GEORGE MASON UNIVERSITY | Center for Government Contracting School of Business

White Paper Series

NO. 10 APRIL 29, 2022

Implementing Responsible AI: Proposed Framework for Data Licensing

Maj. Andrew Bowne and Mr. Benjamin McMartin

ABSTRACT

The Department of Defense (DoD) faces unique data licensing challenges when acquiring artificial intelligence (AI) solutions from the private sector. This is because in addition to standard licensing concerns (e.g., avoiding vendor-lock), acquisition professionals must also consider DoD's Responsible AI (RAI) principles.[1] The responsible use of AI requires license rights to data and software that are not covered by standard procurement clauses. This paper provides an alternative framework to the traditional data licensing strategies to better address the unique challenges of acquiring AI solutions.

In a traditional procurement, DoD identifies data and software to be delivered under the effort, which DoD will fund (at least partially), and thus will expect to obtain a license in that data and software. In this traditional context, the ultimate objective of negotiation is to avoid so-called "vendor lock" in the procurement and sustainment of the developed solution. The data (including software documentation) and software deliverables are typically designed to allow for the government's use in subsequent competitions for further development, production, and sustainment, or, in some situations, for organic testing and modification. The data deliverables are thus meant to provide the information necessary to produce an equivalent part, system, or application.

AI-based solutions present a challenge to this paradigm, however. Cutting-edge AI technologies are largely developed in the commercial sector at private.To achieve the Responsible AI principles, the government customer must understand how a technology works and how the results of the technology are derived, not simply how to produce more units. As a result, the types of data required to assure that delivered AI-enabled technologies are responsible, equitable, traceable, reliable, and governable[2] go beyond the delivery of end-solutions. This presents a challenge for companies attempting to enter the Defense market through commercial business models, standard seat or enterprise licenses, or "As–a–Service models." Providing source code, proprietary datasets, software architecture, or background IP to any entity, let alone the government, is far outside of the norm for these commercial AI firms.

Thus, to attract and integrate cutting-edge AI solutions, while adhering to Responsible AI principles, DoD—specifically DoD acquisition professionals—should implement a framework for assessing license needs, and craft custom use licenses that balance the need for access and use of data and software for RAI purposes against industry partners' desire to protect their IP and maintain standardized commercial licensing practices.

CONTEMPORARY HISTORY OF RESPONSIBLE AI STRATEGY, POLICY, AND GUIDANCE

In the past four years, there have been multiple national-level strategies, policies, and guidelines focused on responsibility, traceability, explainability, and ethics, as well as intellectual property in artificial intelligence.

Executive Order 13859 (February 2019)

Executive Order 13859, entitled "Maintaining American Leadership in Artificial Intelligence," established an American AI Initiative guided by five principles:

1. The United States must drive technological breakthroughs in AI across the federal government, industry, and academia in order to promote scientific discovery, economic competitiveness, and national security.
2. The United States must drive development of appropriate technical standards and reduce barriers to the safe testing and deployment of AI technologies in order to enable the creation of new AI-related industries and the adoption of AI by today's industries.
3. The United States must train current and future generations of American workers with the skills to develop and apply AI technologies to prepare them for today's economy and jobs of the future.

4. The United States must foster public trust and confidence in AI technologies and protect civil liberties, privacy, and American values in their application in order to fully realize the potential of AI technologies for the American people.
5. The United States must promote an international environment that supports American AI research and innovation and opens markets for American AI industries, while protecting our technological advantage in AI and protecting our critical AI technologies from acquisition by strategic competitors and adversarial nations.

In carrying out these principles the order identifies implementing agencies and a set of six strategic objectives for implementing agencies to pursue. Specific to issues of AI ethics and intellectual property, the objectives state that implementing agencies should:

1. Enhance access to high-quality and fully traceable federal data, models, and computing resources to increase the value of such resources for AI research and development, while maintaining safety, security, privacy, and confidentiality protections consistent with applicable laws and policies.
2. Reduce barriers to the use of AI technologies to promote their innovative application while protecting American technology, economic and national security, civil liberties, privacy, and values.
3. Ensure that technical standards minimize vulnerability to attacks from malicious actors and reflect federal priorities for innovation, public trust, and public confidence in systems that use AI technologies; and develop international standards to promote and protect those priorities.

Thus, the order introduces the objectives of traceability and reliability of federal data and models, AI applications, and technical standards as key components of the American AI Initiative.

National AI R&D Strategic Plan (June 2019 update)

The National AI R&D Strategic Plan Update, issued by the Select Committee on Artificial Intelligence of the National Science & Technology Council (hereafter, "the Committee"), identified eight strategic priorities to further the American AI Initiative. Those priorities included:

1. Make long-term investments in AI research
2. Develop effective methods for human-AI collaboration
3. Understand and address the ethical, legal, and societal implications of AI
4. Ensure the safety and security of AI systems
5. Develop shared public datasets and environments for AI training and testing
6. Measure and evaluate AI technologies through standards and benchmarks
7. Better understand the national AI research and development workforce needs
8. Expand public–private partnerships to accelerate advances in AI

Of primary importance for purposes of this paper are the committee's strategic priorities to understand and address the ethical, legal, and societal implications of AI and expand public–private partnerships to accelerate advances in AI.

As identified by the Committee, "maintaining American leadership in AI requires a concerted effort to promote advancements in technology and innovation, while protecting civil liberties, privacy, and American values. More R&D is needed to develop AI architectures that incorporate ethical, legal, and societal concerns through technical mechanisms such as transparency and explainability."[3]

National Defense Authorization Acts (FY20–FY22)[4]

In addition to executive action, Congress has been active in enacting AI-related legislation and programs. Of particular note, Section 256 of the Fiscal Year 2020 NDAA directed the secretary of Defense to develop a strategy for educating service members in relevant occupational fields on matters relating to artificial intelligence. Section 235 of the Fiscal Year 2021 NDAA directed the secretary of Defense to conduct an assessment to determine "whether DoD has the ability, requisite resourcing, and sufficient expertise to ensure that any artificial intelligence technology acquired by the department is ethically and responsibly developed; and (B) how the department can most effectively implement ethical artificial intelligence standards in acquisition processes and supply chains." Finally, Section 226 of the Fiscal Year 2022 NDAA required the secretary of Defense to review the potential application of artificial intelligence to a wide array of applications, including logistics and business applications, amongst others.

National Security Commission on AI **(Final Report, March 2021)**

Amongst numerous recommendations by the Commission, notable for the purposes of this paper, the Commission identified "If AI systems routinely do not work as designed or are unpredictable in ways that can have significant negative consequences, then leaders will not adopt them, operators will not use them, Congress will not fund them, and the American people will not support them. To establish justified confidence, the government should focus on ensuring that its AI systems are robust and reliable…."

The Commission further identified, "The United States lacks the comprehensive IP policies it needs for the AI era…."

Implementing Responsible AI in DoD **(May 2021)**

At the departmental policy level, the Department of Defense in 2021 issued a memorandum for senior pentagon leadership, "Implementing Responsible Artificial Intelligence in the Department of Defense." The memo reaffirms DoD AI ethical principles issued just over a year prior and places the principles under the umbrella of "Responsible Artificial Intelligence (RAI)." The memo identifies that the five principles: responsible, equitable, traceable, reliable, and governable, should be applied holistically across the department, including the AI product and acquisition life-cycle, with the ultimate objective of implementing RAI across the department at scale.

Responsible AI Guidelines in Practice: Lessons Learned From the DIU Portfolio **(November 2021)**[5]

As part of its mission to accelerate the adoption of commercial technology within DoD, the Defense Innovation Unit (DIU) launched a strategic initiative in March 2020 to integrate DoD's Ethical Principles for AI into its commercial prototyping and acquisition programs. Drawing upon best practices from government, non-profit, academic, and industry partners, DIU explored methods for implementing these principles in several of its AI prototype projects. The result is a set of responsible artificial intelligence (RAI) guidelines. In developing these guidelines, DIU identified that while the DIU RAI guidelines are a useful starting point for operationalizing DoD's ethical principles for AI, DIU will continue collaborating with experts and stakeholders from government, industry, academia, and civil society to further develop the RAI guidelines. Through this paper's exploration of the practical intersection of RAI and intellectual property, we hope to further add to this exemplary body of work.

Department of the Air Force/Massachusetts Institute of Technology Artificial Intelligence Accelerator - Artificial Intelligence Acquisition Guidebook **(February 2022)**

In February 2022, the Air Force/MIT AI Accelerator released an AI acquisition guidebook as a starting point, or "catalyst" for discussion around topics related to the intersection of artificial intelligence and defense acquisition. The guidebook generally covers contracting strategies for research and development, prototyping and beyond for AI technologies, along with a general description of the common treatment of software and data deliverables under the DFARS data rights scheme, and recognition of the RAI principles. Similar to the DIU RAI guidelines, the AI acquisition guidebook serves as a catalyst for further examination, which we hope to supplement with the subject matter of this paper.

OWNERSHIP, TITLE, AND RIGHTS IN SOFTWARE AND TECHNICAL DATA

A common misperception of government acquisition professionals is that the government is best served by "owning" or "obtaining title" to software and technical data under government-funded development efforts. Sometimes this is conveyed as a requirement to "give the government its rights" in software or technical data as a condition for the award of federal dollars. This leads to a misperception on the side of commercial entities that they will forfeit title in their intellectual property to the government if they choose to accept development dollars. This idea is not only in error but is contrary to the very policies which established the current legal and regulatory framework for the treatment of intellectual property and data rights. The intent of the IP scheme can be summarized as follows:

> The government typically relies on contractors (or, more broadly, the private sector) to commercialize innovations, even if they are 100 percent government funded. And typically, the government does not require ownership of the intellectual property, but instead relies on licenses from the contractors to meet the agency requirements. The importance of contractor ownership cannot be overstated, because it allows contractors to develop and reuse government-funded technology, thus applying government-funded, contractor-owned IP to create or support commercial markets. Government ownership would, conversely, unnecessarily inhibit such development.[6]

As explained throughout this paper, the current legal and regulatory scheme is a license-based approach, designed to promote industry commercialization of government-funded research efforts. As it relates to cutting-edge technologies that are largely developed in the commercial sector at private expense, such as AI-based solutions, it is imperative that acquisition professionals understand the policy and structure of the IP and data rights legal and regulatory framework. Speaking in terms of government "ownership" or "obtaining title" is sure to have a chilling effect on companies evaluating whether to work with DoD. It is equally important that technology firms understand that the government will seek licenses in their software and technical data to meet government use cases. Such licenses may be negotiated and valued to meet the needs of the parties.

OVERVIEW OF THE CURRENT DFARS LICENSING SCHEME

The current data and software licensing scheme for DoD, as embodied within the Defense Federal Acquisition Regulation Supplement (DFARS), is a scheme based upon the principle that DoD should only require the technical data and computer software, and rights in that data and software, necessary to satisfy its needs.[7] The scope of the license rights is typically determined by the source of funds used to develop the technical data or software.[8] Technology developed under government funding provides the government with broader rights than technology developed with mixed funding or at private expense. The government can use DFARS clauses found in DFARS 252.227 that provide license rights based on this funding scheme, or it can negotiate special terms if required to meet its needs.[9] DoD IP acquisition and licensing policy supports the overarching DFARS licensing scheme as it recognizes the importance of balancing DoD and industry interests.[10]

DELIVERABLES, DELIVERABLES, DELIVERABLES

The DFARS licensing scheme is based upon the funding characteristics (fully-government funded; partially-government funded; or developed exclusively at private expense) of *deliverable* computer software and technical data. This presents a complication for acquisition personnel procuring AI-based technologies, as software and data deliverables (for non-AI-based solutions) are typically limited to the deliverables required to test, deploy, and use the solution for its intended purpose.

When considering the requirements of RAI, acquisition personnel must expand their required deliverables to include software and technical data required to determine that the delivered solution is responsible, equitable, traceable, reliable, and governable. Depending on the technology procured, this may require delivery of background IP and data sets, training data, input/output data, object code, software architectures, and source code, all of which may have been developed exclusively at private expense. For companies with standardized commercial technologies, terms and conditions, and licensing structures, delivery of such software and technical data may be a showstopper. As explained below, acquisition personnel must consider (and be capable of developing) custom use licenses in terms of purpose, scope, and time to assure that procured solutions meet RAI principles, while assuring that the government is able to procure such technologies, and further that the government is not procuring license terms for which it does not have an intended use.

RAI IN DOD OTHER TRANSACTION AUTHORITIES

When using Other Transaction Authorities,[11] DoD acquisition professionals are exempt from the requirements of Bayh-Dole (35 U.S.C. §§ 201-204) for patents, and 10 U.S.C. §§ 3771-3772[12] (including the applicable FAR and DFARS implementation of these statutes) for rights in software and technical data. This exemption provides maximum flexibility for negotiating custom licenses that meet the short-term and long-term needs of the parties to the transaction. DoD other transaction authorites guidance provides:

> In negotiating [intellectual property and license rights in computer software and technical data] under an Other Transaction, it is a best practice for the government and solution provider to identify business plans for the subject technology at 1-year, 3-years, 5-years, and beyond. By establishing the short-term and long-term needs of the parties, a tailored IP scheme can more easily be determined and factored into the government's [overall] IP negotiation strategy.
>
> The negotiated IP terms and conditions should facilitate all parties' business plans and project goals, including any likely production and follow-on support of the prototype developed, and balance the relative investments and risks borne by the parties both in past development of the technology and in future development and maintenance of the technology.[13]

In relation to the development of AI-based technology solutions under DoD other transactions, the principles of Responsible AI still apply, and acquisition personnel should identify the license rights and use terms necessary to assure the RAI principles are met. In this context, "future development and maintenance" includes the ability to assure that development technologies are responsible, equitable, traceable, reliable, and governable. Practitioners should leverage the flexibilities of the Other Transaction Authorities to craft license terms that support RAI requirements.

THE ROLE OF DATA AND SOFTWARE LICENSING IN ASSURING RESPONSIBLE AI

Within the field of AI and the acquisition of AI-based technologies from the private sector, there are critical data and software licensing needs that are not present in typical software or hardware procurement. These requirements emerge as DoD must

be able to access, modify, disclose or otherwise use data to operate the AI application and measure outcomes to achieve the principles of Responsible AI. While the DFARS licensing scheme was created for hardware and services (and later added traditional software programs), the design, development, deployment, and use of AI tools and capabilities, especially those that employ machine learning, may require license rights into data and software that are not typically required by the government.[14] Babak Siavoshy, general counsel at Anduril Industries, asked what type of license terms should apply to an AI algorithm privately developed for computer-vision object detection and adapt it for military targeting or threat-evaluation?[15] Neither commercial software licenses nor standard DFARS data rights clauses adequately answer this question as neither appropriately protects the developer's interest or enable the government to gain the insight into the system to deploy it responsibly. Siavoshy argues that commercial AI developers invest in their technology and should be justly compensated without concern that the government will reverse engineer or share their code with competitors.[16] The government can obtain necessary rights in data and software to implement the Responsible AI principles without requiring rights to reverse engineer or share proprietary IP.

AI solutions involve numerous components working as a system to provide capabilities to assist the end-user in making data-driven decisions faster.[17] The canonical architecture for AI includes data, algorithms, computing, and governance of this system.[18] Understanding the end-to-end pipeline of AI solutions is necessary to acquire the data and license rights required to meet DoD's requirements to operate and sustain the capability, as well as meet its legal obligations and implement the Responsible AI principles.

Data is a foundational component of AI. Thus, large collections of curated datasets are valued by the government and industry alike.[19] The AI architecture starts with data collection and curation, often from multiple sources—government-owned, the contractor's proprietary data, third-party licensed data, open source or a combination thereof. As the National Security Commission on Artificial Intelligence (NSCAI) warned, "the absence of data governance policies (such as contracting best practices) for IP-type protections or ownership rules could undermine the willingness of companies to enter into the public–private partnerships that are crucial for creating cutting-edge technological innovations."[20] The collection of data, whether the data contains metadata, and the cleaning and labeling of data can affect its suitability for modeling and can impact the degree of accuracy and bias of the model.[21]

The conditioned data is then fed into algorithms that convert the input information into model output that is actionable knowledge represented in a form usable by humans.[22] These algorithms can be trained by various techniques, though predominantly through supervised learning (pre-labeled data for input and output), unsupervised learning (unlabeled data), or reinforcement learning (training through reward signals).[23] The model output can be used to make decisions, predictions, relate inputs and outputs, or take actions autonomously.[24] However, many AI applications relevant to DoD require human judgment, thus, the human-machine teaming phase of the pipeline connects the data and algorithms to the end-user.[25] Underlying and enabling this process is computing technology.[26]

While the data–algorithm–human–machine interface pipeline powered by computing hardware creates AI applications, trust in the output is critical to DoD's mission. This concept is known as *robust* or *trusted AI*.[27] In some AI systems, such as decision trees or logistic regression algorithms, humans can explain the way the decision was made and answer why an output was produced for a given input.[28] However, many AI systems today, and indeed the most advanced systems that are relevant to DoD, use machine learning and deep learning by neural networks that make it impossible for any human to fully understand and explain the decision-making by the model.[29] These complex models are often referred to as a "black box" due to the challenge of interpreting and explaining how the model functioned for a given decision, though the ability to explain the model ensures trust.[30]

Some jurisdictions require systems to provide some level of explanation such as the European General Data Protection Regulation (GDPR).[31] For the end-user, a commander, operator, or analyst in DoD context, to trust the algorithm's output, and for the public to trust DoD's use of AI, the output should be explainable (i.e., why an AI algorithm recommends a particular course of action), verifiable and validated, secure, safe, ethical, and responsible.[32] These concepts are reflected in DoD AI Ethical Principles[33] and reaffirmed by the deputy secretary of Defense implementing guidance by the collective concept of "Responsible AI" (RAI)."[34] Implementation of RAI is in accordance with the following tenets:

- RAI Governance
- Warfighter Trust
- AI Product and Acquisition Life-cycle
- Requirements Validation
- Responsible AI Ecosystem
- AI Workforce[35]

The AI pipeline may span multiple programs, contracts, and vendors.

Practically, license rights may be necessary to use, share, display, modify or otherwise access and practice the data or software to ensure interoperability across multiple components in the pipeline; moreover, license rights may be required to implement the RAI principles. Thus, DoD must understand each component in the pipeline and negotiate rights in data and software that permit the operational use of such data and software in alignment with DoD's RAI tenets.

DFARS LICENSING AND IMPLEMENTATION OF RAI

The traditional DFARS licensing scheme does not clearly cover all considerations for AI systems. A traditional DFARS license that defaults to the source of funding for development makes little sense practically in developing a machine learning model that learns and reprograms its software code through the ingestion of training data. For an algorithm developed at private expense, using government-owned data to create a trained model, the DFARS would provide the government with restricted rights in the algorithm, unlimited rights in the data and model. However, if the data used to train the model was proprietary, or third-party licensed, or even open source (as many training datasets are), the model would arguably be developed by costs not allocated to a government contract,[36] and the government would be entitled only to restricted rights in the model predictions.[37]

The government will often require at least government purpose rights and may require unlimited rights in the model predictions as the intent of modeling data is for the user to use, perform, display, reproduce, modify, and release the data downstream to other government users and contractors. Formatted output data that is labeled and machine-readable can be used by other organizations and even used as training data on other models; thus, obtaining sufficient license rights to the output data is valuable to the government. Such rights to that data may also be necessary to conduct regular audits, testing, and verification that the model is functioning as intended.

Further, Specially Negotiated License Rights under the DFARS may limit the ability of the government to attract best-in-class talent and products from industry. Industry concerns about the government treatment of IP is one of the primary reasons companies decide to avoid working with the government.[38] Although negotiating license rights can address some concerns from potential contractors, the government is limited in what rights it can relinquish in the DFARS licensing scheme. The DoD is prohibited from negotiating for less rights in computer software or technical data than what it would be entitled to receive based on the funding scheme.[39] This statutory restriction serves as a limit on the government's ability to conduct tradeoffs in negotiation that could reduce the cost of the contract (as it relinquishes any rights it does not require), attract non-traditional defense contractors that are unwilling to give unlimited rights in their IP, and meet its requirements.

Moreover, the DFARS license rights definitions do not clearly include some components of the AI pipeline. For example, "data," the foundation of AI, may consist of information that is contemplated by the definition of "technical data" under DFARS 252.227-7015, though it may include data that is excluded by that definition. Yet raw data may be relevant to a use case and DoD would need to obtain rights to use that data and order said data as a deliverable. Obtaining a specially negotiated license rights to access, use, reproduce, release, or disclose components of the AI system may require a class deviation under DFARS 201.402, adding months to the process (there is no such requirement when using other transaction authority).

Finally, the never finished nature of a machine learning model challenges the fundamental definition of the DFARS licensing scheme developed. When the computer program or software is developed and by which entity's expense are threshold questions used to determine the allocation of rights under the DFARS. An algorithm may be developed at private expense, but when the algorithm is trained pursuant to a government contract, the model's programming is altered as directed by the government. There may be room to argue the DFARS standards license can be interpreted to include these apparent idiosyncrasies. Nonetheless, a claim against DoD for an ambiguous clause could prove extremely costly to the government.

Acquisition professionals need the flexibility to iterate the model and conduct frequent testing: designing, developing, deploying, and operating a model will likely require numerous training and testing iterations, user feedback, and algorithmic refinements throughout the life-cycle. License rights in the data, algorithm, output, computing, and interface components may be necessary to implement the AI system consistent with the RAI principles. Thus, the contract's statement of objectives, modification clause, and contract line items should permit the necessary flexibility to iterate.

PROPOSED FRAMEWORK FOR DATA LICENSING TO IMPLEMENT RAI

This section provides examples of common use cases and the authors' proposed

framework for developing data licensing terms to implement DoD's ethical principles for RAI. Across each ethical principle and inherent in the concept of RAI is ensuring the use of the AI system is lawful.[40] However, the law is the baseline and RAI requires higher levels of responsibility, trust, governance, and ethics to meet U.S. policy and strategic objectives. In offering a practical approach to implementing a licensing scheme that balances the needs of the parties, while assuring RAI principles are addressed, we are proposing the following method, which should be applied to each of the five principles:

1. Identify the specific use case or cases the government is seeking to address;
2. Identify *what* types of information, data, and software are required to measure whether the principle has been met for the specific use case;
3. Identify *how* the identified information, data, and software will be used to measure whether the principle has been met for the specific use case; and
4. Identify *when* the information, data, and software will be required to measure whether the principle has been met for the specific use case, and when the need will expire.

The considerations offered below are non-exhaustive and each use case and purpose will have additional mission and business considerations that impact the negotiation of license rights.

Principle 1 – Responsible

DoD personnel will exercise appropriate levels of judgment and care, while remaining responsible for the development, deployment, and use of AI capabilities.

Use Case Examples

Target classification; performance prediction using physiological and cognitive metrics; autonomous driving.

Data Licensing Considerations

Information Required to Assure Principle is Achieved:

- Input data to ensure data collection on persons is used appropriately and protects the privacy of those persons the data is collected.
- Model output to conduct continuous testing and verification that the model is performing as designed.

How the Information Is Used to Assure Principle Is Achieved

When collecting and using human subject data, there may be a legal requirement to obtain human subject research approval and publication of system of records notice (SORN).[41] If the model leads to decisions by a human operator or an autonomous system, the information is required to trust that the system is producing reliable, trustworthy predictions. Data necessary to audit the decision-making process of the model, such as logs and any local interpretable model-agnostic explanations (LIME) system models that build an interpretable approximation of the neural network, can verify the model functions and establish due diligence and responsibility in relying on the model under the circumstances present at the time of the decision to act. The primary question that must be answered is who is responsible for the model? While decisions and actions may be executed by the AI system, humans remain responsible under the law. The context and use case may leave the government responsible for unintended consequences regardless of how the contract assigns or indemnifies such acts. Thus, the license rights should cover any data or software necessary for the responsible party to operate with due care and judgment.

When the Information Is Needed and in What Format to Assure Principle Is Achieved

The input data and model output should be provided in a machine- and human-readable format throughout the life-cycle to the government and third-party contractors assisting in the operation, maintenance, or TEVV (Test, Evaluation, Validation & Verification) processes. The license to this data and release to third parties can be restricted to auditing, evaluation, validation, and verification purposes, unless the government is entitled to broader rights.

Principle 2 – Equitable

The Department will take deliberate steps to minimize unintended bias in AI capabilities.

Use Case Examples

Human Resources: From 2014 to 2015, Amazon used historical data from the last 10 years to train an AI model on job applicants. The system automatically screened unqualified applicants. The AI system incorrectly learned that male candidates were preferable and used gender as a screening criteria.[42] Equitable AI establishes rigorous processes to mitigate unintended bias and harm on a class of individuals.

Data Licensing Considerations

Information Required to Assure Principle is Achieved:

- Data used to train the algorithm and provenance of the training data
- Trainable and non-trainable parameters (filters, weights and biases) of the model
- Model output
- Any procedures, such as perfor-

mance tracking over time, used to monitor that the model does not drift during operation or yields biased results.[43]

How the Information Is Used to Assure Principle Is Achieved

Equitable AI starts at data collection. Ensuring the training data is representative of what capability the model is intended to enable is critical to avoiding unfair biases. Care must be taken to avoid historic biases that are non-inclusive of women, minorities, or other underrepresented classes. Mitigating unintended biases can lead to decreased model accuracy,[44] so collaboration between the government and contractor is necessary to understand the trade-off equitable AI may have on reliability.

When the Information Is Needed and in What Format to Assure Principle Is Achieved

If training data or pre-trained models are supplied by the contractor or third party, the government should have access to such information at the planning phase of the project. Access to the weights and balances for government or third-party auditing should start at the development phase and continue throughout the life-cycle to mitigate unintended biases.

Principle 3 – Traceable

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.

Use Case Examples

Autonomous vehicles: In 2018, the first pedestrian was killed by an autonomous vehicle. The operator was an Uber driver that had the Uber driving system in full control until two-hundreds of a second before impact when the driver took the car out of autonomy and into manual mode, but too late to avoid the collision. In the investigation, law enforcement and the National Transportation Safety Board reviewed the model to forensically analyze the event, piecing out each fraction of a second prior to the crash. The question in the investigation, and potential negligent homicide trial, is whether Uber made responsible decisions and exercised judgment throughout the development and deployment of the autonomous driving system.[45] Ultimately, the decisions made by the relevant stakeholders in the development and deployment, and throughout the use of the system, should be auditable.

Data Licensing Considerations

Information Required to Assure Principle is Achieved:

- Logs documenting processes and decisions made by the AI system
- When the AI system includes human-machine interfacing, with some decisions made by the AI system and others made by a human responsible party, the chain of events should be documented[46]
- Data sources

How the Information Is Used to Assure Principle Is Achieved

The question that must be answered for traceable AI is how are the actions of the model recorded? Additionally, can the responsible party (DoD) audit the AI system's actions to understand how that action occurred? Records and processes must be accessible and explainable. The government or third-party auditor should be able to go back to the steps in the decision process to understand why the outcome occurred and identify lessons learned.[47] In some occasions, the processes and records of decision may need to be disclosed to the public. Additionally, because the operation of the AI system is dependent on the data, the provenance of data sources, and even the motivation, composition, and the collection of the data, is required to understand the quality and relevance of the input.[48]

When the Information Is Needed and in What Format to Assure Principle Is Achieved

The information should be accessible by the government during the development and deployment phases of the AI system's life-cycle. In the event of a mishap, the data may be released to regulators, investigations, and the public. The ability to trace back to how decisions were made should demonstrate the values of the stakeholders involved in the development and deployment of the model.

Principle 4 – Reliable

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.

Use Case Examples

Deepfakes and adversarial perturbations: Deepfakes use machine learning algorithms called Generative Adversarial Networks (GANs) that create uncannily realistic fake videos and images.[49] These have been used for many purposes though concerningly by Russia in the early weeks of the war on Ukraine, broadcasting a video depicting Ukrainian President Volodymyr Zelenskyy surrendering Ukraine to Russia.[50] Similarly, models can become corrupted when adding an imperceptible perturbation to the

input image. These are called adversarial examples and can lead to a model to make wrong predictions with very high confidence yet are difficult for humans to detect.[51] Deepfakes and adversarial perturbations that are inserted into training data can make the model unreliable. Model training processes must be evaluated in a reproducible manner and robustly documented and organized to assess reliability of the deployed model.[52]

Data Licensing Considerations

Information Required to Assure Principle is Achieved:

- Training and testing data
- Weights and biases of the model
- Processes, such as robustness tools, and metrics

How the Information Is Used to Assure Principle Is Achieved

Test and Evaluation, Verification and Validation (TEVV) of the model performance. The question that must be answered for reliable AI is whether the model can be trusted. Access to performance data and measurements against performance metrics may be needed to assess the reliability of the model. This information may be shared with third party contractors, though the use of such information may be restricted to TEVV purposes.

When the Information Is Needed and in What Format to Assure Principle Is Achieved

Before deployment and throughout use of the model. Perturbations can occur at any time and can affect the accuracy of the model. Robust processes to assess performance are needed to ensure the model is trustworthy.

Principle 5 – Governable

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.

Use Case Examples:

Criminal justice: Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is an algorithm used to predict recidivism in convicted criminals during sentencing.[53] The statistical results of the algorithm predict black defendants pose a higher risk of reoffending than a true representation.[54] A model that is governable functions as intended and has robust processes that detect and avoid unintended consequences.

Data Licensing Considerations

Information Required to Assure Principle is Achieved:

- Input dataset (data source, raw data, missing values, incorrect labels, appropriate anonymization, representative, relation of labels to intended predictions in supervised models)
- Trainable and non-trainable parameters (filters, weights and biases) of the model

How the Information Is Used to Assure Principle Is Achieved

Understanding the input data and the weights and biases of the model can mitigate unintended consequences of acting on the model's recommendations or predictions. The question that must be answered for Governable AI is how is the model controlled? Data necessary to answer that question may be spread across multiple components of the AI system and may require sharing data with other contractors within that system. Accordingly, failure to obtain the appropriate license rights on one contract that covers a discrete function in the AI system may lead to a failure in the governance or trustworthiness of the entire system.

When the Information Is Needed and in What Format to Assure Principle Is Achieved

For purposes of AI governance, the license required is not intended for competition but verifying the model is performing as intended. An acceptable use agreement and license to access, display, and share data within the government and third parties (i.e., independent TEVV contractors or vendors of other components of the system) can permit DoD to mitigate unintended consequences without diluting the value of the contractor's IP.

CONCLUSION

Implementation of responsible artificial intelligence principles, when acquiring AI solutions from private industry, necessitates a relook at Defense federal acquisition regulations licensing structures, and requires acquisition professionals to consider strategies to balance RAI principles with the government's ability to attract and engage with commercial AI firms.

Acquisition personnel developing/procuring AI-based solutions must plan for software and data (including software documentation) deliverables necessary to support the RAI requirements, which will frequently require specially negotiated licenses that properly identify and limit their use for a specific purpose.

- These specially negotiated licenses are in the interest of both industry and government.
- These licenses permit the government to access and use the data required to assure RAI principles are met and not

overpay for rights without a valid use case.

- Additionally, such limited use affords industry partners confidence that data use will be restricted to valid government uses.

Ultimately, it is a balancing act that is needed from learned professionals to achieve adoption and integration of AI into the Department of Defense, one that we believe may be furthered through application of our proposed framework.

Disclaimer: The views expressed are those of the authors and do not reflect the official guidance or position of the United States government, the Department of Defense or of the United States Air Force.

Statement from DoD: The appearance of external hyperlinks does not constitute endorsement by the United States Department of Defense (DoD) of the linked websites, or the information, products, or services contained therein. DoD does not exercise any editorial, security, or other control over the information you may find at these locations.

ENDNOTES

1. "Ethical Principles for Artificial Intelligence." United States Department of Defense. https://www.ai.mil/docs/ethical_principles_for_artificial_intelligence.pdf.

2. Ibid.

3. National Artificial Intelligence R&D Strategic Plan, June 21, 2019, 20. https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf.

4. See also, FY 22 NDAA §§ 226-228, 232, 247; FY 21 NDAA §§ 234, 235, 241; FY 20 NDAA §§ 230, 232, 256

5. "Responsible AI Guidelines in Practice," Defense Innovation Unit, (November, 2021), available at: 2021_RAI_Report-v3.pdf (ctfassets.net)

6. James G. McEwen, David S. Bloch, Richard M. Gray, and John T. Lucas, "IP and Technology in Government Contracts: Procurement and Partnering at the Federal and State Level," (LexisNexis, 2019), 17.

7. See DFARS 227.7103-1(a), 227.7203-1(a).

8. See DFARS 227.7103-4(a), 227.7203-4(a).

9. For more background on the DoD's IP policy, see DoD Instruction (DODI) 5010.44, Intellectual Property Acquisition and Licensing (October 16, 2019).

10. DODI 5010.44, 4.

11. 10 U.S.C. §§ 4021-4023 (renumbered from 10 U.S.C. §§ 2371, 2371b, 2373 in the FY22 NDAA).

12. Renumbered from 10 U.S.C. §§ 2320-2321 in the FY22 NDAA.

13. "Other Transactions Guide," Office of the Undersecretary of Defense for Acquisition and Sustainment, Version 1.0, 50 (November 2018).

14. Babak Siavoshy, "The DoD Should Pilot a New Category of Software Data Rights," Anduril Blog (March 2, 2022), explaining how the DFARS data rights scheme is "bizarre if seen in the context of modern software industry practices, as they treat software like a physical widget rather than code that can be infinitely copied and decompiled at almost no cost…. However, the primary focus of almost every software license is to dictate appropriate uses of the software, to prevent reverse engineering, modification, or sharing of the core software code, and to protect the developer's hard work—i.e., the precise things that restricted rights do not regulate." https://blog.anduril.com/the-dod-should-pilot-a-new-category-of-software-data-rights-a949cc9aaae4.

15. Ibid.

16. Ibid.

17. Vijay Gadepally et al., *AI Enabling Technologies: A Survey*, Lincoln Laboratory (Lexington, MA: Massachusetts Institute of Technology, April 2019), 1. AI is a broad field of study that evades precise definitions, though is often described as the ability of computers to perform tasks typically requiring human intelligence. Within AI, there are many subsets of methodologies, including expert and knowledge-based systems; however, the state of the art and most relevant AI discipline to DoD is machine learning and its subset of deep learning through neural networks. As discussed in this paper, machine learning presents complications to the traditional IP licensing framework and requires careful planning and clear communication with industry to implement Responsible AI in the design, development, deployment, and operation phases of the AI life-cycle. https://vijayg.mit.edu/sites/default/files/images/enablingtechnologies_042319.pdf.

18. Ibid., 1–2.

19. "Public Views on Artificial Intelligence and Intellectual Property Policy," US Patent and Trademark Office (October 2020), 15. https://www.uspto.gov/sites/default/files/documents/uspto_ai-report_2020-10-07.pdf.

20. National Security Commission on Artificial Intelligence (NSCAI), Final Report (2021), 205. https://www.nscai.gov/2021-final-report/.

21. Gadepally, 2.

22. Ibid.

23. Ibid., 15.

24. Ibid. Because integrating the model with other components and systems that may be provided by third party contractors is common, the government requirements should include an open standard application programming interface (API). An API is critical to interoperability between software, but also protects proprietary IP as the government does not require source code in the core software as a deliverable to be able to connect various software programs in a system.

25. See Ibid., 2.

26. The hardware tuned to AI applications has seen a 300,000-fold increase in performance between 2012 and 2018 with expectations high that techniques and advancements in quantum computing will lead to greater acceleration. Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach*, (Englewood Cliffs, NJ: Prentice Hall, 2021, 4th edition), 15. https://www.academia.edu/39196290/artificial_intelligence_a_modern_approach.

27. Gadepally, 3.

28. Russell & Norvig, 711.

29. Mark Coeckelbergh, *AI Ethics* (2020), 117.

30. Russell & Norvig, 711.

31. Ibid., 712.

32. Ibid.

33. JAIC, Ethical Principles for Artificial Intelligence. https://www.ai.mil/docs/ethical_principles_for_artificial_intelligence.pdf.

34. Kathleen Hicks, "Implementing Responsible Artificial Intelligence in

the Department of Defense" (May 26, 2021), https://media.defense.gov/2021/may/27/2002730593/-1/-1/0/implementing-responsible-artificial-intelligence-in-the-department-of-defense.pdf.

35. Ibid.

36. DFARS 252.227-7014(a)(8). If the costs of development are not allocated to a government contract, development is considered to be accomplished exclusively at private expense even if the contractor did not fund the development. With open-source code and open source data, contractors can assert restrictions on computer software not developed internally provided they obtain a license to use the third-party software and documentation.

37. See also Advanced Technology Academic Research Center, *From Ethics to Operations: Current Federal AI Policy* (January 5, 2022), 20–21, stating, "Current regulation found in the federal Acquisition Regulations (FAR) provides strong intellectual property (IP) protection for vendors, covering their proprietary algorithms, as well as the data they use and the data they generate. These protections may create additional uncertainty for AI system acquisition, as the FAR currently relies on a clear distinction between "software" and "data". When current ML systems are "trained," they generate new data (the refined weights of nodes) and this data is integrated into the new ML model. In effect, the AI system blends customer software with the new data generated. Federal agency procurement policies need to explicitly address how rights to that data and the resulting AI systems are to be distributed, and under what constraints and conditions."

38. Andrew S. Bowne, "Making the Pentagon an Even More Attractive Customer for AI Upstarts," *Contract Management* (March 2021).

39. See 10 U.S.C. § 3771, as regulated by DFARS 252.227-7014(b)(4).

40. As with all systems that can be used in the engagement of hostilities, law applicable to the military's use of AI-enabled systems includes international (law of armed conflict, including the United Nations Charter, Geneva Conventions, Hague Conventions, and Additional Protocol I), and domestic laws (including the U.S. Constitution and statutes, as well as executive orders, regulations, and DoD policies).

41. 32 CFR 219; DoDI 3216.02.

42. Jeffrey Dastin, "Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women," *Reuters* (October 10, 2018), www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-iduskcn1mk08g.

43. Christopher Duckworth et al., "Using Explainable Machine Learning to Characterise Data Drift and Detect Emergent Health Risks for Emergency Department Admissions during COVID-19," *Scientific Reports* 23017 (2021), 11.

44. Coeckelbergh, 131.

45. Lauren Smiley, "I'm the Operator," *Wired* (April 2022), 54, 56–61.

46. S. Kate Devitt & Damian Copeland, "Australia's Approach to AI Governance in Security & Defence," *AI Governance for National Security and Defence: Assessing Military AI Strategic Perspectives*, (2022), 22.

47. Defense Innovation Board, "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense (Support Document)" (October 2019), 33–34. https://media.defense.gov/2019/oct/31/2002204459/-1/-1/0/dib_ai_principles_supporting_document.pdf.

48. Timnit Gebru et al., "Datasheets for Datasets," arXiv (December 1, 2019) https://arxiv.org/pdf/1803.09010.pdf.

49. Sally Adee, "What are deepfakes and how are they created," *IEEE Spectrum* (April 29, 2020), https://spectrum.ieee.org/what-is-deepfake.

50. Bobby Allen, "Deepfake Video of Zelenskyy Could Be 'Tip of the Iceberg' in Info War, Experts Warn," NPR (March 16, 2022), https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia.

51. Sid Ahmed Fezza, Yassine Bakhti, Wassim Hamidouche, and Olivier Déforges, "Perceptual Evaluation of Adversarial Attacks for CNN-based Image Classification," *Quality of Multimedia Experience* (May 2019).

52. Ryan Soklaski et al., "Tools and Practices for Responsible AI Engineering," arXiv:2201.05647 (January 14, 2022).

53. Jeff Larson, Surya Mattu, Lauren Kirchner & Julia Angwin, "Machine Bias," *ProPublica* (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

54. Ibid.

ABOUT THE AUTHORS

Major Andrew Bowne
Chief Legal Counsel
Department of the Air Force /
Massachusetts Institute of Technology
AI Accelerator

Major Andrew Bowne is the chief legal counsel of the Department of the Air Force, Massachusetts Institute of Technology Artificial Intelligence Accelerator (AIA) in Cambridge, Massachusetts. In addition to advising the AIA on contracts, intellectual property, fiscal law, information law, and ethics, Bowne is the Air Force lead on a multidisciplinary team of MIT and Lincoln Laboratory researchers for developing faster, more efficient computing, and more sustainable datacenters through AI solutions.

Bowne previously served as a professor of contract and fiscal law at The Judge Advocate General's School, teaching the future of acquisitions, intellectual property, emerging technology, other transaction authority (OTA), and national security law. He contributed to several DoD policies and reports, including work on the OTA Guide, IP Policy, Air Force Artificial Intelligence Strategy, and Section 809 Panel recommendations.

He covers a wide array of topics including intellectual property, other transactions, security cooperation, artificial intelligence at the nexus of national security, and technical papers on high performance computing, neural differential models for navigation, and data discovery in data lakes.

Bowne earned his commission through direct appointment in 2010. He is a member of the State Bar of California and is admitted to the Air Force Court of Criminal Appeals, the Army Court of Criminal Appeals, and the United States Court of Appeals for the Armed Forces. Prior to commissioning, Bowne was a business and entertainment law attorney in Southern California. He holds a Bachelor of Arts from Pepperdine University, a Juris Doctor from the George Washington University Law School, a Master of Laws from the Judge Advocate General's School, and is a Doctor of Philosophy candidate at the University of Adelaide.

Benjamin McMartin, Esq. CPCM
Senior Fellow
Center for Government Contracting

Benjamin McMartin, Esq. CPCM is a Senior Fellow in the School of Business at George Mason University. His appointment is with the Center for Government Contracting. McMartin is the managing partner of the Public Spend Forum, a firm dedicated to enabling open government markets worldwide. He is a nationally recognized expert, speaker, and author in federal procurement and non-traditional acquisition methodologies, who spent more than a decade developing some of the most unique procurement solutions for the Department of Defense.

McMartin is a regular speaker on topics related to alternative acquisition methods, public procurement policy, and acquisition reform. He previously served as Chief of the Acquisition Management Office for the US Army Combat Capabilities Development Command–Ground Vehicle Systems Center, and prior to that, as a Procuring Contracting and Agreements Officer for the US Army Contracting Command–Warren.

McMartin earned his J.D. from the University of Detroit-Mercy Law School and has been a member of the Michigan State Bar since 2008. He is twice the recipient of the Army Achievement Medal for Civilian Service, and a recipient of the Army Civilian Service Commendation Medal. He is DAWIA Level III certified in Contracting; a Certified Professional Contracts Manager (CPCM); and Fellow of the National Contract Management Association.

The White Paper Series

The purpose of the Center's White Paper Series is to promote research and discussion on topics of interest and importance to the government contracting community.

Comments from the community are welcome and may be sent to govcon@gmu.edu.

NO. 1. *Unintended Consequences of Small Business Contracting,* Craig R. Reed, Ph.D. NOVEMBER 25, 2019

NO. 2. *Pricing Intellectual Property in Defense Competitions: Toward Theoretical and Practical Advice for government Officials and Government Contractors,* James Hasik, Ph.D. NOVEMBER 25, 2016

NO. 3. *The Cost of Saving Money: The Negative Impact of Roller Coaster DoD Funding,* Jennifer Taylor. NOVEMBER 25, 2016

NO. 4. *The Value of Intellectual Property in Government Procurement Auctions,* James Hasik, Ph.D. JULY 14, 2020

NO. 5. *The DoD Budget Process: The Next Frontier of Acquisition Reform,* Eric Lofgren. JULY 29, 2020

NO. 6. *Building Resilience: Mobilizing the Defense Industrial Base in an Era of Great-Power Competition.* Jerry McGinn, Ph.D. SEPTEMBER 28, 2020

NO. 7. *What Future for Remote Work in Federal Contracting?* James Hasik, Ph.D., JANUARY 14, 2021

NO. 8. *Building Industrial Resilience with a Little Help from Our Friends,* Jerry McGinn, Ph.D. JUNE 5, 2021

NO. 9. *Achieving Defense Exportability, Frank Kenlon,* AUGUST 16, 2021

GEORGE MASON | School of Business

Center for Government Contracting

The vision of the Center for Government Contracting is to establish the first-in-the-nation university center to address business, policy, regulatory and other issues in government contracting.

Activities to implement this vision will focus on three lines of effort: ***Research***, ***Education & Training***, and ***Collaboration***.

The George Mason University School of Business is uniquely positioned to create this center by virtue of the composition of our faculty and students as well as our geographic co-location with the Federal Government and many headquarters and major facilities of companies that make up the $500 billion government contracting (GovCon) industry.