Contract AI Risk Engine (CARE) to Reduce Cyber Contracting Risk

Major Y. Brian Lee
Major Dennis Kim
Major Wallace Rollins

INTRODUCTION

The Fiscal Year 2019 National Defense Authorization Act (NDAA) established the National Security Commission on Artificial Intelligence (NSCAI) to consider the methods and means necessary to advance development of artificial intelligence (AI), machine learning (ML), and other associated technologies to address America's national security concerns. NSCAI's final report to the President and Congress identified areas of weakness that the federal government must address to elevate data security as a national security priority. NSCAI recommended the federal government implement a security development lifecycle approach for AI systems, prioritize data privacy and security considerations as part of larger efforts to strengthen foreign investment screening and supply chain intelligence and risk management, and integrate national security considerations into efforts to legislate and regulate data protection and privacy.[1]

Current Department of Defense (DoD) information technology (IT) contracting policies, vehicles, and practices lack definitive language or terms that give due process to national security considerations. Without contracting language specifically tailored to the cyber security threats facing the United States (US), DoD cannot adequately secure the DoD Information Network (DODIN) nor protect it from foreign influence. Contractual languages often favor the vendor. For example, DoD cyber vendors can potentially circumvent DoD prohibited IT equipment or prevent DoD Cyber Protection Teams from inspection or damage assessment during cyber breaches or attacks, citing ambiguous contracting language and proprietary corporate intellectual protection as justifications.[2] Unfortunately, contracting personnel, commanders, and staffs across the DoD lack training and expertise in

Major Y. Brian Lee, U.S. Army, is a Medical Service Corps officer assigned to the Department of Defense Chief Digital and Artificial Intelligence Office, Arlington, VA. He holds a BA from Washington University in St. Louis, a MS from the University of Maryland Global Campus, and a Masters in Operational Studies from the U.S. Army Command and General Staff College. During his career, MAJ Lee served with the Joint Artificial Intelligence Center, 7th Special Forces Group, 2nd Infantry Division, 82nd Airborne Division, and 65th Medical Brigade.

reducing cyber security risk. An objective cyber contract risk score does not exist. DoD should leverage ML in the cyber contract requirements generation process to reduce cyber contract risk and position DoD to better prevent, monitor, and respond to cyber threats.

Issue

Contracts for cyber or IT related products and services present a cyber supply chain risk for the DoD. Cyber supply chain risk stems from a lack of visibility into, understanding of, and control over many of the processes and decisions involved in the development and delivery of cyber products to the Joint Force.[3]

Requirement owners and contract management offices are at the forefront of cyber supply chain risk management (C-SCRM). As the requiring activity, commanders and their staff determine and develop requirements and generate the performance work statement (PWS). Contracting officers, vested with the authority to obligate the US government to legally binding contracts, coordinate and finalize contracting actions to provide the goods or services needed by the requiring activity. Unfortunately, requiring activities and contracting professionals often lack the technical expertise to articulate specific C-SCRM measures within contracts. Further, existing resources that provide guidelines and standards for C-SCRM are inadequate with respect to the granular process of contract writing and are spread across a multitude of DoD policies (Figure 1).

Publications from the National Institute of Standards and Technology (NIST), Defense Acquisition University (DAU), and DoD Instruction documents describe how to conduct C-SCRM, but no publication goes into more nuanced details on contract language, thus creating gaps in cyber supply chains. Current acquisition processes account for various risks, but in-depth technical understanding of the cyber supply chain is required to properly translate mitigation measures into contract language during the requirements generation process.

Major Dennis Kim, U.S. Army, is a Medical Service Corps officer assigned to 65th Medical Brigade, Camp Humphreys, Republic of Korea. He holds a BS from Boston University, an MBA from The College of William and Mary, and a Masters in Operational Studies from the U.S. Army Command and General Staff College. During his career, MAJ Kim served with the 10th Mountain Division, 2nd Infantry Division, and the U.S. Army Medical Materiel Agency.

During a lecture at the U.S. Army Command and General Staff College in April 2021, Brigadier General (BG) Paul Craft, Commandant of the U.S. Army Cyber School, used the cloud migration of Army data as an opportunity to address both the benefits and challenges that data contracting presents. BG Craft acknowledged it is unrealistic to expect all contracting officers to be cyber security experts, but a lack of understanding of cyber security can lead to inadequate language in contracts. This has led to instances where data became lost, mishandled, or the DoD denied access to its own data and required to pay to get data back. BG Craft cautioned that this situation can be especially damaging when there is a breach, and the language of the contract does not authorize DoD Cyber Protection Teams to investigate the breach. This lack of transparency and access erodes the public trust and harms national security.

APPROACH AND SOLUTION

This proposal recommends the use of AI through ML to review draft contracts uploaded by contracting officers and analyze the cyber security risk to the DoD. After review, the Contract AI Risk Engine (CARE) produces recommended clauses most advantageous to DoD for cyber security along with a cyber risk level which measures the level of risk to DoD for the contract as written. The requiring activity reviews the recommendations and adjusts the contract as necessary. The contracting officer subsequently takes the improved contract and obtains a new risk score, with scores above a certain threshold requiring command concurrence by both the requiring activity commander and the supporting contracting commander before moving to contract fulfillment. As a pilot, CARE recommendations are initially based upon the Army Contracting Command's (ACC) repository of previous IT and cyber related contracts. Upon successful testing, the intent will be to incorporate a Joint solution and include data from all services and DoD agencies. CARE relies upon cloud computing and AI platforms, such as the DoD's Advana enterprise

Major Wallace Rollins, U.S. Army, is an Acquisition Corps officer assigned to Program Executive Office Soldier, Fort Belvoir, VA. He holds a BA from Virginia Polytechnic and State University, an MBA from the University of Kansas, and a Masters in Operational Studies from the U.S. Army Command and General Staff College. During his career, MAJ Rollins served with the 82nd Airborne Division, 3rd U.S. Infantry Regiment, "The Old Guard," the 1st Cavalry Division, and 1st Security Force Assistance Brigade.

analytics platform, for data analysis, model generation, and risk score calculation.

Artificial Intelligence Design

Contracting affects DoD agencies and activities, the military services, and Combatant Commands. Using CARE to reduce cyber contracting risk is a feasible ML project with immediate real-world applications and implications where end users can see the benefits of augmenting contracting processes with AI. DoD has partnered with national academic research institutions, such as the MIT Lincoln Laboratory and the Army's AI Task Force at Carnegie Mellon University, to accelerate the research and development of national security AI priorities. While partnerships and national conversations on the research, development, and applications of AI advance the state of DoD AI initiatives, Soldiers, Airmen, and Sailors have yet to experience the transformational benefits promised by AI in daily operations. Incorporating AI into the Joint Force will create a generational shift in how business is conducted. For commanders to champion AI and for the end user to experience the benefits of AI, DoD must bridge the crisis of trust between humans and AI, whether that AI is operating in autonomous-capable weapons systems or as software platforms.[5] Building trust requires repetitive exposure through the rapid development and implementation of small-scale projects rather than conceptual projects that will not mature for years to come. Quick wins that create buy-in from the operational force will advance the state of DoD AI.

The human-machine relationship should be carefully considered when designing AI projects and use cases. Requirement developers and AI practitioners determine the degree of autonomy granted to each AI product. The three degrees of autonomy are commonly referred to as human-in-the-loop, human-on-the-loop, and human-out-of-the-loop. In human-in-the-loop (HITL) operations, the machine performs a task and waits for the human

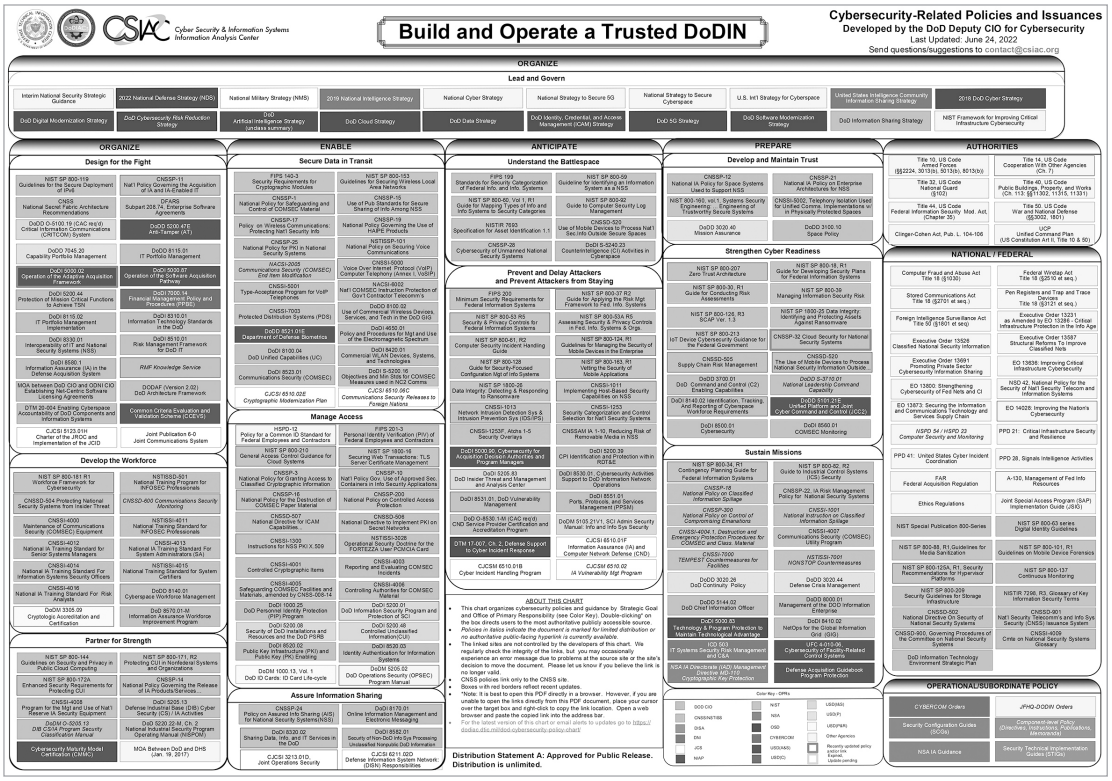

Figure 1. Cybersecurity policies and issuances for the DoD.[4]

user to take an action.[6] In human-on-the-loop (HOTL) operations, the machine decides and acts on its own, but a human user supervises its operations and can intervene if necessary.[7] In human-out-of-the-loop (HOOTL) operations, the machine decides and acts on its own, and the human user cannot intervene in a timely fashion.[8] The risk associated with the degrees of autonomy vary and should be carefully weighed based on the intended applications of the machine, the chances of faulty actions to occur, and the severity caused by faulty actions. Given that the purpose of this project is to reduce the risk associated with DoD cyber and IT contracting, we propose that AI recommended contracting clauses and risk determination require HITL acceptance both in modifying contracting language during the contract support process as well as involving commanders to accept contracts of considerable risk with or without language modification. Once implemented, CARE augments, rather than replaces, the human decision-making process.

To develop DoD end user trust in AI, CARE does not remove human involvement and instead harnesses the efficiency of intelligent automation to best inform the human decision-maker.[9] Trust builds as users throughout the contracting chain see tangible benefits from CARE-assisted contracting compared to the standard human-only contracting process.

ML requires data to improve model performance. DoD contracts in document format cannot provide the necessary data to begin training ML algorithms. Natural language processors combined with numerical scoring of contract features must be developed, and contract scoring does not currently exist. Feature engineering is the determination of the appropriate data variables necessary for ML algorithms to assess what the user requires.[10] In other words, poor feature engineering results in subpar model performance. Prior to any data collection for CARE development, DoD contract stakeholders throughout the contracting process with proper AI education must carefully determine the features that will create the contracting data necessary for ML algorithms to work and with the least amount of data bias (Figure 2).

Contract Num	Type	Unit Type	Cost	Feature 1	Feature 2	Feature 3
2020.02.01	Hardware	Tactical	$5,002	5	2	4
2020.02.02	Software	Service	$22,678	2	3	2
2020.02.03	Hardware	CCMD	$540,555	5	1	1
2020.02.04	Hardware	CCMD	$874,322	3	5	5
2020.02.05	Software	Operational	$54,178	1	3	3

Figure 2. Feature engineering example

Development and Operational Concept

In a case study on Army contracting analytic capabilities, the RAND Corporation piloted an effort to make unstructured historical contract data machine readable to forecast a contract's likelihood to have unliquidated obligations.[11] We propose to utilize similar methodologies as RAND in accessing and scoring cyber and IT contracts over a set number of fiscal years with the inclusion of contract performance and contract closeout reports. Contracts would be analyzed by trained cyber and contracting experts and scored on features developed during feature engineering for the data. We seek to score cyber and IT specific contractual language in a tabular format. Proposed feature categories include, but are not limited to, contract duration, contract language, contract outcome, contract performance, adversarial incursion, DoD cyber response, and contract barriers. Close collaboration with data scientists during contract scoring will reduce introducing biased data into the dataset. While RAND utilized over 300,000 contracts with 150 features over three fiscal years, we are unsure how many Army-specific cyber and IT contracts exist at this time.[12] A period of discovery should be included in the CARE development timeline.

Upon completion of contract scoring, developers perform exploratory data analysis to ensure quality data, build and work with predictive models, evaluate models and receive predictions, and refine outputs. CARE determines a contract's risk to DoD and outputs a risk percentage and recommended changes to reduce the risk. A lower risk means that the contract's language provides DoD with favorable execution outcomes. A higher risk percentage suggests that DoD will potentially meet resistance from contractors in response to adverse security events. CARE will recommend specific contractual language modifications and inform end users where that language should go in the contract. Users explore how CARE recommended modifications af-

fect risk, whereby as modifications are selected in the user interface, the contract would be reassessed and the net result displayed in a live risk meter. Users could choose all recommendations or select recommendations, with selections based on the requiring activity's desired combination of potential cost, time, and scope as considerations for risk acceptance. As a HITL system, CARE must rely upon the contracting officer to accept modifications. Cyber and IT contracts continue to be generated by requiring activities, and CARE will be further refined in the future as new data, including CARE augmented contracts, are introduced into the model.

CARE would be a web-portal ML platform with a file upload and document review user interface (Figure 3). Contracting officers upload draft contracts for analysis and interact with recommendations for decision-making analysis only. To reduce the cost and complexity of developing and maintaining CARE, contracting officers transfer recommendations manually into the original document creation software, most likely Microsoft Word or Adobe Acrobat, prior to contract fulfillment. CARE is decision augmentation only. Contracting officers should consult with the requiring activity before accepting any CARE modifications, and risk scores above a certain percentage would require both the requiring activity and contracting commanders to concur. CARE enables commanders to analyze risk, considering risk to the force and risk to the mission against the perceived benefit of the contract.[13]

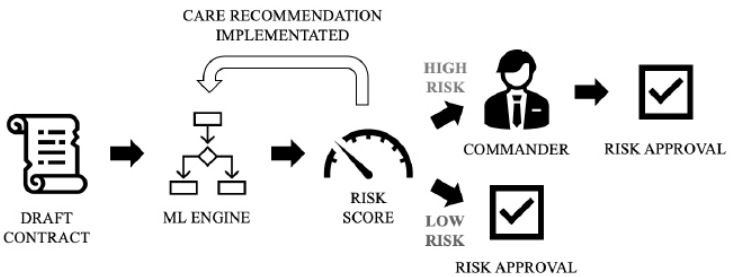

Figure 3. CARE use case

Based upon current development timelines from ML projects being piloted at U.S. Army Forces Command (FORSCOM), we believe that CARE can be rapidly developed with the involvement of data scientists, contract specialists, and cyber security experts in under three months (Figure 4) utilizing the collaborative framework of DevSecOps and agile delivery. We anticipate an additional six to nine months to complete Authorization-To-Operate (ATO) requirements as necessary, working through ML Ops challenges to deploy and maintain models reliably in the production environment, user interface design, and policy decisions. By developing a narrow scope that precisely targets the problem that CARE solves, DoD can responsibly and rapidly prototype and field a platform that decreases contracting risk with immediate and tangible benefits. However, we do acknowledge the risk of the "valley of death" that a successful model development does not guarantee inclusion into a program of record for further sustainment and adoption.

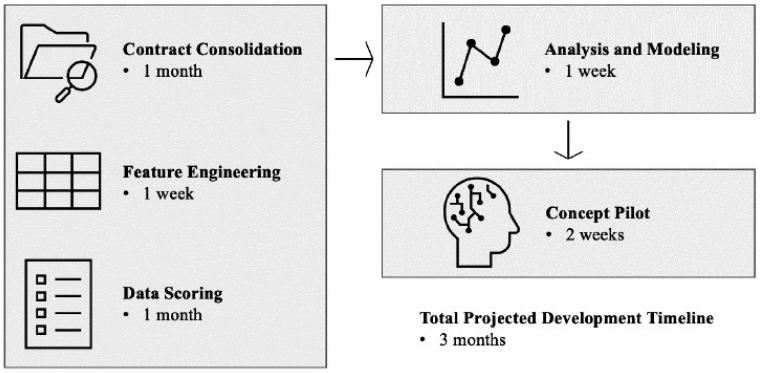

Figure 4. Projected CARE development timeline

CONCLUSION

Cyber-attacks by foreign adversaries and criminal organizations have revealed how the American people and the economy rely on the cyberspace domain. As more DoD operations migrate to the cloud with as-a-service contracting and as DoD activities contract for capabilities to enable a competitive edge in training and in combat, reducing the cybersecurity risk of these contracts is paramount for DoD to defend against and respond to adversarial cyber operations. We recommend that the U.S. Army Materiel Command, assisted by, in coordination with, and potentially developed through the DoD Chief Digital and Artificial Intelligence Office (CDAO), funds and develops CARE. Upon successful pilot testing, it would mandate all cyber and IT contracts to adopt CARE as a critical component in the contract approval process. DoD cannot allow contracting language to cripple America's national security interests. Developing and implementing CARE for DoD cyber contracting will create a more resilient DoD cyber supply chain with the necessary contractual safeguards for DoD to prevent, monitor, and respond to cyber and IT related adversarial events.

NOTES

1. National Security Commission on Artificial Intelligence, Final Report: National Security Commission on Artificial Intelligence (Washington, DC, 2021), 50.
2. Paul G. Craft, personal communication, April 20, 2021.
3. National Institute of Standards and Technology, Information and Communications Technology Supply Chain Risk Management, (Washington, DC: Department of Commerce, 2021), https://csrc.nist.gov/CSRC/media/Projects/Supply-Chain-Risk-Managements/documents/nist_ict-scrm_fact-sheet.pdf.
4. "Build and Operate a Trusted DoDIN," Defense Technical Information Center, 2022, https://dodiac.dtic.mil/wp-content/uploads/2022/07/2022-06-24-csiac-dod-cybersecurity-policy-chart.pdf.
5. Dan G. Cox, "Artificial Intelligence and Multi-Domain Operations: A Whole-of-Nation Approach to Success," *Military Review*, 101, no. 3 (May-June 2021): 76-91, https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MJ-21/MJ21-Whole-Book-2.pdf.
6. Paul Scharre, *Army of None* (New York: Norton, 2018), 29.
7. Ibid., 29.
8. Ibid., 30.
9. Ge Wang, "Humans in the Loop: The Design of Interactive AI Systems," Stanford University, 2019, https://hai.stanford.edu/news/humans-loop-design-interactive-ai-systems.
10. "Feature Engineering," DataRobot, 2021, https://www.datarobot.com/wiki/feature-engineering/.
11. William Marcellino et al., *Army Analytic Capabilities: A Case Study Within Army Contracting Command and Its Implications*, RR-A106-1 (Santa Monica, CA: Rand, 2021), 1, https://www.rand.org/pubs/research_reports/RRA106-1.html.
12. Ibid., 5.
13. Department of the Army, *Mission Command: Command and Control of Army Forces*, ADP 6-0 (Washington, DC: Department of the Army, 2019), 1-13.