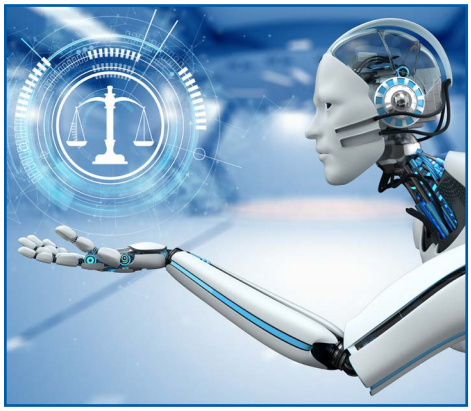
CATEGORY Operations and International Law

Text Excerpt from JAG Reporter Post: 27 Oct 2021 https://www.jagreporter.af.mil

Attorney's Guide to AI

Primer: A Practicing Attorney's Guide to Artificial Intelligence



BY MAJOR DAVID F. JACOBS AND CAPTAIN FLEMING E. KEEFE

Use of AI in both warfare and military administration is poised to increase dramatically, and a DoD that embraces AI and its potential will gain a strategic advantage over its competitors in the future.

Artificial intelligence (AI) can be defined as "the ability to perform tasks that normally require human intelligence."[1] This definition encompasses technology that has been around for nearly a century as well as decades-old technology already embedded throughout the Department of Defense (DoD) such as: aircraft autopilots, missile guidance, signal processing systems, and even our human resource systems.[2] While the emergence of AI may be aged, recent advancements in large data sets, increased computing power, improved machine learning algorithms, and open source code libraries have led to a considerable increase in real-world applications for AI.[3] These advances are already drastically revolutionizing our gadgets and lives towards a more AI-centric future.[4] A more AI-centric future can offer increased accuracy, increased capability, reduced human capital requirements, and a distinct advantage in future military operations. There is even evidence that AI can make us happier and healthier.[5] However, the promise of a more AI-centric future also brings unfamiliar threats driven by the speed of development and technological sophistications within the field of AI. The legal profession is poised to play a vital role in shaping how AI impacts our lives, the DoD, and society. But, before it is possible to know how Air Force and other DoD attorneys can succeed in such an endeavor, it is necessary to understand how AI works.

EXPAND YOUR KNOWLEDGE

External Link to Additional Resources

- Video: Artificial Intelligence (DVIDS)
  https://www.dvidshub.net/video/793739/artificial-intelligence
- U.S. and UK Research Labs Collaborate on Artificial Intelligence
  https://www.afrl.af.mil/News/Article-Display/Article/2816018/

3D illustration. Humanoid robot presenting legal scales.
(Illustration © iStock.com/style-photography)

MAJOR GROUPS OF ARTIFICIAL INTELLIGENCE SYSTEMS

Rule-Based Systems

Generally speaking, AI is split into two large groups, *rule-based (RB) systems* and *machine learning (ML) systems*, conditioned upon on how the machine "learns."[6] The first group is comprised of systems which learn from rule-based techniques; these machines are called *rule-based systems* or *handcrafted knowledge systems*.[7] RB systems learn through a process of reducing knowledge to if-then statements—known as a *rule set* or *rule sets*—whereby each rule obliges a specific output that is predetermined by the given input.[8] A human operator "teaches" the machine using traditional software programming.[9] A "classic" example of a RB system is International Business Machines Corporation's (IBM) Deep Blue® chess playing computer.[10] The RB Deep Blue® system bested reigning World Chess Champion Gary Kasparov in New York City on 11 May 1997.[11] This victory was the result of IBM's extensive collaboration with chess champions to develop if-then rule sets that the computer system would follow when countering a chess move made by a human player.

> A human operator "teaches" the machine using traditional software programming.

To put this in everyday terms, the concept of RB learning can also be seen in the use of e-mail inbox rules.[12] For example, an operator may teach a system to automatically move e-mails sent by airmansnuffy@us.af.mil from an inbox folder to a subfolder labeled "Office." More complex inputs by the operator illustrate multilayered rules, like when an operator "teaches" a system to forward e-mails received from airmansnuffy@us.af.mil *which also* contain the word "invoice" to a different e-mail address altogether (i.e., financeworkflow@us.af.mil.) Depending on a system's function, the rule set(s) will be more or less complex and can even be used in conjunction with the second major group, machine learning systems, to form an AI system.[13] As such, RB learning systems will continue to remain relevant for the DoD.[14]

> ML systems "learn" by interactions with a real environment, a simulated environment, and/or training data sets.

Machine Learning Systems

Machine learning systems include machines which learn through adaptive capabilities.[15] In contrast to RB systems which are human-programed and have fixed rule sets, ML systems "self-program" by creating rules which the system may later discard, modify, and/or create new rules—to varying degrees.[16] ML systems "learn" by interactions with a real environment, a simulated environment, and/or training data sets. In simple terms, the primary distinction between RB and ML systems is: when an AI system executes the same task(s) on the same data population, a RB system will have the same output every time but a ML system *should* produce a more efficient and effective output at each subsequent interval. To achieve more efficient and effective outputs, ML systems use a mathematical algorithm built with software code that gives varying values to data which is imputed into the system.[17] ML systems are further divided into four subsets, called *learning methods*, determined by how the algorithm handles data. The four learning methods—*supervised learning, unsupervised learning, semi-supervised learning,* and *reinforcement learning*—are differentiated by learning algorithm and input data characteristics, as discussed below.[18]

Supervised, Unsupervised, Semi-supervised, & Reinforcement Learning Methods

Supervised learning uses input data[19] known as *training data*. Training data is data which has been labeled, often by a human supervisor, to the correct data class.[20] This type of input training data is called *labeled data*.[21] In other words, supervised learning involves the process of identifying raw data (images, text files, videos, etc.) and adding one or more meaningful labels to provide context that is used by a software algorithm.[22] The goal of each learning method is to teach a machine a particular function and the human supervisor must keep that goal in mind if labeling input

training data. For example, if the goal is to identify F-16 aircraft from overhead imagery, the human supervisor should collect a sample group of aircraft photographs and assign the photographs to a particular class (F-1s, F-14s, F-15s, F-16s, F-22s, etc.*)*. When shown a new image, the model will predict the correct aircraft classification of the new image by comparing the new image to the training data.

Supervised learning systems tend to have higher performance levels than unsupervised systems; however, they are more time-intensive to build and require sizable training data sets.

In contrast, ***unsupervised learning*** uses unlabeled training data and assigns a particular data class based on detected patterns.[23] Unsupervised learning occurs most frequently when there is not enough expert knowledge to assign correct class labels, when the training data is so large that it is economically or temporally impractical to label the data, or when researchers are asking questions without already knowing the correct answer. Supervised learning systems tend to have higher performance levels than unsupervised systems; however, they are more time-intensive to build and require sizable training data sets.[24] *Semi-supervised learning*, as the name suggests, uses a combination of both labeled and unlabeled training input data.[25] Once trained, a ML system may use labeled or unlabeled data regardless of whether it is a supervised, semi-supervised, or unsupervised system.[26]

The last type of learning method is ***reinforcement learning***. Reinforcement learning uses feedback obtained through trial-and-error; whereby, a machine is tasked to make a decision (action), receives a reward or punishment (feedback) based on whether the action was consistent with the machine's predefined goal(s), and then applies feedback to influence subsequent decisions.[27] Reinforcement learning appears to be the most complicated learning method at first glance but it is easily demonstrated through a real-world example. When teaching a dog to sit (action) you provide the dog feedback based the action aligning, or not aligning, with your predefined goal (the dog sitting). If the dog sits, you give the dog a treat (positive feedback). If the dog does not sit, you do not give the dog a treat (negative feedback). The dog uses the earlier feedback (receiving a treat or not) to decide whether to comply the next time you ask the dog to sit. By repeating the action and feedback loop, the output accuracy should increase at each subsequent interval and, eventually, the dog *should* sit every time you ask.

It is worth noting that ***deep learning***, also known as *deep neural networks*, is a ML technique that can be applied to any of the abovementioned learning methods[28] but the technical details are beyond the scope of this article. While an in-depth discussion of deep learning and neural networks goes beyond the scope of this article, it is important to highlight a common AI technique that is frequently used to improve the software algorithm. *General adversarial models* (GANs) use two sub-models to train an AI system—*a generator* and a *discriminator*.[29] A generator produces a plausible example, such as an image of a fake person, and a discriminator compares the plausible example against a real image to determine which is real.[30] [31] The generator improves its algorithm model based upon its ability to trick the discriminator.[32]

LEGAL PRINCIPLES FOR ARTIFICIAL INTELLIGENCE

In general, an AI system consists of two components: the software that makes up an algorithm and the data that interacts with the software algorithm. Humans remain critical to the development and deployment of AI systems through choosing algorithms, formatting data, setting learning parameters, and troubleshooting problems.[33] Potential legal challenges relating to AI systems are numerous and daunting, so having a baseline understanding of AI systems is crucial to a successful legal review. In this regard, the ability to distinguish between RB and ML systems and various ML models is essential for attorneys practicing in this field. For one reason, a legal review of RB systems does not require addressing training data because RB systems do not use training data. Similarly, understanding the kind of ML model that is being used can help the legal practitioner

identify which types of data should be evaluated in the legal review. When examining an AI system, it is best to analyze the system from three distinct perspectives: first, an overarching view imposed by the DoD, called *AI ethics*; second, a view from the point of data, which interacts with the software algorithm to produce an output; and third, a view from the point of software, which is primarily comprised of the algorithm on which the AI system operates.

Ethical Principles

The DoD published five ethical principles for guiding the ethical development of combat and non-combat AI capabilities on 24 February 2020, as part of its efforts to be a leader in the fields of AI and AI regulation.[34] Those principles, listed below, encompass the general areas of responsibility, equity, traceability, reliability, and governability. However, ethical principles exist beyond those listed below, and legal reviewers must exercise due diligence and remain cognizant of the fact that there may be other controlling regulations, dependent on the customer or audience.[35]

> *1. Responsibility: DoD personnel will exercise appropriate levels of judgment and care, while remaining responsible for the development, deployment, and use of AI capabilities.*
>
> *2. Equity: The Department will take deliberate steps to minimize unintended bias in AI capabilities.*
>
> *3. Traceability: The Department's AI capabilities will be developed and deployed such that relevant personnel possess an appropriate understanding of the technology, development processes, and operational methods applicable to AI capabilities, including with transparent and auditable methodologies, data sources, and design procedure and documentation.*
>
> *4. Reliability: The Department's AI capabilities will have explicit, well-defined uses, and the safety, security, and effectiveness of such capabilities will be subject to testing and assurance within those defined uses across their entire life-cycles.*
>
> *5. Governability: The Department will design and engineer AI capabilities to fulfill their intended functions while possessing the ability to detect and avoid unintended consequences, and the ability to disengage or deactivate deployed systems that demonstrate unintended behavior.*[36]

On 26 May 2021, Deputy Secretary of Defense issued a memorandum affirming DoD's commitment to the DoD ethical principles and implementing responsible AI (RAI) in the DoD.[37] Despite the potential confusingly similar name to first ethical principle of AI ethics, RAI is DoD's implementing strategy for all five of the ethical principles.[38] The Joint Artificial Intelligence Center (JAIC) serves as DoD's coordinator for development and implementation of RAI strategy, guidance, and policy.[39] As of the date of this article, the JAIC has not published any official policy for interpreting AI ethical principles outside of the data strategy document and the RAI implementation memorandum; however, legal practitioners employed in the development or deployment of AI should check the JAIC website for updated material.[40] Until the DoD formalizes such policies, attorneys should, at a minimum, note in their legal reviews that the DoD AI ethical principles were considered prior to procurement, development, or deployment of an AI system.

Data Principles

As AI becomes increasingly more widespread, complex legal questions will naturally arise regarding acquisition, development, use, and ownership of the underlying data used to train or operate an AI system. Just as the underlying software is becoming more advanced, so too is the availability of data and the complexity associated with handling that data. Forbes reported that 2.8 quintillion bytes of data were created each day and over 90% of the world's data had been created over the preceding two years, at the time of the article in 2018.[41] Since 2018, the amount, type, and availability of data has only been increasing.[42] The DoD is starting to recognize the power that vast data can have in AI development and deployment.[43] To that end, the DoD has begun to focus on becoming a more data-centric organization that uses data at speed and scale for operational advantage and

increased efficiency.[44] The transformation of the DoD to a data-centric organization created the need to re-think the importance of data throughout the organization and acquisition life-cycle.[45] The result is that the federal government and the DoD now consider data a strategic asset.[46]

The legal practitioner should consider what is happening to the data at each of these stages with a general understanding of personnel with access, how the data will be used, and constitutional or other legal implications.

Three Distinct States

Data is a strategic asset that does not exist in a single state, but exists across three distinct states—*in use*, *at rest*, and *in transit*, also called *data in motion*. *Data at rest* is all data in computer storage that is not currently being accessed or transferred, *data in motion* is data that is moving or being transferred between locations within or between computer systems, and *data in use* is data that is currently being updated, processed, accessed and read by a system.[47] The legal practitioner should consider what is happening to the data at each of these stages with a general understanding of personnel with access, how the data will be used, and constitutional or other legal implications. This is particularly important in areas where there are restrictions, controls, or privacy implications to data access. For example, data containing personally identifiable information (PII)[48] may require: a privacy impact assessment,[49] system of records notice (SORN),[50] contractor approval for handling,[51] a public affairs review,[52] or constitutional considerations for how the data is being used.[53] Additionally, multiple contractor approvals may be necessary if different contractors handle the data at different states. For example, one contractor may work on storage and handling of the data and another contractor may work on handling the data when used to train an AI system.

In order for data to be usable for AI systems it must be properly formatted across all three states.[54] For the DoD, proper formatting means that the data is visible, accessible, understandable, linked, trustworthy, interoperable and secure across each state.[55] The Department of the Air Force, Chief Data Office, is responsible for the Air Force's policies and procedures for handling data.[56] As of the date of this article, the Chief Data Office has not published an official policy on how to satisfy DoD's formatting requirements, but education and training standards are already being implemented.[57] As such, legal practitioners should continue to monitor this area for new developments in standard licensing terms, formatting requirements, and other data-structuring procedures.

When evaluating an AI system, a legal practitioner must determine the underlying rights, if any, associated with the data.

Data-Use Rights

The DoD is directed to maximize data sharing and data-use rights.[58] In fact, ownership in technical data is essential for ensuring Department of the Air Force systems remain affordable and sustainable.[59] However, data suitable for AI training and use may carry various restrictions or terms which could limit the application of an AI system in its intended end state. Just because a Department of the Air Force unit has access to the data, does not necessarily mean the unit owns the data itself.[60] Failure to properly account for ownership and future use of underlying data can have drastic implications for usability of an AI system down the line.[61] When evaluating an AI system, a legal practitioner must determine the underlying rights, if any, associated with the data. For example, some licenses limit data use to educational or research purposes only.[62] Additionally, to ensure the data is ultimately usable, attorneys should be careful to clarify new rights, or changes to existing rights, if formatting data. If a contractor formats or makes changes to the data, associated licenses or contracts must identify what rights are attached to the formatted data. The role of a legal practitioner when examining data for AI

systems should be to ensure the Department of the Air Force is using data consistent with any terms or restrictions and that data acquired for AI enables the greatest flexibility well into the future.

Data Ethics

The DoD Data Strategy calls out *data ethics* as a legal consideration distinct and unique from AI ethics.[63] While the DoD Data Strategy does not provide a clear definition for data ethics, the federal data ethics framework defines *data ethics* as "norms of behavior that promote appropriate judgments and accountability when acquiring, managing, or using data, with the goals of protecting civil liberties, minimizing risks to individuals and society, and maximizing public good."[64] Therefore, a legal review assessing data which was acquired for, or is used by, an AI system should regard ethical implications of the data itself as a discrete consideration. While most ethical considerations relating to data are self-evident, such as civil liberties, some are not as readily apparent. One less obvious consideration involves actions that may qualify as human subject research.[65] Many AI systems utilize or analyze data containing PII, but such use may qualify as human subject research under applicable DoD regulations regardless of whether the data was training or operational data.[66] For example, surveillance cameras outside of a Base Exchange (BX) capture images of its customers and those images likely contain PII, or information which could be used to distinguish or trace the identities of those customers. If a ML system uses the BX live camera feeds, it may qualify as human subject research at two distinct times—as the machine trains on the data before operational use and as the machine adapts once it becomes operational.

Software Principles

Similar to data, ownership in software is essential for ensuring Department of the Air Force systems remain affordable and sustainable.[67] Legal practitioners must determine software ownership and applicable restrictions, if any. While this task may seem relatively straightforward, it can quickly become complicated if employing a software suite that contains software code incorporating several different license structures.[68] Aside from the complexities of multi-layered license structures, even so called "open-source" software sets may convey restrictions and terms on software use.[69] Addressing this early on can pay dividends for the command and mission in the long run.

While weapon systems are the obvious choice for legal review requirements, non-weapon systems may also violate policy and law at state or federal levels.

Review of the software is also where a legal practitioner should look to federal and state laws addressing how AI systems can and will be used. Department of Air Force attorneys may quickly jump to the weapons review process and considerations outlined in Department of Defense Directive (DoDD) 3000.09, *Autonomy in Weapon Systems*, and Air Force Instruction 51-401, *The Law of War*; however, those policies address legal considerations for some, but not all, of the uses of AI in a weapons system. Indeed, there currently exists a noted gap in DoDD 3000.09 for AI weapon system considerations[70] and an ongoing debate outside the scope of this article on how to address that gap. While weapon systems are the obvious choice for legal review requirements, non-weapon systems may also violate policy and law at state or federal levels. The ability of DoD-compliant AI systems to lawfully operate is not at all assured. For example, several states have laws restricting or banning the use of biometrics,[71] which would directly affect the feasibility of an AI facial recognition system for detecting intruders. Many of the data considerations will affect the software considerations and vice versa; however, a review should analyze considerations separately given that data and software may operate independently from each other in an AI system.

SUMMARY

The use of AI in both warfare and military administration is poised to increase dramatically, and a DoD that embraces AI and its potential will gain a strategic advantage over its competitors in the future. However, a number of challenges related to technology, policy, process, and data will continue to challenge those working in the dynamic field of AI. To address these challenges, the DoD published ethical principles and an implementation memorandum for responsible AI, but the absence of formal policies related to data formatting and data rights is another obstacle to the successful integration of AI in military applications. Collectively, these factors represent an enormous charge for legal practitioners. To help navigate this ever-changing field we established key concepts and provided a framework for legally examining AI from three viewpoints—data, software, and ethics. In evaluating data, an attorney must determine data ownership and whether there are restrictions, controls, or privacy implications all while considering how the data is being used and accessed against all three states. Similarly, an attorney must evaluate governing licenses and laws to determine whether there are restrictions on software and its use. Lastly, there is an obligation to ensure ethical use of data, software, and AI systems generally. As such, practicing attorneys must examine an AI system from three distinct views in order to ensure the system as a whole is legal.

Layout by Thomasa Huffstutler

ABOUT THE AUTHORS

Major David F. Jacobs, USAF

(B.S., Arizona State University; J.D., Stetson University College of Law, Patent Attorney, Certified Information Systems Security Professional) is currently assigned as the Staff Judge Advocate at USMTM, Riyadh, Saudi Arabia.

Captain Fleming E. Keefe, USAF

(B.S., Pennsylvania State University; M.P.S., Pennsylvania State University; J.D., University of South Carolina School of Law) is currently assigned as the Chief of Operations Law, Training, & Readiness for the 11th Wing, Joint Base Anacostia–Bolling, Air Force District of Washington, Washington, D.C.

ENDNOTES

[1] Greg Allen, Chief of Strategy and Communications, Joint Artificial Intelligence Center, Department of Defense, *Understanding AI Technology*, April 2020, available at: https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf [hereafter Allen, *Understanding AI Technology]*.

[2] *Id.* at Executive Summary.

[3] *Id.* at 8 - 9.

[4] Tom Simonite, *The Wired Guide to Artificial Intelligence*, Wired, 2 January 2018, available at: https://www.wired.com/story/guide-artificial-intelligence/.

[5] *Id.*

[6] *Rule Based System*, DeepAI, available at: https://deepai.org/machine-learning-glossary-and-terms/rule-based-system.

[7] Allen, *Understanding AI Technology.*

[8] Robert Smith, *The Key Differences Between Rule-Based AI and Machine Learning,* 14 July 2020, available at: https://becominghuman.ai/the-key-differences-between-rule-based-ai-and-machine-learning-8792e545e6.

[9] *Id.*

[10] Chris Higgins, *A Brief History of Deep Blue, IBM's Chess Computer,* Mental Floss, 29 July 2017, available at: https://www.mentalfloss.com/article/503178/brief-history-deep-blue-ibms-chess-computer.

[11] *Id.*

[12] *What is a Rule-based System? What is it Not?,* Think Automation, 21 May 2021, available at: https://www.thinkautomation.com/eli5/what-is-a-rule-based-system-what-is-it-not/.

[13] Catalytic, *Machine Learning vs. Rule-Based Systems, Explained,* Digital Transformation, 21 May 2021, available at: https://www.catalytic.com/blog/machine-learning-vs-rules-based-systems.

[14] Department of Defense, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*, 2018, available at: https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/summary-of-dod-ai-strategy.pdf.

[15] Allen, *Understanding AI Technology.*

[16] *Id.*

[17] *Id.* at 2 - 4.

[18] *Id.* at 11 - 15.

[19] Data can be defined as "the representation of information in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means, and is concerned with the encoding of information for repeatability, meaning, and proceduralized use." *See* Chief Information Officer, Department of Defense, *Glossary,* available at: https://dodcio.defense.gov/Library/DoD-Architecture-Framework/dodaf20_info_data/.

[20] International Journal of Advanced Research in Artificial Intelligence, Vol. 2, No. 2, 2013, available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.278.5274&rep=rep1&type=pdf#page=41 [hereafter IJAIA, Vol. 2, No. 2].

[21] Allen, *Understanding AI Technology* at 7.

[22] Amazon Web Services*, What is Data Labeling for Machine Learning?*, available at: https://aws.amazon.com/sagemaker/groundtruth/what-is-data-labeling/.

[23] IJAIA, Vol. 2, No. 2.

[24] *Id.* at 12.

[25] *Id.* at 14.

[26] Allen, *Understanding AI Technology* at 11 - 13.

[27] Huang (Steeve) Kung-Hsiang, *Introduction to Various Reinforcement Learning Algorithms,* 11 January 2018, available at: https://towardsdatascience.com/introduction-to-various-reinforcement-learning-algorithms-i-q-learning-sarsa-dqn-ddpg-72a5e0cb6287.

[28] *Id.* at 16.

[29] Jason Brownlee, *A Gentle Introduction to Generative Adversarial Networks (GANs),* 17 June 2019, available at: https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/ [hereafter Brownlee, *A Gentle Introduction to Generative Adversarial Networks (GANs)*.

[30] *Id.*

[31] An example of a GAN in action can be seen on the website https://www.whichfaceisreal.com/. Here, a computer system trained through a GAN presents an image of a fake person. The visitor to the website, working as a discriminator, selects which image is a photograph of a real person. Each selection by the human improves the accuracy.

[32] Brownlee, *A Gentle Introduction to Generative Adversarial Networks (GANs).*

[33] Allen, *Understanding AI Technology* at 3.

[34] Department of Defense, *DOD Adopts Ethical Principles for Artificial Intelligence,* 24 February 2020, available at: https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/ [hereafter DoD, *DOD Adopts Ethical Principles for Artificial Intelligence]*.

[35] Office of the Director of National Intelligence, *Principals of Artificial Intelligence Ethics for the Intelligence Community,* available at: https://www.dni.gov/files/ODNI/documents/Principles_of_AI_Ethics_for_the_Intelligence_Community.pdf (noting the Intelligence Community has its own set of AI Ethical Principles).

[36] DoD, *DOD Adopts Ethical Principles for Artificial Intelligence*.

[37] Deputy Secretary of Defense, *Implementing Responsible Artificial Intelligence in the Department of Defense*, 27 May 2021, available at: https://media.defense.gov/2021/may/27/2002730593/-1/-1/0/implementing-responsible-artificial-intelligence-in-the-department-of-defense.pdf.

[38] *Id.*

[39] *Id.*

[40] *See* JAIC website, available at: https://www.ai.mil/.

[41] Bernard Marr, *How Much Data Do We Create Everyday? The Mind-Blowing Stats Everyone Should Read,* Forbes, 21 May 2018, available at: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=145195f660ba.

[42] Jacquelyn Bulao, *How Much Data is Created Every Day in 2021?,* TECHJURY, updated 18 May 2021, available at: https://techjury.net/blog/how-much-data-is-created-every-day/#gref.

[43] C. Todd Lopez, *If DoD Wants AI in its Future, It Must Start Now, Officials Say,* DoD News, 23 March 2021, quoting Lt. Gen. Michael Groen "these enterprises are sitting on massive amounts of data. It's a natural target for AI implementation to create more efficiencies and economics and effectiveness in those large scale enterprises," available at: https://www.defense.gov/Explore/News/Article/Article/2547622/if-dod-wants-ai-in-its-future-it-must-start-now-official-says/.

[44] Department of Defense, *Executive Summary: DoD Data Strategy,* 20 September 2020, available at: https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DoD-data-strategy.pdf [hereafter DoD, *Executive Summary: DoD Data Strategy].*

[45] *Id.*; Deputy Secretary of Defense, *Memorandum for Senior Leadership – Creating a Data Advantage,* 5 May 2021, available at: https://media.defense.gov/2021/May/10/2002638551/-1/-1/0/deputy-secretary-of-defense-memorandum.pdf [hereafter Deputy Secretary of Defense, *Memorandum for Senior Leadership – Creating a Data Advantage].*

[46] *Id.*

[47] Nate Lord, *Data Protection: Data In transit vs. Data At Rest*, DIGITAL GUARDIAN, 15 July 2019, available at: https://digitalguardian.com/blog/data-protection-data-in-transit-vs-data-at-rest; TECHOPEDIA, *Dictionary: Data in Use,* available at: https://www.techopedia.com/definition/29515/data-in-use.

[48] As defined in DoDI 5400.11, *DoD Privacy Program.*

[49] Department of Defense Instruction 5400.16, *DoD Privacy Impact Assessment Guidance*, 15 July 2015, incorporating change 1, 11 August 2017, available at: https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/540016p.pdf.

[50] Defense Privacy, Civil Liberties, and Transparency Division, Department of Defense, *System of Records Notice (SORN),* available at: https://dpcld.defense.gov/Privacy/SORNs/.

[51] Office of Small Business Programs, Department of Defense, *Cybersecurity – What Small Businesses Need to Know*, viewed on 16 January 2021, available at: https://business.defense.gov/Small-Businesses/Cybersecurity/.

[52] Air Force Instruction 33-360, *Publications and Forms Management,* 1 December 2015, available at: https://static.e-publishing.af.mil/production/1/saf_aa/publication/dafi33-360/dafi33-360.pdf.

[53] 4th Amendment, United States Constitution.

[54] Gregory Vial, Jinglu Jiang, Tanya Giannelia, and Ann-Frances Cameron, *The Data Problem Stalling AI,* MIT SLOAN MANAGEMENT REVIEW, 8 December 2020, available at: https://sloanreview.mit.edu/article/the-data-problem-stalling-ai/.

[55] Deputy Secretary of Defense, *Memorandum for Senior Leadership – Creating a Data Advantage.*

[56] SSgt Rusty Frank, *AF Chief Data Officer: Data is the Future of the Force*, Secretary of the Air Force Public Affairs, 23 February 2018, available at: https://www.af.mil/News/Article-Display/Article/1448828/af-chief-data-officer-data-is-the-future-of-the-force/.

[57] Secretary of the Air Force Public Affairs, *DAF Chief Data Office Launches Unprecedented Data Governance Training, Certification Program,* 21 January 2021, available at: https://www.af.mil/News/Article-Display/Article/2478289/daf-chief-data-office-launches-unprecedented-data-governance-training-certifica/.

[58] Deputy Secretary of Defense, *Memorandum for Senior Leadership – Creating a Data Advantage.*

[59] Defense Acquisition University, *Air Force Data Rights Guidebook,* available at: https://www.dau.edu/tools/t/Air-Force-Data-Rights-Guidebook [hereafter Defense Acquisition University, *Air Force Data Rights Guidebook]*.

[60] Sandra Erwin, *Intellectual Property Fights Par for the Course in F-35 Program,* National Defense Magazine, 8 Sept. 2016, available at: https://www.nationaldefensemagazine.org/articles/2016/9/8/intellectual-property-fights-par-for-the-course-in-f-35-program.

[61] *Id.*

[62] University at Buffalo, The State University of New York, *IBM-UB Handwritten Database Sub-License Agreement,* available at: https://cubs.buffalo.edu/hwdata/license-agreement (an example of a data license which restricts use to research and educational purposes only).

[63] DoD, *Executive Summary: DoD Data Strategy.*

[64] General Services Administration, *Federal Data Strategy – Data Ethics Framework*, available at: https://resources.data.gov/assets/documents/fds-data-ethics-framework.pdf.

[65] 32 CFR 219, *Protection of Human Subjects;* DoDI 3216.02, *Protection of Human Subjects and Adherence to Ethical Standard in DoD-Conducted and Supported Research*, 15 April 2020.

[66] DoDI 3216.02.

[67] Defense Acquisition University, *Air Force Data Rights Guidebook*.

[68] Philip A. Albert, *Dual Licensing: Having Your Cake and Eating it Too,* LinuxInsider, 16 November 2004, available at: https://linuxinsider.com/story/dual-licensing-having-your-cake-and-eating-it-too-38172.html.

[69] Open Source Initiative, *Open Source Licenses by Category*, available at: https://opensource.org/licenses/category.

[70] Brenda Marie Rivers, *National Security Experts Cite Need to Revise DoD Directive on Autonomous Weapons Dev't,* 13 December 2019, available at: https://www.executivegov.com/2019/12/natl-security-experts-cite-need-to-revise-dod-directive-on-autonomous-weapons-devt/.

[71] Susan Crawford, *Facial Recognition Laws Are (Literally) All Over The Map*, Wired, 16 December 2019, available at: https://www.wired.com/story/facial-recognition-laws-are-literally-all-over-the-map/.