

Viewpoint

# Ethical Applications of Artificial Intelligence: Evidence From Health Research on Veterans

Christos Makridis<sup>1,2,3\*</sup>, PhD; Seth Hurley<sup>1\*</sup>, PhD; Mary Klote<sup>4</sup>, MD; Gil Alterovitz<sup>1,5\*</sup>, PhD

<sup>1</sup>National Artificial Intelligence Institute, Department of Veterans Affairs, Washington, DC, United States

<sup>2</sup>Stanford Digital Economy Lab, Stanford University, Stanford, CA, United States

<sup>3</sup>WP Carey School of Business, Arizona State University, Tempe, AZ, United States

<sup>4</sup>Office of Research & Development, Department of Veterans Affairs, Washington, DC, United States

<sup>5</sup>Boston Children's Hospital, Harvard Medical School, Boston, MA, United States

\*these authors contributed equally

**Corresponding Author:**

Christos Makridis, PhD

National Artificial Intelligence Institute

Department of Veterans Affairs

810 Vermont Avenue NW

Washington, DC, 20420

United States

Phone: 1 2022977787

Email: [christos.makridis@va.gov](mailto:christos.makridis@va.gov)

## Abstract

**Background:** Despite widespread agreement that artificial intelligence (AI) offers significant benefits for individuals and society at large, there are also serious challenges to overcome with respect to its governance. Recent policymaking has focused on establishing principles for the trustworthy use of AI. Adhering to these principles is especially important for ensuring that the development and application of AI raises economic and social welfare, including among vulnerable groups and veterans.

**Objective:** We explore the newly developed principles around trustworthy AI and how they can be readily applied at scale to vulnerable groups that are potentially less likely to benefit from technological advances.

**Methods:** Using the US Department of Veterans Affairs as a case study, we explore the principles of trustworthy AI that are of particular interest for vulnerable groups and veterans.

**Results:** We focus on three principles: (1) designing, developing, acquiring, and using AI so that the benefits of its use significantly outweigh the risks and the risks are assessed and managed; (2) ensuring that the application of AI occurs in well-defined domains and is accurate, effective, and fit for the intended purposes; and (3) ensuring that the operations and outcomes of AI applications are sufficiently interpretable and understandable by all subject matter experts, users, and others.

**Conclusions:** These principles and applications apply more generally to vulnerable groups, and adherence to them can allow the VA and other organizations to continue modernizing their technology governance, leveraging the gains of AI while simultaneously managing its risks.

(*JMIR Med Inform* 2021;9(6):e28921) doi: [10.2196/28921](https://doi.org/10.2196/28921)

**KEYWORDS**

artificial intelligence; ethics; veterans; health data; technology; Veterans Affairs; health technology; data

## *Ethical Applications of Artificial Intelligence in Veterans' Health Research*

There is increasing recognition that artificial intelligence (AI) offers significant potential to help or harm the world. Much like other technologies, ranging from the internet to computers, AI

is neither bad nor good: the impact of AI depends on how its users wield it. Already, there is an emerging body of AI use cases in health care [1,2], including for vulnerable groups and veterans [3], that are increasingly originating from populations the federal government considers to be “potentially vulnerable patient populations” [4]. These groups can be especially sensitive to adoption of technology; therefore, additional

scrutiny is required around the ethical underpinnings and likely causal effects on these groups. In this sense, the question is not whether the federal government should engage AI for broader social benefit, but how it can do so using a values-based framework to guide AI applications and their continued research and development.

At least since the publication of the Belmont Report [5], there has been general recognition in the federal government of three principles that guide the introduction of new technologies to this day. First, *respect for persons* details that individual autonomy and privacy must be protected. Second, *beneficence* states that technologies should be designed to maximize the potential net benefits to society, safeguarding against potential harms and long-term consequences. Third, *justice* ensures that there are equitable benefits from research. That is, when individual data is collected, it must be used to benefit those individuals. Although the Belmont Report focused on biomedical technologies, they exhibit many similarities with AI, particularly in terms of their ethical implications and long-lasting impacts.

The primary contribution of this commentary is to explore the ethical applications of AI by building on the Belmont Report and relating it with the principles established in the recent executive order on trustworthy AI. Although there has been a recognition of data ethics and privacy within the US federal government, a new challenge has emerged: how can the federal government balance between the competing priorities of stewarding sensitive data and using AI to analyze it to drive veteran outcomes?

To answer this question, we apply the perspective of the Veterans Health Administration, within the Department of Veterans Affairs (VA), which has the largest integrated health care system in the United States and has pioneered several technological aspects now widely seen in this country, such as electronic health records (EHRs). Additionally, more than half of physicians training within the United States receive some training at a VA medical center. By evaluating uses for AI and implications in health care, veteran input and priorities can be proactively developed to enhance care. For example, the VA is already using AI to facilitate early detection of cancer [3], detection of acute kidney injury [6], and prediction of loneliness and declines in mental health [7,8]. These examples all highlight the ways that AI can be used to advance patient outcomes; however, they also point toward data privacy and trust considerations.

The recent executive order, “Promoting the Use of Trustworthy Artificial Intelligence in Government” [9], provides a framework for the VA to move forward with using AI to improve veteran health on a larger and more systematic scale. We focus on three principles that are especially relevant to the advancement of the health and well-being of veterans: (1) purposeful and performance-driven; (2) accurate, reliable, and effective; and (3) understandable.

## 1. Purposeful and Performance Driven

*...seek opportunities for designing, developing, acquiring, and using AI, where the benefits of use significantly outweigh the risks and the risks are assessed and managed [9]*

The VA is working to employ AI in high priority areas where there is robust opportunity to advance veteran health outcomes. Recent work indicates that difficulty in transitioning to civilian life is a critical factor underlying negative mental health outcomes in veterans. For example, Makridis and Hirsch [10] have documented a deterioration in labor market outcomes among veterans over the past decade, showing that veterans are increasingly concentrated in metropolitan areas with lower wage and employment growth. Moreover, Makridis et al [11] show that socioeconomic factors are the largest predictors of mental health outcomes among veterans, dwarfing the contribution of location and demographic-specific features. Intuitively, because a significant amount of time is allocated toward work activities, the absence of purpose and self-efficacy in the workplace, especially after coming from a mission-driven environment in the armed services, will impact veterans’ mental health.

AI can be part of the solution. To the extent that veteran records from combat are combined with self-assessments of skills and career preferences, and these data could be comprehensively gathered and harmonized, researchers could use methods from AI to provide veterans with personalized recommendations regarding not only potential job fits but also counseling over the course of their careers. One of the sources of low engagement among employees is a feeling of plateauing and helplessness; therefore, AI-driven recommendations regarding how to optimize career mobility and human capital development would provide veterans with actionable steps to continuously acquire and apply new skills at work.

Another prime example involves personalizing feedback to veterans about how to live healthier lives. End-of-life care is one of the largest sources of health care expenditures. For example, Riley and Lubitz [12] estimate that a quarter of all Medicare spending goes toward care for people during their last year of life. These resources could be more impactful if they were allocated more toward preventative care earlier in life. Using deep learning methods, Ahadi et al [13] illustrate how biological data can be used for longitudinal profiling. Implementing this algorithm, combined with EHRs at the VA, offers the potential to provide practical advice about how to live more productive and happier lives, raising both economic and social well-being.

Veterans in rural areas face challenges accessing care due to a paucity of rural treatment facilities. AI, implemented along with smart devices (eg, smart wearables), could allow for remote monitoring of rural veterans’ health and enable smart devices to alert veterans of health concerns. Recent evidence indicates that AI may be able to predict a person’s mental state, including the likelihood of suicide, raising the likelihood that smart devices could be used for predicting and intervening in veteran suicide [7,8].

However, the benefits of AI depend on ethical implementation. Risks associated with AI implementation need to be thoroughly assessed and managed. If, for example, privacy is disrespected, public trust and confidence, particularly among those who have already sacrificed so much for their country, would be undermined. This is extremely important at the VA, where sensitive data, which is under continuous reassessment and review, is routinely collected from veterans. Moreover, researchers must be cognizant of the potential for replicating sources of bias when training their AI algorithms. That is, researchers must investigate the data and model to, at least qualitatively, assess whether there are potential biases that could lead to error replication through the AI-driven recommendations. For example, one possibility is that samples are not representative of the entire population of veterans [14], particularly those who do not feel comfortable using technology. Researchers must also ensure that AI-driven insights are derived from representative samples that reflect the diversity of experiences, attitudes, ethnic, and gender composition among veterans. Recent evidence, for example, highlights the lack of diversity in many health care databases as a major limitation [15].

## 2. Accurate, Reliable, and Effective

*...ensure that their application of AI occurs in well-defined domains, and is accurate, reliable, effective, and fit for intended purposes [9]*

The VA is well-equipped to ensure the accuracy, reliability, and effectiveness of AI applications in health and well-being. The VA has collected and catalogued over two petabytes of data, including data on veteran health, prescription data, and inpatient and outpatient services, among others. Further, the VA established the Million Veteran Program, which characterizes, through a consented cohort of subjects, the confluence of genes, lifestyle, and military exposure on veteran health outcomes. This breadth of data paves the way for algorithms that promote personalized medicine based upon life experience and genetic factors. In particular, the plethora of data at the VA can be leveraged to train high-quality algorithms to serve veteran needs.

Concerns have been raised over whether AI algorithms will be effective and generalize beyond the training set originally used to develop machine learning (ML) algorithms [14]. Importantly, the VA's data sets are generated from VA centers across the country and, in principle, data should accurately capture the diverse spectrum of veterans. Therefore, AI algorithms trained on these data should prove to be reliable even when implemented in varied VA centers throughout the United States. However, cautious implementation and monitoring is necessary to ensure that each developed AI algorithm is beneficial at VA centers.

Although the VA database spans millions of veterans, there are still many veterans who are not included in the system. For example, homelessness is a large challenge for the veteran population, and if these veterans are not included within the VA system, they cannot receive the available benefits and treatment [16]. Our internal calculations from the American Community Survey conducted by the Census Bureau suggest

that there are roughly 18 million veterans in the United States, whereas the VA only covers roughly 9 million of them [17]. To ensure that AI applications produce reliable recommendations for all veterans, it is important to ensure that the data being fed into predictive models is representative.

In addition to the importance of maintaining a representative sample, researchers and clinicians must use appropriate AI techniques. One particularly large challenge with clinical decision support tools and the use of electronic health records is the presence of missing data and small sample sizes. While sample size is less of a challenge within the VA because of the size of its EHR database, missing data can be a source of bias if they are not missing at random [18]. Some ML techniques, such as gradient boosting and decision trees, can deal well with missing data; however, researchers need to be careful about applying ML and automation in these environments. There is also a well-known bias that can emerge against specific groups, whether by race or even socioeconomic status, which can be propagated at scale if ML algorithms are not trained and “de-biased” properly [19]. However, it is becoming clear that researchers developing predictive models for clinical use need to transcend traditional conversations about algorithmic bias and think harder about the broader and structural forces that are at play in the observed phenomena [20].

## 3. Understandable

*...ensure the operations and outcomes of their AI applications are sufficiently interpretable and understandable by subject matter experts, users, and others as appropriate [9]*

A concern for AI development is the necessity for algorithms to be explainable. Explainability is the concept that users should be able to understand how algorithms function, and it is conceptualized along a continuum where relatively simple algorithms based upon branching decision trees and linear regression are feasible to understand [21]. However, the use of deep neural networks (DNNs), where decision-making is spread across multiple layers of interconnected decision-making nodes, currently produces results that are difficult to accurately interpret. Although DNNs provide great utility in analyzing complex data sets, there is concern over the “black box” nature of DNNs, although new methods are being developed to provide explainability to DNNs [22]. Explainable algorithms will foster trust in AI by both clinicians and patients at the VA.

These new principles for the promotion of trustworthy AI build upon an existing framework developed in the VA, Ethical Principles for Access to and Use of Veteran Data [23], that safeguards veterans and their data and ensures that veterans benefit from research. In other words, research is not an end in and of itself—it is a means toward delivering value to veterans. Moreover, these principles are rooted in the legacy of the Belmont Report from 1979 [5], which emphasized privacy, beneficence, and justice in applications of technology. At its root, technology exists to help improve well-being, whether through heightened productivity or quality of the services provided. Together, these principles provide a signpost for clinicians and researchers to work collaboratively so that AI is

developed and deployed for social good, especially for vulnerable groups and veterans.

Moreover, these ethical principles developed and operationalized within the VA can be extended across the broader health care sector. For example, large university hospital systems that exist within the research ecosystem can adopt these ethical principles to guide their strategic investments and the development and deployment of AI tools. In fact, these university ecosystems have many similarities to the VA because they bring together a combination of researchers and clinicians under a common umbrella and institutional resources. Researchers and clinicians can work hand-in-hand to ensure that research and development investments are fundamentally driven by areas of great need and potential impact.

These processes for the development and application of ethical AI extend beyond veterans. In particular, members of any vulnerable group are beneficiaries of adherence to these processes because, by definition, they may find it harder to benefit from AI. For example, while AI is also leading to the invention of new jobs and tasks in the labor market, AI also

reduces the demand for other skills that are more routine and manual, which may affect veterans more if they are concentrated in those types of jobs and occupations. In this sense, applications of AI aimed at improving the transition of service members into the civilian sector could not only help veterans directly by, for example, providing them with tools to more efficiently match into jobs that suit their preferences and abilities, but could also improve trust and confidence in the benefits of AI. Moreover, other vulnerable groups likely face similar challenges; therefore, processes for the development and application of AI would help them too.

Our paper explains some of the most important ingredients for ensuring that AI advances are applied in ways that promote improved veteran outcomes. Furthermore, the VA could serve as a model organization, protecting VA patient data and leveraging it for their good and ultimately cutting health care costs, increasing efficiency, and enhancing health care for veterans. If the United States can successfully scale AI under a technology governance structure using these principles, the possibilities are limitless.

## Conflicts of Interest

None declared.

## References

1. Yu K, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018 Oct;2(10):719-731. [doi: [10.1038/s41551-018-0305-z](https://doi.org/10.1038/s41551-018-0305-z)] [Medline: [31015651](https://pubmed.ncbi.nlm.nih.gov/31015651/)]
2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019 Jan 7;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [Medline: [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/)]
3. Borkowski AA, Wilson CP, Borkowski SA, Thomas LB, Deland LA, Grewe SJ, et al. Comparing artificial intelligence platforms for histopathologic cancer diagnosis. *Fed Pract* 2019 Oct;36(10):456-463 [FREE Full text] [Medline: [31768096](https://pubmed.ncbi.nlm.nih.gov/31768096/)]
4. Overview of VA research on Health Equity. US Department of Veterans Affairs, Office of Research & Development. URL: [https://www.research.va.gov/topics/health\\_equity.cfm](https://www.research.va.gov/topics/health_equity.cfm) [accessed 2021-05-25]
5. The Belmont Report. US Department of Health and Human Services. 1979. URL: <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html> [accessed 2021-04-01]
6. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019 Jul 31;572(7767):116-119. [doi: [10.1038/s41586-019-1390-1](https://doi.org/10.1038/s41586-019-1390-1)]
7. Badal VD, Graham SA, Depp CA, Shinkawa K, Yamada Y, Palinkas LA, et al. Prediction of loneliness in older adults using natural language processing: exploring sex differences in speech. *Am J Geriatr Psychiatry* 2020 Sep 12:1-14 [FREE Full text] [doi: [10.1016/j.jagp.2020.09.009](https://doi.org/10.1016/j.jagp.2020.09.009)] [Medline: [33039266](https://pubmed.ncbi.nlm.nih.gov/33039266/)]
8. Fonseka TM, Bhat V, Kennedy SH. The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. *Aust N Z J Psychiatry* 2019 Oct 26;53(10):954-964. [doi: [10.1177/0004867419864428](https://doi.org/10.1177/0004867419864428)] [Medline: [31347389](https://pubmed.ncbi.nlm.nih.gov/31347389/)]
9. Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. Executive Office of the President of the United States. URL: <https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government> [accessed 2021-05-25]
10. Makridis C, Hirsch B. Labor Market Earnings of Veterans: Is Time in the Military More Valuable or Less than is Civilian Experience? Social Sciences Research Network Working Paper. 2019 Oct 18. URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3466518](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3466518) [accessed 2021-05-25]
11. Makridis C, Zhao D, Bejan C, Alterovitz G. Leveraging machine learning to characterize the role of socio-economic determinants on physical health and well-being among veterans. SSRN Preprint posted online on October 19, 2020. [doi: [10.2139/ssrn.3686845](https://doi.org/10.2139/ssrn.3686845)]
12. Riley GF, Lubitz JD. Long-term trends in Medicare payments in the last year of life. *Health Serv Res* 2010 Apr;45(2):565-576 [FREE Full text] [doi: [10.1111/j.1475-6773.2010.01082.x](https://doi.org/10.1111/j.1475-6773.2010.01082.x)] [Medline: [20148984](https://pubmed.ncbi.nlm.nih.gov/20148984/)]
13. Ahadi S, Zhou W, Schüssler-Fiorenza Rose SM, Sailani MR, Contrepolis K, Avina M, et al. Personal aging markers and ageotypes revealed by deep longitudinal profiling. *Nat Med* 2020 Jan 13;26(1):83-90. [doi: [10.1038/s41591-019-0719-5](https://doi.org/10.1038/s41591-019-0719-5)]



14. Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Health* 2020 Sep;2(9):e489-e492. [doi: [10.1016/s2589-7500\(20\)30186-2](https://doi.org/10.1016/s2589-7500(20)30186-2)]
15. Kaushal A, Altman R, Langlotz C. Geographic distribution of US cohorts used to train deep learning algorithms. *JAMA* 2020 Sep 22;324(12):1212-1213 [FREE Full text] [doi: [10.1001/jama.2020.12067](https://doi.org/10.1001/jama.2020.12067)] [Medline: [32960230](https://pubmed.ncbi.nlm.nih.gov/32960230/)]
16. Peterson R, Gundlapalli AV, Metraux S, Carter ME, Palmer M, Redd A, et al. Identifying Homelessness among Veterans Using VA Administrative Data: Opportunities to Expand Detection Criteria. *PLoS ONE* 2015 Jul 14;10(7):e0132664 [FREE Full text] [doi: [10.1371/journal.pone.0132664](https://doi.org/10.1371/journal.pone.0132664)]
17. Department of Veterans Affairs. 2021. Veterans Health Administration (VHA). URL: <https://www.va.gov/health/aboutvha.asp> [accessed 2021-01-04]
18. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine* 2018 Nov 01;178(11):1544-1547. [doi: [10.1001/jamainternmed.2018.3763](https://doi.org/10.1001/jamainternmed.2018.3763)]
19. Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA* 2019 Dec 24;322(24):2377-2378. [doi: [10.1001/jama.2019.18058](https://doi.org/10.1001/jama.2019.18058)] [Medline: [31755905](https://pubmed.ncbi.nlm.nih.gov/31755905/)]
20. Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. *Journal of Biomedical Informatics* 2021 Jan;113:103621. [doi: [10.1016/j.jbi.2020.103621](https://doi.org/10.1016/j.jbi.2020.103621)]
21. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Benetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inform Fusion* 2020 Jun;58:82-115. [doi: [10.1016/j.inffus.2019.12.012](https://doi.org/10.1016/j.inffus.2019.12.012)]
22. Xie N, Ras G, van Gerven M, Doran D. Explainable deep learning: a field guide for the uninitiated. *ArXiv Preprint* posted online on April 30, 2020 [FREE Full text]
23. The DOVA. Ethical principles for access to and use of veteran data. Department of Veterans Affairs. 2020. URL: <https://vawww.oit.va.gov/oit/office-technical-integration/ethical-data-use/> [accessed 2021-01-04]

## Abbreviations

- AI:** artificial intelligence  
**DNN:** deep neural network  
**EHR:** electronic health record  
**ML:** machine learning  
**VA:** Department of Veterans Affairs

*Edited by C Lovis; submitted 18.03.21; peer-reviewed by J Liew, A Amritphale; comments to author 03.04.21; revised version received 23.04.21; accepted 27.04.21; published 02.06.21*

*Please cite as:*

Makridis C, Hurley S, Klote M, Alterovitz G

*Ethical Applications of Artificial Intelligence: Evidence From Health Research on Veterans*

*JMIR Med Inform* 2021;9(6):e28921

URL: <https://medinform.jmir.org/2021/6/e28921>

doi: [10.2196/28921](https://doi.org/10.2196/28921)

PMID:

©Christos Makridis, Seth Hurley, Mary Klote, Gil Alterovitz. Originally published in *JMIR Medical Informatics* (<https://medinform.jmir.org>), 02.06.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.