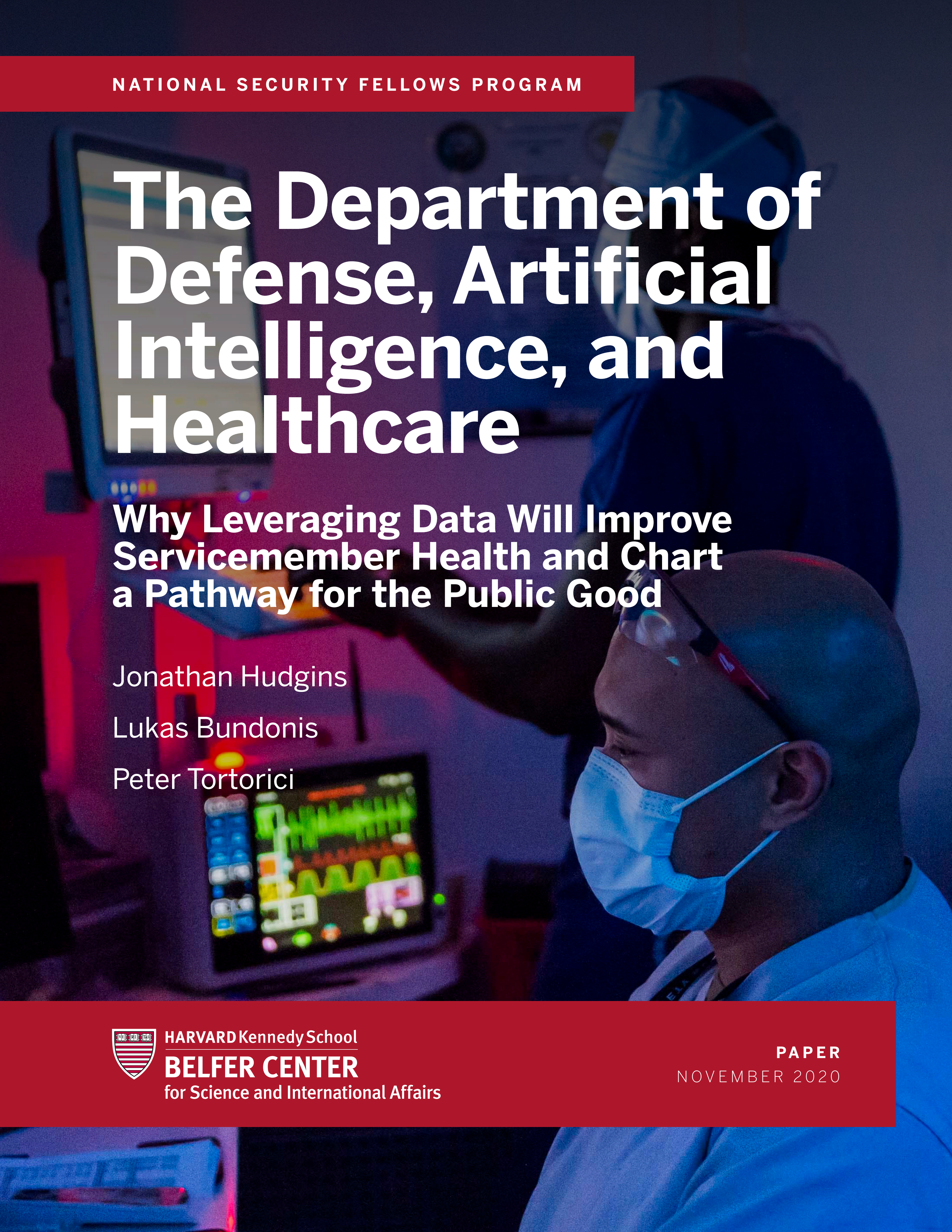


NATIONAL SECURITY FELLOWS PROGRAM

The Department of Defense, Artificial Intelligence, and Healthcare

Why Leveraging Data Will Improve Servicemember Health and Chart a Pathway for the Public Good

Jonathan Hudgins

Lukas Bundonis

Peter Tortorici

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs

PAPER
NOVEMBER 2020

National Security Fellowship Program
Belfer Center for Science and International Affairs
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138

www.belfercenter.org/NSF

Design and layout by Andrew Facini

Printed in the United States of America

NATIONAL SECURITY FELLOWS PROGRAM

The Department of Defense, Artificial Intelligence, and Healthcare

Why Leveraging Data Will Improve Servicemember Health and Chart a Pathway for the Public Good

Jonathan Hudgins

Lukas Bundonis

Peter Tortorici

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs

PAPER
NOVEMBER 2020

Authors

Jonathan Hudgins, Harvard Kennedy School, MPA

Lukas Bundonis, The Fletcher School of Law and Diplomacy, MALD

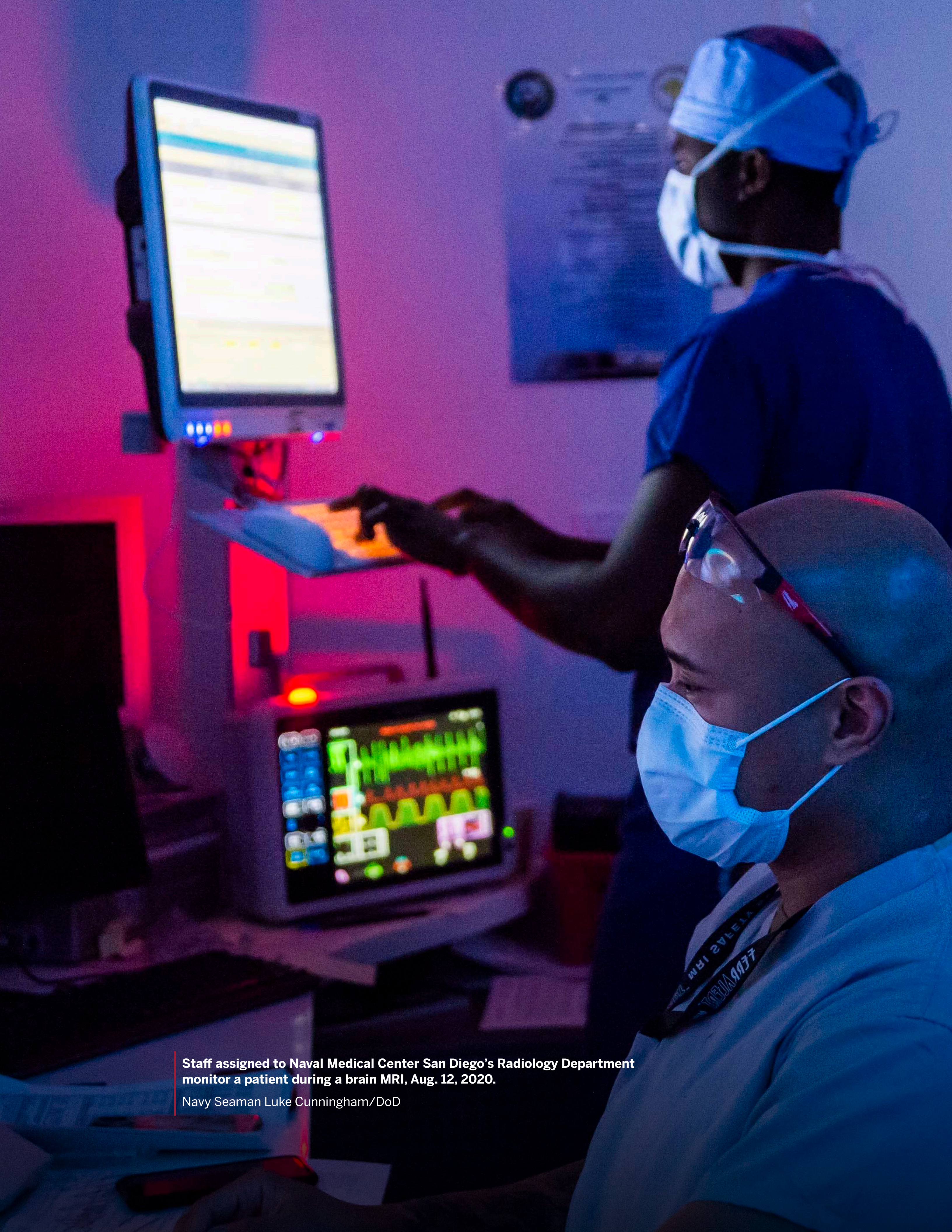
Peter Tortorici, Harvard Kennedy School, National Security Fellow

Acknowledgements

The authors would like to thank GEN (ret.) Joseph Dunford, LTG (ret.) Nadja West, MD, MG (ret.) Bill Rapp, Col Mike McGinley, CAPT Hassan Tetteh, MD, COL Sang Han, LtC. (ret.) Travis 'Wilf' Burton, LtC. (ret.) Ryan Kappedal, Capt Ryan Solis, Dr. Daniel McFarlane, Dr. Joseph Frassica, Dr. Danny Park, Nathanael 'Nat' Higgins, Allison Lazarus, Greg Woolf, and Michael Davies, and many more members of the Defense Innovation Unit's (DIU) Cambridge Project and AI/ML Portfolio for helping bring this project to life.

> "When health is absent, wisdom cannot reveal itself, art cannot manifest, strength cannot fight, wealth becomes useless, and intelligence cannot be applied."
>
> \- Herophilus

Table of Contents



Staff assigned to Naval Medical Center San Diego's Radiology Department monitor a patient during a brain MRI, Aug. 12, 2020.

Navy Seaman Luke Cunningham/DoD

Abstract

Recent growth in artificial intelligence (AI) and machine learning (ML) is poised to revolutionize industries and economies, while a supporting drive to leverage large data sets is a crucial component of that effort. The Department of Defense (DoD) currently owns one of the most extensive collections of medical data in the world. By harnessing this data, the DoD could make significant advances in diagnostic care, pathology research, and force readiness. However, questions of privacy and patient rights remain a concern when accessing and analyzing the data. This research paper examines the intersection of privacy, artificial intelligence, policy, and medical advancement. Through interviews with senior military leaders, technologists, medical professionals, and academia, the research examines the potential impacts of leveraging artificial intelligence and machine learning within DoD healthcare. The researchers also examine public policy documents and relevant technology research to discover policy solutions and highlight innovative remedies for this problem set. Specifically, could researchers harness the DoD's large medical datasets while offsetting risk to patient privacy? Significant opportunities exist to make lasting impacts using DoD medical data. Recent national technology and defense policies provide a clear mandate to explore methods to properly and safely leverage this data. Still, DoD leadership must equip its AI researchers with unambiguous command authority to action this mandate. Finding ways to employ AI ethically remains a priority for DoD leaders, and advances in de-identification technology can protect patient privacy while achieving substantial research gains for broad societal benefit.

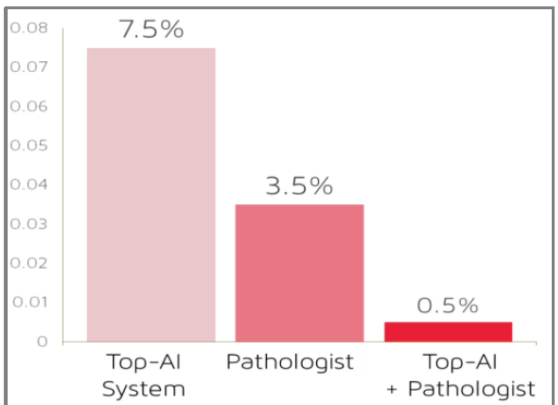
Introduction

The combination of computing power advancements, cloud storage, and sophisticated machine learning algorithms has allowed for significant opportunities to achieve breakthroughs in healthcare using Department of Defense (DoD) medical data. The Joint Pathology Center (JPC), an organization owned by the DoD, controls the world's largest tissue archive, located in physical warehouses. Recent national technology and defense policies provide a clear mandate to explore methods to safely and adequately leverage this data for medical researchers and practitioners. Large data storage and artificial intelligence have provided the tools to digitize, analyze, and use this data for servicemembers' welfare. There are clear benefits to military medicine (increase readiness, reduce costs), but the value of this opportunity goes beyond the benefits it can provide for the military. The United States has the chance to take on a moonshot project that could cure many diseases. However, the US government cannot do it alone. Though the Department of Defense owns the data, it simply does not possess the requisite resources and expertise to leverage it for medical benefit. The DoD will need to seek external assistance from private sector innovators to get the job done. DoD leadership must equip its AI researchers with unambiguous command authority to action this mandate and should immediately partner with competent commercial parties to leverage modern technology on DoD medical datasets.

Examining US and Department of Defense AI Policy

Artificial Intelligence has been a significant focal point within innovation circles for the last decade. However, substantive policy efforts to understand and harness this technology have only emerged within the last three years. The Obama administration released two documents through the National Science and Technology Council (NSTC) identifying focus areas for AI employment. The National Artificial Intelligence Research and Development Plan pushed for a government strategy regarding AI/ML research and development (R&D) and emphasized seven strategies that should be federal priorities. The plan speaks to specific opportunities for AI employment within healthcare, noting: "AI techniques can allow assessments across multi-dimensional data to study public health issues and to provide decision support systems for medical diagnoses and prescribe treatments.[1]"

Diagnostic Error Rate



AI significantly reduces pathologist error rate in the identification of metastatic breast cancer from sentinel lymph node biopsies. (Pathai, 2016)

The NSTC's early exploration of the policy mechanisms for AI recognized the importance of human-AI collaboration within medical diagnostics and pathology research. The top performer in the Camelyon16 Grand Challenge demonstrated a 0.5% error rate when an AI system and pathologists partnered during the diagnostic pathology of lymph node biopsies (Pathai, 2016) to detect metastatic breast cancer.[2] This partnership between AI technology and health professionals is only useful when datasets are available in an environment that fosters AI training, testing, and collaboration.

1 NSTC, "The National Artificial Intelligence Research and Development Strategic Plan," 10.

2 NSTC, 16.

By 2017, the Trump administration released several key national policy documents that pressed the need for greater access to data to promote AI/ML innovation across the national security enterprise. The National Security Strategy (NSS) noted the need for US decision-makers to prioritize emerging technology, specifically data science and AI, to maintain strategic advantage.[3] In order to lead in research, technology, invention, and innovation, the NSS outlines four priority actions for success:[4]

- Understand worldwide science and technology trends
- Attract and retain inventors and innovators
- Leverage private capital and expertise to build and innovate
- Rapidly field inventions and innovation

These priorities are particularly relevant and essential within AI-enabled healthcare, where the training of AI on military datasets would have impacts far beyond the Department of Defense.

The 2018 National Defense Strategy (NDS) correctly assessed that rapid technological advancement, lower barriers to entry, and acceleration of development would profoundly change society and the nature of global competition.[5] Defense priorities for sustaining military advantage, delivering performance with affordability, and supporting the National Security Innovation Base have ties to fielding a ready and capable force while leveraging data to those ends. AI-enabled healthcare will provide many opportunities for advancement in care inside and outside the Department of Defense if applied ethically and deliberately. The NDS called for greater performance and affordability within the Department by pushing these benchmarks:[6]

- Deliver performance at the speed of relevance
- Organize for innovation

3 "2017 National Security Strategy of the United States of America," 20.

4 "2017 National Security Strategy of the United States of America," 21.

5 GEN Mattis, "Summary of the 2018 National Defense Strategy," 3.

6 Mattis, 10.

- Drive budget discipline and affordability to achieve solvency
- Streamline rapid, iterative approaches from development to fielding
- Harness and protect the National Security Innovation Base

Under Secretary James Mattis, the National Defense Strategy continued to push for the DoD innovation started under Secretary Ash Carter. By 2019, several DoD policy documents began to articulate what Defense efforts would look like in the realm of AI/ML. The 2019 DoD Artificial Intelligence Strategy focuses on core principles to accelerate and integrate AI development into the fabric of a modernizing force. Human-centric AI adoption is balanced with rapid, iterative delivery of AI to allow scalable use across the defense enterprise.[7] The plan identified the need for a cutting-edge AI workforce and fostered partnerships with academia, commercial tech, and international partners.[8] The strategy highlighted how AI would change the character of conflict and the pace of threats in future crises. In turn, this meant AI would internally impact all areas of defense, from training and operations to acquisitions and healthcare.[9]

The DoD Digital Modernization Strategy (DDMS) established a roadmap for digital modernization in cloud, AI, command, control and communications, and cybersecurity.[10] The DoD Chief Information Officer described the strategy as "the cornerstone for advancing our digital environment to afford the Joint Force a competitive advantage in the modern battlespace."[11] The Digital Modernization Strategy has four pillars of modernization: cloud technology, artificial intelligence, command, control, and communications (C3), and cybersecurity. A cloud enterprise will enable widespread information sharing across the Department and enhance efficiency and scaled data sharing to build AI/ML capability.[12] Most important to this paper, this modernization strategy established an objective to treat data as a strategic asset.

7 "New Strategy Outlines Path Forward for Artificial Intelligence."

8 "DOD AI Strategy," 7.

9 "New Strategy Outlines Path Forward for Artificial Intelligence."

10 "DOD Digital Modernization Strategy 2019," 4.

11 "DOD Digital Modernization Strategy 2019," 3.

12 "Digital Modernization to Benefit Warfighters, DOD CIO Says."

"The full benefits of decision-making, information sharing, cloud migration, AI, and other DoD objectives stated in this strategy are dependent upon DoD's data being visible, accessible, understandable, trusted, and interoperable."[13]

The Digital Modernization Strategy cites the 2018 National Defense Authorization Act establishing Congressional expectations on the treatment of data:[14]

- Establish policy and governance for Common Enterprise Data related to business operations and management
- Conduct pilot programs to extract this data from relevant systems
- Analyze that data to generate insights that answer critical operational and business questions
- Evolve the pilots into a Data Management and Analytics Shared Service
- Develop an Implementation Plan

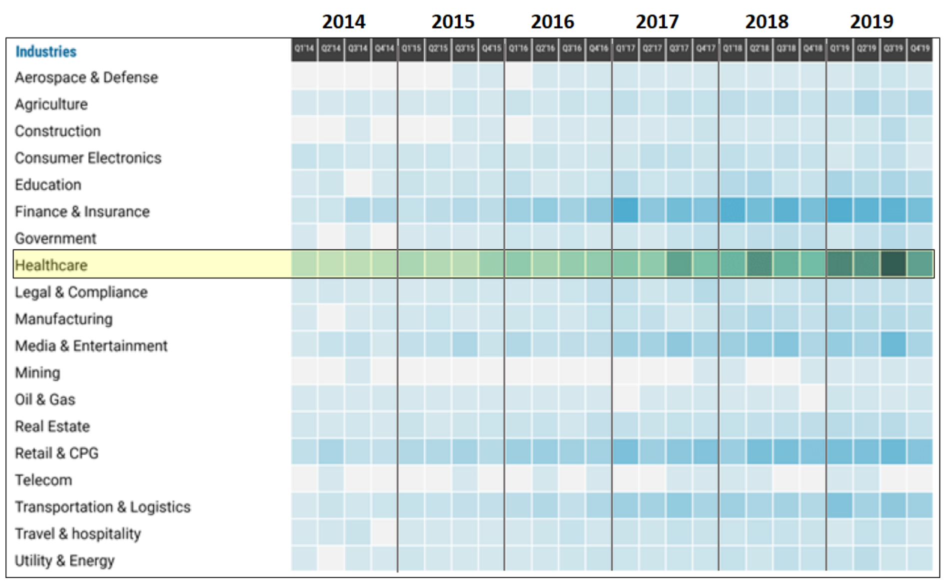
The policy documents outlined in this section provide a clear path to advancing effective data practices within DoD healthcare to prioritize AI-enabled initiatives. Working with industry on best practices for de-identifying data, utilizing available datasets that provide the highest return on investment, and prioritizing safe and ethical data use could be critical first steps towards revolutionizing diagnostic care within the DoD.

13 "DOD Digital Modernization Strategy 2019," 17.

14 "DOD Digital Modernization Strategy 2019," 17.

What Can AI/ML Bring to Military Healthcare?

AI Equity Deals from 1Q14 to 4Q19



Source: CB Insights, State of AI 2020

Healthcare led all AI equity deals from 1st quarter 2014 to 4th quarter 2019. Clearly, the private sector has recognized the game-changing potential of AI in healthcare. Section 714 of the 2020 National Defense Authorization Act instructs the Department of Defense and Veterans Administration to "establish an environment that will enable and encourage the adoption of innovative technologies for health care delivery.[15]" In addition to improving health outcomes, focusing on the military health system by utilizing modern breakthroughs in artificial intelligence can create new, productive public-private partnerships for achieving shared goals. By working together to solve tough healthcare challenges through the implementation of emerging technologies, the Department of Defense and the private sector can partner in the shared goals of saving lives, lowering costs, and strengthening national security.

Artificial intelligence needs three things to be successful: computing power, big data, and talented AI/ML engineers. Technology has reached a point

15 2020 National Defense Authorization Act for Fiscal Year 2020 (H.R.2500, 116th Congress)

where computing power is robust enough to efficiently process extremely large datasets. Additionally, through partnerships with private sector companies and academia, the DoD has access to talented AI engineers, and many useful AI algorithms already exist in open source. However, unless the DoD allows access to large datasets, the benefits of AI and machine learning will be rendered nonexistent. "Collectively, AI coupled with the unique abundance of data stored on service members may help to understand disease at its fundamental level to facilitate better treatment and therapeutics to help patients more efficiently.[16]"

The military health system currently holds one of the largest medical datasets in the world. Combining DoD's robust medical data with medical textbooks, literature, and other unique datasets, such as the more than 5,000 diagnosis questions from the American College of Physicians' Doctor's Dilemma competition, would provide unique training data from which to build new machine learning algorithms. Additionally, advances in current technology allow for de-identifying medical records, well above compliance, while applying the most robust privacy standards. The technical infrastructure and expertise exist within the DoD to apply them both and allow the DoD to solve critical, readiness-minded healthcare challenges at the bleeding edge of technological achievement. A smaller, but equally wide subset of private sector companies are currently applying machine learning methods toward operational and clinical issues in healthcare. As machine learning transitions from the age of discovery to the age of widespread implementation, nearly all of the best machine learning methods and algorithms are available in open source. Also, computing power has become commoditized to the point where most businesses have what they need to run sophisticated algorithms on large datasets. However, at present, none of these companies combine:

1. A large dataset of over 10 million medical records, 55 million pathology slides containing unique "edge cases" and lengthy time series across a patient's lifespan.
2. The command authority to implement the largest digitization of health data.

16 Interview with Dr. Hassan Tetteh, JAIC Health Mission Chief

3. The immense resource base to work across public-private partnerships in support of a national mission.

The primary purpose of the United States military health system (MHS) is to support medical readiness. The medical readiness mission involves promoting "a healthy and fit fighting force that is medically prepared to provide the Military Departments with the maximum ability to accomplish their deployment missions throughout the spectrum of military operations."[17] The MHS also serves to "create and maintain high morale in the uniformed services by providing an improved and uniform program of medical and dental care for members and certain former members of those services, and for their dependents."[18]

Our research suggests that the application of artificial intelligence and modern machine learning methods in the Department of Defense, and within the military health system specifically, will provide three key enabling functions for military commanders and Department of Defense leadership:

- Drive down costs
- Increase personnel readiness
- Optimize personnel performance

Drive Down Costs

The Fiscal Year 2020 budget for the Military Health System is $50 billion, or 6.9% of DoD's total budget.[19] This amount is $4 billion more than the entire budget of the U.S. Marine Corps. Additionally, DoD's budget request for FY2020 includes a plan to reduce its active-duty medical force by 22% or 17,991 personnel. What's more, DoD medical costs are likely to rise due to the recent decision to no longer exempt the military from medical malpractice claims. "A provision tucked into the massive defense-spending bill

17 Bryce H P Mendez, "Defense Primer: Military Health System," n.d., 1.

18 Mendez, 1.

19 Mendez, 2.

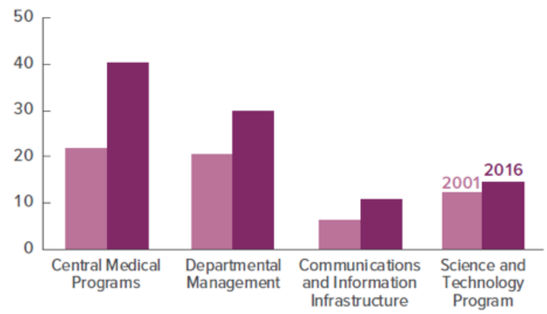
creates a medical-malpractice exception to a 1950 US Supreme Court decision in a case called Feres v. United States, which bars active-duty military personnel from suing the government for damages."[20]

According to the Congressional Budget Office, healthcare was not only the most expensive support function for the DoD in 2001 and 2016, it also saw the greatest increase in cost over the same time. In an era of "do more with less," AI can take on menial tasks currently performed by medical personnel, freeing up time for doctors and nurses to focus on patients and spend more time working at the top of their license. AI can help transfer time-consuming human tasks to machines while enabling patients to self-service their care needs where possible.

Figure 2.

Funding for Selected Support Functions, 2001 and 2016

Billions of 2016 Dollars



Source: Congressional Budget Office, using data from DoD.

CBO examined recent trends in spending among these four subcategories because their growth in spending from 1980 to 2016 outpaced the overall growth in support spending during that period.

CBO's analysis examines funding in DoD's base budget; it does not include funding for overseas contingency operations such as the current Operation Freedom's Sentinel in Afghanistan or Operation Inherent Resolve in Iraq and Syria.

DoD = Department of Defense.

Medical errors are the third leading cause of death in the United States after heart disease and cancer.[21] A 2008 study places the accrued costs of measurable human-errors in the US at $20.1 billion, adjusted for inflation.[22] These costs can be further broken down into direct and indirect costs. Direct costs are medical costs of providing inpatient, outpatient, and prescription drug services to individuals affected by medical errors. Indirect costs are related to increased mortality rates among individuals who experience medical errors as well as lost productivity due to

20 David P. Sheldon and Corey D. Bean, "Explainer: Can You Now Sue the Military for Medical Malpractice?," Navy Times, December 26th, 2019, https://www.navytimes.com/news/your-navy/2019/12/26/explainer-can-you-now-sue-the-military-for-medical-malpractice/.

21 CNBC on the U.S. third leading cause of death: medical errors

22 Health Affairs Project HOPE – Van Den Bos et. al., "The $17.1 Billion Problem: The Annual Cost of Measurable Medical Errors"

short-term disability.[23] Artificial intelligence can be utilized to support human decision-making and dramatically reduce medical errors, thereby driving down costs. A study conducted by Harvard researchers showed an Israeli AI startup named MedAware showed promise in preventing incorrect prescriptions and other common medical errors. The company claims it can detect instances of incorrect drug prescriptions, or a prescription assigned to the wrong patient. It can also raise a warning for time-dependent events, such as changes in a patient's vitals or lab work that could make a drug riskier. The study, published in the Joint Commission Journal on Quality and Patient Safety, estimated existing systems would not have detected 68.2 percent of errors flagged by MedAware's platform.

> "The retrospective study looked at records from 747,985 patient encounters at Brigham and Women's Hospitals' outpatient clinics between 2012 and 2013. From that, MedAware's system flagged 10,668 potential errors, 300 of which researchers sampled for further review. According to the study, 92 percent of warnings generated by MedAware's system were accurate based on the available data, and 79.7 percent were clinically valuable. The study estimated Brigham and Women's would have saved $1.3 million in healthcare-related costs, extrapolating on the results from this sample.[24]"

In addition to utilizing AI to drive down costs via reduced medical errors, it can also be effective in driving down costs through greater efficiency as it is estimated that 30 to 50 percent of the 80 million CT scans in the United States are unnecessary.[25]

23 Shreve et. al., "The Economic Measurement of Medical Errors"

24 Elise Reuter for MedCityNews, "Study shows reduction in medication errors using health IT start-up's software"

25 Rebecca Smith-Bindman, "Use of Advanced Imaging Tests and the Not-So-Incidental Harms of Incidental Findings," JAMA Internal Medicine 178, no. 2 (February 1st, 2018): 227–28, https://doi.org/10.1001/jamainternmed.2017.7557.

Increase Personnel Readiness

Former NFL safety Brian Dawkins once said, "The best ability is availability." Put another way, professional sports teams depend on athlete availability to be successful. A professional athlete is useless to the team if they are injured or sick, and professional athletic teams are investing heavily in AI and machine learning to assist with recruitment and increased athlete availability. Companies like Sparta Science use software to assist teams like the Cleveland Cavaliers, Colorado Rockies, San Francisco 49ers, Detroit Lions, and Pittsburgh Steelers in aiding recruiting decisions and improving athlete availability. Machine learning and big data analytics have become critical in helping some professional teams identify fitness levels to recruit and select the best talent, prevent injuries, speed up the return from injury, and assess future injury risk.

A key impediment to athlete availability is musculoskeletal (MSK) injuries, and the military is no exception. Musculoskeletal injuries among active duty soldiers result in over 10 million limited duty days each year and account for over 70% of the medically non-deployable population.[26] Additionally, military training-related musculoskeletal injuries have been called "the single most significant medical impediment to military readiness" and musculoskeletal injuries and their long-term effects are a leading cause for medical disability and discharge.

Currently, the US military relies on a "one-size-fits-all" approach to readiness with a physical fitness test once every six months to a year combined with a once per year physical with a doctor to determine a service member's availability for duty. Modern machine learning applications can complement these requirements and provide more precision, providing ongoing, weekly analysis of servicemembers' availability along with insights for injury prevention. Machine learning provides an innovative way to tackle readiness efficiently while also enabling the objective use of data to make evidence-based decisions.

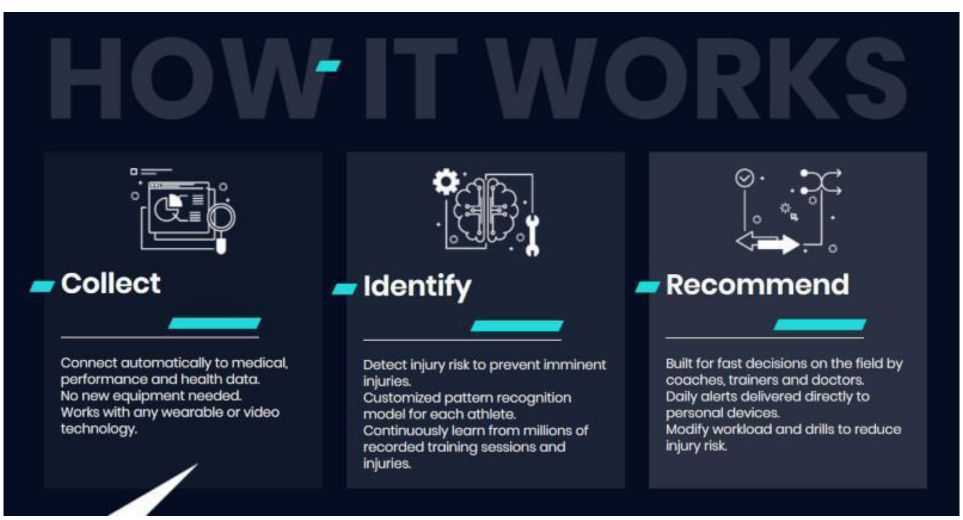
When a servicemember does fall ill, artificial intelligence can provide immediate support to the servicemember while simultaneously informing

26 APHC Fact Sheet on Musculoskeletal Injuries

the local military treatment facility and supporting local decision making. In his National War College research paper titled *Artificial Intelligence and Military Medicine: Strategic Application of the Technology of our Era*, current JAIC Health Mission Chief Dr. Hassan Tetteh provides a clear example of how this might work:

> "Consider the example in which a service member wakes up one morning feeling ill. He 'explains' to his AI-enhanced home platform that he is not feeling well. A series of sensors in his home 'interact' with him to determine that he has a fever, an elevated heart rate, and a slightly elevated respiratory rate. His AI-enhanced platform is 'trained' to recognize his voice and detects that his voice is slightly distorted. His AI-enhanced platform calls upon a telehealth connection, and a virtual visit is established with a nurse practitioner at the military treatment facility that confirms the service member's symptoms. Additional diagnostic tests during the encounter, including a rapid strep test processed through the AI-enhanced platform 'amplify' the nurse practitioner's assessment and the results return positive for streptococcus infection."

Another example of AI being used to help sports teams increase performance and athlete availability is Zone 7 AI. Zone 7 uses machine learning algorithms to enable high levels of athlete performance and availability. Zone 7 claims to have achieved up to "70% reduction in days lost to injury and 75% reduction in injury rates.[27]"



27 Zone7.ai "About" page

Former Chairman of the Joint Chiefs of Staff General Joseph Dunford estimates that 18,000 troops are unfit for duty at any given time because of

medical-related issues.[28] This number is the equivalent of an entire Army Division or 10% of the Army's frontline fighting force as there are currently 10 Divisions in the Active Duty Army.

AI has the potential to increase combat medical capability through combat health informatics. Countries like Estonia are already using AI-based healthcare systems in the form of apps designed to provide paramedics information about patients before reaching their homes.[29] The US military could utilize similar capabilities for field medicine and combat search and rescue, providing higher probabilities for survival within the "golden hour" — the critical period after a traumatic injury where immediate medical treatment decreases fatalities.

Optimize Personnel Performance

As the Department of Defense shifts its focus toward great-power competition, and as the recent COVID-19 pandemic highlights, the ability for military medical units, both deployed and home station, to operate at a continuous operational tempo while making minimal errors will be critical in optimizing warfighter/emergency responder performance and sustaining power projection. It might seem counterintuitive at first, but one of the integral parts to an AI-enabled DoD will be healthy service members who are fit to fight. In order to maximize warfighting potential in the future, while maintaining true to American values of liberty and privacy, the Department of Defense needs to lead in the application of artificial intelligence in all domains and particularly in healthcare.

One healthcare area that has already received a lot of attention with respect to artificial intelligence is in assisting physicians with medical diagnosis. In the book *Thinking Fast and Slow*, Daniel Kahneman describes the

28 Interview with GEN Joseph Dunford (ret.)

29 Eric J. Topol, Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again, First edition (New York: Basic Books, 2019), 275.

dichotomy between two modes of thought: "System 1" is fast, instinctive, and emotional; "System 2" is slower, more deliberative, and more logical. Artificial intelligence and machine learning can be used to support System 2 thinking or the slow, reflective process that involves much more analytical effort than quick, intuitive judgments. "There are about 10,000 human diseases, and there's no doctor who could recall any significant fraction of them. If doctors can't remember a possible diagnosis when making up a differential, then they will diagnose according to the possibilities that are 'mentally' available to them, and an error can result. This is called the availability bias.[30]" Machine learning algorithms have no problem recalling and analyzing 10,000 different human diseases and provide a strong defense against the "availability bias" while providing doctors regular feedback. As mentioned previously, AI is not an elixir; it is an enabler. We don't foresee AI/ML replacing physicians or operating under a human out-of-the-loop construct. Instead, similar to the concept of sensor fusion in 5th generation fighter aircraft such as the F-35, we see AI/ML as a tool for greater situational awareness for medical professionals. In the end, the doctor still makes the final decisions regarding diagnosis and treatment, just as an F-35 pilot makes the final determination to employ ordnance. However, like the F-35 analogy, we imagine a future of highly efficient and effective healthcare delivery informed by data integration and machine learning.

Artificial intelligence can automate many routine tasks, taking work off the shoulders of an already overworked military health system. Artificial Intelligence is unlikely to replace most healthcare professionals. Healthcare will still require the wisdom and judgment required by trained healthcare professionals as "proper interpretation and use of computerized data will depend as much on wise doctors as any other source of data in the past."[31] The promise of artificial intelligence in healthcare is to improve doctor decision-making, reduce errors in misdiagnosis and recommend proper treatment. Unlike humans who get tired, sick, or emotional, machines can work twenty-four hours a day and seven days a week without a decrease in performance. Allowing AI to handle menial, repetitive tasks while providing more accurate diagnosis will free up time for humans to work on higher-order, more complex issues that demand extensive medical

30 Topol, Deep Medicine

31 R.M. Califf, "The Doctor and the Computer," Western Journal of Medicine, 1981, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1273186/.

training along with creative and innovative thinking. Michael Recht and Nick Bryan, writing in the Journal of the American College of Radiology said: "We believe that machine learning and AI will enhance both the value and the professional satisfaction of radiologists by allowing us to spend more time performing functions that add value and influence patient care and less time doing rote tasks that we neither enjoy nor perform as well as machines.[32]"

Another use case for AI/ML in personnel optimization is in personalized nutrition. Advances in machine learning now provide individuals with highly personalized advice catered to their bodies and designed to maximize human performance. Through advances in machine learning and big data, researchers have been able to more precisely design nutrition plans capable of maximizing human performance. "Extensive monitoring of a human cohort for variations in dietary intake, lifestyle, host phenotype, and the gut microbiome has enabled the development of a machine-learning algorithm that accurately predicts the individual glycemic response to meals, providing an important first step toward personalized nutrition."[33] Determining specific foods that are best for any given person could provide the DoD with a healthier, high-performing war-fighting force. Instead of relying on a red/yellow/green color chart at a dining facility buffet, soldiers, sailors, airmen, and marines could utilize wearable devices connected to their health data to provide the precise nutrition and fitness information designed specifically for them. "With the assistance of AI, adjustments in diet and behavior can be suggested *before* a service member slips from healthy into early illness.[34]"

32 Michael Recht and R. Nick Bryan, "Artificial Intelligence: Threat or Boon to Radiologists?"

33 Reiner Jumpertz von Schwartzenberg and Peter J. Turnbaugh, "Siri, What Should I Eat?," Cell 163, no. 5 (2015): 1051–52, https://doi.org/10.1016/j.cell.2015.11.012.

34 Tetteh, Artificial Intelligence and Military Medicine: Strategic Application of the Technology of our Era

Health Data is a Strategic Asset

The Department of Defense Digital Modernization Strategy states that "the DoD CIO is tasked with improving the combat power of the Department—as well as its security and efficiency—by ensuring that the Department treats information as a strategic asset and that innovative information capabilities are available throughout all areas of DoD supporting warfighting, business, and intelligence missions.[35]" To be sure, the Department has made the innovation mission a centerpiece of its efforts in recent years, standing up the Defense Innovation Unit (DIU), the Joint Artificial Intelligence Center (JAIC), and other units and task forces well-equipped to continue the implementation of the Modernization Strategy's vision.

However, this Strategy's intent must be better applied at places in the DoD where actionable data is being "left on the table." As this paper was in draft, the Defense Innovation Board (DIB) published a timely report on the Joint Pathology Center's long-term plan to digitize its staggeringly large record base, which includes over 55 million glass pathology slides, 31 million paraffin-embedded tissue blocks, and 500,000 wet tissue samples - 7.4 million cases in all - collected over the past 100 years.[36] This digitization is a noble effort; the above-cited Defense Innovation Board (DIB) report acknowledges that "if the entire repository were leveraged to its fullest potential, it would advance diagnosis and treatment for thousands of illnesses, saving lives across the DoD and the global population."

This digitization effort also represents a dangerous thinking trap. The US has a remarkable opportunity to lead *now* in the fight to address the world's most pressing public health issues. Our largest, near-term AI development competitor - China - owns a comparatively shallow pool of civilian and military health data to develop predictive health algorithms. However, it is growing exponentially due in equal parts to civil-military fusion, granting its own bleeding-edge scientists access to its most data-rich health

35 Department of Defense, DoD Digital Modernization Strategy. Washington, DC: Department of Defense, 2019. Available online at https://media.defense.gov/2019/Jul/12/2002156622/-1/-1/1/DOD-DIGITAL-MODERNIZATION-STRATEGY-2019.PDF.

36 Lander et. al. "Defense Innovation Board Joint Pathology Center (JPC) Repository Enhancement." Department of Defense Office of Prepublication and Security Review. March 4th, 2020. Available online at https://media.defense.gov/2020/Mar/05/2002259783/-1/-1/0/DIB_JPC_RECOMMENDATIONS_FINAL.PDF.

repositories and with a greater risk tolerance for privacy costs in the development of these algorithms on the part of the country's defense leadership. China's own New Generation of Artificial Intelligence Development Plan (AIDP) flatly sets 2030 as the country's deadline to become "the major artificial intelligence innovation center of the world.[37]"

To be clear, some of what the Chinese scientific community has done with its growing pool of health data and healthcare in general is wholly incompatible with American values. Using CRISPR to edit embryonic genomes *in vivo* (editing the genes of soon-to-be-living babies)[38] and conducting genomic studies of ethnic minorities with the direct backing of and use of data collected by state intelligence agencies[39] are just two examples. However, the clear need for an ethical, privacy-protective approach to scientific inquiry in the AI healthcare space is no excuse for a failure to innovate. Without concurrent efforts to leverage the most impactful repository of American military health data for predictive health research, the current Joint Pathology Center digitization plan risks planning to solve problems on a timeframe in which our country's most formidable competitor will already have them solved.

For example, under the Defense Innovation Board report's second major recommendation, titled "Pilot on Linkage to Medical Records," the argument begins that "the Joint Pathology Center should initiate a pilot to assess the ability to add value to the physical slide and tissue block collections by connecting a representative collection of slides to medical record information on a more automated and digital basis, to determine the extent of information challenges." Such technology *has already been deployed within the DoD* to organize and link de-accessed cancer biopsy data across a large test group of military medical records. The Joint Artificial Intelligence Center/Defense Innovation Unit Predictive Health

37 Translated by Sapio, Flora et. al. "A Next Generation of Artificial Intelligence Development Plan." State Council Document No. 35. 2017 Available online at https://flia.org/notice-state-council-issuing-new-generation-artificial-intelligence-development-plan/.

38 Greely, Henry T. "CRISPR'd babies: human germline genome editing in the 'He Jiankui affair.'" Journal of Law and the Biosciences, Volume 6, Issue 1. October 2019. Pages 111-183. Available online at https://academic.oup.com/jlb/article/6/1/111/5549624.

39 Jing Jia et. al. "Developing a Novel Panel of Genome-Wide Ancestry Informative Markers for Bio-Geographical Ancestry Estimates." Forensic Science International, Volume 8, Issue 1, pages 187-194. January 1st, 2014. Available online at https://www.fsigenetics.com/article/S1872-4973(13)00193-2/fulltext.

collaboration has been working for the better part of a year on several supporting components for *exactly* this proof-of-principle, including:

- A highly customizable de-identification solution for free-text medical records. This doubles as a sorting mechanism pairing relevant clinical notes with a de-accessed cancer biopsy slide based on a unique de-accession number (akin to a library catalog number)
- An augmented-reality microscope which, in prior partnership with expert pathologists, can assist clinicians in flagging areas of interest for potential cancer diagnosis.

Empowering Concurrent Predictive Health Initiatives

The scope of the JPC's biopsy data extends far beyond cancer diagnoses and their relevant clinical notes. As currently developed, the predictive health technology deployed by DIU's Predictive Health team would eventually need access to a large amount of data to provide reliably small amounts of degradation across - or, effectiveness in de-identifying - a broader set of medical record types. However, the DIB paper's working assumption that the Protected Health Information (PHI) contained in comparable Joint Pathology Center records is currently not equipped to aggregate and optimize the development of predictive health algorithms is fundamentally flawed.

The requirement to store Protected Health Information (PHI) in a secure fashion arises from the Health Insurance Portability and Accountability Act (HIPAA), first passed by the Clinton administration in 1996.[40] HIPAA was also the first bill to *define* PHI as follows:

Individually Identifiable Health Information is information, including demographic data, that relates to:

- the individuals' past, present or future physical or mental health or condition,
- the provision of health care to the individual, or
- the past, present, or future payment for the provision of health care to the individual, and that identifies the individual or for which there is a reasonable basis to believe it can be used to identify the individual.[41] Individually identifiable health information includes many common identifiers (e.g., name, address, birth date, Social Security Number - see Annex A).

40 CDC, "Health Insurance Portability and Accountability Act of 1996 (HIPAA)," Available online at https://www.cdc.gov/phlp/publications/topic/hipaa.html

41 US Department of Health and Human Services (HHS). "Summary of the HIPAA Privacy Rule." Available online at https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html.

HIPAA further defines De-Identified Health Information as verifiably de-identified in two ways:

- By the formal determination of a qualified expert; ("expert determination")
- The removal of specified identifiers of the individual and the individual's relatives, household members, and employers. ("Safe Harbor")

Either method is also "adequate only if the covered entity has no actual knowledge that the remaining information could be used to identify the individual."

However, recent advances in natural language processing - some of which are explicitly employed in the DIU/JAIC proof-of-principle mentioned above - have made it possible to de-identify clinical datasets well above federal compliance with these standards, first, by labeling all words and phrases that could identify an individual, and then replace them with surrogate data or context-specific labels.[42]

For example:

- Original record:
    - "John London complains of chest pain that started on January 1st, 2012"
- De-identified record:
    - "[Person] complains of chest pain that started on [Date]."

The de-identification process further requires high recall (sensitivity) because the public release of PHI represents a legal and ethical liability. It also requires reasonable precision; the unnecessary removal of

42 Hassidim et. al. "Customization Scenarios for De-Identification of Clinical Notes. BMC Medical Informatics and Decision-Making, 2020. Available online at https://research.google/pubs/pub48935/.

non-identifying text would limit the data's usefulness to researchers. Finally, the de-identification system must be context-aware: London, which is usually known as a city, is accurately labeled as a name based on the sentence structure.

Defense Innovation Unit's machine learning solution provides this context-awareness through what is known as a bootstrap confidence interval. The algorithm the team has developed will take three passes at a sample, first scanning for obvious identifiers like names (to include special characters), Social Security Numbers, and common lead-in phrases that a doctor might use to describe a given health condition, such as the chest pain suffered by "John London" in the example above. In the second pass, the algorithm will use those lead-in phrases to preserve what aspect of the note is relevant (the "chest pain") while scrubbing the clinical note of the remaining 10-15 identifiers based on which are present. The second pass will also check to see if any clinically relevant information (in a cancer record, this might be the name Brown, which is also the name of a class of thyroid tumor) that may have been accidentally excluded. The final pass checks the accuracy of each preceding step, resulting in a sensitivity of 99.9917% before manual verification, reducing the cost of de-identification by 20x across the cancer records to which it has currently been applied.

This already-low error rate must be placed in the appropriate context. These cancer biopsy records were data on their way to "de-accession," i.e. an incinerator for their pending expiry. The Defense Innovation Unit team used these de-accessed records to develop the de-identification approach described above, then selected cases from the records to develop a deep learning system that classified prostate cancer from slide images alone.[43] The team selected cancer for its stable and relatively easy-to-identify clinical presentation. If their efforts can achieve this error rate on such a limited corpus or body of clinical notes, it directly follows that access to more data would not only improve the efficacy of their de-identification algorithm, but resulting improvements would directly address other concerns presented in the Defense Innovation Board's recommendations to the Joint Pathology Center.

43 Nagpal et. al. "Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer." Nature, 2019. Available online at https://www.nature.com/articles/s41746-019-0112-2.pdf.

From the Defense Innovation Board report, under "Data Security & Access:"

> Once the Joint Pathology Center has determined its data storage strategy, the next challenge will involve the security and access to that stored data. Security is particularly critical in the case of the JPC, given the wealth of personal data that will result from the slide scanning process. Access to the data will inherently be correlated to security. The Joint Pathology Center is not unique in being a caretaker of sensitive personal data: other medical organizations, government organizations, financial institutions, and others bear the same responsibility and have security products that defend that data accordingly. There are a broad range of data architecture strategies and commercial products that can ensure the security of the Joint Pathology Center's data, as well as control access and retrieval of that data. There are a variety of "data de-identification" products that separate Personally Identifiable Information (PII) for Protected Health Information (PHI), **creating a category of data that cannot be tied back to an individual while maintaining the richness of the Joint Pathology Center's data. This process also provides more storage flexibility, as de-identified data does not fall under HIPAA.**

The first bolded sentence is exactly the challenge that Defense Innovation Unit's de-identification solution is designed to solve. The current scheme of pairing an individual case's de-identified clinical notes with a unique de-accession number allows for tagging the anonymized data and sorting it for ease-of-access while still preserving its richness. Similar pseudonymized identifiers (whether barcoded or otherwise) could be utilized in the case of the Joint Pathology Center. The sorting aspect turns millions of records from a useless aggregate into a searchable database.

The second bolded sentence refers indirectly to the following data classification scheme, which is used to determine access controls for DoD information depending on the kind of data the job requires. It also groups these data by impact level (IL), assigning a numerical value based on which leak or spill would cause the greatest damage to national security.

IMPACT LEVEL	INFORMATION SENSITIVITY	SECURITY CONTROLS	LOCATION	OFF-PREMISES CONNECTIVITY	SEPARATION	PERSONNEL REQUIREMENTS
2	PUBLIC or Non-critical Mission Information	FedRAMP v2 Moderate	US / US outlying areas or DoD on-premises	Internet	Virtual / Logical PUBLIC COMMUNITY	National Agency Check and Inquiries (NACI)
4	CUI or Non-CUI Non-Critical Mission Information Non-National Security Systems	Level 2 + CUI-Specific Tailored Set	US / US outlying areas or DoD on-premises	NIPRNet via CAP	Virtual / Logical Limited "Public" Community Strong Virtual Separation Between Tenant Systems & Information	US Persons ADP-1 Single Scope Background Investigation (SSBI)
5	Higher Sensitivity CUI Mission Critical Information National Security Systems	Level 4 + NSS & CUI-Specific Tailored Set	US / US outlying areas or DoD on-premises	NIPRNet via CAP	Virtual / Logical FEDERAL GOV. COMMUNITY Dedicated Multi-Tenant Infrastructure Physically Separate from Non-Federal Systems Strong Virtual Separation Between Tenant Systems & Information	ADP-2 National Agency Check with Law and Credit (NACLC) Non-Disclosure Agreement (NDA)
6	Classified SECRET National Security Systems	Level 5 + Classified Overlay	US / US outlying areas or DoD on-premises CLEARED / CLASSIFIED FACILITIES	SIPRNET DIRECT With DoD SIPRNet Enclave Connection Approval	Virtual / Logical FEDERAL GOV. COMMUNITY Dedicated Multi-Tenant Infrastructure Physically Separate from Non-Federal and Unclassified Systems Strong Virtual Separation Between Tenant Systems & Information	US Citizens w/ Favorably Adjudicated SSBI & SECRET Clearance NDA

In our interview with the Joint Artificial Intelligence Center's Colonel Sang Han - a member of the team spearheading the development of the Joint Common Foundation (JCF), a solution that would allow for collaboration by cleared analysts with uncleared researchers that has been delayed due to a contract fairness suit against Microsoft by Amazon[44] - the team discussed what it would take to bring data from IL-4 (the current rating for PHI) to IL-2 or lower, which has lower investigative requirements for access control. While Colonel Han explained that anonymization "could objectively lower the impact level of a given dataset,[45]" the DIB report points at another benefit. Anonymized data falls outside HIPAA, and if new policy is drafted to provide leadership guidance for concerns regarding the exact definition of "anonymized," the anonymization mechanism by itself provides a direct path forward to proofs-of-principle which are concurrent with the Joint Pathology Center digitization effort.

44 Shaw, Frank X. Microsoft On the Issues, "Bid high, lose, try again. Amazon continues to push for a JEDI re-do." Available online at https://blogs.microsoft.com/on-the-issues/2020/05/07/amazon-jedi-re-do-dod/

45 Interview with Colonel Sang Han, Department of Defense

COVID-19 and the Policy Mandate it Provides

As also partially argued, the COVID-19 pandemic provides an unprecedented opportunity for the US to lead the global conversation on predictive health. Real or perceived shortcomings in the federal response to the pandemic have also not been directly attributed to the military, which bodes well for its highest public trust rating among federal institutions. However, that trust must be properly leveraged to provide forward-thinking leadership on this still-futuristic dimension of AI research.

The JAIC's Project Salus - an effort to integrate national supply chain data into the military's COVID-19 response plan - is an excellent start. Described as "data-crunching" by a recent FedScoop report,[46] it is intended to help Northern Command and the National Guard understand where critical supplies are in short supply, and aims to unify data from retailers with data from government sources to create a better Common Operating Picture of the overall American supply chain. During previous/comparable disaster response efforts, Guard units specifically would have to count supplies of things like water bottles in stores, according to Nand Mulchandani, acting Director of the JAIC. Mr. Mulchandani is further quoted in a Roll Call piece: "The hope if we do this right, this becomes the supply and logistics platform for all the services. You just plug in different datasets." His words echo this team's confidence in broader data access to maximize the performance of predictive health algorithms that map more directly to health conditions.

A similar "horizon" dataset from the private sector comes from a start-up called C3.ai. The company has developed a COVID-19 data lake that provides COVID researchers access to integrated datasets across disparate sources. This data lake can be accessed through a simple API call with R Studio or Python packages. This ease-of-access to useful data is incredibly important in the context of machine learning. In the hours it might typically take a researcher to wrangle and clean data, C3.ai's integrated

46 Barnett, Jackson. "JAIC's new Project Salus links up supply-chain data in fight against coronavirus." FedScoop. April 15th, 2020. Available online at https://www.fedscoop.com/supply-chain-jaic-project-salus-coronavirus/.

data lake allows a researcher to do the same data wrangling in about 7 minutes. The DoD could utilize similar data unification methods similar to those used by C3.ai to significantly reduce time spent data wrangling and increase time on data analysis. The urgency of the COVID-19 problem is forcing companies to innovate in providing efficient data utilization methods and the DoD should capitalize on this going forward. "Previously disconnected data sources can easily be integrated – including time series data – with a single, simple API request. Navigating, pulling, and aggregating data like country statistics, patient diagnosis, age, and even time-based location movements can be done in minutes instead of days or weeks. This is incredibly valuable when time is critical, and breakthroughs can mean thousands of lives saved.[47]"

47 Interview with Conor Makowski, MIT Cave Lab

Recommendations

1. Unambiguous Command Authority

The Secretary of Defense needs to make the digitization of DoD-owned medical data a top priority and should direct the JAIC and DHA to seek and plan for additional datasets immediately. We agree with the National Security Commission on AI's recommendation to create a Steering Committee on Emerging Technology tri-chaired by the Deputy Secretary of Defense, the Vice Chairman of the Joint Chiefs of Staff and the Principal Deputy Director of National Intelligence. We also agree that the Joint Artificial Intelligence Center should be moved up the chain of command, reporting directly to the Secretary of Defense or delegated to the Deputy Secretary of Defense. It should continue to be led by a three-star General or Flag Officer with significant operational experience. Concerning artificial intelligence, we suggest that the steering committee include the Director of the JAIC as well as the Assistant Secretary of Defense for Health Affairs, Director of the Defense Health Agency, JAIC's Warfighter Health Lead, and the CIO of the Defense Health Agency (DHA) who currently oversees the Authority to Operate (ATO) process for IT systems within the Defense Health Agency. Under strategic guidance from the Steering Committee on Emerging Technology, the JAIC Warfighter Health Lead should partner with the Defense Health Agency CIO to execute a plan for utilizing DoD medical data to train machine learning algorithms.

2. Industry and Academic Partnerships

The DoD should partner with industry and academia to effectively store, maintain, and utilize digitized medical data. We recommend that the Joint Artificial Intelligence Center and Defense Health Agency jointly partner with Defense Innovation Unit to use the Predictive Health project as a proof of principle for future projects and scaling within the DoD. The JAIC can leverage DIU's well-established presence in Silicon Valley and Boston (an established biotech hub) to engage with commercial vendors and identify innovative technologies, methodologies, and algorithms that can be used to enhance military health. The JAIC can also leverage a proven Commercial Solutions Opening (CSO) process to develop prototype projects with customer government organizations, including the Defense Health Agency and Joint Pathology Center. Relevant government partners would supply prototype funding with the Joint Artificial Intelligence Center providing a transition plan to ensure successful prototype adoption. We also recommend detailing key Defense Health Agency personnel to the JAIC to help steer this process.

In addition to the above, we also recommend the Joint Artificial Intelligence Center continue to establish official partnerships with top AI research institutions like MIT and Carnegie Mellon. Leveraging established DoD partnerships like the Air Force/MIT AI Accelerator to grow partnerships with MIT Principal Investigators leading discoveries in AI and healthcare. Currently, the Air Force/MIT partnership is funded for ten projects. The Joint Artificial Intelligence Center could collaborate with the Air Force, provide additional personnel and funding, and partner with researchers like Dr. Luis Soenksen at MIT Proto Ventures and the J-Clinic who is focused on AI and healthcare and leading research around federated machine learning, differential privacy, digital phenotyping, and adversarial re-programming. Academic institutions like MIT have experience working with datasets similar to the Joint Pathology Center's and are intimately familiar with the methods for properly releasing data for research. Additionally, the Air Force's AFWERX small business technology transfer funding could be utilized to jumpstart DoD/industry/academia collaboration on DoD AI and healthcare projects.

3. Data Digitization

Using the DIU/JAIC Predictive Health project as a proof-of-principle, the Joint Artificial Intelligence Center and Defense Health Agency should cooperate to develop new digital banks through partnerships with existing Medical Treatment Facilities. Naval Medical Center San Diego (NMCSD) has been collaborating with the Defense Innovation Unit to digitize local pathology slides. The team have collaborated to digitize and store over 400,000 pathology slides to date. The Joint Artificial Intelligence Center and Defense Health Agency can use this project as a proof-of-principle to scale the adoption of data digitization across select Military Treatment Facilities. This recommendation has already seen baseline promise in a recent contract award to Google Cloud for digitization efforts at West Coast Veterans Affairs and Military Treatment Facilities.[48]

While the ability to call a de-identification algorithm directly from Google Cloud does not solve the issue of connecting Google-quality engineers with health data at military facilities, it *does* represent an important launchpad for future collaborative efforts. For example, once the Joint Common Foundation is operational, engineers and researchers at companies across Silicon Valley, the Boston Health Community, the National Capital Region, and elsewhere could compete per this paper's prior recommendations for the best way to maintain low degradation rates in de-identification and sorting solutions which can be applied across other types of medical records. A successful shift left in cancer diagnosis timelines can pre-empt equal or greater shifts in the earlier prediction of degenerative or even infectious diseases with the right mix of data and talent to fuel them.

48 "Defense Innovation Unit Selects Google Cloud to Help U.S. Military Health System with Predictive Cancer Diagnoses." Press Release. Google. September 2nd, 2020. Available online at https://cloud.google.com/press-releases/2020/0902/predictivehealthdiu

Final Note:

In the end, the recommendations in this paper provide a roadmap to transform the digital data landscape within DoD healthcare. The goal is better diagnostic care, healthier service members, and increased readiness across the force. However, here is a more important opportunity at play. The size and uniqueness of the DoD dataset means advances in healthcare and medicine with national and global impact. The opportunities for advancements in healthcare are endless if we allow AI/ML researchers and medical professionals access to this data. It only requires the belief that data is a strategic asset and the desire to leverage that dataset for the betterment of all people.

National Security Fellows Program
Belfer Center for Science and International Affairs
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138

www.belfercenter.org/NSF

Printed in the United States of America