Developing the Human Machine Teaming (HMT) Ecosystem

Anastacia MacAllister, Rey Nicolas
General Atomics
San Diego, CA
anastacia.macallister@ga-asi.com, rey.nicolas@ga-asi.com

Patrick Rupp, George Hellstern, Miguel Morales
Lockheed Martin Corporation
Fort Worth, TX
patrick.h.rupp@lmco.com, george.hellstern@lmco.com, miguel.morales@lmco.com

Dan Javorsek
Air Force Operational Test and Evaluation Center
Nellis AFB, NV
daniel.javorsek@us.af.mil

Louis Dube
EpiSci
Poway, CA
louis.dube@episci.com

ABSTRACT

Department of Defense (DoD) investment in artificial intelligence (AI) and machine learning (ML) related efforts has amounted to hundreds of billions of dollars over the past five years. Driving this high level of investment is AI/ML's potential to unlock revolutionary new capabilities for the warfighter. Use cases explored by the DoD include Project Maven's imagery analysis, CDAO's Smart Sensor, and DARPA's Air Combat Evolution (ACE). While these programs have shown AI/ML's utility and promise, they have only been deployed in isolated or developmental environments. Moving forward, to truly realize the potential of operationalized AI/ML technology, more complex and real-world representative use cases are required.

One such use case within the DoD is Human Machine Teaming (HMT) with Unmanned Combat Aerial Vehicles (UCAV). The goal of HMT in this context is to develop and deploy autonomous aircraft to augment and enhance today's resource constrained fighter aircrews. While the DoD has successfully deployed AI/ML on a more limited scale, HMT represents a large increase in mission and system complexity. The effort will require seamless integration of uncrewed aircraft, crewed aircraft, advanced sensing methods, high performance edge computing, crewed-uncrewed teaming, AI/ML aircraft control algorithms, simulation-based training for operators and rapid AI/ML algorithm retraining for securely updating autonomy at the speed of relevance. As a result, the HMT ecosystem will require government, crewed aircraft providers, uncrewed aircraft providers, and autonomy software developers to work together in an agile and open way. The work in this paper aims to begin the HMT ecosystem development process by decomposing and describing critical components. After defining and describing these elements, the authors then begin exploring policies, standards, and management practices required to ensure effective HMT deployment. Ultimately, the work presented in this paper helps begin defining the HMT ecosystem's components and their interfaces to ensure delivery of powerful new warfighter capabilities.

ABOUT THE AUTHORS

Anastacia MacAllister, Ph.D., is Technical Director of Autonomy and Artificial Intelligence for General Atomics Aeronautical Systems (GA-ASI). Her work focuses on prototyping novel machine learning algorithms, developing machine learning algorithms using sparse, heterogenous or imbalanced data sets, and exploratory data analytics. Throughout her career she has provided key contributions in a number of interdisciplinary areas such as prognostic health management, human performance augmentation, advanced sensing, and artificial intelligence aids for future warfare strategies. This work has received several awards and commendations from internal technical review bodies. Dr. MacAllister has published over two dozen peer reviewed technical papers, conference proceedings, and journal articles across an array of emerging technology concepts. Dr. MacAllister holds a bachelor's degree from Iowa State University (ISU) in Mechanical Engineering. She also holds a Master's and PhD from ISU in Mechanical Engineering and Human-Computer Interaction.

Rey Nicolas is Director of Software, Autonomy and AI Solutions and leads GA-ASI's Autonomy and AI teams developing next generation Autonomous Command and Control (C2) systems, Autonomous Processing, Exploitation, Dissemination (PED) systems, and AI edge processing for flight and sensor autonomy, air-to-air combat, and air-to-ground Intelligence, Surveillance and Reconnaissance (ISR) missions. Rey is also an Executive Committee Leader for SAE Standard Works that is developing AI in Aviation Safety Standards. Previously, Rey worked for Intel Corporation and Motorola/Google leading AI R&D and product development for multiple product lines in the smart and connected home, autonomous driving, and healthcare applications. Rey holds a bachelor's degree in Mechanical Engineering from University of California San Diego, Master's in Computer Science from the University of Illinois, and MBA from San Diego State University.

Patrick Rupp is an AI/ML Engineer at Lockheed Martin's Skunk Works®. Mr. Rupp specializes in developing deployable ML & Data Analytics solutions across a wide variety of problems ranging in size from small prototypes to enterprise solutions. Mr. Rupp is currently developing AI autonomy solutions including ML training, evaluation, deployment, pipelines, and processes in support of the warfighter. He has previously supported Lockheed's work in Defense Advanced Research Projects Agency's (DARPA) Systems of Systems Integration Technology and Experimentation (SoSITE) program and Air Combat Evolution (ACE). Mr. Rupp received his Masters in Data Science from Johns Hopkins University.

George Hellstern has 25 years of experience with systems design, including AI solutions for air-to-air combat and sustainment. He is a program manager for autonomy and AI, uncrewed air systems C2 and human performance. Previous experience includes operational, programmatic and technical experience from the Air Mobility Command, the Office of the Secretary of Defense and Lockheed Martin Advanced Development Projects also known as Skunk Works®.

Miguel Morales is a full-time research engineer at Lockheed Martin, Skunk Works and a Computer Science Ph.D. student at the Georgia Institute of Technology. He's the author of "Grokking Deep Reinforcement Learning" and the content developer for the Deep Reinforcement Learning Nanodegree's Actor-Critic lectures. His research interests center around multi-agent learning, reinforcement learning, machine learning, artificial intelligence, deep learning, and robotics.

Dan Javorsek, Ph.D., is the Commander of Detachment 6, Air Force Operational Test and Evaluation Center (AFOTEC), Nellis Air Force Base, Nev., and Director, F-35 United Operational Test Team (UOTT). AFOTEC's Detachment 6 plans, conducts, and reports on realistic, objective, and impartial operational test and evaluation of fighter aircraft. The detachment evaluates the operational effectiveness, suitability, and mission capability of the A-10, F-16, F-15C/E/EX, F-22, F-35, and Next Generation Air Dominance (NGAD) aircraft, and reports results in support of major acquisition program milestone decisions and combatant command fielding decisions.

Louis Dube currently serves as Principal Research Engineer at EpiSci. Prior to his arrival at EpiSci, Mr. Dube was the Chief Engineer for the F-35 U.S. Operational Test Team, where he provided technical advisership and oversight of tri-service operational test and evaluation designed to accurately represent modern joint doctrine concepts. In addition to his involvement in operational test and evaluation, Mr. Dube spent eleven years in a variety of functions on the F-16, F-15, and F35 developmental test programs. Notably, Mr. Dube was involved in the integration and developmental test of the Automatic Ground Collision Avoidance System on both the F-16 and F-35 platforms, ushering aviation into the age of autonomy.

Developing the Human Machine Teaming (HMT) Ecosystem

Anastacia MacAllister, Rey Nicolas	Patrick Rupp, George Hellstern, Miguel Morales	Dan Javorsek	Louis Dube
General Atomics	Lockheed Martin Corporation	Air Force Operational Test and Evaluation Center	EpiSci
San Diego, CA	Fort Worth, TX	Nellis AFB, NV	Poway, CA
anastacia.macallister@ga-asi.com rey.nicolas@ga-asi.com	patrick.h.rupp@lmco.com george.hellstern@lmco.com, miguel.morales@lmco.com	daniel.javorsek@us.af.mil,	louis.dube@episci.com

INTRODUCTION

The United States Department of Defense (DoD) has invested hundreds of billions of dollars over the past five years into artificial intelligence (AI) and machine learning (ML) related technologies (United States Government Accountability Office, 2022). Driving this high level of investment is AI/ML's potential to unlock revolutionary new capabilities for the warfighter. Use cases explored by the DoD include Project Maven's imagery analysis, Chief Digital and AI Office's (CDAO) Smart Sensor resource manager, and Defense Advanced Research Projects Agency's (DARPA) Air Combat Evolution (ACE) autonomous within visual range dogfighting program (U.S. Department of Defense, 2023); (DARPA, 2023). While these projects have demonstrated the utility and potential of AI/ML to leadership, they focused on foundational AI/ML research and development (R&D) efforts. As a result, wider considerations like integration into networked systems or interoperability with other vendors have yet to be robustly considered. Without addressing these wider considerations, effectively transitioning this impactful R&D technology into the hands of frontline warfighters will be challenging. To truly operationalize AI/ML systems, the technology needs to be integrated into real-world representative use cases aligning with the DoD's future network centric warfighting strategies (Department of Defense, 2023).

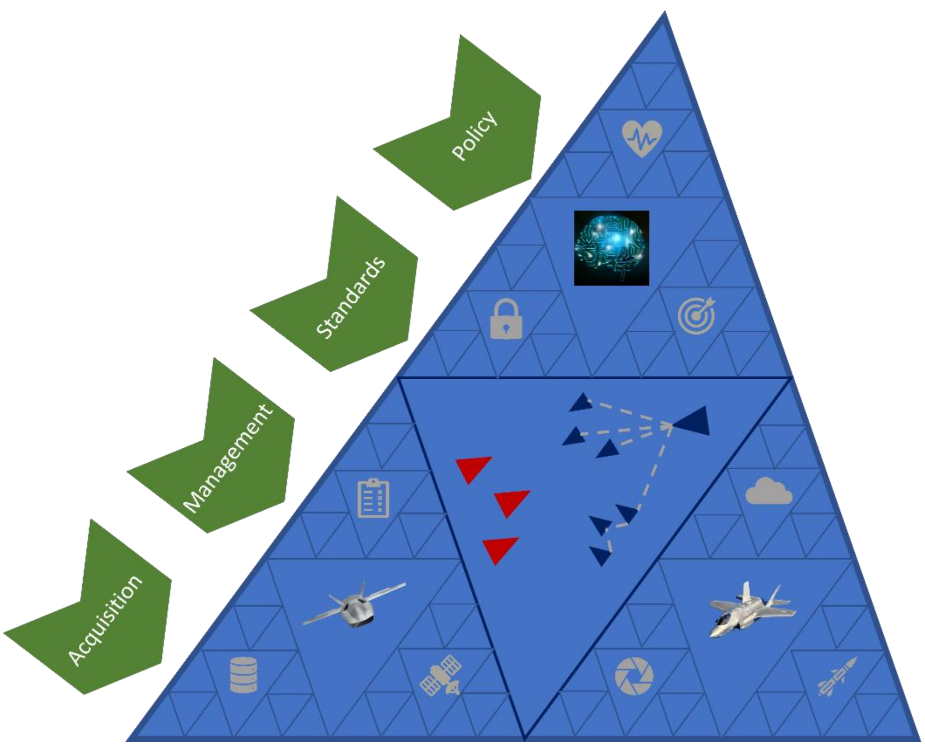

Figure 1. Operational vignettes for HMT often present futuristic capabilities without considering the significant underlying ecosystem required for deployment

One such use case is the development of Human Machine Teaming (HMT) using Unmanned Combat Aerial Vehicles (UCAV) as part of the Next Generation Air Dominance (NGAD) Family of Systems (FoS). UCAVs are initially planned as drones that will pair with crewed platforms to augment all types of tactical aircraft, not just those associated with the NGAD system (Tirpak, Collaborative Combat Aircraft Will Join the Air Force Before NGAD, 2023). The United States expects to field over one thousand of these aircraft in the coming decade (Secretary of the Air Force Public Affairs, 2023). Operational vignettes often present a futuristic blend of crewed-crewed and crewed-uncrewed systems working in perfect coordination to overcome near peer adversaries in complex contested environments. While these vignettes present an impressive vision of future warfare strategies, they often neglect to consider all the facilitating infrastructure required to connect systems.

For example, Figure 1 shows a simplified mission vignette where one crewed blue entity directs six uncrewed platforms against a team of three red crewed fighters. The dotted lines represent a two-way connection between the crewed and uncrewed fighters to maintain situational awareness. This mission, while simplistic by operational standards, is quite complex in terms of the technological relationships required between entities. Unpacking technological dependencies between entities reveals a multitude of supporting components all interacting to accomplish what might be considered a relatively simple mission for a human team. At surface level there are uncrewed aircraft, crewed aircraft, and AI/ML aircraft control algorithms seamlessly working in harmony to achieve the desired outcome. In the caricature of the vision, these three elements are all stitched together with supporting components that are not often visualized in such Operational Viewpoint (OV-1). This underlying necessary structure includes unseen and often under-resourced features such as disparate data sources, advanced sensors, health monitoring, security protocols, data aggregation services, and munition management. In addition, traditional OV-1s do not consider infrastructure logistics elements such as aircrew training, security, or autonomy behavior development. Ultimately, diving further into the vignette's technical details reveals more and more subsystem components required to make even a simplistic mission a reality.

In the end, fielding HMT will require seamless integration of uncrewed aircraft, crewed aircraft, advanced sensing methods, high performance edge computing, crewed-uncrewed teaming, artificial intelligence aircraft control algorithms, simulation-based training for operators, and rapid AI algorithm training. To operate effectively at the speed of relevance, these subsystems will require seamless integration and data sharing across disparate operating frameworks provided by a variety of vendors (Penney, 2022). While the DoD has developed and deployed pieces of HMT on an isolated scale, fully integrating systems will require increased emphasis on wholistic system engineering. As a result, the ecosystem will require government, crewed aircraft providers, uncrewed aircraft providers, and autonomy developers to work together in an agile and open way.

To accomplish this seamless integration, robust policies, standards, architectures, and management principals will be required. This paper begins by describing the high-level ecosystem of technologies required to make HMT an operational reality at scale. Then discusses how sound policies, standards, management, and acquisition (PSMA) practices might be applied to effectively integrate these pieces within a vendor agnostic ecosystem that democratizes software development and breaks traditional vendor locked solutions. Ultimately, work presented in this paper describes how to effectively use PSMA practices to operationalize AI/ML technology for future 21st Century warfighting constructures.

HUMAN MACHINE TEAMING ECOSYSTEM

In order to be successful, the HMT ecosystem requires a democratized heterogeneous technology environment often found at data centric commercial software providers such as Amazon, Facebook, Google, Netflix, Twitter. This requirement is in direct tension with common DoD acquisition processes that reward contractors for tightly coupling their hardware with their software. Designing an acquisition process that supports the needs of HMT requires understanding each ecosystem component and how sound policies, standards, and management practices can be applied to ensure vendor agnostic democratization. While there are many parts of the ecosystem required to make HMT an operational reality depending on the mission, there are several common unifying elements, as shown in Figure 2, required to make it a success. This section begins to define and discuss these high-level ecosystem components.

Data Sources

To effectively execute envisioned missions, ample amounts of data are required to provide situational awareness and crewed-uncrewed teaming capability. This means facilitating architectures will require some of the most comprehensive manicured information ingestion yet seen in DoD programs. Data for the system will be required from a wide variety of DoD sources such as radar, electronic warfare, communications, electro-optical, infrared, and other sensing methods. In addition, to create a complete picture of situational awareness information sources will have to

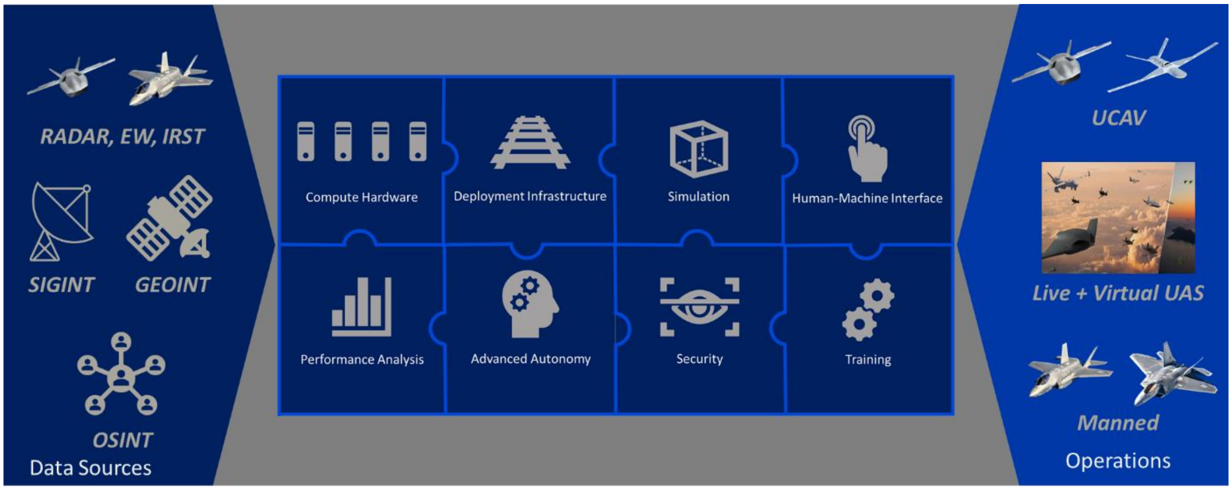

Figure 2. The HMT ecosystem consists of equally important subsystems that require well defined interfaces and policies to operate effectively

extend beyond those that are organically generated by the platform alone, to include other traditional sources of intelligence, such as signals, open-source and geospatial intelligence passed via data links. This information will be synthesized to inform global awareness on the adversary in a way to inform and enable command and control (C2) but will also provide synthesis of specific observations of adversary capabilities. This data aggregation is complicated by the fact that all these sensing mediums come with their own heritage and legacy that extends not only to the basics of data formats but even to the semantics and nomenclature human developers use to discuss the different "skills" or activities they are trying to elicit. In the end, a variety of challenges in accommodating different formats, sampling, data storage, postprocessing, and overall data curation will be required to ensure the success of the endeavor. This also extends to much larger challenges such as the aggregation of data with different classifications that results in reduced accessibility, along with the need for some sanitization or downgrade capabilities beyond standard cleaning and filtering, so that it can all be analyzed and converted into meaningful knowledge complete with context and validity.

As an example, traditionally detailed own-ship data collection is reserved to the aircraft participating in developmental and operational test. Small subsets of data can be recorded on operational aircraft, but this data is typically not comprehensive enough to enable troubleshooting of technical or performance issue – much less enabling the continued development of a combat autonomy. The HMT ecosystem requires a new data sourcing paradigm, where detailed own-ship data can be retrieved from any asset, while also providing unprecedented accessibility and synthesis of external data sources to enable continued development and maturation of all facets of the weapon system. While some basic flight sciences data will remain in the developmental test domain, mission systems data that used to be limited to the operational test domain will need to be available from fielded aircraft as well. Such expansions in capability often come with additional challenges since interactions with UCAV vehicles will present the need for a closer look at cyber protection and risks not needed when communicating with weapons and sensors via an ethernet, fiber, 1760 or 1553 bus.

Connectivity and Data Standards

To achieve operationalization of crewed-uncrewed teaming, aircraft must be able to connect and share data. However, these crewed and uncrewed aerial vehicles will consist of many heterogenous systems developed by different providers. If providers are not managed accordingly, data sharing and connectivity challenges will prevent effective deployment of collaborative command and control, autonomy, and artificial intelligence. A successful HMT ecosystem will address connectivity challenges with data link commonality, adherence to data standards, and data transformation.

To enable communication between aerial vehicles, a data link must exist that both parties can utilize. While seemingly a trivial consideration, currently 5th generation aircraft such as the F-35 and F-22 cannot communicate due to incompatible data links (Everstine, 2018). If uncrewed aircraft are to effectively team with 5th Generation aircraft, there must be a common data link standard across providers. Assuming a data link has been implemented to ensure connectivity between assets, the data content and structure must be handled to ensure on-board systems can understand and act on the data. To provide this common understanding in the ecosystem will require agreed upon messaging standards, unfortunately in the past these standards are often driven by transmitting data link providers. In a heterogeneous provider and requirements environment this may present a challenge since different data structures will hinder cohesive system of systems integration. To mitigate these issues a method of third party commonization may be required. Past programs have successfully demonstrated this 3rd party approach such as DARPA's System of Systems Integration Technology and Experimentation (SoSITE) (DARPA, n.d.). This program demonstrated the ability to integrate and operate heterogenous systems across domains by auto-generating extremely low latency and high throughput middleware between systems without needing to upgrade hardware or breaking into existing system software. This was done through a government owned (non-proprietary) technology called System-of-systems Technology Integration Tool Chain for Heterogeneous Electronic Systems (STITCHES). Taking this approach does not force a common interface standard; rather it rapidly creates the needed connections based on existing fielded capabilities obviating the need to upgrade to interoperate (DARPA, 2020). While ideally all systems in the ecosystem would adhere to Open Missions Systems (OMS) architectures, the combination of provider heterogeneity and legacy platforms may hinder this goal. As a result, government translation toolchains like STITCHES will be required to facilitate development and integration of components into the greater ecosystem.

Compute Hardware

Compute hardware will be a critical component of the HMT ecosystem. This hardware will drive simulations used to train humans, train autonomy, and run critical software elements within the ecosystem. As a result, careful consideration needs to be given to the compute power required to make HMT an operational reality. For the ecosystem compute can be split into two categories, ground vs airborne compute requirements. These environments, as shown in Figure 3, will differ based on physical constraints in airborne platforms. Ground compute requirements will be analogous to traditional compute paradigms seen today for high demand environments. Ground compute nodes will be relied upon to do some of the heavy lifting in terms of running simulations, agent training, platform design, and data analytics possibly making use of cloud technology to achieve this (National Science and Technology Council, 2022). As a result, ground compute will leverage high-quality high-performance computers (HPC) nodes with large amounts of storage capacity. These nodes can be costly to procure and maintain. Some of these costs may be cumbersome for small business to incur. As a result, consideration should be given to compute infrastructures required for developing ecosystem functionality. Large companies and the government may have to consider ways smaller performers can have access to their large HPC resources, so testing and validation can take place at the speed of relevance. To do this, standards will have to be established that makes sharing data, simulations, software, and autonomy models as seamless as possible to avoid unnecessary integration hurdles.

Figure 3. Ground compute infrastructure vs airborne infrastructure needs differ due to SWaP constraints

The most important hardware consideration for the HMT ecosystem, however, will be onboard or airborne compute requirements. These types of considerations are found in the DoD because of what is often referred to as deployment on the edge on low size, weight, and power (SWaP) platforms. Edge AI/ML refers to processing data and running algorithms on distributed platforms rather than centrally. This concept allows AI/ML to run without connectivity and with minimal response latency, an important feature for military platforms that can experience degraded communications or in Electronic Warfare attacks (Qiu, Kung, & Gai, 2020; Ahamed, Tariq, & Nusir, 2019). In addition, cost minimization and modernization challenges also play a role in driving the DoD to edge computing in

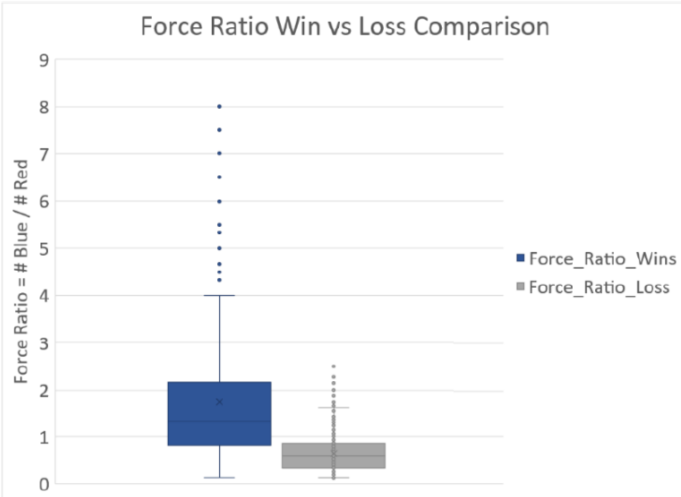

Figure 4. Performance analytics enables targeted exploration of behaviors associated with high performing crewed-uncrewed teams (MacAllister, et al., 2021)

low SWaP environments. Due to the prohibitive cost of upgrading systems across the DoD, current legacy hardware within the HMT ecosystem cannot all be expected to have state of the art computing power to run AI/ML algorithms. These compute, bandwidth, and power limitations need to be understood as pieces of the ecosystem are connected. For example, if legacy platforms do not have compute power required to perform current auto-target recognition (ATR) algorithms. Image data may need to be sent to an off-board platform via data link to receive a positive target identification. Ultimately, to effectively deploy HMT and make it tactically relevant, edge computing and low SWaP deployment problems need to be considered. Solutions need to focus on enabling ease of integration with existing infrastructure, seamless software migration and effortless deployment via open mission system interfaces.

Performance Analysis

Two main questions summarize the goal of performance analytics: 1) does blue win? and 2) if not, why? Answering these two questions allows adaptation of tactics and procedures over time to increase warfighter effectiveness. As HMT-enabled combat moves the DoD away from monolithic specialized systems to new, flexible, distributed, connected, and reconfigurable warfighting strategies, performance analytics will be imperative to developing tactics and procedures that effectively use this technology (O'Donougue, 2021; Grana, 2021; Sweetser & Bexfield, 2020; Deptula & Penny, 2019; Jensen & Paschkewitz, 2019). To do this requires data, particularly autonomy and human performance analysis both isolation and in concert. Autonomy performance analysis will need to be considered before an algorithm is even released into the battlespace. This analysis will focus on benchmarking agent performance in isolation under varying battlefield conditions as in Figure 4. This will help determine when certain agents are effective or ineffective, informing the development of autonomy enabled tactics and procedures during crewed-uncrewed exercises. In addition to understanding autonomy performance in isolation, human performance teamed with autonomy also needs to be understood and analyzed. Data from humans conducting operations with autonomy needs to be collected to determine if they were effectively teaming with uncrewed entities and where tactics might be improved.

To conduct this analysis, data are required. The HMT ecosystem will require significant low level data collection capabilities such as flight control logs, autonomy actions, human actions, and human cognitive loading. As a result, all ecosystem components need to provide access to lower-level information and it needs to be fused into a common format for processing. If data are disaggregated and unable to be linked either via time or event, they will become unusable for analytics, limiting the ability to drill down to the root cause of suboptimal performance (MacAllister, et al., 2021).

Advanced Autonomy

The modern-day combat cockpit is more about information and battle management than it is about airmanship in the traditional sense. Today’s missions often require operators to manage a wealth of information and timelines, quickly leading to task saturation. To redistribute cognitive workload in crewed cockpits, management of autonomy must be kept to a minimum – instead, high-level supervision of autonomous entities should be the goal as it is with crewed wingmen. However, pilots still must trust and understand how autonomous platforms are operating in the battlespace to effectively manage it. Therefore, autonomy must be derived from today’s tactics and techniques to streamline integration of uncrewed assets through familiarity.

To effectively operationalize HMT autonomy must be trusted to develop and execute solutions in tactical situations with relatively little human guidance. Building trust in combat autonomy is critical because lack of trust and cultural reluctance are the primary reasons behind delayed adoption of disruptive emerging technologies (Birkey, Deptula, & Struzriem, 2018). If the autonomy can tackle tactical problems in ways that are familiar to current aircrews, it will likely lower entry barriers. Artificial intelligence explainability will also play a key role in the trust-building process; in parallel, human-machine interfaces that efficiently convey not only current behaviors, but future actions and overall intent, will be necessary to enable efficient battle management of combat autonomy. For example, it has been demonstrated in Centaur Chess matches and other early adoption of teaming, minimizing friction at the interface is critical to overall success and is often even more important than the individual quality of human or machine (Muller, 2022). Ultimately, as autonomy is integrated into the ecosystem clear mission requirements and evaluation criteria for the technology must be established. This will allow anticipation of what data autonomy will require to operate and how it will be evaluated with operators to ensure trust.

Security

The security environment surrounding HMT will be challenging. Because of anticipated mission diversity, multitudes of information will be required by different pieces within the ecosystem. However, due to current DoD and intelligence community policies information required is heavily fragmented and will be challenging to move between platforms at the speed of battle (Javorsek, Rose, Marshall, & Leitner, 2015). As a result, new policies and standards will be required to exchange information between various parts of the system at appropriate classification levels.

Once such data sharing construct the HMT ecosystem could draw inspiration from is the Open-Air Battle Shaping (OABS) efforts utilized for Initial Operational Test and Evaluation (IOT&E) of prior 5th Generation aircraft. The OABS datalink pathway is based on the one used by Air-to-Air Range Infrastructure (AARI) and the Common Range Integrated Instrumentation System (CRIIS) that incorporates an intermediary to adjudicate outcomes between systems at different classification levels (Collins Aerospace, 2019). As a result of this system, respective up and downlinks for each vehicle can be preserved at the individual platform level but allow them to share a communized set of information across these boundaries to an approved common system. However, this comes at the cost of a required ground station that operates at the system high level to perform pairing and validate shots as well as outcomes based on combined information. Naturally, disintermediation of these messages would be desired, and a framework could be realized that offloads some of this to a software trusted guard to address the multi-level security challenge. Unfortunately, the current antiquated and analog security apparatus is not postured from a policy perspective to handle such technical solutions. However, if security policies were to change to enable the introduction of translators and abstraction layer they could be worked in concert with the connectivity and data standards concepts introduced above.

Training

Within the ecosystem the concepts of training and simulation will be very tightly coupled. Training specifically is the gradual accumulation of skills by humans, by autonomy, and by humans working with autonomy to accomplish certain goals. Much of this practice will be facilitated by simulation. The focus of this section is to describe the training process for both humans and autonomy. The following section then describes simulation's role in facilitating this training. Human operator training, specifically, refers to practice repetition operators will require flying with uncrewed platforms before leaving on a mission. This training will usually be conducted with a previously trained and validated AI/ML autonomy capability (Hunter, 2021). Human operator training in crewed-uncrewed mission constructs will be imperative to ensure operators are adept at using autonomy tools available to execute missions. Pilot training will include using the system to conduct missions both in simulation and live fly exercises. This training will require practicing hand off command and control or tasking of uncrewed platforms during a fight as well as building trust in their anticipated decisions when communications are degraded or non-existent. Also, as HMT is introduced it's not just pilots that will require training. Ground operators and other mission support personnel will require training as well. For example, mission support personnel will have to train developing, releasing, testing, and redeploying autonomy behaviors. This training will not only help familiarize personnel with the system so they can train like they fight, but it will also be a source of data for performance analytics. As such, data from training should be collected at the same or greater fidelity during missions so performance can be analyzed to improve future mission success probabilities.

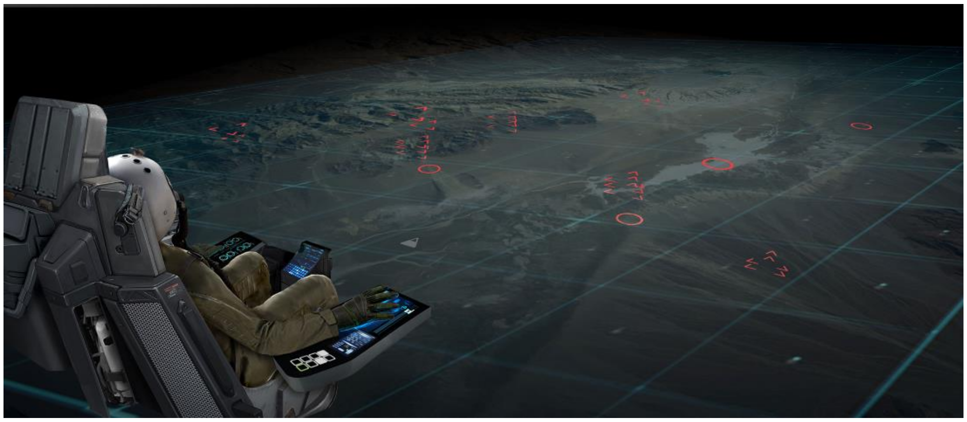

Figure 5. Simulation and training will play a key role in the autonomy ecosystem by allowing humans to experience executing missions with new autonomous teammates

Autonomy training refers to the continuous and incremental process of incorporating information and improving the autonomous AI/ML components supporting the HMT ecosystem. Autonomy training with a disciplined and bounded build-up approach will be essential for the performance of crewed-uncrewed teams, mission success, and risk mitigation and safety. Well-trained autonomous agents can contribute to mission success and operational efficiency by relieving humans from mundane duties and supporting them in critical battle management or command and control decisions. Training helps agents recognize and respond to risks and safety concerns inherent in military operations and fosters effective teamwork and mutual understanding, enhancing the team's overall performance. Training agents will require a continuous process using diverse data sources and experiences, such as simulated environments, historical military operations, real-world exercises, and mission briefs and debriefs. This process will allow AI agents to learn and adapt to new information rapidly, enabling them to keep pace with evolving threats and changing operational environments. This adaptability is crucial in dynamic and unpredictable military scenarios. Ultimately, for both human and autonomy training clearly proscribed missions and performance acceptance criteria will be required to effectively develop capabilities for the ecosystem.

Simulation

Simulation within the ecosystem refers to any computer-generated digital environment whose goal is to provide a testing ground for humans or machines. Simulation will touch on everything from using digital twins for validation of system components before flight, to training operators and autonomy agents together as one team. For humans, simulation, like shown in Figure 5, will allow them to practice and develop tactics in safe and cost-effective environments not possible for security reasons (such as reserve capabilities that rely on secrecy or surprise to be effective) or due to threat replication limitations. Simulations generated for humans will be realistic practice environments that allow aircrew to prepare for live missions. In addition, it's anticipated that humans will need to train with their autonomy so they can be ready to utilize the technology effectively. Due to the distributed systems required for the HMT ecosystem, a large emphasis will need to be given to live, virtual, and constructive (LVC) training and experimentation. Such training will permit different aspects of the mission to be blended in cost effective ways. It will also ensure training can occur using digital twins of technology as developed, providing operators the opportunity to co-evolve tactics with autonomy as well as affording them valuable opportunities to shape the new capabilities during prototyping.

Autonomy focused simulation will allow training, validation, verification, and ultimately accreditation of AI agents. However, simulation environment design will differ from traditional human simulation design. While there will be high fidelity human-like mission scenarios, simulation for autonomy may be specially designed to promote or test certain trustworthy behaviors. For example, autonomy simulations may start simple and get progressively more challenging to ensure agents are learning the right behaviors. In addition, verification and validation of autonomy may require specialized simulations where certain artifacts are introduced to try and generate a failure. These specialized techniques such as domain randomization and evolutionary search will be designed to help autonomy manage the critical simulation-to-real transition. Ultimately, synthetic environments and simulation will play an important role in training and validating autonomy in the HMT ecosystem. As a result, simulations and other digital twins used in the system will have to be well understood to ensure artifacts or overfitting to the synthetic environment are not introduced, negatively impacting autonomy behavior in the field. In the end, interoperability and open mission system standards will be key for constructing LVC simulation environments for humans and autonomy. These open standards will ensure extensible data collection and analysis to ensure both humans and autonomy are effective once deployed into the real world.

Human-Machine Interface

The HMT ecosystem envisions a new era of tactical-level collaboration between operators and uncrewed vehicles in which the barriers between operators and machines are eroded, allowing them to interact in a more natural manner to increase effectives of HMT. However, lack of trust in intelligent systems has slowed technology adoption within the DoD. Like the horse mounted cavalry of the 21st Century aircrew have been slow to adopt even the simplest of automated systems. For example, the three-decade delay between when Automatic Ground Collision Avoidance Systems (AGCAS) were technically feasible in the mid-1980’s to their deployment on the F-16 in 2014 demonstrates this lack of trust.

To effectively operationalize UCAVs there needs to be effective human-machine interfaces that work to bridge the trust gap and help humans effectively battle manage under stressful dynamic conditions. Traditionally HMIs have focused on translation between speech/touch to machine language, with less focus on deep semantic understanding on both ends of the interaction. To make crewed-uncrewed teaming effective HMIs will have to be more dynamic and responsive both to humans and machines. This will require fusion of new cockpit technologies and data to orchestrate exchanges between humans and autonomy. To accomplish this, interfaces need to combine three models when presenting information. First, a physical model of the human; second, a battlespace model; and third, a cognitive model of crewed-uncrewed teaming. These three models will combine to adaptively show human operators only the information they need during different phases of the mission. This will help alleviate some of the task saturation and allow them to effectively battle manage parts of the operation requiring their attention. However, to ensure humans, trust autonomy to operate on their behalf in the background, operators will have to be given contextual cues to current autonomy status and intent via the interface. Providing this context aware representation will require fusing information from all three states mentioned above, the necessity of this information for trust again demonstrates the importance of data and the associated standards to ensure it travels to the point of need.

Deployment Infrastructure

Deployment infrastructure relates to the end-to-end process of designing, developing, and deploying pieces of the ecosystem. To ensure the technology is effective, robust testing, validation, and tracking procedures must be in place for all components (U.S. Department of Transportation Federal Aviation Administration, 2015). To complicate matters, components of the ecosystem will all have varying requirements due to role and provider heterogeneity. For example, hardware, software, and humans will all require different testing procedures. As the DoD comes closer to making HMT an operational reality, the logistics of validating AI/ML behaviors, certifying pilots to operate with autonomy, and methods for updating autonomy or data interfaces will be required. To do this effectively the DoD needs to understand how all these pieces fit together so appropriate organizations can be established and empowered to establish strong policy, standards, and management practices between different providers contributing to the ecosystem. With out these steering organizations, pieces of the puzzle will not fit in a harmonious way.

Operations

Ultimately, autonomy tied to UCAVs must be verified in a wide variety of environments to be deployed in an operational setting on live platforms. From test and evaluation to training and tactical operations UCAV platforms will have to show fitness and trustworthy behavior before they have any hope of making it to the field. For background, crewed platforms always benefit from the presence of the aircrew with semi-autonomous modes and the inherent override capabilities in the system. The mere concept that the aircrew can take over if trust or performance is ever brought into question is reassuring to those performing the risk acceptance on such systems. To alleviate concerns about uncrewed platforms they will have to undergo rigorous testing before they are deployed.

These challenges will likely be mitigated using many of the systems currently employed in both the test and training communities already. Significant resources are often devoted to live, virtual, and constructive (LVC) environments that allow for emulation of adversary capabilities, virtual deployment of high-value, low-availability assets, and refinement of tactics and techniques on an experimental basis. The HMT ecosystem will greatly benefit from ongoing LVC investments so long as the autonomy and other key aspects are positioned in a way to capitalize on them. LVC environments must support the appropriate level of data collection and run-time environments to integrate uncrewed assets, whether they are live or virtual autonomous entities participating in either test events or exercises.

Once tested and validated, working at the deployment level such autonomous systems will be scaled to much larger and more complex environments. There will be natural extensions of trusted versions of the autonomy, especially if it is possible to maintain trust at higher levels of abstraction. A significant benefit of the hardware-enabled, software-defined autonomy within the HMT ecosystem is that underlying approaches are often agnostic to specific applications. This means that in principle, autonomous agents developed for UCAVs, might also find use in higher levels of the Joint All Domain Command and Control (JADC2) framework.

While highlighting the potential applications beyond just UCAVs, trust in these algorithms will enter increasingly difficult scenarios harder to script and test using traditional methods of bounding and building from simple to more complex individual engagements. This is because battle management, command and control (BMC2) complexity jump quickly to a level often incomprehensible and beyond human imagination. However, as mentioned before, the willingness to accept risk and move quickly into other domains or accept non-intuitive cross-domain solutions will likely be heavily dependent on the threat context in which DoD finds itself. In the end, the HMT use case for autonomy is a vast jump in complexity for AI/ML technology. To develop and eventually deploy HMT technology in an operational setting, it will take considerable coordination between suppliers and the government to adhere to common standards. By seeking to define and understand high level components of the ecosystem during development sound policy, standards and management practices can be implemented from the onset, helping move this valuable technology closer to the warfighter.

Table 1. Key PSMA Ecosystem Impacts

PSMA Category	Dominant Factor	Ecosystem Impact
Policy	Challenge	Security classification architecture limits information sharing between ecosystem components, potentially negatively impacting performance
Standards	Insight	DoD is moving towards open mission system architectures; this transformation will help aid ecosystem standards development
Management	Risk	Requirements based management practices are not suited for the software centric ecosystem, failure to adapt will lead to costly delays

EFFECTIVE POLICY, STANDARDS, AND MANAGEMENT PRACTICES FOR HMT ECOSYSTEM

With the stage set for all the major elements of the ecosystem defined, it is helpful to discuss the challenges, insights and risks surrounding some of the key policies, standards, and management practices that will undoubtedly influence HMT success. To avoid exceeding the scope of the current paper, we have chosen to limit the discussion to the dominant factors influencing each category in order. In what follows, we characterize each as a challenge, an insight, or a risk depending on our subjective assessment of the current state of the nascent HMT ecosystem. These ecosystem impacts are summarized in Table 1.

Policy – Challenge: The current security classification architecture and its associated processes, will likely prove to be the largest policy challenge for the effort to overcome if it is to realize success. The inherent nature of compartmented programs with fundamentally different structures as applied to Title 10 (military) and Title 50 (intelligence community) authorities, has plagued the successful employment of combat systems for decades. In prior acquisitions programs, the 20th Century fixation on exquisite platforms, operating in isolation, served as a benchmark of western defense departments and limited the extent of the negative effects from differing Title 10 and Title 50 approaches. In fact, the near duplication of structures (Special Access Program Facilities, SAPF, for Title 10 and Sensitive Compartmented Information Facilities, SCIF, for Title 50) extends well beyond their just their facilities and associated programs. More importantly, it greatly influences the information technology (IT) and networks upon which the data reside. Even though the facilities are often built to nearly the same specifications they have different accreditation processes, risk acceptance personnel, and rarely occupy the same spaces. In the data-driven 21st Century, this already poses a formidable challenge that will certainly be amplified by the HMT ecosystem. Designed from the beginning to require a level of connectivity, automation, and collection capability never before realized in an operational platform, HMT with UCAVs will be particularly susceptible to these policies. The public statements that teaming will occur with multiple crewed aircraft, that include the F-35, will be particularly problematic since that aircraft is unique in its own right. More than any other aircraft in history, the F-35 is influenced not only by the US Air Force, the US Navy, and the US Marine Corps but also several foreign countries. Coupled with the state-of-the-art collection capabilities of modern multi-function arrays and processors, even the most basic of software-defined radio frequency (RF) systems today blur the traditional line separating Title 10 and Title 50 collection. Without some willingness to modernize the policies surrounding security classification management and governance, HMT with UCAVs will be limited to extremely narrow applications of autonomy as a rather simple weapons truck or decoy. If this is the first instantiation of the platforms, some of the early development will undoubtedly not be affected as much it otherwise would be, if the missions were extended to Electronic Warfare or Distributed Sensing. Since the early stages of the effort will take a more traditional, 20th Century, platform-centric approach, the impact of current security classification management policies will not be as great as they will later in the evolution of the program. The silver lining on all of these security classification policy challenges is that once DoD obtains the right leadership, with the courage to tackle the security establishment and modernize these policies, the stroke of a pen will open up a remarkable array of possibilities for adaptability and data sharing.

Standards - Insight: The defense industrial base's adoption of open architectures across a wide array of applications has been an extremely refreshing and positive move for DoD acquisitions. Since the fall of the Berlin wall, DoD has conditioned government contractors to tightly couple, and vendor lock their hardware with their software. This effectively removed the competition so critical for market economies and was exacerbated by the "Last Supper" event hosted in 1993 by then Secretary of Defense Les Aspin. Amidst the budget cuts following the end of the Cold War, Secretary Aspin invited the CEOs of America's largest defense contractors to a secret dinner where he urged his guests to consolidate to stay in business (Tirpak, The Distillation of the Defense Industry, 1998). This led to a hollowing out of the defense sector that has left us with the few remaining prime contractors we see today. In that process, these companies began to look at any opportunity to merge and began the 5th Generation game of winner take all, large aircraft acquisitions programs. To the credit of the Rapid Capabilities Office (RCO), they made considerable strides at reversing this trend by instituting a government-owned Open Mission Systems (OMS) and Universal Command and Control Interface (UCI) architecture on their platforms that helped address this vendor lock problem. Programs at DARPA such as the CONCERTO and STITCHES efforts mentioned above, helped accelerate the adoption of government reference architectures that will enable the democratization of software on uncrewed, and other, future aircraft. This is one of the most important developments in the 21st Century and will greatly benefit not only the HMT ecosystem, but it also sets the stage for a healthier defense industrial base characterized by a small number of hardware providers and complemented by a rich software development marketplace.

Management Practices - Risk: The 20th Century acquisition approach emphasizing platforms made of atoms was a hallmark of the Cold War and conditioned us to serial development, testing, and fielding of systems. In its defense, this process worked well for things that come off a physical assembly line. It was also well suited to most of our earlier aircraft. In this acquisitions model, requirements came to rule the day, but as complexity rose, so too did costs and such changes often also led to schedule slips. These programmatic slips in schedule were terrible for a system designed against the projection of a possible threat, since higher costs translated into fewer platforms making us less willing to accept attrition. Coupled with longer timelines, this resulted in the projection of more uncertainty of threat assessments. Unfortunately, such changes in threat potential, then led to changes in the requirements that often sets up a vicious cycle of perpetual slips and cost overruns that have become commonplace on modern programs. One strategy to combat such features, is to design an acquisitions program that is focused on visionary attributes instead of traditional hard requirements. When coupled with the delegation of acquisitions risk acceptance to the combined test force commander at the Field Grade Officer (FGO) level, this is a remarkable opportunity to break the traditional acquisition cycle discussed above. With the strong support from Secretary of the Air Force Frank Kendall, embracing management practices such as this will produce the best chance the program will have to survive. Like all organizations with institutional inertia and cognitive entrenchment, this approach will be viewed as threatening from within because of the uniqueness in the method and the challenge it poses to the status quo (Gvindarajan & Trimble, 2010). This represents considerable risk to the success of the effort since, if allowed to follow a more conventional acquisition path, HMT with UCAVs will undoubtedly be in jeopardy. However, if pursued as described above, the current political environment and senior leader support suggests that there is no better time than now for HMT ecosystem to be successful.

CONCLUSIONS AND FUTURE WORK

Even with billions of dollars in investment the DoD has yet to effectively operationalize AI/ML at a meaningful level for warfighters. To truly realize the potential of this technology it needs to be applied to a complex-real world operational use case. Towards this goal the US DoD is beginning to develop uncrewed assets which will leverage autonomy and interconnected crewed-uncrewed teaming to operate effectively in denied contested environments. This paper categorized and described each of the key elements required to make a HMT ecosystem using UCAVs to enable futuristic missions. Elements described such as uncrewed aircraft, crewed aircraft, advanced sensing methods, high performance edge computing, crewed-uncrewed teaming, artificial intelligence aircraft control algorithms, simulation-based training for operators, and rapid AI algorithm training will require seamless integration to make these visions a realty. A common thread connecting pieces within the system is the need for standardized flow of data and communized tracking of performance metrics. Goals that can only be accomplished for such a heterogeneous system through communized PSMA principals. While PSMA principals like Open Mission Systems (OMS) architectures has helped push standardization in the right direction, there exist other institutional roadblocks to sharing data between entities in the ecosystem. To overcome these roadblocks as well as antiquated winner take all management practices leadership will have to change existing dysfunctional practices to ensure success of the program. Moving forward, the removal or entrenchment of these practices and their impact on the program should be tracked to demonstrate how PSMA principals directly impact programs success or failure. Ultimately, concepts and recommendations highlighted in this paper help leadership understand steps required to move promising AI/ML closer to the hands of the warfighter.

REFERENCES

Ahamed, A., Tariq, U., & Nusir, M. (2019). Real-Time Methodology for Improving Cyber Security in Internet of Things Using Edge Computing During Attack Threats. *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT).* IEEE.

Birkey, D., Deptula, L. G., & Struzriem, M. G. (2018, 07). Manned-Unmanned Aircraft Teaming: Taking Combat Airpower to the Next Level. *Mitchell Policy Papers, 15*. Retrieved from https://defense.info/wp-content/uploads/2018/08/Manned-Unmanned-Teaming.pdf

Case, N. (2018). *How To Become A Centaur*. Retrieved from JoDS: https://jods.mitpress.mit.edu/pub/issue3-case/release/6

Collins Aerospace. (2019, 08 08). *Collins Aerospace to enhance CRIIS software and operator consoles for Air Force, Navy and Army test ranges*. Retrieved from Collins Aerospace: https://www.collinsaerospace.com/news/news/2019/08/collins-enhance-criis-software-operator-consoles-air-force-navy-army-test-ranges

DARPA. (2020, 09 18). *Creating Cross-Domain Kill Webs in Real Time*. Retrieved from DARPA: https://www.darpa.mil/news-events/2020-09-18a

DARPA. (2023, 06 04). *Air Combat Evolution((ACE)*. Retrieved from DARPA: https://www.darpa.mil/program/air-combat-evolution

DARPA. (n.d.). *System of Systems Integration Technology and Experimentation (SoSITE)*. Retrieved from DARPA: https://www.darpa.mil/program/system-of-systems-integration-technology-and-experimentation

Department of Defense. (2023). *Strategy for a Net-Centric, Service Oriented DoD Enterprise.* DoD CIO. Retrieved from https://dodcio.defense.gov/portals/0/documents/dod_netcentricservicesstrategy.pdf

Deptula, L., & Penny, U. (2019). *Mosaic Warfare*. Retrieved from https:www.airforcemag.com-article-mosaic-warfare

Everstine, B. W. (2018, 01 29). *The F-22 and the F-35 Are Struggling to Talk to Each Other … And to the Rest of USAF*. Retrieved from Air & Space Forces Magazine: https://www.airandspaceforces.com/article/the-f-22-and-the-f-35-are-struggling-to-talk-to-each-other-and-to-the-rest-of-usaf/

Grana, J. (2021). *FINDINGS ON MOSAIC WARFARE FROM A COLONEL BLOTTO GAME.* Nielsen Book Data.

Gvindarajan, V., & Trimble, C. (2010). *The Other Side of Innovation: Solving the Execution Challenge.* Harvard Business Review Press.

Hunter, J. (2021, 08 03). *The Truth About The Air Force's Biggest Changes To Pilot Training Since The Dawn Of The Jet Age*. Retrieved from The Drive: https://www.thedrive.com/the-war-zone/41789/the-truth-about-the-air-forces-biggest-changes-to-aircrew-training-since-the-dawn-of-the-jet-age

Javorsek, D., Rose, J., Marshall, C., & Leitner, P. (2015, 08 05). A Formal Risk-Effectiveness Analysis Proposal for the Compartmentalized Intelligence Security Structure. *International Journal of Intelligence and Counter Intelligence, 28*(4), 734-761. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/08850607.2015.1051830

Jensen, B., & Paschkewitz, J. (2019). *Mosaic Warfare: Small and Scalable are Beautiful*. Retrieved from https://warontherocks.com/2019/12/mosaic-warfare-small-and-scalable-are-beautiful/

MacAllister, A., Rupp, P., Garrison, J., Hellstern, G., Javorsek, C. D., & Chu, P. (2021). Using Machine Learning for Battle Management Analysis. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, (p. Paper No. 21150 Page 1 of 14).

Muller, E. (2022). *How AI-Human Symbiotes May Reinvent Innovation and What the New Centaurs Will Mean for Cities.* Retrieved from https://www.scirp.org/journal/paperinformation.aspx?paperid=115255

NATIONAL SCIENCE AND TECHNOLOGY COUNCIL. (2022). *LESSONS LEARNED FROM FEDERAL USE OF CLOUD COMPUTING TO SUPPORT ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT.* MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2022/07/07-2022-Lessons-Learned-Cloud-for-AI-July2022.pdf

O'Donougue, N. (2021). *DISTRIBUTED KILL CHAINS: Drawing insights for mosaic warfare from the immune system and from the navy.* RAND Corporation.

Penney, H. (2022, 12 02). *Crewed-Uncrewed Teaming*. Retrieved from Air & Space Forces: https://www.airandspaceforces.com/article/crewed-uncrewed-teaming/

Qiu, M., Kung, S.-Y., & Gai, K. (2020). Intelligent security and optimization in Edge/Fog Computing. *Future Generation Computer Systems*.

Secretary of the Air Force Public Affairs. (2023, 03 07). *Kendall outlines clear path for modernizing Air, Space Forces & confronting peer competitors*. Retrieved from Air Force: https://www.af.mil/News/Article-Display/Article/3321908/kendall-outlines-clear-path-for-modernizing-air-space-forces-confronting-peer-c/

Sweetser, A., & Bexfield, J. (2020). *Lessons from the MORS Workshop on Advancing Campaign Analytics on JSTOR.* Military Operations Research Society.

Tirpak, J. A. (1998, 07 01). *The Distillation of the Defense Industry*. Retrieved from Air & Space Forces Magazine: https://www.airandspaceforces.com/article/0798industry/

Tirpak, J. A. (2023, 03 29). *Collaborative Combat Aircraft Will Join the Air Force Before NGAD*. Retrieved from Air & Space Forces: https://www.airandspaceforces.com/collaborative-combat-aircraft-ngad-timeline/

U.S. Department of Defense. (2023, 06 04). *Project Maven to Deploy Computer Algorithms to War Zone by Year's End*. Retrieved from U.S. Department of Defense: https://www.defense.gov/News/News-Stories/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/

U.S. Department of Transportation Federal Aviation Administration. (2015). *Advanced Verification Methods for Safety-Critical Airborne Electronic Hardware.* Retrieved from https://www.faa.gov/aircraft/air_cert/design_approvals/air_software/media/TC-14-41.pdf

United States Government Accountability Office. (2022). *Artificial intelligence: DOD Should Improve Strategies, Inventory, Process, and Collaboration Guidance.* Retrieved from https://www.gao.gov/assets/gao-22-105834.pdf