR-005 (Pittsburgh, PA: Carnegie Mellon University, 2019), 7. https://resources.sei.cmu.edu/asset_files/TechnicalReport/2019_005_001_633597.pdf.

for learning that provides a foundation for the model to execute its given task.[114] Poisoning works by manipulating the data set used for learning or even by changing the algorithm itself to create "back doors" into a seemingly functioning system.[115] If the algorithm can be augmented through introducing changes in a neural network, these will merge, essentially becoming a part of the model. As the learning process depends on the integrity of the algorithm, any alterations can render the model practically useless.[116] If undetected, these attacks can be used to deceive the opponent, trigger responses, or degrade trust in a system.

Manipulating the initial data training set or the algorithm in production is obviously the most effective way, since the adversary can more or less input any number of backdoors or totally alter the behavior of the model.[117] Unfortunately, direct access to the initial learning data is not always required, since some types of ML also "learn on the job." The reoccurring training of models offers possibilities for attack through novel methods by introducing data sets aimed at dislocating the model's boundaries in the ongoing learning process. If the network from which the model learns is known and data can be infused into the stored data sets used for relearning, then the boundaries of the model can be skewed to misinterpret information, or even allow for a back door.[118] For instance, if a system learns to always classify a certain string of data as benign, this string can then be attached to any malicious hardware in the future to hamper the effectiveness of the system further. Although not as effective a method as directly controlling the algorithm or data, it offers

---

[114] Stephen Marsland, *Machine Learning: An Algorithmic Perspective, Second Edition* (Boca Raton, FL: CRC Press, 2015), 4–6.

[115] Moisejevs, "Poisoning Attacks on Machine Learning," Spam or Ham?

[116] Moisejevs, Spam or Ham?

[117] Moisejevs, Attacker Capabilities and Poisoning Attacks.

[118] Alexander Polyakov, "How to Attack Machine Learning (Evasion, Poisoning, Inference, Trojans, Backdoors)," Towards Data Science, August 2019, https://towardsdatascience.com/how-to-attack-machine-learning-evasion-poisoning-inference-trojans-backdoors-a7cb5832595c.

an adversary a more obtainable option to interfere with a potential target's decision chains.[119]

Defense against poisoning has proven hard, but the most common way is through outlier identification, or adversarial-noise resilient regression.[120] The basic idea of this method is that the adversarial training example can be classified as an outlier, and therefore be disregarded in the training process.[121] Another approach that shows signs of efficiency is the use of micro-models, where small sets of the larger training data are used to trains several models, which then assess the large data together, basically sanitizing the large data set from anomalies.[122]

b. Evasion Attacks

Even though the learning data and the algorithm might be intact, and the system conducts the assigned duties flawlessly, the possibility to exploit flaws in the algorithms to evade accurate categorization exists. Through minor alterations in the input data, a model can be tricked into misclassifying data. Known as evasion attacks or adversarial examples, these can produce spectacular results.[123] One famous example of this is when an AI system was tricked into categorizing a STOP sign as a 45-miles per hour sign through the simple method of applying some tape to the road sign.[124] More typical use, though, is the inclusion

---

[119] Moisejevs, "Poisoning Attacks on Machine Learning," Attacker Capabilities and Poisoning Attacks.

[120] Moisejevs; Matthew Jagielski et al., "Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning," in *2018 IEEE Symposium on Security and Privacy (SP)* (Piscataway, NJ: IEEE, 2018), 24. https://doi.org/10.1109/SP.2018.00057.

[121] Jagielski et al., "Manipulating Machine Learning," 24.

[122] Gabriela F. Cretu et al., "Casting out Demons: Sanitizing Training Data for Anomaly Sensors," in *2008 IEEE Symposium on Security and Privacy (Sp 2008)* (Piscataway, NJ: IEEE, 2008), 94. https://doi.org/10.1109/SP.2008.11.

[123] Polyakov, "How to Attack Machine Learning," Evasion (Adversarial Examples).

[124] Kevin Eykholt et al., "Robust Physical-World Attacks on Deep Learning Models," *Computer Vision and Pattern Recognition (CVPR 2018)* (Kolkata, India: CVPR, April 2018), 2. http://arxiv.org/abs/1707.08945v5.

of certain pictures, phrases, or words in an email to trick spam filters into classifying the mail as safe.[125]

There are several categories of evasion attacks, relying on different methods: *Gradient attacks* leverage the knowledge of the algorithm and *score based attacks* utilize the predicted scores of a model, while *transfer based attacks* require knowledge of the learning data sets.[126] Although most of these attacks require some knowledge of the model attacked, recent studies suggest that systems can be attacked with limited or no knowledge of the model.[127] Furthermore, one of the difficulties related to defending against these attacks is that it is not totally clear why these attacks work.[128]

Although there are many existing methods to increase resistance against evasion attacks—such as regulating the effect an input has on the gradient of the model, or a multi-model defense where the model is supported by several models of the same family to increase success—the most common is through different methods of adversary training.[129] This method is based on presenting adversarial examples in a controlled way to teach the model to avoid misclassifications. Although the method can make the models more robust, it has also proven to make models less reliable.[130] As more effort is put into the research of defense against these attacks, they will surely evolve; however, so will the attacks.

---

[125] Yuxin Ma et al., "Explaining Vulnerabilities to Adversarial Machine Learning through Visual Analytics," *IEEE Transactions on Visualization and Computer Graphics* 26, no. 1 (Piscataway, NJ: IEE, January 2020): 1–2. https://doi.org/10.1109/TVCG.2019.2934631.

[126] Wieland Brendel, Jonas Rauber, and Matthias Bethge, "Decision-Based Adversarial Attacks: Reliable Attacks Against Black-Box Machine Learning Models," in *ICLR 2018 Conference* (Vancouver, BC: ICLR, February 2018), 2, http://arxiv.org/abs/1712.04248.

[127] Pin-Yu Chen et al., "ZOO: Zeroth Order Optimization Based Black-Box Attacks to Deep Neural Networks without Training Substitute Models," *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security - AISec '17* (Dallas, TX: ACM, November 2017) 15–26. https://doi.org/10.1145/3128572.3140448; Brendel et al., "Decision-Based Adversarial Attacks," 3.

[128] Moisejevs, "Evasion Attacks on Machine Learning," Why do Adversarial Examples Exist?

[129] Chen et al., "Adversarial Attack and Defense in Reinforcement Learning-from AI Security View," 15–18.

[130] Aditi Raghunathan et al., "Adversarial Training Can Hurt Generalization," in *ICML 2019 Workshop on Identifying and Understanding Deep Learning Phenomena* (Long Beach, CA: ICML, August 2019), 1, http://arxiv.org/abs/1906.06032.

c. Summary of Risks Related to ML

With the possibilities offered through poisoning and input attacks of tricking ML models into missing or failing to detect threats, the consequences can be catastrophic, especially in a military context. As several experiments have proven, even minute changes in the data used to train machines can render models totally unreliable.[131] The need to ensure the integrity of algorithms and training data becomes clear; if there are any faults in the training data set, the interpretation of incoming data by the machine will be affected.[132]

2. Ensuring Secure Military ML

Since producing novel results unobtainable through human brain power is one of the obvious strengths of ML, paradoxically, this also seems to be one of the major weaknesses of ML. Even though there has been much development in the field of understanding how deep learning really works, there is still much left to understand, making attacks to the systems harder to re-trace and detect the point of failure.[133] Nevertheless, as most of the attacks just mentioned rely on knowledge of either the data sets or the algorithm used in the model, and even though "black-box" attacks are becoming more practical and frequent, limiting access to and information about the model used, seems crucial to ensure the integrity of the model. Furthermore, the need for quality assurance of both the algorithms and the training data is apparent.

a. Secrecy

With both poisoning and evasion attacks, the effect and possibilities to attack are closely related to the knowledge of the models. Even though direct access to or knowledge of a model is not required for all types of attacks, knowledge facilitates tailoring the attacks

---

[131] Moisejevs, "Poisoning Attacks on Machine Learning," Attacker Capabilities and Poisoning Attacks.

[132] Spring et al., "Machine Learning in Cybersecurity: A Guide," 12.

[133] Matt Turek, "Explainable Artificial Intelligence (XAI)," Defense Advanced Research Projects Agency (Arlington, VA: DARPA, August 2020), https://www.darpa.mil/program/explainable-artificial-intelligence.

and opens the way for more direct invasions into the integrity of the models employed.[134] Therefore, following the logic of security from obscurity, maintaining strict secrecy regarding the existence and features of the model's algorithms and the data used for training should prevent some of the attacks.[135] One example of this could be the use of secure enclaves to ensure data confidentiality, where the data can be monitored and only accessed by authorized personnel, as proposed in Stoica et al.'s paper "A Berkley View of Systems Challenges for AI" from 2017.[136] On the other hand, since security through obscurity is highly contested as a method, security should not solely rely on obscurity but work as an extra layer in conjunction with other security measures.[137] Therefore, given the military's ability to classify sensitive information, and the array of possible attacks faced, limiting an adversary's knowledge of models and data used seems to be a possible method to enhance security when paired with regular cybersecurity.

b. Quality Assurance

Traditionally, the data required for the initial learning has been provided by the organization producing the product, thereby ensuring the quality of the data used to train the model. However, increased demand for big data sets has led to cloud learning and data sharing becoming more common.[138] Even though, as a thumb rule, a larger data set produces a more accurate model, this also increases the risk for manipulation.[139] Therefore, since poisoning and input attacks can occur during the later stages of learning

---

[134] Polyakov, "How to Attack Machine Learning," Knowledge restriction.

[135] Dafydd Stuttard, "Security & Obscurity," *Network Security* 2005, no. 7 (July 2005): 11–12. https://doi.org/10.1016/S1353-4858(05)70259-2.

[136] Stoica et al., *A Berkeley View of Systems Challenges for AI*, 5.

[137] Eric Diehl, *Ten Laws for Security* (Culver City, CA: Springer International Publishing, 2016), 78, https://doi.org/10.1007/978-3-319-42641-9.

[138] Stoica et al., *A Berkeley View of Systems Challenges for AI*, 6.

[139] Lydia Kostopoulos, *The Role of Data in Algorithmic Decision-Making* (Geneva, Switzerland: The United Nations Institute for Disarmament Research (UNIDIR), 2019), 9. https://www.unidir.org/files/publications/pdfs/the-role-of-data-in-algorithmic-decision-making-en-815.pdf.

as well, the need to monitor and evaluate the data used for teaching the ML-models exists over time.[140]

As the military increases its use of data in operations, it will face a greater need for quality assurance, thus increasing the need for qualified personnel.[141] Even though there are a number of private companies engaged in quality assurance, due to the nature of military use of ML in warfare, this will require a military-specific perspective to allow for accountability and ethical use.[142] Therefore, while many of these tasks can be performed by civilian contractors, there is a need to allow for adequate resources to be allocated to this perform task within the military organizations themselves.

3. Best Practices

As noted in the previous section, ML requires the need for assurance and trust in the supply chain, as well as an ability to certify and scrutinize the integrity of the products. The following paragraphs explore the ways and means to achieve the desired goals of developing and implementing robust systems, capable of adhering to the confidentiality, integrity, and accessibility demands of the military related to the posed threats.

a. Ways

Apart from the standard cyber protection put in place by nations around the world to keep the models and data used secure, secrecy seems to be the best way to further strengthen the security of military ML applications. As the Russian hack of the Ukraine Artillery application reported in 2016 blatantly showed, rushing unsecure products into use

---

[140] Benjamin Cheatham, Kia Javanmardian, and Hamid Samandari, "Confronting AI Risks," McKinsey & Company, April 26, 2019, https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence.

[141] Singer et al., *Cyberspace Solarium Commission: A Warning From Tomorrow*, 43–45.

[142] Kostopoulos, *The Role of Data in Algorithmic Decision-Making*, 13.

might cause organizations considerable harm.[143] Recognizing that cooperation with the commercial sector is required, however, three ways are suggested for maintaining secrecy.

Own production: This is decidedly the best option, as control of the product is maintained from the idea stage, all the way to actual implementation. China has already started an ambitious program to heighten the expertise within its military, aiming to leave the rest of the world in its wake.[144] Even though America is competitive now, China's ability to leverage its state control of the commercial sector and academia might see America falling behind in the future.[145]

While "in house" production might be the best way from a security standpoint, due to the inherent financial demands and know-how needed for this option, it seems unlikely that this approach is appropriate for other than the most affluent nations. In truth, even an AI superpower like the United States has problems to fill the knowledge gap within the military, as pointed out by the 2020 Solarium Commission report.[146]

Close supervised cooperation with the commercial sector: Since most nations have special rules for government and military contracts, these contracts can enforce strict regulations on companies regarding the release of information. This practice is commonplace today and will probably be the standard procedure for ML products used by the military as well.[147] Nonetheless, even if the companies provide the essential information for the ML algorithms and security for the data sets used in training, the need

---

[143] Dustin Volz, "Russian Hackers Tracked Ukrainian Artillery Units Using Android Implant: Report," Reuters, December 22, 2016, https://www.reuters.com/article/us-cyber-ukraine-idUSKBN14B0CU.

[144] Kania, "Chinese Military Innovation in the AI Revolution," 34.

[145] Allen, *Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*, 22.

[146] Singer et al., *Cyberspace Solarium Commission: A Warning From Tomorrow*, 43.

[147] Derek Braddon, "New Challenges for Defense, Security and the Public Finances: Guest Editor's Introduction to the Symposium," *Public Finance and Management* 7, no. 3 (2007): 233–35. http://search.proquest.com/docview/236991459/citation/6C7C4EFAB9E64B31PQ/1.

for scrutiny of the effectiveness of the product still resides with the military, requiring educated personnel.[148]

An example of this solution is the U.S. effort to incorporate AI full-motion video into drones, project MAVEN.[149] Apart from the security aspects, where the military can have full control over the software, hardware, and training data, this approach also allows the military personnel involved to gain insights into and understanding of the product being produced.[150] Working in close connection with trusted companies bound by secrecy agreements ensures product quality as well as limits knowledge of the products being developed. Even if there are a few projects implementing this approach, however, these tend to be both knowledge-intensive and costly.[151] Furthermore, as it turned out in the case of project MAVEN, partnering with the military can raise ethical issues within the civilian sector that might impact the security of the project.[152]

Commercial of the Shelf (COTS): Accepting that the commercial sector is at the forefront of development, one possibility for the military is to acquire already developed products from the commercial sector. Since a product has already been implemented in the civilian market, this can serve as a quality assurance factor, allowing the military to buy a mature and tested product. The benefits of buying an existing product are not only financial; it also reduces the need to incorporate an ample research and development

---

[148] Singer et al., *Cyberspace Solarium Commission: A Warning from Tomorrow*, 75.

[149] Deputy Secretary of Defense, "Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven)" (Official memorandum, Washington, DC: The Pentagon, 2017), 1. https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf.

[150] Cheryl Pellerin, "Project Maven to Deploy Computer Algorithms to War Zone by Year's End," *U.S. Department of Defense News*, July 21, 2017. https://www.defense.gov/Explore/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

[151] Zachary Fryer-Biggs, "The Pentagon Plans to Spend $2 Billion to Put More Artificial Intelligence into Its Weaponry," *The Verge*, September 8, 2018. https://www.theverge.com/2018/9/8/17833160/pentagon-darpa-artificial-intelligence-ai-investment.

[152] Scott Shane and Daisuke Wakabayashi, "'The Business of War': Google Employees Protest Work for the Pentagon," *The New York Times*, April 4, 2018, sec. Technology, https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

organization specifically aimed at ML.[153] On the other hand, the organization's insight into the model purchased and control of information might not be the same as if developed in close supervised cooperation with the company.[154]

Even though several firms are making their products available for sale to the military, the potential risks of security breaches are evident. Merely the fact that the product is available on the market implies that all potential buyers will be able to acquire some insight into the product.[155] To adopt this approach, the military must consider only trustworthy and well-documented companies. Furthermore, when employing a COTS product in a sensitive fashion, such as intelligence gathering, the military should also consider the risk of production breaches in security.[156] Although a viable option, COTS products might be best implemented in less sensitive and demanding areas, where breaches in security might have lesser consequences.

b. Means

In order to achieve robust and resilient enough structures capable of ensuring the CIA triad, there is a need to create an organization capable of continuously evaluating data and algorithms. Two ways of strengthening the military's position in ML are by creating institutions capable of conducting quality assurance and increasing the knowledge within the organization.

Regulatory institutions: As previously stated by Akerlof, one way to limit the uncertainties in products is to create institutions.[157] Considering the consequences of

---

[153] Thomas J. Kelly, "The Shift to Standards-Based Hardware for Military Communications: What Role Will COTS Systems Play?" *Military Embedded Systems*, December 9, 2014, http://mil-embedded.com/articles/the-cots-systems-play/.

[154] Alan D. Campen, "COTS Is Only as Good as the Shelf," *Signal* 55, no. 5 (January 2001): 31, http://search.proquest.com/docview/216182045/abstract/A360F9BAE0A14A6APQ/1.

[155] Don Peeples et al., "COTS versus GOTS: Risk Management Is the Key," *Journal of Electronic Defense* 22, no. 1 (January 1999): 53, https://search-proquest-com.libproxy.nps.edu/docview/216182045/fulltextPDF/854FF996729F4223PQ/1?accountid=12702.

[156] Josh Mayo, "DOD IG Flags COTS Purchases With Cyber Risks," *MeriTalk* (blog), July 30, 2019, https://www.meritalk.com/articles/dod-ig-flags-cots-purchases-with-cyber-risks/.

[157] Akerlof, "The Market for 'Lemons'," 499–500.

buying an ML "lemon," there is a vital need for regulation and scrutiny of ML models prior to the implementation into military applications. Since most countries already have organizations and regulations for cyber security, these organizations should also include branches dedicated to developing both national and international AI standards and ethical norms for the military to adhere to.[158]

Furthermore, institutions charged with scrutinizing and regulating ML must have the proper knowledge to be able to make informed decisions.[159] Therefore, it is not only imperative that the military are a part of these institutions and can offer their perspective, but also focus efforts on building their own institutions in order to provide proper guidance and expertise.

Cooperative institutions: As previously mentioned, institutions able to draw expertise from both the military and civilian sector are much needed. In the United States, the construction of the Joint Artificial Intelligence Center (JAIC) is a good example of institutionalizing collaboration between the military and the civilian sector.[160] By not leveraging the knowledge of ML in the civilian sector and academia in a proper way, military organizations and schools utilizing and teaching AI run the risk of becoming obsolete and incapable of monitoring such systems.

Creating institutions where the military and commercial sector interact not only leverages the knowledge base in a financially responsible way but also allows for the development of trust.[161] Building on these foundations, perhaps the military can gradually

---

[158] Singer et al., *Cyberspace Solarium Commission: A Warning From Tomorrow*, 47–48.

[159] Singer et al., 72.

[160] Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity* (Washington, DC: U.S. Department of Defense, 2018), 9. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

[161] Singer et al., *Cyberspace Solarium Commission: A Warning From Tomorrow*, 105.

mitigate problems such as Google's refusal to cooperate with the Department of Defense (DOD) in project MAVEN.[162]

c. Goals

Faced with the threats posed by AI-augmented attacks, the military needs to implement AI within defense. Paraphrasing John Arquilla's statement: "It takes networks to fight networks," the same seems to hold true for AI as well.[163] Even though Symantec's review of defense against cyber-attacks presents that a 100% success rate in the defense against ransomware attacks could be achieved when using a cloud-based ML platform, setting goals where all attacks can be averted could seem unrealistic.[164] As new forms of defense are invented, inevitably, so are new methods of attack. It is more likely that organizations will have to focus on creating robust and resilient structures, based on a layered defense as described in the CSC report.[165] In conjunction with a more layered defense strategy, militaries should also focus on setting achievable goals, in order to increase productivity and manage expectations.

Although research has proven the benefits in effectiveness of setting both short- and long-term goals, policy and guidelines for AI are often too ambitious.[166] As the RAND article "Pentagon's Ambitious Vision and Strategy for AI Not Yet Backed by Sufficient Visibility or Resources" points out, the JAIC would be better off focusing on building a structure able to cope with contemporary problems rather than overshooting its

---

[162] Daisuke Wakabayashi and Scott Shane, "Google Will Not Renew Pentagon Contract That Upset Employees," *The New York Times*, June 1, 2018, sec. Technology, https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.

[163] John Arquilla, "It Takes a Network," *The RAND Blog* (blog), August 25, 2002, https://www.rand.org/blog/2002/08/it-takes-a-network.html.

[164] Dick O'Brien, "Internet Security Threat Report: Ransomware 2017," Special Report (Mountain View, CA: Symantec, July 2017), 29. https://docs.broadcom.com/doc/istr-ransomware-2017-en.

[165] Singer et al., *Cyberspace Solarium Commission: A Warning from Tomorrow*, 23–30.

[166] Henry L. Tosi, Edwin A. Locke, and Gary P. Latham, "A Theory of Goal Setting and Task Performance," *The Academy of Management Review* 16, no. 2 (April 1991): 237–38, https://doi.org/10.2307/258875.

capacity.[167] By setting achievable goals, not only can effectiveness in organizations be improved but expectation management can also be achieved.[168] Therefore, in order to avoid dissonance, the available ways and means must correlate with the goals. Instead of striving for far-reaching goals, military institutions should focus on expectation management, providing the foundation for facilitating the implementation of AI. In a business where a mistake easily translates into loss of life, there is little room for wishful thinking or worrying about the consequence's later type of mindset.

4. Conclusion Analysis of Security

Given the threats posed by poisoning and evasion attacks to AI, and ML-models in particular, the need for controlled implementation of ML into the military becomes apparent. Of course, the problems do not cease to exist once implemented but require competent supervision throughout its useful lifespan. By investing in institutions capable of evaluating and monitoring AI and leveraging existing protocols when working with the commercial sector, the military can achieve several benefits. First, an increased collaboration with academia and the private sector allows for trustworthy partners to be identified and a mutual understanding regarding the needs for security to develop. Regardless of the procurement method chosen, to highlight the need for secrecy when necessary, another layer of security can be added to ensure the integrity of sensitive data and algorithms used. Second, institutions do not only provide quality assurance but also the necessary platforms to deepen and develop knowledge sharing between different agencies and organizations. Thus, institutions and cooperative institutions, such as the JAIC and the CSC, improve the military's ability to make informed decisions and decide on regulations. Lastly, by setting short term and achievable goals, it is possible to manage expectations effectively, and provide a foundation for acceptance of new technology into the military.

---

[167] RAND Corporation, "Pentagon's Ambitious Vision and Strategy for AI Not Yet Backed by Sufficient Visibility or Resources," *RAND Corporation*, December 17, 2019, https://www.rand.org/news/press/2019/12/17.html.

[168] Kizilcec, "How Much Information?," 2393.

Although AI promises changes to society as a whole, it also promises to introduce challenges. Although many institutions and regulations do already exist within the overarching cybersecurity construct, there is a need to focus on the present problems, rather than assume that AI will change the world by itself. Given the potential impact of poisoning and evasion on ML-models working relatively autonomously in life-or-death situations, building trust in the public eye becomes imperative for a nation's military. Lack of institutions assuring proper use and functionality of military AI, along with poor cooperation with the private sector and academia, might very well push a nation's military to the sidelines in the AI-arms race.

C. DATA

According to Joint Intelligence Publication 2–0, "raw data by itself has relatively limited utility. However, when data is collected from a sensor and processed into an intelligible form, it becomes information and gains greater utility."[169] From an intelligence point of view, we can tell the data itself has limited value. It needs, therefore, examination within a broader scope and in conjunction with the information. Arguably, data has essential value within the context of AI since it is the initiator for the result. The Organization for Economic Cooperation and Development (OECD) Glossary of Terms provides several academic definitions of data. Still, the generalized definition is "data – characteristics of information, usually numerical, collected through observation."[170] It loops back to our earlier statement that data is a polymorphic problem. Whether or not the data collection is accomplished via observation, it is highly dependent on the system that is collecting it.

The definition emphasizes both collection and numerical value. The latter becomes particularly relevant in AI if we examine the data from a computational core function, the binary code. Data is not a simple, standalone entity; it is a syntax of several complex or non-complex parts. Thus, a syntax explanation describes the function better than a semantic

---

[169] Joint Chiefs of Staff, *Joint Intelligence*, JP 2-0 (Washington, DC: Joint Chiefs of Staff, 2007), I–1, https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp2_0.pdf.

[170] *OECD,* s.v. "data," accessed September 8, 2021, https://stats.oecd.org/glossary/search.asp.

definition. When the researchers Agnar Aamodt and Mads Nygård refer to data and its syntax, they explain the following: "data are syntactic entities. Data are patterns with no meaning; they are input to an interpretation process, i.e., to the initial step of decision making."[171] They simplify the model by explaining an input [environment] – decision step – output [environment] model. Their foundation for this model is the use of data in a decision-making process. By explaining data from a syntactic perspective rather than a semantic view, we generate an understanding of data's place in AI. The decision step within the model can be replaced with a black box, or a CPU processor, or an AI function. Whatever the preferred option, there is still an input of data and an output of something.

Aamodt and Nygård leave one more challenge for us to explain—the environment. If one derives data from the environment, how can we explain that? The setting is all the things or possible things in one's surroundings that could produce data inputs. Thus, it is probably more valid to discuss what types of collectors of data there are, rather than explain all the things surrounding us. Data inputs can be collected from the environment by different sources, such as observations, surveys, interviews, polls, reports, optical means, and other data sources. Almost everything one can observe, hear, smell, or feel can with different methods translate into data.

Rama Vedashree, the CEO of the Data Security Council of India, explains the importance of data in the article "How Data Grids Will Power the Economy and Influence Our Future." According to Vedashree, "data is more than seemingly random collections of ones and zeros. It is information, currency, social fabric, safety, knowledge, confidence, and innovation."[172] In light of the broad explanation of data given by Rama Vedashree, the syntax of how it all connects in the digital world becomes more apparent. That also implies the necessity of categorizing data to meet its different demands better. On the other hand, it is not necessary or sufficient to divide data for AI into primary and secondary

---

[171] Agnar Aamodt and Mads Nygård, "Different Roles and Mutual Dependencies of Data, Information, and Knowledge — An AI Perspective on Their Integration," *Data & Knowledge Engineering* 16, no. 3 (September 1995): 6, https://doi.org/10.1016/0169-023X(95)00017-M.

[172] Palo Alto Networks et al., *Navigating the Digital Age: The Definitive Cybersecurity Guide for Directors and Officers* (Santa Clara, CA: Palo Alto Networks, 2018), 33, https://books.google.com/books?id=bvvZvAEACAAJ.

data.[173] Data is also, arguably by many, becoming a strategic resource or value. Arquilla and Ronfeldt expand on the subject by stating, "information is becoming a strategic resource that may prove as valuable and influential in the post-industrial era as capital and labor have been in the industrial age."[174] Once again, we note the importance of both monetary value and computing value for decision making that data and information bring.

1. Types of Data

Tom Taulli expands on the subject and clarifies where data emerges and what types of data there are, concluding that data can come from various sources such as the web/ social media, biometric data, point of sale systems, the internet of things, cloud systems, and corporate databases.[175] And, because different types of data demand different kinds of structures, Taulli suggests four ways to organize data: structured data, unstructured data, semi-structured data, and time-series data.[176]

a. Structured Data

Structured data, often formatted in Excel files and Structured Query Language (SQL) databases, includes lists of phone numbers, social security numbers, and other financial information.[177] Relational databases or spreadsheets are some further examples of structured data formats.[178] Due to its structured and organized nature, Daniel Nelson argues that structured data lends itself to training AI, referred to as supervised learning.[179]

---

[173] "Data Collection Methods: Definition, Examples and Sources," Question Pro, September 8, 2021, https://www.questionpro.com/blog/data-collection-methods/.

[174] John Arquilla and David Ronfeldt, "Cyberwar Is Coming!," *Comparative Strategy* 12, no. 2 (April 1993): 25, https://doi.org/10.1080/01495939308402915.

[175] Taulli, *Artificial Intelligence Basics*, 21.

[176] Taulli, 21–23.

[177] Daniel Nelson, "Structured vs. Unstructured Data," *Unite.AI* (blog), November 14, 2019, https://www.unite.ai/structured-vs-unstructured-data.

[178] Taulli, *Artificial Intelligence Basics*, 21.

[179] Nelson, "Structured vs. Unstructured Data."

b. Unstructured Data

Unstructured data represents about 80% of all the data on the internet, making it a source for feeding an AI system. There is no defined relationship between the data points, and they reside in a data lake or a large pool of unstructured data. Taulli mentions NoSQL databases as the next-generation tool to hold this type of data. Examples of unstructured data include images, videos, and social media information.[180]

c. Semi-structured Data

Semi-structured data is a hybrid consisting of structured and unstructured data, according to Taulli. Data of this kind includes Extensible Markup Language (XML) files and JavaScripts Object Notation (JSON), and CSV files, or Application Programming Interfaces (API).[181]

d. Time-Series Data

Time-series data, as the name indicates, is data ordered in time. It can be structured or unstructured.[182] Taulli explains it as pieces of information about interactions utilized to track a customer journey. As he stated, "When a user goes to a website, uses an app, or even walks in a store."[183]

2. Data and AI

When attempting to define data, it seems more prudent to utilize a syntax explanation rather than the semantic meaning. Thus, data represents the input of something (points of information) into a computational system (the specific function of the AI

---

[180] Taulli, *Artificial Intelligence Basics*, 22.

[181] "Structured Data vs. Semi Structured Data," Tealium Learning Center, September 23, 2016, https://community.tealiumiq.com/t5/Customer-Data-Hub/Structured-Data-vs-Semi-Structured-Data/ta-p/15617.

[182] Marco Peixeiro, "The Complete Guide to Time Series Analysis and Forecasting," Towards Data Science, May 19, 2020, https://towardsdatascience.com/the-complete-guide-to-time-series-analysis-and-forecasting-70d476bfe775.

[183] Taulli, *Artificial Intelligence Basics*, 22.

capacity), generating a predetermined output based on the data input and the computational calculations within the AI function. Data quantity, quality, and representativeness are therefore essential to a customer of an AI function.[184] There are some challenges with data that we need to address to understand the problems it could cause. As a customer, one must define the data to which an AI system should have access. One should assess whether the information is structured, unstructured, or semi-structured because that will affect the decision of a specific AI function. Furthermore, several reports and articles suggest a problem with biases, such as sexism and racism. James Zou and Londa Schiebinger write in one of their articles that "a major driver of data biases in AI is in the training data."[185] This implies that if an AI system or function trains on the wrong sets of data, there can be consequences. Lastly, as data has become a valuable commodity, the question of cost must be addressed.

a. Biases

Biases can occur both in models and in data; while this section only focuses on data, one should be aware that biases can occur on several occasions within an AI system. The challenge as well as the solution to the problem is how and where the collection of data takes place. If, as suggested by Zou and Schiebinger, training data gets scraped from the internet, including Google images and Google News, the collected data sets come from what is most popular on those websites at that particular time. Furthermore, that method, as Zou and Schiebinger highlight, "could unintentionally produce data that encode gender, ethnic and cultural biases."[186] ImageNet, for example, provides datasets for many machine learning systems to learn and train on to function. It is one of the world's largest producers of imagery for training sets.[187] The dataset connects to WordNet, a database that provides

---

[184] "DOD AI Symposium 2020," JAIC, September 9, 2020, https://jaic.6connex.com/event/DOD-AI-symposium2020/en-us#!/Auditorium/n634136.

[185] James Zou and Londa Schiebinger, "AI Can Be Sexist and Racist — It's Time to Make It Fair," *Nature* 559, no. 7714, July 2018, 324–326, https://doi.org/10.1038/d41586-018-05707-8.

[186] Zou and Schiebinger. "AI Can Be Sexist and Racist — It's Time to Make It Fair."

[187] "ImageNet," Imagenet, March 11, 2021, http://image-net.org/.

synonyms, constituted by several words or phrases. Each synonym set has thousands of pictures associated with it to describe the word. All the over 14 million images on ImageNet have been tagged and categorized by Amazon Mechanical Turk, a crowdsourcing project.[188] Here Zou and Schiebinger point to a statistical problem 45% of ImageNet's pictures stemming from the United States. However, the United States is only home to 4% of the world's population.[189] That points to a possible bias when the selection is mainly from the United States. How will that affect overseas usage? The tendencies in this particular example are known, and its potential biases are known.

With the awareness that data can be biased, one solution often offered is to keep humans in the loop as a controller. However, humans too have biases. Therefore, since both data and humans might create or imply bias in an AI algorithm, perhaps as David Danks—a renowned researcher within the field of ethics—suggests, the focus should be on biases that cause harm.[190] Accepting and understanding that biases exist within data, our first step is to identify and mitigate potential problems. The best solution to the problem, then, is to educate and build awareness within organizations utilizing AI, while continuously monitoring and re-assessing algorithms in order to identify shortfalls.

b. Access

Implementing an AI solution implies that some data will have to be assessed or examined by the function. Earlier, we mentioned that data could come in several forms, at rest, in process, or in transit: all of these demands' thoughtful consideration. The previous discussion also assumes that one has access to data. To further explore the information for use as an input to an AI function, it is essential to categorize the data. Taulli utilizes three standard bins to do that: in-house data, open-source data, and third-party data.[191] *In-house*

---

[188] "Case Study: Bias in ImageNet," *Sonder Studio* (blog), accessed September 16, 2020, https://sonderscheme.com/topic/case-study-bias-in-imagenet/.

[189] Zou and Schiebinger, "AI Can Be Sexist and Racist — It's Time to Make It Fair."

[190] David Danks, "Why Ethical A.I. Needs a Focus on the Fundamentals w/ David Danks," Catalyst-Go, last modified September 2019, https://www.catalyst-go.com/thinkingthroughautonomy/2019/9/09/prof-david-danks.

[191] Taulli, *Artificial Intelligence Basics*, 31.

data emerges from the sensors at the user's disposal. *Open-source data* is the type of information one can collect and access from other sources, open to all over the internet or different networks. *Third-party data* is points of information a user buys from a company or enterprise. An example could be to purchase data from Facebook, Google, or similar enterprises. The next step is to define how to store the data at rest.

Regardless of how access to data is acquired, all data collected still needs to be assessed, evaluated, and stored, requiring a sufficient infrastructure at the end user. Furthermore, when handling data for a military purpose, security may bring further limitations. It could mean that a COTS option or cloud storage is not viable. Hence, the need for trustworthy sources of data and in-house ability to maintain the security of and assure the integrity of the data used, collected, and disseminated is essential.

c. Cost

Our findings of calculations of costs are based on Amit Kumar Dutta and Ragib Hasan's research paper, "How Much Does Storage Really Cost? Towards a Full Cost Accounting Model for Data Storage," and another by D. Sculley et al., "Hidden Technical Debt in Machine Learning Systems." The two papers take different approaches to calculating costs; Dutta and Hasan propose a more traditional model, while Sculley et al. propose a set of questions to ask when trying to estimate the specific costs of implementing an ML functionality. Dutta and Hasan have assessed and mathematically modeled the cost in their research on the topic.[192] They suggest a full cost accounting approach to develop a realistic model of the actual price. The model that Dutta and Hasan present includes direct private costs, indirect personal costs, and social and environmental costs.[193] Some of the factors that need consideration are the following: initial value, floor rent, energy, service,

[192] Amit K. Dutta and Ragib Hasan, "How Much Does Storage Really Cost? Towards a Full Cost Accounting Model for Data Storage," *Economics of Grids, Clouds, Systems, and Services*, vol. 8193 (2013), 2, https://doi.org/10.1007/978-3-319-02414-1_3.

[193] Dutta and Hasan, 3.

disposal cost, and environmental costs, to mention a few, as per the authors' suggestion.[194] The cost model to account for the total cost would be the following:

$$Total\ Cost = Initial + Floor + Energy + Service + Disposal + Environmental$$ [195]

In addition to these costs, military units may face organizational changes to meet new demands. In the end, the ability to provide the necessary infrastructure will always determine whether AI is a plausible solution. Failure to accommodate adequate security, maintenance, and assessment exposes the system to adversarial attacks and potential degradation of efficiency. An additional point to the total costs of ML is made in the article "Hidden Technical Debt in Machine Learning Systems." In contrast, the authors argue that there are several initially hidden costs with managing machine learning systems. The authors know, since all of them are long time employees within Google Inc.[196] The authors mention an article from Ward Cunningham in 1992 that uses the term technical debt. In short, calculations of the costs when moving fast into new software engineering risks missing out on the long-term costs that a quick move can bring. An exciting discussion arises about the encapsulation of data or the system. We earlier mentioned that AI systems depend on data, internal or external, which brings us to the conclusion that it is difficult or nearly impossible to encapsulate a larger ML system if one wants it to work at its full potential. Scully et al. named this the erosion of boundaries.[197] Eroded boundaries occur because models in ML AI systems are interchangeable, and to seek a solution to one problem could introduce another problem or opportunity. As we have earlier said, one way of accessing data to fuel the function within the technological operation is by buying data.

---

[194] Dutta and Hasan, 4–5.

[195] Dutta and Hasan, 6.

[196] D. Sculley, Gary Holt, D. Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo and Dan Dennison, "Hidden Technical Debt in Machine Learning Systems," *NIPS* (2015), 1, https://www.semanticscholar.org/paper/Hidden-Technical-Debt-in-Machine-Learning-Systems-Sculley-Holt/1eb131a34fbb508a9dd8b646950c65901d6f1a5b.

[197] Sculley et al., 2.

Scully et al. call this data dependency,[198] which also creates complexity when calculating an AI functionality's costs. According to Sculley et al., the precise cost assessment is hard to estimate and demands a holistic view of the system. A human-centric approach that asks the following questions could then provide a supportive contribution to Dutta and Hasan's costs. Suggested questions, according to Scully et al., are:

- How easily can an entirely new algorithmic approach be tested at full scale?
- What is the transitive closure of all data dependencies?
- How precisely can the impact of a new change to the system be measured?
- Does improving one model or signal degrade others?
- How quickly can new members of the team be brought up to speed?[199]

When pursuing the introduction of AI functionality into the organization, it is necessary to make several calculations and estimates to assess the actual cost. And, because those calculations are challenging to carry out effectively, we recommend that individuals with specific knowledge be brought onto the acquisition team for this purpose.

3. Conclusion Analysis of Data

Before the decision is made to implement AI functions within an organization, decision makers must consider several issues, revealing one of the core challenges with data; it cannot be examined and judged as a single problem, but must be viewed as a part of the AI discussion as a whole. Thus, it has the similarities of a wicked problem, since the function or the outcome of an AI capability is dependent on the type of data used to train, feed, and mine it with. Therefore, as soon as AI presents an option to solve a problem, the complications that data mining, storage, and access associated with the solution present must be considered. Lack of sufficient infrastructure, knowledge, or possibilities to acquire and examine data might put an end to the project, as well as harm an organization's credibility.

---

[198] Sculley et al., 3.

[199] Sculley et al., 8.

Since understanding and identifying the potential pitfalls related to data is paramount, educated personnel are vital to examining and scrutinizing the process. In order to mitigate these problems, education, trustworthy sources for data, and reliable checks and balances must be used. These considerations hold true for all aspects of AI, and cannot be neglected at any time; rather, these should provide the foundation for all AI discussions within an organization and be weighed against current monetary limitations.

D. INTERVIEWS

The interviews for our research were conducted with persons within the academic, business, and military sectors. All the interviewed persons were chosen for their knowledge within the field of AI and their differing perspectives on the use of AI in military applications, as described in Appendix B. By contrasting the different drivers of development of AI through their respective positions, the intent was twofold; identify the validity of transparency, security, and data as considerations for acquiring AI, as identified in the SDRA document "Possibilities and Challenges for Artificial Intelligence in Military Applications." Furthermore, the aim was to examine whether there were other aspects that were reoccurring and might be incorporated to enhance the current model, adapting it to the specifics of SOF procurement.

1. Findings

All interviewed persons acknowledge the previous mentioned pillars of transparency, security, and data as a sound foundation for the application of AI into an existing organization. Some aspects touched upon were more specific to smaller organizations and were therefore particularly interesting for this research. The reoccurring themes were those of speed, flexibility, and front-end vs. back-end usage of AI. During the interviews, cost was never addressed as a problem and none of the interviewed individuals brought up price as a go- or no-go criterion. Perhaps this is because the interviews focused on U.S. individuals and personnel, who are accustomed to the U.S. defense budget, which in 2019 accounted for 38% of the entire world's total military spending according to the

Stockholm International Peace Research Institute (SIPRI).[200] In the case of a smaller country with limited resources, cost must be considered, regardless of whether the goal sometimes might justify any means.

a. Speed

The general opinion during the interviews was that the primary gain of introducing AI into an organization is that processes can be sped up, thereby shortening the decision cycle. At the same time, with this speed comes an inherent need to know what the AI is capable of, and what the aim of acquiring a specific system is. Furthermore, one of the interviewed subjects expressed that the speed of decision making, enhanced by AI, will eventually only be hampered by the humans.[201] Therefore, there is a risk that AI will be called upon not only to analyze but also make decisions in order to outperform an adversary. If this becomes a reality, there is an apparent risk that humans lose control over the systems, actively losing accountability for the decisions taken during a war.[202] With this in mind, several of the subjects expressed concerns that many organizations currently are buying into AI enhanced products without the proper knowledge of the limitations and capacities of the systems.[203] Hence, all AI procurements must be based on educated decisions, clearly understanding the implications of incorporating AI systems into organizations.

In order to manage the speed offered through AI enhancement, there is a need to educate users of the potential shortfalls and gains offered. Allowing AI to make decisions just to increase performance might seem like a good solution; however, there must always be a system of checks and balances in place to mitigate potentially dangerous outcomes of

---

[200] "Global Military Expenditure Sees Largest Annual Increase in a Decade—Says SIPRI—Reaching $1917 Billion in 2019," SIPRI, April 27, 2020, https://www.sipri.org/media/press-release/2020/global-military-expenditure-sees-largest-annual-increase-decade-says-sipri-reaching-1917-billion.

[201] Source #30, Capstone Interview AI Source #30, in-person meeting, September 17, 2020.

[202] Source #30.

[203] Source #10, Capstone Interview AI Source #10, Microsoft Teams, June 8, 2020; Source #20, Capstone Interview AI Source #20, Microsoft Teams, June 10, 2020; Source #30, Capstone Interview AI Source #30.

the systems.[204] As AI is a field still under development and constantly promising new implementations, it might also grow outside its originally intended boundaries as new functions are developed. Especially the dangers of bias in the training and learning of an AI system were highlighted more than once in the interviews.[205] Nevertheless, the benefit of using AI will always transcend the potential shortcomings, and using it for a specific purpose and maintaining humans in the loop can mitigate most of these problems. In the current situation on the modern battlefield, as long as AI lacks some aspects of human decision-making capability based on the information received, there will probably remain a need for a human in the loop.

b. Flexibility

One of the clearest points made during the interviews was one of flexibility; that is, the need for organizations to be able to quickly adapt to, incorporate, and understand AI in order to truly reap the benefits in the future.[206] Maintaining an edge on the enemy, however, requires more than just "getting the right stuff." It requires a chain where procurement decisions are based on knowledge about current systems, well thought through use, and capable operators of the equipment. As mentioned previously regarding the need for understanding, the whole chain within an organization must be equipped to trust, understand, and utilize the products for their intended purposes to be effective. Furthermore, organizations must take the emergence of AI seriously, developing protocols and procedures to speed up the process of certifying the use of new technology.[207]

With a constantly developing and lucid field such as AI, procurement processes must adapt to the increasing speed of development. Five-year procurement plans might work for huge systems, such as replacing an entire network; however, as recent fighting in the Nagorno-Karabakh region suggests, failing to incorporate new technology at the right

---

[204] Source #20, Capstone Interview AI Source #20.

[205] Source #20; Source #50, Capstone Interview AI Source #50, Microsoft Teams, August 11, 2020.

[206] Source #20; Source #30; Source #50.

[207] Source #50.

time might turn out to be very costly.[208] Therefore, when it comes to procuring AI, flexible procurement procedures in the SOF community lend themselves well to maintaining an edge on the enemy. By contrast, the lack of knowledge within organizations must be addressed at all levels through building networks where the full potential of the defense industry, private industry, and academia can be put to full use.[209] Given the constant developments within the field, it is hardly possible for a single organization to maintain a working knowledge regarding current developments. Expanding the network outside the defense industry, such as the efforts within NATO's invention program, will also allow for the necessary trust to be built between organizations.[210]

c. Front-end vs. Back-end AI

There are certain distinctions between the different AI products produced and used for military applications that were highlighted during the interviews. The main differences mentioned were what amount of data the model requires and how it learns. Most decision enhancing models, based on analyzing and comparing data, require huge amounts of data and interconnected networks to produce the best outcome. These systems often "learn on the job," requiring scrutiny of incoming data and recertification of the model.[211] Contrasting this, there are several systems based on a small learning database employed to perform a certain task, such as surveillance systems augmented by shape recognition AI, which uses a controlled data set to learn to recognize and categorize different shapes prior to deployment. Once in use, these models work on the already trained data and do not learn on the job.[212] A suggested way to differentiate between these different AI models might

---

[208] Shaan Shaikh and Wes Rumbaugh, "The Air and Missile War in Nagorno-Karabakh: Lessons for the Future of Strike and Defense," Center for Strategic & International Studies, last modified December 8, 2020, https://www.csis.org/analysis/air-and-missile-war-nagorno-karabakh-lessons-future-strike-and-defense.

[209] Source #30.

[210] NATO, "About Innovation Hub," NATO Innovation Hub, accessed March 13, 2021, https://www.innovationhub-act.org/about.

[211] Source #10.

[212] Source #40, Capstone Interview AI Source #40, Microsoft Teams, August 14, 2020.

be to label the data heavy and learning models as "back-end" products—since they require large infrastructure both in regard to the amounts of data as well as support and monitoring—whilst the more field adapted products that are trained once on a limited data set are labeled "front-end" products.

A division between front-end and back-end products and concepts can be found in a multitude of disciplines, such as management, innovation, and computer science. Although no previous division of military front-end and back-end AI exists to the authors' knowledge, the proposed division builds on a mixture of interpretations of management philosophy and innovation theory. In management, the division of front-end and back-end activities often focuses on efficiency gains by separating high consumer contact (front-end) activities—requiring close contact with customers—from less consumer intense (back-end) activities aimed at maximizing efficiency of the organization.[213] Another interpretation is offered by James P. Ignizio in the foreword to Adolfo Crespo Márquez's book "Dynamic Modeling for Supply Chain Management: Dealing With Front-end, Back-end, and Integration Issues":

> The front end addresses those portions of the organisation and its business processes that deal with sales and marketing, organised according to customer type. The back end portion of the model encompasses the units that deal with research, development, and the methods and processes of manufacturing.[214]

In the case of military AI, front-end solutions can then be understood as systems incorporating AI that task are oriented and can be controlled, understood, and evaluated by an end-user without a large support infrastructure. Conversely, back-end AI are products that require large support structures, research, are less visible to the end-users, and might imply potential structural changes affecting an organization.

---

[213] Jawwad Z. Raja, Mehmet Chakkol, Mark Johnson, Ahmad Beltagui, "Organizing for Servitization: Examining Front- and Back-End Design Configurations," *International Journal of Operations & Production Management* 38, no. 1 (2018): 252–53, http://dx.doi.org.libproxy.nps.edu/10.1108/IJOPM-03-2016-0139.

[214] Adolfo Crespo Márquez, *Dynamic Modelling for Supply Chain Management* (London, UK: Springer London, 2010), viii. https://doi.org/10.1007/978-1-84882-681-6.

According to this perspective, the front-end solutions are what is visible to the client, or in a military context, the user. This also allows the front-end concept to imply the interpretation regarding front-end innovations suggested in the paper "From Preliminary Ideas to Corroborated Product Definitions: Managing the Front End of New Product Development," published in the *California Management Review*: "The front end can be understood as three key activities: idea and concept development, idea and concept alignment, and idea and concept legitimization."[215] Hence, the concept of front-end military AI lends itself well to the SOF mindset of rapid testing and concept development.

Back-end products, on the other hand, might be more suitable for use within larger organizations—since operating, maintaining, and updating them usually involves a higher economic cost and need for highly skilled personnel within the organization. However, since these types of systems, once implemented, will be used for a foreseeable future, the effects of a breach in security can have more long-term strategical effects. In turn, this emphasizes the need for long-term planning, research, and careful consideration prior to procurement of back-end AI products. In contrast, the front-end products can rely on data and models that can be physically monitored by the users, lowering the requirements for a larger security infrastructure while the products are in use. Since the aim of a front-end AI product is to speed up more time critical tasks, which are constantly evolving and changing by their nature, one factor of these products should be that the risk of a breach in security would only cause tactical effects. Hence, the risk mitigation considerations should factor in whether an AI application falls under the front-end or back-end category of AI. To simplify, if the data used and gathered can be physically owned and controlled by individual operators while in use, and the loss of data or AI model would only produce effects at the tactical level, this should indicate that the product is to be considered as a front-end application. Hence, front-end products should not require an organization to maintain them, but rather be considered as more or less COTS products, where a company delivers a product that is ready to use with minimal training.

---

[215] Henrik Florén and Johan Frishammar, "From Preliminary Ideas to Corroborated Product Definitions: Managing The Front End of New Product Development," *California Management Review* 54, no. 4 (Summer 2012): 21–22, https://doi.org/10.1525/cmr.2012.54.4.20.

Given the changing nature of missions and relative lack of infrastructure of SOF, front-end AI should provide a more relevant focus for the organization seeking to maintain an edge on the battlefield. Since these systems are easier to maintain and might even be bespoke for certain missions, the need for rapid procurement seems more necessary. The rapid procurement chains of the SOF should provide the necessary flexibility to keep up with development. In contrast, back-end applications are constantly learning—relying on masses of data that needs to be monitored and evaluated by trained technicians—and the security of such products can have strategical consequences, making a larger organization more suitable to acquire and maintain them. Specializing on a certain type of system could also prove beneficial for division of labor between the regular forces and SOF, leading to synergy between the two as one can benefit of the developments of the other.

E. CONCLUSION OF ANALYSIS

Every military has its own acquisition evaluation matrices. Nonetheless, when it comes to incorporating AI, lengthy processes and lack of knowledge, might result in missed opportunities in the battlefield.[216] Although it is true that security concerns on a strategical level require all the preparation and precautions necessary to safeguard vital information and people, most militaries around the world currently lack educated personnel with the required skills to assess AI; there must be other ways to move forward.

As discussed in the previous chapters about trust and AI, it is not only transparency that can mitigate trust issues related to AI, but equally important is the human trust between persons and a product's ability to perform its given task. Hence, when procuring front-end products for applications in the field, the need for trust in the provider of the product will often be more important, especially in the cases where the organization itself will not own the actual data used to train the models. Unless an organization has the economic wherewithal to augment its resources and constantly attract the most skilled personnel in the field of AI, is the organization has only two options to move forward: use trustworthy

[216] Lindsey Sheppard, "Accelerating the Defense Department's AI Adoption," Council on Foreign Relations, last modified April 9, 2020, https://www.cfr.org/report/accelerating-defense-departments-ai-adoption.

suppliers that adhere to present military standards, regulations, and security requirements or get left behind on the AI battlefield. To maintain a forward momentum in the procurement of AI, the military cannot merely seek to reproduce the knowledge that is available in academia and private enterprises. Instead, the military must create reliable networks, such as the NATO Innovation Hub initiative, that can assist in the innovation and implementation of AI into the organization.[217] This is especially true for SOF, since available manpower and funding are often limited within such organizations.

When it comes to acquisition, AI should not be treated much differently from other hardware. First, a problem that needs to be solved should be identified, before even considering a purchase. Second, every plausible solution should be considered and evaluated before deciding that AI must be involved. Lastly, the economic consequences must always have the last say before a new system is decided upon. There is a parenthesis to the problem statement; sometimes, a community or organization needs to be presented with an opportunity to realize they have a problem. However, the common and preferred way should be to identify the problem or the concern to be solved. If the best solution involves AI, there should be a model for evaluating the potential for a successful implementation of a system within the organization. While the NATO model provides an excellent foundation for considerations regarding military applications for AI, when adapting it to SOF-specific considerations two additional pillars are suggested based on the analysis: Application and Risk.

a. Application

The previous front-end vs. back-end discussion, including the implications for *where* and *how* AI is intended to be used, can guide the considerations for the need for data, security, and transparency. Answering these questions will give an idea of whether the acquisition will entail a lengthy process of coordination with other agencies or if there is a need for further ethical discussions before implementing the product.

---

[217] NATO, "About Innovation Hub."

- *Where* the product is intended to be employed physically will have consequences for the type of security needed, both concerning the physical product as well as for the data used; how might the AI model and data be exposed to adversaries?

- *How* the model is used indicates the need for ethical considerations and training needed for the operators based on whether there will be a human in the loop or if the product is intended to operate autonomously.

If the risks associated with where the model is employed can be mitigated, and the ethical and educational requirements of the how factor are known and acceptable, then, there should be no real hindrance as to why the product should not be treated as any other COTS system. Although there is often talk about bias in algorithms, it is only an issue if the algorithm has been programmed to behave in a particular way—that is, the designer has projected his or her bias, or the data used for training is faulty or biased in its composition—rather than any inherent bias in algorithms themselves.[218] Therefore, by using trusted manufacturers, voicing these concerns during the development of products, and scrutinizing the data used for training, organizations can address much of the ethical issues before implementation.[219] Furthermore, as argued by several scholars; as long as there is a human in the loop, the system itself can hardly be blamed for any action it suggests.[220] Awareness of the potential problem with biases will also create sufficient knowledge of the problem within the community; after all, biases could potentially exist in different forms in many acquisitions and implementations.

---

[218] Manish Raghavan and Solon Barocas, "Challenges for Mitigating Bias in Algorithmic Hiring," Brookings, last modified December 6, 2019, https://www.brookings.edu/research/challenges-for-mitigating-bias-in-algorithmic-hiring/.

[219] James Manyika, Jake Silberg, and Brittany Presten, "What Do We Do About the Biases in AI?," *Harvard Business Review*, October 25, 2019, https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai.

[220] Ryan Daws, "Pentagon Is 'falling behind' in Military AI, Claims Former NSWC Chief," *AI News*, October 23, 2019, https://artificialintelligence-news.com/2019/10/23/pentagon-military-ai-former-nswc-chief/.

b. Risk

Even though security of information has already been mentioned as a part of the concept, mission security adds a new dimension to considerations regarding AI. Since some systems incorporate the AI algorithm in the physical product—as in the case of the assassination of the Iranian nuclear scientist, which sparked speculations that a facial recognition algorithm was used to identify the target—protecting the product might require certain restrictions to implementation.[221] Although a loss of an algorithm or model might not impede the success of the current mission, it might turn out to be problematic if the same solution is to be used again. As previously stated, if the algorithm is known it becomes easier to manipulate and exploit its weaknesses. Therefore, especially when considering the use of front-end AI, the organization must establish protocols and procedures to mitigate potential information loss.

Although AI can support a mission, it can also cause severe harm to the organization if not properly managed.[222] If data and algorithms are stored in a place that the enemies have access to, there is always a possibility that the tactics, technology, and knowledge of a target can be compromised. Therefore, limiting the use of sensitive systems to smaller and specialized units such as SOF, where security can be monitored and guaranteed, seems preferable. Yet, even if the SOF community would serve well as a platform to test, evaluate, and develop procedures for mitigating the risks before a large-scale implementation of front-end AI is implemented, if the risk to mission is deemed too great, other options must be considered. Hence, *Risk*, as in long-term risk to mission, should be a factor to consider in the early stages of procurement, along with other security issues connected to AI.

---

[221] Zoe Kleinman, "Mohsen Fakhrizadeh: 'Machine-Gun with AI' Used to Kill Iran Scientist," *BBC News*, December 7, 2020, sec. Middle East, https://www.bbc.com/news/world-middle-east-55214359.

[222] Sheppard, "Accelerating the Defense Department's AI Adoption."

V. RESULTS

The interviews conducted for this research not only cemented the previous factors of *Transparency*, *Security*, and *Data* as a foundation for considerations when incorporating AI into the military, but also added several insights for enhancing the model for application within the SOF community. The inherent flexibility in procurement available for many SOF provides the foundation for a divide in focus between the usually more rigid regular army and the SOF. As the author David Egerton explains in his book *The Shock of The Old: Technology and Global History Since 1900*, "although we can stop projects, it is often said that we cannot uninvent technologies, usually meaning we cannot get rid of them," implying that we need to move forward into a future that incorporates the technology whether we like it or not.[223] While AI most likely will be incorporated throughout most organizations in the foreseeable future, the flexibility needed to remain in the front-line of development must be ensured within the SOF community, leading the way in more field-oriented applications where the need for infrastructure is not as great. Connected to the need to utilize the current developments in an ethical and an orderly fashion, it is apparent that organizations must develop networks capable of providing both trustworthy products and the required knowledge. Developing networks, such as the JAIC's project *Tradewind*, which spans the military, private, and the academic sectors, is required to maintain an edge within the field, since they all interact to reshape the current battlefield.[224]

Focusing on front-end applications allows for a more rapid incorporation and use of systems and reduces the security considerations related to more data heavy applications. Furthermore, even though front-end applications mitigate many of the long-term support and maintenance issues, the mere fact that operators use and encounter AI on a regular

---

[223] David Edgerton, *The Shock of The Old: Technology and Global History since 1900* (New York, NY: Oxford University Press, 2011), 211, https://books.google.com/books?id=IdVGikvzIHoC.

[224] Jared Serbu, "DOD's JAIC Rolling out New Contracts to Speed up AI Acquisition," *Federal News Network*, February 11, 2021, https://federalnewsnetwork.com/artificial-intelligence/2021/02/dods-jaic-rolling-out-new-contracts-to-speed-up-ai-acquisition/.

basis will most likely help build trust in the systems that work, while other systems can be evaluated and sent back to the drawing board.[225] Building this trust and knowledge in products, manufacturers, and the technology should surely benefit the entire armed forces in the end. Regardless, since technology can decide the outcome on the battlefield in today's modern war, a testbed for emerging and existing technology must exist within every modern army. SOF, through its culture and small scale, has traditionally provided this testbed for other military hardware, such as weapons, communications, and tactics. Given the wide and undefined nature of AI, at the end of the day, there is nothing to prevent much of what is labeled AI to be part of this. In addition, borrowing from the "lean start up" community, we suggest that it is important to start small and be prepared to scale or fail fast and then move on.[226] Hence, it is important to start with some AI functionality now, rather than wait for the perfect solution or ending up in a large and time-consuming procurement process.

Admittedly, unproven technology might pose risk to mission, and that risk should be addressed through continued testing, evaluating, and disseminating lessons learned. However, with the current development within the field of AI, armies cannot afford the luxury of long-term evaluation processes before tapping the perceived benefits of simple products that are readily available to the public. Instead, tests, implementation, and evaluation must be considered as a continuous process. Since this has been a hallmark of SOF in the past, and considering the ambiguity of the concept of AI and the continuously evolving tasks of SOF, there seems to be little to hinder the future use of SOF as a testbed in this area as well.[227] While some systems might prove likely to endanger the integrity of

---

[225] M. L. Cummings, *Artificial Intelligence and the Future of Warfare*, International Security Department and U.S. and the Americas Programme (London, UK: Chatham House, 2017), 1, https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf.

[226] Eric Ries, *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses* (New York, NY: Crown Business, 2011). https://books.google.com/books?id=r9x-OXdzpPcC.

[227] Andrew White, "How NATO's Special Operations Can Take Advantage of the Tech Boom," *C4ISRNET*, July 16, 2019, https://www.c4isrnet.com/artificial-intelligence/2019/07/26/how-natos-special-operations-can-take-advantage-of-ai/.

an entire country, there are several instances where these risks can be considered local and outweighed by the potential benefits versus the risks of not applying AI as a solution. With this in mind, we advance a simple model to evaluate whether the proposed technology is suitable for rapid test and development, or if a more comprehensive analysis is needed.

A. PROPOSED MODEL

As a result of our research, building on the pillars of transparency, security and data, we propose a model for organizations to use when aiming to acquire AI technology. Its intention is to support decision makers in assessing common factors. In addition to the three earlier mentioned criteria, our findings suggest adding *risk* and *application* as primary considerations, along with the *front-end/back-end* technologies construct for cooperative considerations between conventional forces and SOF. Following the assessment of the parameters in the tentative model, the proposed model provides a broad foundation for considerations before acquiring and adopting AI into military organizations. Although the model is intended as a model for determining the suitability for rapid implementation and testing of systems within the SOF structure, it is generic in its nature and would serve as an AI acquisition evaluation tool for organizations of any size or military affiliation. Following are the model and suggested considerations connected to *Risk* and *Application*.

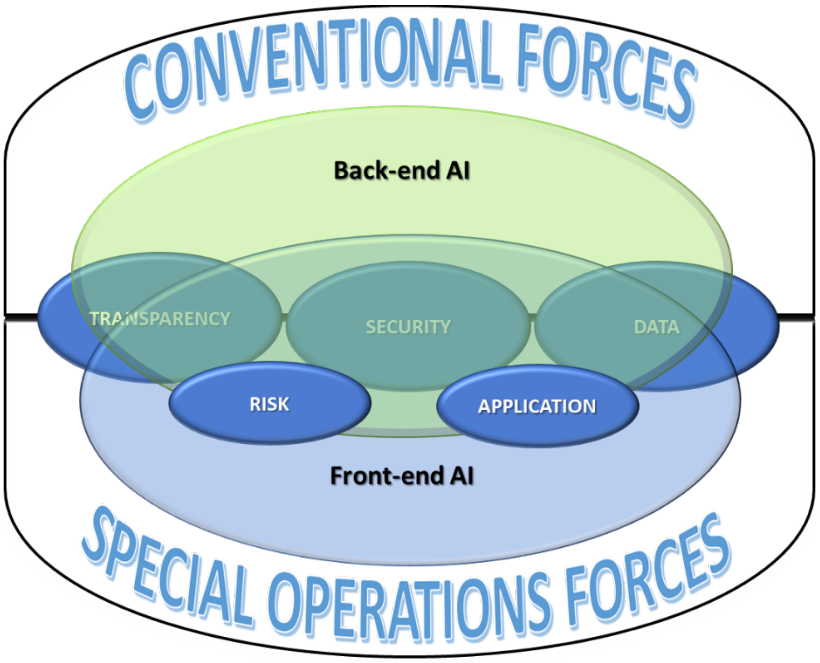

Figure 1. Proposed Model for Analysis

Risk — Since AI in its current applications seldom appears as a single entity, but rather as incorporated into existing systems and machines to enhance capabilities and speed up processes, it is more important to focus on the products' combined features than on the specifics of a particular AI-model. In many cases, AI need not be treated as something that always requires special attention, considering that facial recognition, pattern analysis, and self-driving cars exist, and the technology has existed for quite some time. When AI proposes to fundamentally change the decision process or actively make changes to the information provided for decision makers, however, a larger, more holistic approach is required. Hence, looking at the potential risks of the proposed solution in a suitable context, while acknowledging the aspect of AI as a facilitator for the solution, it is possible to avoid unnecessary considerations that might end up slowing down procurement. Suggested parameters for selection could be as simple as:

- *Mission, Tactical, or Strategic*: Will the loss of data or algorithm inflict irreparable damage, or will the damage be of a negligible character?

- *Life, System, or Product:* Will a malfunction of the AI component cause bodily harm, disruption of a larger dependent system, or simply lower trust in the product?

- *Autonomous, Controlled, or Supporting:* How will the AI component affect decision making; will action be autonomous, controlled by a human, or only used to support human decision making?

If the answers to these questions are acceptable, can be mitigated, and controlled locally, just as for any other equipment, there should be little reason as to why AI-driven solutions could not be implemented as well. Although these parameters are only suggestions, and organizations will surely adapt their own framework of risk assessment, they can provide a quick way to assess whether the product poses an acceptable risk for quick implementation, or if a more thorough assessment and procurement procedure is needed.

Application — Depending on how the product is configured and its intended use, the application of the product will indicate if it is suitable for rapid acquisition or requires a larger support structure to perform the expected tasks. While there are many factors to consider, the main issues center around the need for training, physical configuration of the product, and where it needs to be deployed. In a similar fashion as the proposed risk assessment, a few simple questions can be applied to determine whether the product can be tried, tested, and assessed within the organization. Suggested parameters could be:

- *Plug and Play, Organizational Support, or Outside Expertise*: What kind of expertise is needed to operate the product? Will a single untrained operator be able to use and maintain the product or does it require the support from existing experts within or outside the organization?

- *Standalone, Integrated, or Foundational*: What kind of effects will the product have on other parts of the organization? Is the product intended to work as a separate entity, be integrated as a part of an existing system, or even provide the foundation for a new system?

- *High Risk, Controlled, or Secure*: Where will the product be applied? Will the product be deployed unsupervised in the field, in a controlled office environment, or implemented in a setting where a high level of security is, or can be, maintained.

Considering which other systems are affected by the application, it is crucial to determine the overall effects of the implementation. Determining the second and third order effects of the system will give guidance as to whether the product is to be considered a front- or back-end AI implementation. If the solution is provided by a trustworthy source, the data used is controlled and confined to the organization, and there is little or no need for expert maintenance, it is likely that the system can be considered a front-end solution.

The combined results of the initial analyses of the risks and the proposed application, as well as the resulting implications of implementing a specific AI-system, will help decide the remaining questions related to the need for transparency, security, and data. Furthermore, the result should also give an indication in an early stage whether the product is suitable for SOF-specific acquisition or if a more formal process is needed.

B. SUMMARY OF RESULTS

As has previously been discussed, there is no real consensus on what AI entails; therefore, it is unnecessary for organizations to fear implementing a product, just because it contains an AI element. Instead, what becomes interesting is what the product proposes to solve, how it solves the problem, and what implications this has for the product's usability in a certain application. By expanding the previously proposed existing parameters for evaluation of AI in the military context—data, transparency, and security—along with risk and application considerations, the model proposes a simple tool for initial analysis for SOF AI procurement. Although there is little novel in doing a quick mission or risk analysis before implementing a product, the novelty in the proposed method consists of looking at AI as a system of systems instead of focusing on the possible dilemmas implied simply by the name AI. Paradoxically, looking beyond the current discussion of what AI is seems to be the only way to keep up in the current AI race.

VI. CONCLUSION

Support from the empirical studies and interviews used in this research suggests that it is more important for SOF to invest in available front-end AI systems today, rather than more extensive and complex back-end solutions tomorrow. Since many intricate functions and designs are available on the market today, speed in acquisitions is vital to develop an organization's understanding of how and what AI can bring to the future battlefield. Although there are certain risks and ethical considerations associated with adopting any new inventions, many potential pitfalls can be mitigated through simple measures such as maintaining a mechanism that allows for human intervention or for working with trustworthy partners. Only through constant exposure, development, adaptation, and experimentation will an organization be able to maintain the edge in AI tomorrow. Instead of focusing on the term AI itself, as with all future acquisitions, organizations need to focus on the problem or system that needs to be addressed and apply the best solution available to solve it. Realizing that the main benefits of AI functions are improved efficiencies in an organization, rather than a silver bullet to eliminate all challenges faced by the organization, is a good starting point. Based on the conducted research, the proposed model offers a suggested framework of considerations to be incorporated into the regular decision-making process and procurement process. Additionally, three guidelines related to SOF acquiring and implementing an AI functionality can be discerned; the need to *assess organizational capability, focus on front-end solutions, and manage expectations.*

First, assessing organizational capability is critical to make sure that the proposed solution will render the desired effects. When conducting the assessment, it is essential to address the question of capability; does the organization have the necessary knowledge to run and maintain an AI functionality, or what is needed to acquire these critical capabilities? By applying the proposed model, an organization should be able to rapidly determine whether outside support or new skills are needed to implement a solution. Although knowledge and skill are vital to address the issue of adopting AI, these skills take

a long time to acquire and maintain. Therefore, a smaller organization should focus on its strengths, which in the case of SOF is the faster acquisition process rather than a high degree of technical expertise. For more complicated applications, the SOF will likely need to collaborate with organizations capable of educating proficient AI professionals. By determining the problem to be solved and applying the model, if the suggested solution contains an AI element, the organization should quickly realize if the capacity exists within the organization to implement the suggested solution or if outside support is needed.

Second, a front-end focus is also essential. Starting with a front-end solution will enable a quicker acquisition process, and the investment cost will likely be lower. The risk of creating an economic black hole might be mitigated by promoting the benefit of gaining a first understanding of the technology. An early adoption within the organization might not only give an edge to the organization but also act to increase trust in future AI solutions. As recommended within the lean start-up industry, start small and scale, or in other words, fail fast to get organizational understanding. Remember, AI is software, and all software can in some way be upgraded as technology progresses and user development continues. To reduce the risk posed by the technology, it might be desirable to field test products in a less contested environment, such as counterterrorism operations, that is closely associated with SOF operations. However, once the proper understanding of possible security concerns has been established and tested, the knowledge acquired can be used to gain an edge against a more conventional adversarial competitor.

The third factor is managing expectations. AI can, and will, bring drastic change within society and the future battlefield. Nonetheless, here and now, AI is primarily focused on replacing human conversation within call centers and facilitating information exchange with Alexa and Siri while struggling to make self-driving cars acceptable on the freeway; hence, it seems unlikely that an autonomous SkyNet option á la *The Terminator* movie will gain acceptance in the near future. To be clear, that is our main point; although there is almost no limitation to what an AI-powered technology could do or perform, human thought and culture will still be in the driver's seat of societal adoption. It is critical to accept this fact and focus on the possible solutions available, instead of getting stuck investing in the next AI winter. This mindset should help the SOF community hedge

against projects that might never be, while taking advantage of opportunities to buy in at a low cost when they become available. In short, adhering to the proposed model and these recommendations will not in itself produce an AI revolution, but will keep the SOF organization agile, ready, and adaptive on the ever-changing battlefield of tomorrow.

Although there is much debate regarding the implementation of AI in military products and solutions, AI is already implemented directly or indirectly in our everyday work right now. Working with trustworthy partners to test new ideas and concepts aimed at solving present and future challenges is the only way forward for the SOF community. As an organization constantly on the edge of development of new skills and tactics, the SOF community seems to provide a perfect platform to implement new technology that promises increased efficiency and capabilities to front-end users, while pushing the need to develop back-end solutions to further the gains made. With little to suggest that society will become less reliant on AI implementations in the future, the military needs to accept and embrace AI now in order to remain relevant in the future. In the world of AI, it seems, that failing to try is trying to fail.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. HISTORY OF AI

This literature review aims to provide a historical foundation for understanding the possibilities, weaknesses, and potential of AI. The review consists of three parts. First, the focus is on the origins and emergence of AI; second, an explanation of the differences within AI, machine learning, deep learning, supervised learning, and unsupervised learning are explored. Finally, a short explanation on data is provided.

A. THE EMERGENCE OF AI

Prof. John Haugeland defines AI as "the idea that thinking and computing are radically the same,"[228] while Tom Taulli describes AI as, "where computers can learn from experience, which often involves processing data using sophisticated algorithms."[229] When examining AI, it is worth highlighting the differences between humans and machines, as this could otherwise be confusing when talking about intelligence. One overarching way to compare the differences is presented by the Human Intelligence website, shown in the Table 1.[230]

---

[228] Haugeland, *Artificial Intelligence*, 2.

[229] Taulli, *Artificial Intelligence Basics*, 179.

[230] Jean Brown, "Difference Between Artificial Intelligence and Human Intelligence" Difference Between Net, August 30, 2018, http://www.differencebetween.net/science/difference-between-artificial-intelligence-and-human-intelligence/.

Table 1. Human and Artificial Intelligence[231].

Artificial intelligence	**Human Intelligence**
Created by human intelligence	Created by inherent Intelligence
Process information faster	Process information slower
Highly objective	May be subjective
More accurate	May be less accurate
Uses 2 watts	Uses 25 watts
Cannot adapt to changes well	Can easily adapt to changes
Cannot multitask that well	Can easily multitask
Below average social skills	Excellent social skills
Still working towards self-awareness	Has self-awareness
Optimization	Innovation

Academia suggests that the first study on AI began in 1955. In the spring of 1956, there was a seminar, an early year tech convention, that first introduced AI: the McCarty's Dartmouth Summer Research Project on Artificial Intelligence. During the convention, the three fathers of AI emerged, Allen Newell, Cliff Shaw, and Herbert Simon.[232] They assumed that intelligence is the ability to solve problems, and that solving problems is the same as finding solutions.

Both computing and implementations of AI have different explanations within academia. Haugland explains in *Artificial Intelligence*: *The Very Idea*, "a computer is an interpreted automatic formal system. A formal system can be described as a game in which tokens are manipulated according to rules, to see what configurations can be obtained."[233] To further understand this, it can help to imagine tokens as characters in a game, like chess. If a player moves the "king," the player is only allowed to move him according to previously agreed on rules. As pieces are moved around on the board, the situation changes, but the rules remain constant, resulting in different configurations. This mind game will help the reader to understand how a basic automatic formal system works.

---

231 Adapted from J. Brown, "Difference Between Artificial Intelligence and Human Intelligence."

232 Haugeland, *Artificial Intelligence*, 176.

233 Haugeland, 48.

An example of a basic formal system is the calculator; we have agreed upon the rule that 1 + 1 = 2. When a calculator user presses the keys 1, + , 1, and finally the key =, the configuration of 2 is presented, according to the rules already set and the automated answer it produces. It describes a digital system following a formal set of rules, such as a computer, and the essence of artificial intelligence.

The measurement of computational intelligence, explored in one way by academia, is the Turing Test. Alan Turing suggested that a philosophical approach to intelligence would impose deep discussions on thinking and intelligence. Turing proposed that machine intelligence could be tested based on a simple game called the imitation game. The game was developed in the 1950s and is now known as the Turing Test.[234] The imitation game consists of three characters, two witnesses, and one interrogator. With the characters separated in space, the interrogator tries to guess which of the witnesses is a machine and which is the human. The interrogator only knows the witnesses by the letters X and Y. The interrogator can ask a question of the type, "Will X please tell me whether X plays chess?"[235] The answers are presented on a screen without sound to avoid the problem with voices; the machine has achieved intelligence if it can persuade the interrogator to perceive the other witness as a machine. Presently, there are other ways of testing machines, but the Turing Test is still used to test whether and to what degree a machine can think.[236] However, as Haugeland suggests, the essence is talk; does the machine talk like a human, and is the machine perceived as a human in the eyes of an observer? [237]

One could ask if something is intelligent, why not test it with an IQ test? Haugeland has a good answer to this: IQ tests test the degree of intelligence on the assumption that it has some intelligence to measure. This, Haugland argues, is the issue itself; we must know whether it makes sense to attribute intelligence to machines at all before we can ask how

---

[234] Haugeland, *Artificial Intelligence*, 6.

[235] Graham Oppy and David Dowe, "The Turing Test," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Spring 2019 (Metaphysics Research Lab, Stanford University, 2019), https://plato.stanford.edu/archives/spr2019/entriesuring-test/.

[236] Taulli, *Artificial Intelligence Basics*, 17.

[237] Haugeland, *Artificial Intelligence*, 6.

much of it they have.[238] We rather need a test that can determine whether a computer is intelligent at all, invoking the need for something like the Turing Test. One of the major differences between humans and computers is the speed of calculation. A human is capable of one calculation per second, while a computer can do one billion calculations per second, implying that we can hardly measure intelligence as the rate of calculations over time.

To summarize, something is artificial if humans make it and it does not occur as a result of natural growth. Something is intelligent in the computer realm if it can pass the Turing Test (or similar) rather than discussing it from a philosophical view. Artificial intelligence is, therefore, something created that can adopt knowledge and pass a test deceiving humans. We should, therefore, not misinterpret strict human intelligence and behaviors with machines passing the Turing Test. Just because a machine could have intelligence and deceive a human in one area does not mean that it is valid over several areas or realms.

B. ARTIFICIAL INTELLIGENCE AND ITS SUBFIELDS

Within AI several subfields exist, and these subfields are graphicly displayed in Figure 2, and further explained in the following paragraphs. Figure 2 shows the connections and a suggested hierarchy of the different subsets; this capstone focuses on the ones described in Figure 2.[239] The visualization aims to clarify how different types and subsets within AI are connected or interrelated with each other.

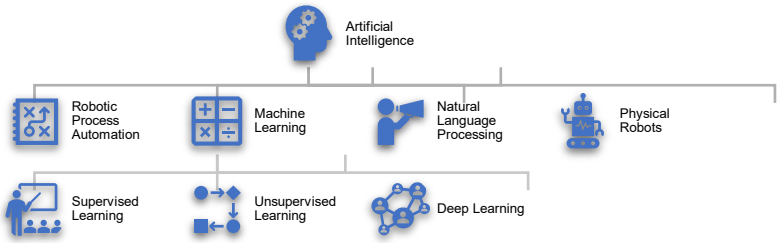

Figure 2. Graphical Overview of AI.

[238] Haugeland, 7.

[239] The figure is a visualization of the corresponding chapters in Taulli, *Artificial Intelligence Basics*.

1. Machine Learning

Machine learning (ML) is how a computer can learn without being programmed.[240] ML is an area with several subcategories: supervised learning (SL), unsupervised learning (UL), and deep learning (DL), ML is the most well-known part of AI.[241]

ML algorithms use statistics or pre-defined images to identify patterns in large amounts of data. According to Karen Hao, typically ML encompasses numbers, words, images, and other data, which is processed and enabled by many of the services that we use today, exemplified by Netflix, YouTube, and Spotify.[242] In these services, the aim is to give the user a more tailored experience. Within SL, a user will train the network to learn a certain input-output function, given a large dataset of examples. The outputs work as labels for the inputs, and the network will classify them.[243] In SL, the computer trains on labeled data. By doing so, the computer will know what pattern to look for.[244] UL is somewhat the opposite to SL, as the raw data has no markers or tags; the machine searches and develops patterns by itself. Some, like Hao, would argue that is why UL is not equally popular, and the applications are not obvious.[245] The difference between SL and UL is that SL uses labeled data, and UL does not.[246]

2. Deep Learning

Deep learning is considered by some scholars as the revolution in AI, as further explained in this paragraph. The argument maintains that even for a lot of data, DL can find a solution by the algorithms within DL. Taulli explains, "systems … can process huge

---

[240] Taulli, *Artificial Intelligence Basics*, 66.

[241] Chaudhary, "Artificial Intelligence."

[242] Karen Hao, "What Is Machine Learning?," *MIT Technology Review*, November 17, 2018, https://www.technologyreview.com/s/612437/what-is-machine-learning-we-drew-you-another-flowchart/.

[243] Peter Denning, "A Hierarchy of AI Machines" (class notes for CS4000: Harnessing AI, Naval Postgraduate School, October 2, 2019).

[244] Hao, "What Is Machine Learning?"

[245] Hao.

[246] Taulli, *Artificial Intelligence Basics*, 67.

amounts of data to find relationships and patterns that humans are often unable to detect."[247] This argument is reinforced by Hao when she states, "it [DL] uses a technique that gives machines an enhanced ability to find – and amplify – even the smallest patterns, this technique is called deep neural networks."[248] DL sometimes uses artificial neural networks; these networks are based more on a human brain and are not logic-based like other ML techniques.[249] DL is dependent on computational power, and therefore, its rise has come somewhat newly. DL sometimes is described as using ML models of multiple layers of non-linear processing units.[250]

3. Robotic Process Automation

Robotic Process Automation (RPA) is the possibility for companies to automate rule-based processes. Taulli describes RPA when he says "robots do not mean physical robots; it is about software-based robots or bots."[251] These bots can replace humans performing rule-based tasks; an example of typical tasks can be found in the Deloitte report on RPA, such as "opening email and attachments, moving files and folders, filling in forms, and reading and writing to databases."[252] RPA is the use of software to ease the burden of the human workforce. The aim is to replace "simple" tasks that could be done by machines rather than using human resources to do them. Taulli concludes that RPA can lead to major savings. RPA is therefore an attractive option for implementing basic AI within a community or company.

---

[247] Taulli, *Artificial Intelligence Basics*, 71.

[248] Hao, "What Is Machine Learning?"

[249] Selmer Bringsjord and Naveen Sundar Govindarajulu, "Artificial Intelligence," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2019 (Stanford, California: Metaphysics Research Lab, Stanford University, 2019), 62–74, https://plato.stanford.edu/archives/win2019/entries/artificial-intelligence/.

[250] Svenmarck et al., "Possibilities and Challenges for Artificial Intelligence in Military Applications."

[251] Taulli, *Artificial Intelligence Basics*, 93.

[252] Deloitte U.K., *The Robots Are Ready – Are You?*, (London, 2017), 3, https://www2.deloitte.com/cn/en/pages/strategy-operations/articles/the-robots-are-ready.html.

To summarize, RPA is a software bot (robot) tool that could enhance effectiveness within an organization by replacing human resources assigned to rule-based tasks. Achieving efficiency, however, demands proper planning and assessments of the tasks performed by humans that can be easily replaced by software bots.

4. Natural Language Processing

Natural Language Processing (NLP) is an area described by Michael J. Garbade as "a branch of artificial intelligence that deals with the interaction between computers and humans using the natural language."[253] An interaction between the computer and the human client is utilizing NLP. People often encounter NLP in daily situations without realizing it is powered by an AI functionality. Taulli explains this as "the first interaction with NLP is with virtual assistants."[254] Virtual assistants have become popular over the last couple of years. The technology can, just like RPA, provide substantial savings for a company. Insurance companies and customer service interfaces, such as Amazon's chat bot, are examples of this technology and its functionality will likely continue to flourish. Taulli continues with the argument that "NLP will become increasingly important for e-commerce and customer service."[255] Both Garbade and Taulli refer to ML when describing NLP; however, not all NLP techniques rely on ML.[256] NLP is typically used in modern services like Alexa, Cortana, and Siri.[257]

5. Physical Robots

Physical Robots (PR) are, according to Taulli, characterized by four main parts, Physical, Act, Sense, and Intelligence. Taulli suggests that the physical aspects of a robot

---

[253] Michael J. Garbade, "A Simple Introduction to Natural Language Processing," *Medium*, October 15, 2018, https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32.

[254] Taulli, *Artificial Intelligence Basics*, 123.

[255] Taulli, 123.

[256] Taulli, 48; Garbade, "A Simple Introduction to Natural Language Processing."

[257] Taulli, *Artificial Intelligence Basics*, xii.

shape range from massive industrial systems and underwater vessels to flying machines.[258] A robot does not necessarily have to look like the famous Star Wars character R2-D2, but it could. When describing the act characteristic, Taulli defines it as the actions a robot can perform, including moving items and talking.[259] For a robot to act, it must be able to sense and understand its environment; thus, it requires sensors and a feedback system. Haugeland goes even deeper to describe feelings related a robot. He lists seven kinds of feelings that might or might not affect the behavior of robots and AI.[260] Finally, a robot needs intelligence to be programmed. Robots are an emerging technology within AI and are moving more and more towards autonomy; an example of a daily use robot is the Roomba vacuum cleaner.[261]

To summarize, AI has several subsets with different possibilities and constraints. Machine learning is the foundation of modern AI and its subsidiaries; supervised learning, unsupervised learning, and deep learning enable other subsets of AI.

C. DATA EXPLAINED

Bits describes data, and is short for a binary digit, and described in binary language as a one [1] or a zero [0]. When we define amounts of data, the term "bytes" describes the size of stored data. Data is in one of three states: in transit, in process, or at rest. Data at rest is data stored in a computer's memory, a hard drive, or a cloud. The opposite is when a user works on his or her client (computer), where the data is in process. Data in transit is when data is sent between a host and a client; for example, when a user sends an email. When examining AI, it is inevitable to mention storage space and the process rate of bits and bytes to get a better understanding of the concept, and the speed at which a computer conducts computation. In Table 2, the authors have borrowed from Taulli to clarify data volumes.

---

[258] Taulli, *Artificial Intelligence Basics,* 126.

[259] Taulli, 126.

[260] Haugeland, *Artificial Intelligence*, 232.

[261] Taulli, *Artificial Intelligence Basics*, 140.

Table 2. Data Volumes Explained[262]

Unit	Value	Use case
Megabyte	1,000 kilobytes	A small book
Gigabyte	1,000 megabytes	About 230 songs
Terabyte	1,000 gigabytes	500 hours of movies
Petabyte	1,000 terabytes	Five years of earth observing systems (EOS)
Exabyte	1,000 petabytes	The entire Library of Congress 3,000 times over
Zettabyte	1,000 exabytes	36,000 years of HD-TV video
Yottabytes	1,000 zettabytes	This would require a data center the size of Delaware and Rhode Island combined

262 Adapted from Taulli, *Artificial Intelligence Basics*, 21.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. INTERVIEWS

When conducting interviews, the authors chose to use close-ended questions and open-ended questions. The function or the aim of close-ended questions is described by Lorène Fauvelle as "closed-ended questions are designed and oriented to follow a pattern and framework predefined by the interviewer."[263] A close-ended question offers, therefore, a limited number of answers, often a simple yes or no.[264] The intention is to provide the interviewer with a clear and easily measured answer. Although a comparative study could be done with the yes and no answers alone where conclusions could be drawn and shown on graphs, Fauvelle notes that "an open-ended question is a question that allows the respondent to express himself or herself freely on a given subject."[265] It allows the interviewer to ask the interviewee to elaborate or explain his or her standpoint on a specific subject. Rather than comparing the number of yes and no answers, a researcher can compare what specifics within a topic stands out or are similar between different thematic areas.

In this report, the use of open-ended questions as per the *SAGE Encyclopedia on Qualitative Research Methods* is motivated by the following: "Open-ended questions also called open, unstructured, or qualitative questions, refer to those questions for which the response patterns or answer categories are provided by the respondent, not the interviewer."[266] As the researchers' knowledge of the AI area is not deep enough to allow for a comprehensive question bank, open-ended questions allow the range of possible answers to exceed what could be provided in a questionnaire.[267]

---

[263] Lorene Fauvelle, "Qualitative Research: Open-Ended and Closed-Ended Questions," *Market Research Consulting* (blog), April 19, 2019, https://www.intotheminds.com/blog/en/qualitative-research-open-and-closed-ended-questions/.

[264] Fauvelle.

[265] Fauvelle.

[266] Kathryn J. Roulston, "Open-Ended Question," in *The SAGE Encyclopedia of Qualitative Research Methods*, ed. Lisa M. Given (Thousand Oaks: SAGE Publications, Inc., 2008), 583, https://doi.org/10.4135/9781412963909.

[267] Roulston, 583.

A. INTERVIEWEES

The authors chose to interview five individuals from private corporations, scholars within the area of AI, and finally, military personnel with user experiences.

a. Source # 10

Source #10 is an employee within the commercial sector in Silicon Valley.

b. Source # 20

Source #20 works within an organization close to the U.S. SOCOM that helps develop and leverage new technologies.

c. Source # 30

Source #30 works within a defense support organization in Silicon Valley.

d. Source # 40

Source #40 is the CEO of an A.I. company that delivers easy-to-use front end A.I. solutions.

e. Source # 50

Source #50 is a professor at the Naval Postgraduate School in Computer Science.

B. QUESTIONS FOR INTERVIEWS

Since our framework for this research is transparency, security, and data, these words frame some of the questions. The questions build upon the following main categories, justification, experience, presence, security, transparency, data, procurement, and future application. There are two sets of questions, one that focuses on the individual and the individual's opinions. The other set of questions focuses on the organizational perspective and a corporate viewpoint. The interviewee was contacted beforehand and had the option to answer from a regulatory/organizational perspective or a personal perspective.

The interview questions, along with the rationale for asking the question and its category for coding, are presented in Table 3.

Table 3. Interview Questions

#	Individual Opinion	Motive	Coding
1.1	What qualifies you to answer questions on AI?	It is an essential question because it will help us as researchers to justify the answers of the interviewee.	Background
1.2	How would you define AI?	Will help us when coding the response.	
1.3	How did you come in contact with AI?		
1.4	Do you or your organization currently use AI?		
1.5	If yes, how and in what ways, means, or types do you and your organization utilize AI?		
2.1	What do you consider to be AI?	Different individuals and organizations have diverse definitions of AI.	Framing
2.2	What is NOT AI?	It is sometimes easier to describe what	

		something is not instead of what it is.	
2.3	Within AI, which of the subsets are most common or familiar?	Help validate answers on the application and transparency of AI.	
2.4	From your point of view, what are the three most important drivers of AI?	It will give us a direction of potential biases in the later answers.	
2.5	What are the possible adverse outcomes with AI?	Could give a hint on moral and ethical perspectives.	
2.6	When acquiring AI, was it a quick fix or a long-term implementation?	It is interesting because, often SOF has a different acquisition track, which leads to faster shopping.	Timeframe
2.7	Depending on the answer, successful or not?		
2.8	After how long did your organization have an impact on the acquired AI system?	Interesting to compare acquisition times with effectiveness.	
3.1	How would you describe transparency in procurement or acquisition?	Will show if the interviewee has considered different factors on transparency and procurement.	Transparency
3.2	Did your organization acquire AI because of need or because of an introduction to a product or solution?		
3.3	When acquiring AI, what should one question?	This question will give guidance on the mindset of acquisition.	
3.4	When acquiring AI, give us your thoughts on what an enterprise should consider?	See above.	
3.5	Do you think it is vital to get access to the source code of an algorithm when acquiring AI, please develop your answer?	See above. It will also tell the willingness to question or present – Transparency.	

3.6	What infrastructural considerations have you discovered when acquiring AI?	Is it a demanding infrastructural project or not?	Infrastructure
3.7	Has infrastructure inflicted any limitations on your procurement or the acquiring process?	Plan ahead.	
3.8	Has the lack of infrastructure made you or your organization turn down an AI product?	Why didn't or did the organization buy a product.	
3.9	Have you made organizational changes?	Have the personnel been replaced by machines or have human resources been able to focus on other areas?	
4.1	What is security, and is it important?	How does the person define security? Will help us to get the context of the answers on the security questions.	Security
4.2	From your point of view, describe possible security issues when acquiring and implementing AI?	It will give examples of security flaws.	
4.3	How can these possible security issues be mitigated?	See above.	
4.4	Have you heard of, or have you experienced any security flaws directly connect to the implementation of AI?	Possible straightforward answers.	
5.1	When discussing AI, talking data is essential: what is data?	How do the interviewees define data?	Data
5.2	What does the availability of data mean to you?	Will give the perspective on data and AI.	
5.3	Is it essential to define what data one possesses before acquiring AI?	It will provide us with a hint on the focus area of AI.	
5.4	Are there specific areas within AI that address data?	Data in transit or process. Data at rest, or is data just what's being observed by a UAV that is using AI to navigate?	

5.5	If data is not essential to AI, what is?	A more philosophical angle.	
6.1	What can we expect from AI in the future?	A prediction of future application.	Future
6.2	What will we NOT see from AI in the future?	Interesting to get the person's angle.	
6.3	What ethical considerations did you or your organization have?	Ethical discussions are vast when it comes to AI. Therefore, the position on this subject is interesting.	Ethics
6.4	Do you or your organization have an ethical framework for AI and its utilization?		
6.5	Would there be ethical decisions that prevent your organization from buying an AI system?		
6.6	How did you implement the ethical discussion, the framework into your organization?		
7.1	Have you evaluated the effectiveness of your AI system?		Measure of effect
7.2	How were you able to evaluate this?		
7.3	Has AI affected your overall budget?	The answer can be diverse; yes, in the short term, it has been costly, but we think we will have a return on an investment after x years.	
8.1	What, from your perspective, are the three main vital takeaways we should consider when we continue our pursuit of solving our capstone?	A chance to stress what's essential for the individual.	Key points
8.2	Any final considerations and anything you would like to remove, change, or add?	The last chance to change or take back any answers or add information.	Final

#	Organizational Perspective	Motive	Coding

1.1a	What qualifies your position within the organization to answer questions on AI?	It is an essential question because it will help us as researchers to justify the answers of the interviewee.	Background
1.2a	How would your organization define AI?	Will help us when coding the response.	
1.3a	How did your organization come in contact with AI?		
1.4a	Does your organization currently use AI?		
1.5a	If yes, how and in what ways, means, or types does your organization utilize AI?		
2.1b	What does your organization consider to be AI?	Different organizations have diverse definitions of AI.	Framing
2.2b	What is NOT AI?	It is sometimes easier to describe what something is not instead of what it is.	
2.3b	Within AI, which of the subsets are most common or familiar, according to your organization?	Help validate answers on the application and transparency of AI.	
2.4b	From your organizational point of view, what are the three most important drivers of AI?	It will give us a direction of potential biases in the later answers.	
2.5b	What are the possible adverse outcomes with AI?	Could give a hint on moral and ethical perspectives.	
2.6b	When your organization acquired AI, was it a quick fix or a long-term implementation from an organizational point of view?	It is interesting because often SOF has a different acquisition track, which leads to faster shopping.	Timeframe
2.7b	Depending on the answer, successful or not?		

2.8b	After how long did your organization have an impact on the acquired AI system?	Interesting to compare acquisition times with effectiveness.	
3.1c	How would your organization describe transparency in procurement or acquisition?	Will show if the interviewee has considered different factors on transparency and procurement.	Transparency
3.2c	Did your organization acquire AI because of need or because of an introduction to a product or solution?		
3.3c	When purchasing AI, what should one question according to your organization?	This question will give guidance on the mindset of acquisition.	
3.4c	When acquiring AI, give us your organization's thoughts on what an enterprise should consider?	See above.	
3.5c	Does your organization think it is vital to get access to the source code of an algorithm when acquiring AI, and please develop the answer?	See above. It will also tell the willingness to question or present – Transparency.	
3.6c	What infrastructural considerations has your organization discovered when acquiring AI?	Is it a demanding infrastructural project or not?	Infrastructure
3.7c	Has infrastructure inflicted any limitations on your procurement or the acquiring process?	Plan ahead.	
3.8c	Has the lack of infrastructure made your organization turn down an AI product?	Why didn't or did the organization buy a product.	
3.9c	Has your org. made organizational changes?	Have the personnel been replaced by machines or have human resources been able to focus on other areas?	

4.1 d	What is security, and how is it essential to your organization?	How does the person define security? Will help us to get the context of the answers on the security questions.	Security
4.2d	From your organization's point of view, describe possible security issues when acquiring and implementing AI?	It will give examples of security flaws.	
4.3d	How can these possible security issues be mitigated according to your organization?	See above.	
4.4d	Has your organization heard of, or has it experienced any security flaws directly connect to the implementation of AI?	Possible straightforward answers.	
5.1e	When discussing AI, talking data is essential: what is data according to your organization?	How do the interviewees define data?	Data
5.2 e	What does the availability of data mean to your organization?	Will give the perspective on data and AI.	
5.3e	Is it essential to define what data one possesses before acquiring AI into an organization?	It will provide us with a hint on the focus area of AI.	
5.4e	Are there specific areas within AI that address data?	Data in transit or process. Data at rest, or is data just what's being observed by a UAV that is using AI to navigate?	
5.5e	If data is not essential to AI, what is?	A more philosophical angle.	
6.1f	What can we expect from AI in the future, according to your organization?	A prediction of future application.	Future
6.2f	What will we NOT see from AI in the future, according to your organization?	Interesting to get the person's angle.	
6.3f	What ethical considerations did your organization have?	Ethical discussions are vast when it comes to AI. Therefore, the	Ethics

		position on this subject is interesting.	
6.4f	Does your organization have an ethical framework for AI and its utilization?		
6.5f	Would there be ethical decisions that prevent your organization from buying an AI system?		
6.6f	How did your organization implement the ethical discussion, the framework into your organization?		
7.1g	Has your organization evaluated the effectiveness of your AI system?		Measure of effect
7.2g	How was your organization able to evaluate this?		
7.3g	Has AI affected your organization's overall budget?	The answer can be diverse; yes, in the short term, it has been costly, but we think we will have a return on an investment after x years.	
8.1h	What, from your organization's perspective, are the three main vital takeaways we should consider when we continue our pursuit of solving our capstone?	A chance to stress what's essential for the individual.	Key points
8.2h	Any final considerations and anything your organization would like to remove, change, or add?	The last chance to change or take back any answers or add information.	Final

SUPPLEMENTALS: QUICK REVIEW AND ANIMATION

A. CAPSTONE QUICK REVIEW

The supplemental briefly describes major findings of the research and potential applications for the model discussed in the main body of the capstone report and is presented in a more accessible format. The supplement uses parts of the main body text and is intended for presentational use only. Access to the report is possible either via link available within the main thesis catalog entry in the NPS Institutional Archive, Calhoun, or by contacting the NPS library.

B. SHORT ANIMATION AND INTRODUCTION

A short animation in Doodley depicts the research results as three guidelines. The guidelines aim to help an organization interested in acquiring AI and serves to nurture interest in the research report/capstone. It is intended for presentational use only via social media. The animation refers to the capstone report. Access to the animation is possible either via link available within the main thesis catalog entry in the NPS Institutional Archive, Calhoun, or by contacting the NPS library.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Aamodt, Agnar, and Mads Nygård. "Different Roles and Mutual Dependencies of Data, Information, and Knowledge—An AI Perspective on Their Integration." *Data & Knowledge Engineering* 16, no. 3 (September 1995): 191–222. https://doi.org/10.1016/0169-023X(95)00017-M.

Akerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84, no. 3 (August 1970): 488–500. https://doi.org/10.2307/1879431.

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

Arquilla, John. "It Takes a Network." *The RAND Blog* (blog), August 25, 2002. https://www.rand.org/blog/2002/08/it-takes-a-network.html.

Arquilla, John, andDavid Ronfeldt. "Cyberwar Is Coming!" *Comparative Strategy* 12, no. 2 (April 1993): 141–65. https://doi.org/10.1080/01495939308402915.

Bell, Jason. *Machine Learning: Hands-On for Developers and Technical Professionals*. Indianapolis, IN: John Wiley & Sons, 2014. http://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=1818248.

Boell, Sebastian K., and Dubravka Cecez-Kecmanovic,. "A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches." *Communications of the Association for Information Systems* 34, no. 1 (2014): 258–86 https://doi.org/10.17705/1CAIS.03412.

Braddon, Derek. "New Challenges for Defense, Security and the Public Finances: Guest Editor's Introduction to the Symposium." *Public Finance and Management* 7, no. 3 (2007): 227–41. http://search.proquest.com/docview/ 236991459/citation/6C7C4EFAB9E64B31PQ/1 .

Brendel, Wieland, Jonas Rauber, and Matthias Bethge. "Decision-Based Adversarial Attacks: Reliable Attacks Against Black-Box Machine Learning Models." In *International Conference on Learning Representations*, 1–12. Vancouver, BC: ICLR, 2018 http://arxiv.org/abs/1712.04248.

Bringsjord, Selmer, and, Naveen SundarGovindarajulu . "Artificial Intelligence." In *The Stanford Encyclopedia of Philosophy* (Winter 2019): 62–74. https://plato.stanford.edu/archives/win2019/entries/artificial-intelligence/.

Brown, Gene. “Difference Between Artificial Intelligence and Human Intelligence | Difference Between.” Difference Between Net. Last modified August 30, 2018. http://www.differencebetween.net/science/difference-between-artificial-intelligence-and-human-intelligence/.

Burt, Andrew. “The AI Transparency Paradox.” *Harvard Business Review*, December 13, 2019. https://hbr.org/2019/12/the-ai-transparency-paradox.

Byrum, Joseph. “Taking Advantage of the AI Revolution.” *ISE ; Industrial and Systems Engineering at Work* 50, no. 6 (June 2018): 28–32. http://search.proquest.com/docview/2054133919/abstract/27A128515D48480DPQ/1.

Campen, Alan D. “COTS Is Only as Good as the Shelf.” *Signal* 55, no. 5 (January 2001): 31–32. http://search.proquest.com/docview/216182045/abstract/A360F9BAE0A14A6APQ/1.

Chakravorti, Bhaskar. “The New Rules for Bringing Innovations to Market.” *Harvard Business Review*, March 2004. https://hbr.org/2004/03/the-new-rules-for-bringing-innovations-to-market.

Chaudhary, Rafay. “Artificial Intelligence: More than Machine Learning.” *Medium*, September 30, 2019. https://towardsdatascience.com/artificial-intelligence-more-than-machine-learning-fc95a1f8c2f5.

Cheatham, Benjamin,Kia Javanmardian, and Hamid Samandari. “Confronting AI Risks.” McKinsey & Company, April 26, 2019. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence.

Chen, Pin-Yu, Huan Zhang, Yash Sharma, Jinfeng Yi, and Cho-Jui Hsieh. “ZOO: Zeroth Order Optimization Based Black-Box Attacks to Deep Neural Networks without Training Substitute Models.” *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security - AISec ’17*, 2017, 15–26. https://doi.org/10.1145/3128572.3140448.

Chen, Tong, Jiqiang Liu, Yingxiao Xiang, Wenjia Niu, Endong Tong, and Zhen Han. “Adversarial Attack and Defense in Reinforcement Learning-from AI Security View.” *Cybersecurity* 2, no. 1 (December 2019): 11. https://doi.org/10.1186/s42400-019-0027-x.

Ciupa, Martin, and Keith Abney. “Conceptualizing AI Risk.” In *Computer Science & Information Technology*, 73–81. Academy & Industry Research Collaboration Center (AIRCC), 2018. https://doi.org/10.5121/csit.2018.80307.

Cohen, Elliot A. *Commandos and Politicians: Elite Military Units in Modern Democracies.* Harvard Studies of International Affairs 40. Cambridge, MA: Center for International Affairs at Harvard University, 1978.

Comiter, Marcus. *Attacking Artificial Intelligence: AI's Security Vulnerability and What Policymakers Can Do About It*. Belfer Center Paper. Cambridge, MA: Belfer Center for Science and International Affairs, 2019. https://www.belfercenter.org/sites/default/files/2019-08/AttackingAI/AttackingAI.pdf.

Cook, Karen S., Russel Hardin, and Margaret Levi. *Cooperation Without Trust*. Vol. IX. New York, NY: Russell Sage Foundation, 2005.

Crespo Márquez, Adolfo. *Dynamic Modelling for Supply Chain Management*. London: Springer London, 2010. https://doi.org/10.1007/978-1-84882-681-6.

Cretu, Gabriela F., Angelos Stavrou, Michael E. Locasto, Salvatore J. Stolfo, and Angelos D. Keromytis. "Casting out Demons: Sanitizing Training Data for Anomaly Sensors." In *2008 IEEE Symposium on Security and Privacy (Sp 2008)*, 81–95. Piscataway, NJ: IEEE, 2008. https://doi.org/10.1109/SP.2008.11.

Cummings, M L. *Artificial Intelligence and the Future of Warfare*. International Security Department and U.S. and the Americas Programme. London, UK: Chatham House, 2017. https://www.chathamhouse.org/sites/default/files/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf.

Danks, David. "Why Ethical A.I. Needs a Focus on the Fundamentals w/ David Danks." Catalyst-Go, September 6, 2019. https://www.catalyst-go.com/thinkingthroughautonomy/2019/9/09/prof-david-danks.

Daws, Ryan. "Pentagon Is 'falling behind' in Military AI, Claims Former NSWC Chief." *AI News*, October 23, 2019. https://artificialintelligence-news.com/2019/10/23/pentagon-military-ai-former-nswc-chief/.

Denning, Peter. "A Hierarchy of AI Machines." Lecture at Naval Postgraduate School, Monterey, CA. October 2, 2019. https://nps.edu/web/ai-group/harnessing-ai.

Department of Defense. Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity. Washington, DC: U.S. Department of Defense, 2018. https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF.

Deputy Secretary of Defense. "Establishment of an Algorithmic Warfare Cross-Functional Team (Project Maven)." Official memorandum, Washington, DC: The Pentagon, 2017. https://www.govexec.com/media/gbc/docs/pdfs_edit/establishment_of_the_awcft_project_maven.pdf.

Diehl, Eric. *Ten Laws for Security*. Culver City, CA: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-42641-9.

Donahole, Stephanie. "A New Tool For Hackers – AI in Cybersecurity." *GlobalSign Blog* (blog), August 28, 2019. https://www.globalsign.com/en/blog/New-Tool-For-Hackers-AI-Cybersecurity.

Dutta, Amit Kumar, and Ragib Hasan. "How Much Does Storage Really Cost? Towards a Full Cost Accounting Model for Data Storage." In *Economics of Grids, Clouds, Systems, and Services*, 8193 (September 2013):29–43. https://doi.org/10.1007/978-3-319-02414-1_3

Edgerton, David. *The Shock of The Old: Technology and Global History since 1900*. New York, NY: Oxford University Press, 2011. https://books.google.com/books?id=IdVGikvzIHoC.

Eykholt, Kevin, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. "Robust Physical-World Attacks on Deep Learning Models." *Computer Vision and Pattern Recognition (CVPR 2018)*, Kolkata, India, April 10, 2018. http://arxiv.org/abs/1707.08945v5.

Fauvelle, Lorene. "Qualitative Research: Open-Ended and Closed-Ended Questions." *Market Research Consulting* (blog), April 19, 2019. https://www.intotheminds.com/blog/en/qualitative-research-open-and-closed-ended-questions/.

Feng, Zhu, Hu Xiaofeng, Wu Lin, He Xiaoyuan, and Guo Shengming. "Inspiration for Battlefield Situation Cognition from AI Military Programs Launched by DARPA of USA and Development of AI Technology." In *Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems*. Singapore: Springer Singapore, 2016. https://doi.org/10.1007/978-981-10-2666-9_57.

Ferrario, Andrea, Michele Loi, and Eleonora Viganò. "In AI We Trust Incrementally: A Multi-Layer Model of Trust to Analyze Human-Artificial Intelligence Interactions." *Philosophy & Technology*, October 23, 2019. https://doi.org/10.1007/s13347-019-00378-3.

Florén, Henrik, and Johan Frishammar. "From Preliminary Ideas to Corroborated Product Definitions: Managing The Front End of New Product Development." *California Management Review* 54, no. 4 (Summer 2012): 20–43. https://doi.org/10.1525/cmr.2012.54.4.20.

Fryer-Biggs, Zachary. "The Pentagon Plans to Spend $2 Billion to Put More Artificial Intelligence into Its Weaponry." *The Verge*, September 8, 2018. https://www.theverge.com/2018/9/8/17833160/pentagon-darpa-artificial-intelligence-ai-investment.

Garbade, Dr Michael J. "A Simple Introduction to Natural Language Processing." *Medium*, October 15, 2018. https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32.

Geist, Edward M.. "It's Already Too Late to Stop the AI Arms Race—We Must Manage It Instead." *Bulletin of the Atomic Scientists* 72, no. 5 (September 2, 2016): 318–21. https://doi.org/10.1080/00963402.2016.1216672.

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. "Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts." *Journal of Artificial Intelligence Research* 62 (July 31, 2018): 729–54. https://doi.org/10.1613/jair.1.11222.

Guion, Lisa A., David C. Diehl, and Debra McDonald. "View of Triangulation: Establishing the Validity of Qualitative Studies." *EDIS*, no. 8 (August 2011): 3. https://doi.org/10.32473/edis-fy394-2011.

Hao, Karen. "AI Is Helping Triage Coronavirus Patients. The Tools May Be Here to Stay." *MIT Technology Review*, April 23, 2020. https://www.technologyreview.com/2020/04/23/1000410/ai-triage-covid-19-patients-health-care/.

Hardin, Russell. *Trust and Trustworthiness*. 4th ed. New York, NY: Russell Sage Foundation, 2002.

Hasselbalch, Gry. "Why Trust in AI Is Not Enough." *Dataethics* (blog), September 3, 2019. https://dataethics.eu/why-trust-in-ai-is-not-enough/.

Haugeland, John. *Artificial Intelligence: The Very Idea*. 1st paperback ed. Bradford Books. Cambridge, MA: MIT Press, 1989. https://ieeexplore-ieee-org.libproxy.nps.edu/xpl/ebooks/bookPdfWithBanner.jsp?fileName=6302870.pdf&bkn=6276821&pdfType=chapter.

Hendler, James. "Avoiding Another AI Winter." *IEEE Intelligent Systems* 23, no. 2 (March 2008): 2–4. https://doi.org/10.1109/MIS.2008.20.

Hoadley, Daniel S., and Nathan J. Lucas. "Artificial Intelligence and National Security." CRS Report No. R45178. Washington, DC: Congressional Research Service, 2018. https://digital.library.unt.edu/ark:/67531/metadc1157028/m1/1/.

Holzinger, Andreas, Peter Kieseberg, Edgar Weippl, and A. Min Tjoa. "Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction: From Machine Learning to Explainable AI." In *Machine Learning and Knowledge Extraction*, 11015:1–8. Cham, Switzerland: Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-99740-7_1.

Hon, Linda Childers, and Grunig, James E. *Guidelines Measuring Relationships.* Gainesville, FL: Institute for Public Relations, 1999. http://painepublishing.com/wp-content/uploads/2013/10/Guidelines_Measuring_Relationships.pdf.

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

Hurlburt, George. "How Much to Trust Artificial Intelligence?" *IT Professional* 19, no. 4 (2017): 7–11. https://doi.org/10.1109/MITP.2017.3051326.

Jagielski, Matthew, Alina Oprea, Battista Biggio, Chang Liu, Cristina Nita-Rotaru, and Bo Li. "Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning." In *2018 IEEE Symposium on Security and Privacy (SP)*, 19–35, Piscataway, NJ: IEEE, 2018. https://doi.org/10.1109/SP.2018.00057.

JAIC. "DoD AI Symposium 2020," September 9, 2020. https://www.ai.mil/ai2020.html

Joint Chiefs of Staff. *Doctrine for the Armed Forces of the United States, Joint Intelligence*. JP 2-0. Washington DC: Joint Chiefs of Staff, 2007. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp2_0.pdf.

———. *Doctrine for the Armed Forces of the United States, Special Operations.* JP 3-05. Washington, DC: Joint Chiefs of Staff, 2014. http://edocs.nps.edu/2014/July/jp3_05.pdf.

Kania, Elsa B. "Chinese Military Innovation in the AI Revolution." *The RUSI Journal* 164, no. 5–6 (September 19, 2019): 26–34. https://doi.org/10.1080/03071847.2019.1693803.

Kelly, Thomas J. "The Shift to Standards-Based Hardware for Military Communications: What Role Will COTS Systems Play?" *Military Embedded Systems*, December 9, 2014. http://mil-embedded.com/articles/the-cots-systems-play/.

Kemp, Richard. *Legal Aspects of Artificial Intelligence*. London, UK: Kemp It Law, 2016. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiw_sa90JrwAhUqGTQIHQcDDVoQFjABegQIChAD&url=http%3A%2F%2Fwww.kempitlaw.com%2Fwp-content%2Fuploads%2F2016%2F11%2FLegal-Aspects-of-AI-Kemp-IT-Law-v1.0-Nov-2016-2.pdf&usg=AOvVaw39ZNW7TSS26Me8bUhxsIgg.

Kizilcec, René F. "How Much Information?: Effects of Transparency on Trust in an Algorithmic Interface." In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2390–95. San Jose, CA: ACM, 2016. https://doi.org/10.1145/2858036.2858402.

Kleinman, Zoe. "Mohsen Fakhrizadeh: 'Machine-Gun with AI' Used to Kill Iran Scientist." *BBC News*, December 7, 2020, sec. Middle East. https://www.bbc.com/news/world-middle-east-55214359.

Kollock, Peter. "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust." *American Journal of Sociology* 100, no. 2 (September 1994): 313–45. https://doi.org/10.1086/230539.

Korolov, Maria. "Why It's Time to Prepare for AI Wielding Hackers." Data Center Knowledge, June 3, 2019. https://www.datacenterknowledge.com/machine-learning/why-its-time-prepare-ai-wielding-hackers.

Kostopoulos, Lydia. *The Role of Data in Algorithmic Decision-Making*. Geneva, Switzerland: The United Nations Institute for Disarmament Research (UNIDIR), 2019. https://www.unidir.org/files/publications/pdfs/the-role-of-data-in-algorithmic-decision-making-en-815.pdf.

Lake, David A. "Why 'Isms' Are Evil: Theory, Epistemology, and Academic Sects as Impediments to Understanding and Progress." *International Studies Quarterly* 55, no. 2 (2011): 465–80. http://www.jstor.org/stable/23019696.

Larsson, Stefan, and Fredrik Heintz. "Transparency in Artificial Intelligence." *Internet Policy Review* 9, no. 2 (May 5, 2020). https://doi.org/10.14763/2020.2.1469.

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

Lipton, Zachary C. "In Machine Learning, the Concept of Interpretability Is Both Important and Slippery." *Queue* 16, no. 3 (May 2018):1–28. https://dl.acm.org/doi/pdf/10.1145/3236386.3241340.

Lykke, Arthur F. "Defining Military Strategy." *Military Review* 77, no. 1 (February 1997): 183–86. http://search.proquest.com/docview/225319163/abstract/B68261794C094016PQ/1.

Ma, Yuxin, Tiankai Xie, Jundong Li, and Ross Maciejewski. "Explaining Vulnerabilities to Adversarial Machine Learning through Visual Analytics." *IEEE Transactions on Visualization and Computer Graphics* 26, no. 1 (January 2020): 1075–85. https://doi.org/10.1109/TVCG.2019.2934631.

Manyika, James, Jake Silberg, and Brittany Presten. "What Do We Do About the Biases in AI?" *Harvard Business Review*, October 25, 2019. https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai.

Marsland, Stephen. *Machine Learning: An Algorithmic Perspective, Second Edition*. Boca Raton, FL: CRC Press, 2015.

Mayo, Josh. "DoD IG Flags COTS Purchases with Cyber Risks." *MeriTalk* (blog), July 30, 2019. https://www.meritalk.com/articles/dod-ig-flags-cots-purchases-with-cyber-risks/.

McKay, Samantha. "Building a Lean Startup in 3 Steps." RIC Centre. Last modified August 9, 2016. https://riccentre.ca/2016/08/building-a-lean-startup-in-3-steps/.

Merriam, Sharan B., and Elizabeth J. Tisdell. *Qualitative Research: A Guide to Design and Implementation*. Newark, NJ: John Wiley & Sons, 2015. http://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=2089475.

Moisejevs, Ilja. "Evasion Attacks on Machine Learning (or "Adversarial Examples")." Towards Data Science, July 2019. https://towardsdatascience.com/evasion-attacks-on-machine-learning-or-adversarial-examples-12f2283e06a1.

———. "Poisoning Attacks on Machine Learning: A 15-Year-Old Security Problem That's Making a Comeback." Towards Data Science, July 2019. https://towardsdatascience.com/poisoning-attacks-on-machine-learning-1ff247c254db.

NATO. "About Innovation Hub." Innovation Hub. Accessed March 13, 2021. https://www.innovationhub-act.org/about.

Naval Postgraduate School. "Human Research Protection Program Office & Institutional Review Board (IRB)." Accessed May 4, 2021. https://nps.edu/web/research/irb-home.

Nelson, Daniel. "Structured vs. Unstructured Data." *Unite.AI* (blog), November 14, 2019. https://www.unite.ai/structured-vs-unstructured-data/.

Ramsay, Sherri, John Siân, Mischel Kwon, Thomas Ria, Tom Farley, Paul Jackson, Mark Gosling, et al. *Navigating the Digital Age: The Definitive Cybersecurity Guide for Directors and Officers*. Santa Clara, CA: Palo Alto Networks, 2018. https://books.google.com/books?id=bvvZvAEACAAJ.

O'Brien, Dick. *Internet Security Threat Report: Ransomware 2017*. Special Report. Mountain View, CA: Symantec, July 2017. https://docs.broadcom.com/doc/istr-ransomware-2017-en.

Opdenakker, Raymond. "Advantages and Disadvantages of Four Interview Techniques in Qualitative Research." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 7, no. 4 (September 2006): 1–14. https://doi.org/10.17169/fqs-7.4.175.

Oppy, Graham, and David Dowe. "The Turing Test." In *The Stanford Encyclopedia of Philosophy*. Palo Alto, CA: Stanford University Metaphysics Research Lab, Spring 2019. https://plato.stanford.edu/archives/spr2019/entriesuring-test/.

Orwell, George. "You and the Atom Bomb." The Orwell Foundation, 2021. https://www.orwellfoundation.com/the-orwell-foundation/orwell/essays-and-other-works/you-and-the-atom-bomb/.

Ostrom, Elinor. "Coping with Tragedies of The Commons." *Annual Review of Political Science* 2, no. 1 (June 1999): 493–535. https://doi.org/10.1146/annurev.polisci.2.1.493.

Owens, Mackubin T. "Strategy and the Strategic Way of Thinking." *Naval War College Review* 60, no. 4 (2007): 111–124. https://digital-commons.usnwc.edu/nwc-review/vol60/iss4/10

Peeples, Don, April Rimpo, T. Craig Peacock, and Frances E. Connors. "COTS versus GOTS: Risk Management Is the Key." *Journal of Electronic Defense* 22, no. 1 (January 1999): 51–57. http://search.proquest.com/docview/206591512/abstract/3E58ACF2973C458CPQ/1.

Peixeiro, Marco. "The Complete Guide to Time Series Analysis and Forecasting." *Towards Data Science*, May 19, 2020. https://towardsdatascience.com/the-complete-guide-to-time-series-analysis-and-forecasting-70d476bfe775.

Pellerin, Cheryl. "Project Maven to Deploy Computer Algorithms to War Zone by Year's End." *U.S. Department of Defense News*, July 21, 2017. https://www.defense.gov/Explore/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

Petrescu, Florian Ion Tiberiu, and Relly Victoria Petrescu. *The Aviation History*. Norderstedt, Germany: Books on Demand, 2012. https://doi.org/10.13140/RG.2.1.1273.9285.

Pfarrer, Michael A. "Optimizing Procurement of Special Operations Weapons and Equipment." Master's thesis, Naval Postgraduate School, 2000. https://calhoun.nps.edu/bitstream/handle/10945/7758/optimizingprocur00pfar.pdf?sequence=1&isAllowed=y.

Polyakov, Alexander. "How to Attack Machine Learning (Evasion, Poisoning, Inference, Trojans, Backdoors)." Towards Data Science, August 2019. https://towardsdatascience.com/how-to-attack-machine-learning-evasion-poisoning-inference-trojans-backdoors-a7cb5832595c.

Question Pro. "Data Collection Methods: Definition, Examples and Sources." *QuestionPro* (blog), June 27, 2019. https://www.questionpro.com/blog/data-collection-methods/.

Raghavan, Manish, and Solon Barocas. *Challenges for Mitigating Bias in Algorithmic Hiring.* Washington, DC: Brookings, December 6, 2019. https://www.brookings.edu/research/challenges-for-mitigating-bias-in-algorithmic-hiring/.

Raghunathan, Aditi, Sang Michael Xie, Fanny Yang, John C. Duchi, and Percy Liang. "Adversarial Training Can Hurt Generalization." *ArXiv:1906.06032 [Cs, Stat]*, August 26, 2019. http://arxiv.org/abs/1906.06032.

Raja, Jawwad Z., Mehmet Chakkol, Mark Johnson, and Ahmad Beltagui. "Organizing for Servitization: Examining Front- and Back-End Design Configurations." *International Journal of Operations & Production Management* 38, no. 1 (2018): 249–271. http://dx.doi.org.libproxy.nps.edu/10.1108/IJOPM-03-2016-0139.

RAND Corporation. "Pentagon's Ambitious Vision and Strategy for AI Not Yet Backed by Sufficient Visibility or Resources." RAND Corporation, December 17, 2019. https://www.rand.org/news/press/2019/12/17.html.

Rasch, R., A. Kott, and K.D. Forbus. "Incorporating AI into Military Decision Making: An Experiment." *IEEE Intelligent Systems* 18, no. 4 (July 2003): 18–26. https://doi.org/10.1109/MIS.2003.1217624.

Raval, Mehul S., Ratnik Gandhi, and Sanjay Chaudhary. "Insider Threat Detection: Machine Learning Way." *Versatile Cybersecurity, Advances in Information Security*, 72 (October 2018): 19–53. https://doi.org/10.1007/978-3-319-97643-3_2.

Rawlins, Brad R. “Measuring the Relationship between Organizational Transparency and Employee Trust.” *Public Relations Journal* 2, no. 2 (2008): 1–21. https://scholarsarchive.byu.edu/facpub/885?utm_source=scholarsarchive.byu.edu%2Ffacpub%2F885&utm_medium=PDF&utm_campaign=PDFCoverPages

Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. “‘Why Should I Trust You?’: Explaining the Predictions of Any Classifier.” *ArXiv:1602.04938 [Cs, Stat]*, August 9, 2016. http://arxiv.org/abs/1602.04938.

Ries, Eric. The Lean Startup: How Today’s Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. New York, NY: Crown Business, 2011.

Roulston, Kathryn J. “Open-Ended Question.” In *The SAGE Encyclopedia of Qualitative Research Methods*, edited by Lisa M. Given, 583–583. Thousand Oaks, CA: SAGE Publications, Inc., 2008. https://doi.org/10.4135/9781412963909.

Samonas, Spyridon, and David Coss. “The CIA Strikes Back: Redefining Confidentiality, Integrity And Availability In Security.” *Journal of Information System Security* 10, no. 3 (2014): 21–45. http://www.proso.com/dl/Samonas.pdf.

Sayler, Kelley M. *Artificial Intelligence and National Security.* CRS Report No. R45178 Washington, DC: Congressional Research Service, 2019. https://fas.org/sgp/crs/natsec/R45178.pdf.

Schramm, Harrison, and Jeff Kline. “Can Warfighters Remain the Masters of AI?” *War on the Rocks*, February 6, 2020. https://warontherocks.com/2020/02/can-warfighters-remain-the-masters-of-ai/.

Sculley, D., Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo, and Dan Dennison. "Hidden Technical Debt in Machine Learning System*s*." In *Proceedings of Advances in Neural Information Processing Systems* (*NIPS),* 2015. https://www.semanticscholar.org/paper/Hidden-Technical-Debt-in-Machine-Learning-Systems-Sculley-Holt/1eb131a34fbb508a9dd8b646950c65901d6f1a5b.

Serbu, Jared. “DoD’s JAIC Rolling out New Contracts to Speed up AI Acquisition.” Federal News Network, February 11, 2021. https://federalnewsnetwork.com/artificial-intelligence/2021/02/dods-jaic-rolling-out-new-contracts-to-speed-up-ai-acquisition/.

Shaikh, Shaan, and Wes Rumbaugh. “The Air and Missile War in Nagorno-Karabakh: Lessons for the Future of Strike and Defense.” Center for Strategic & International Studies, December 8, 2020. https://www.csis.org/analysis/air-and-missile-war-nagorno-karabakh-lessons-future-strike-and-defense.

Shane, Scott, and Daisuke Wakabayashi. "'The Business of War': Google Employees Protest Work for the Pentagon." *The New York Times*, April 4, 2018, sec. Technology. https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

Sheppard, Lindsey. "Accelerating the Defense Department's AI Adoption." Council on Foreign Relations, April 9, 2020. https://www.cfr.org/report/accelerating-defense-departments-ai-adoption.

Siau, Keng, and Weiyu Wang. "Building Trust in Artificial Intelligence, Machine Learning, and Robotics." *Cutter Business Technology Journal* 31, no. 2 (2018): 47–53. https://www.cutter.com/sites/default/files/itjournal/2018/cbtj1802.pdf.

Singer, Peter, and August Cole. *Cyberspace Solarium Commission: A Warning From Tomorrow.* Arlington, VA: Cyberspace Solarium Commision, 2020. https://drive.google.com/file/d/1ryMCIL_dZ30QyjFqFkkf10MxIXJGT4yv/view.

SIPRI. "Global Military Expenditure Sees Largest Annual Increase in a Decade—Says SIPRI—Reaching $1917 Billion in 2019." April 27, 2020. https://www.sipri.org/media/press-release/2020/global-military-expenditure-sees-largest-annual-increase-decade-says-sipri-reaching-1917-billion.

Sonder Studio. "Case Study: Bias in ImageNet." September 16, 2020. https://sonderscheme.com/topic/case-study-bias-in-imagenet/.

Spring, Jonathan M., Joshua Fallon, April Galyardt, Angela Horneman, and Leigh Metcalf. *Machine Learning in Cybersecurity: A Guide*. Report No. CMU/SEI-2019-TR005. Pittsburgh, PA: Carnegie-Mellon University Software Engineering Institute, September 2019. http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=633583.

Stanford Vision Lab. "ImageNet." 2020. http://image-net.org/.

Stephens, Neil. "Collecting Data from Elites and Ultra Elites: Telephone and Face-to-Face Interviews with Macroeconomists." *Qualitative Research* 7, no. 2 (May 2007): 203–16. https://doi.org/10.1177/1468794107076020.

Stoica, Ion, Dawn Song, Raluca Ada Popa, David Patterson, Michael W. Mahoney, Randy Katz, Anthony D. Joseph, et al. *A Berkeley View of Systems Challenges for AI*. Report No. UCB/EECS-2017-159. Berkley, CA: Electrical Engineering and Computer Sciences, University of California at Berkeley, 2017. https://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-159.pdf.

Stuttard, Dafydd. "Security & Obscurity." *Network Security* 2005, no. 7 (July 2005): 10–12. https://doi.org/10.1016/S1353-4858(05)70259-2.

Svenmarck, Peter, Linus Luotsinen, Mattias Nilsson, and Johan Schubert. "Possibilities and Challenges for Artificial Intelligence in Military Applications." *NATO Big Data and Artificial Intelligence for Military Decision Making*. Stockholm: Swedish Defence Research Agency, May 2018. https://www.researchgate.net/publication/326774966_Possibilities_and_Challenges_for_Artificial_Intelligence_in_Military_Applications.

Taulli, Tom. Artificial Intelligence Basics: A Non-Technical Introduction. Monrovia, CA: Apress, 2019.

Tealium Learning Center. "Structured Data vs. Semi Structured Data," September 23, 2016. https://community.tealiumiq.com/t5/Customer-Data-Hub/Structured-Data-vs-Semi-Structured-Data/ta-p/15617.

Deloitte. *The Robots Are Ready – Are You?*. London, UK: Deloitte, 2017. https://www2.deloitte.com/cn/en/pages/strategy-operations/articles/the-robots-are-ready.html.

Tosi, Henry L., Edwin A. Locke, and Gary P. Latham. "A Theory of Goal Setting and Task Performance." *The Academy of Management Review* 16, no. 2 (April 1991): 480–483. https://doi.org/10.2307/258875.

Trappl, Robert, ed. Impacts of Artificial Intelligence: Scientific, Technological, Military, Economic, Societal, Cultural, and Political. 1st ed. New York, NY: Elsevier Science Publishers B.V., 1986.

Turek, Matt. "Explainable Artificial Intelligence (XAI)." DARPA, August 5, 2020. https://www.darpa.mil/program/explainable-artificial-intelligence.

Volz, Dustin. "Russian Hackers Tracked Ukrainian Artillery Units Using Android Implant: Report." Reuters, December 22, 2016. https://www.reuters.com/article/us-cyber-ukraine-idUSKBN14B0CU.

Wakabayashi, Daisuke, and Scott Shane. "Google Will Not Renew Pentagon Contract That Upset Employees." *The New York Times*, June 1, 2018, sec. Technology. https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.

Walker, James, and Elinor Ostrom. "Trust and Reciprocity as Foundations for Cooperation." In *Whom Can We Trust? How Groups, Networks, and Institutions Make Trust Possible*, edited by Karen S. Cook, Margaret Levi, and Russell Hardin, 91–124. Trust. New York, NY: Russell Sage Foundation, 2009.

White, Andrew. "How NATO's Special Operations Can Take Advantage of the Tech Boom." C4ISRNET, July 16, 2019. https://www.c4isrnet.com/artificial-intelligence/2019/07/26/how-natos-special-operations-can-take-advantage-of-ai/.

Wichowski, Dawn Emsellem, and Laura E. Kohl. "Establishing Credibility in the Information Jungle: Blogs, Microblogs, and the CRAAP Test." In *Online Credibility and Digital Ethos: Evaluating Computer-Mediated Communication,* edited by Joel Gamon, 229–51. Hershey, PA: IGI Global, 2013. http://doi:10.4018/978-1-4666-2663-8.ch013.

Wildemuth, Barbara M. Applications of Social Research Methods to Questions in Information and Library Science. 2nd ed. Santa Barbara, CA: ABC-CLIO, 2017.

Xin, Yang, Lingshuang Kong, Zhi Liu, Yuling Chen, Yanmiao Li, Hongliang Zhu, Mingcheng Gao, Haixia Hou, and Chunhua Wang. "Machine Learning and Deep Learning Methods for Cybersecurity." *IEEE Access* 6 (2018): 35365–81. https://doi.org/10.1109/ACCESS.2018.2836950.

Zou, James, and Londa Schiebinger. "AI Can Be Sexist and Racist — It's Time to Make It Fair." *Nature* 559, no. 7714 (July 2018): 324–26. https://doi.org/10.1038/d41586-018-05707-8.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California