Paper Title: Separating Buzz from Method: The AI Revolution in Military Training

Jenna Tuck
Bohemia Interactive Simulation
Orlando, FL
Jenna.Tuck@bis.im

ABSTRACT

In an era where technological advancement is pivotal to maintaining strategic superiority, the application of Artificial Intelligence (AI) in military training systems represents a novel frontier of limited exploitation. The use of the term AI often obscures the complexity and considerations of implementation, creating a gap between promised and realized capabilities. This paper describes the accepted classifications of AI and suggests tangible applications to enhance training efficacy, inform future requirements, leverage big data, reduce training overhead, increase modeling and simulation event cadence, and assess echelon readiness. It delineates the strengths of different AI technologies and explores their practical application to close modeling and simulation (M&S) gaps that current capabilities cannot address.

Drawing from existing literature and anecdotal insights, this paper provides an overview of AI's practical applications within modeling and training systems. By leveraging numerous examples, it illustrates AI's potential to enhance operational readiness and advance training objectives. From adaptive training modules that personalize experiences for individual learners to simulating complex scenarios for collective training and evaluation, the paper examines the impacts that AI can deliver to the M&S domain. It clarifies common misconceptions about AI's capabilities while emphasizing the paradigm-shifting potential of these technologies to inform procurement requests and responses.

Indiscriminate use of imprecise terminology must be shunned in favor of precise terminology, achievable only by furthering understanding of AI's capabilities. By demystifying AI, advocating for its purposeful application, and providing industry-accepted terminology, this paper showcases the transformative power of AI in military training. It advocates a strategic approach prioritizing meaningful, outcome-driven application of AI, aiming to be a reference source for stakeholders seeking to leverage AI as a powerful mechanism to achieve enhanced training outcomes, reduce M&S overhead, and better position friendly forces for the complexities of modern warfare.

ABOUT THE AUTHORS

Jenna Tuck
Jenna Tuck, in her role as the Senior Director of Partnerships for Bohemia Interactive Simulations, leverages a deep understanding of AI/ML, enterprise architecture, and Industry 4.0 in shaping defense and strategic development. Jenna has devoted over a decade to pursuing revolutionary technologies, identifying and fostering strategic relationships, and working on the bleeding edge of the Modeling and Simulation community. Jenna has a demonstrated understanding of the current and prospective landscape of the domain. As a respected voice in her field, Jenna actively participates in high-level industry events, as well as working groups, to contribute and learn within the industry. Jenna actively works to bridge the gap between theoretical and real-world applications, consistently offering a grounded, yet forward-thinking perspective.

Paper Title: Separating Buzz from Method: The AI Revolution in Military Training

Jenna Tuck
Bohemia Interactive Simulation
Orlando, FL
Jenna.Tuck@bis.im

THE REALITY OF CURRENT AFFAIRS

AI stands poised to redefine every facet of Modeling and Simulation (M&S). AI presents an exponential potential for new and/or improved capabilities, and with the possibility of supporting two simultaneous military fronts, in Eastern Europe and the Indo-Pacific region, the US Military must increase the speed and effectiveness of training to maintain overmatch over our adversaries. This imperative, combined with the potential that AI can bring to bear, has sent the M&S community into a frenzy of ideas, possibilities, research, and (often primarily) buzzwords.

As the industry continues to grapple with the power of AI, the recent attention on AI has taken the industry by storm. Yet much of the effort has yielded only limited results, often with compliance linked to a poorly defined requirement, which was defined based on little more than a desire to 'sprinkle AI' on the problem. Additionally, procurement and marketing efforts are often focused more on wedging some mention of AI into a solicitation/product/solution and less on identifying how to use AI to provide quantifiable benefits with tangible outcomes by properly analyzing the problem. The assertion of this paper is that the current environment is due in large part to a fundamental misunderstanding of AI, when and how AI can provide benefit(s), and their respective limitations.

In 10 pages this paper cannot delve deeply into all of the jargon and nuances about AI. Instead, this paper seeks to convey a baseline understanding of different AI techniques and their potential application in military simulation and training. To achieve that objective, AI types and techniques will be described, as well as thoughts for developing a coherent AI Strategy. This paper will explore the interplay between AI and Data strategies to establish a methodology to deliver capabilities that offer measurable improvements to training capabilities. Finally, this paper will provide actionable recommendations for taking AI from a buzzword and a checkbox to something that provides quantifiable, real, benefits. AI isn't a magic fix for unsolved problems; however, AI provides tools to revolutionize the world. Our adversaries are actively pouring research into this field. It's time AI stopped being a buzzword and started being applied properly.

AI THE BUZZWORDS

AI is a term that has quickly become a ubiquitous buzzword in the technology and defense sectors. However, However, amidst the excitement and speculation, few parts of the industry truly understand the fundamentals of AI technology. An 'expert model' is fundamentally different from a 'machine learning' algorithm, but both are called 'AI'. For example, what is called 'AI' in a video game is often just an algorithm and a hierarchical behavior tree, and most AI that people interact with is simply a rule-based system behaving exactly as a human defined it. This epitomizes the AI juxtaposition: marketing has created a facade of what AI is and to conform to that mirage, companies claim nearly anything is 'AI'.

AI

Simply put, AI is "the simulation of human intelligence in machines that can perform tasks usually requiring human intellect." (Institute of Data, 2023)

Capability

Perhaps the simplest categorization, AI can be categorized by *its capability* in three (3) ways, of which only ***one*** is realized to date on a scale which could be used widely:

- **Artificial Narrow Intelligence (ANI):** also referred to as 'weak' AI, and currently the only AI achieved to date; ANI systems "execute *specific* tasks using human-like abilities" (Kahn, 2021).
- **Artificial General Intelligence (AGI):** completely hypothetical to date, AGI systems would operate like human beings, to include the ability to "learn, function, understand, and operate like human beings" (Kahn, 2021).
- **Artificial Super Intelligence (ASI):** also completely theoretical, ASI systems would "surpass human intelligence in virtually every aspect." (Institute of Data, 2023)

Functionality

AI can also be categorized by the functionality (the scope and capability) of the AI system. There are four types as defined by Arend Hintze (Knight, 2016), and range from the least to the most complex.

- **Reactive Machines** are the most basic type of AI system and are designed to respond to specific inputs with predefined outputs. These systems do not have memory-based functionality or the ability to use past experiences to inform future decisions**.** ***For example,*** Automated turrets that detect enemy vehicles using sensors and immediately fire at them. These turrets operate based on predefined rules and do not adapt or learn from previous encounters.
- **Limited Memory Machines** are slightly more sophisticated than reactive machines in that they can store past 'experiences' or data for a short period to improve decision-making capabilities. However, these systems cannot store large amounts of data to build an 'experience base' the way for example a human would**.** ***For example,*** **s**urveillance drones that remember the last few minutes of flight data to adjust their path and avoid obstacles. These drones can make better short-term decisions based on recent information but do not retain knowledge from past missions.
- **Theory of Mind** is largely theoretical to date and is characterized by systems that would be able to understand emotions, beliefs, and other human mental states of being, and in turn be able to utilize this understanding to interact with humans more effectively and be able to more accurately predict human behavior. ***For example,*** virtual training assistants that adapt to the emotional and cognitive states of soldiers. If a trainee shows signs of frustration, the AI adjusts its training approach to provide additional support or change its teaching style.
- **Self-aware AI** is completely theoretical to date, and is the most advanced form of AI, possessing self-awareness and consciousness. These systems can understand their state and potentially have a sense of self. ***For example,*** in the future we may see advanced command and control systems that autonomously coordinate operations while being aware of its' capabilities and limitations, with the ability to make strategic decisions, delegate tasks, and request additional resources if necessary–all while understanding its' role and impact on mission success.

Functionality Types

Not to be confused with functionality, AI systems can also be categorized by their functionality *types*, and are rooted in the data analytics methodology which is applied as part of the AI. One exception, however, is Generative AI.

- **Descriptive AI** aligns closely with traditional data analytics, as it involves summarizing and interpreting past data to provide insights into what has occurred. ***For example***, the collection and summary of data from a soldier's past marksmanship training sessions could provide insights into shooting accuracy, reaction times, and consistency under various conditions.
- **Diagnostic AI** uses data analytics to identify problems and determine why they occurred by analyzing data patterns and relationships. ***For example,*** the analysis of soldier performance data to identify specific issues, such as a tendency to miss targets under low-light conditions or difficulty maintaining accuracy after prolonged sessions, pinpointing potential underlying causes like hand steadiness or sight alignment problems.
- **Prescriptive AI** suggests actions to achieve desired outcomes, combining insights from descriptive, diagnostic, and predictive analytics. ***For example, r***ecommending a customized training regimen for a soldier suggesting specific drills to improve low-light shooting accuracy, and exercises to enhance endurance and hand steadiness. It might also recommend incremental exposure to moving targets to build proficiency gradually.

- **Cognitive AI** mimics human thought processes, including learning, reasoning, and problem-solving. ***For example,*** a virtual coach during the soldier's training sessions provides real-time feedback on form and technique, offering tips or links to reference material or training aids.
- **Predictive AI** uses historical data to forecast future events/trends. ***For example,*** the use of historical performance data to forecast future challenges that the soldier might face, such as predicting the soldier has a high probability of struggling with moving targets based on past difficulties with similar exercises and allowing trainers to anticipate and prepare targeted interventions.
- **Generative AI*:*** Generative AI emerged on the scene in 2014 with the invention of Generative Adversarial Networks (GANs) by Ian Goodfellow and is the stand-out from the other five (5) functionality types as it is not based primarily on traditional data analytics methods. Instead, it focuses on creating new content by learning patterns from existing data. "Generative AI uses GANs and transformers" (LeewayHertz, n.d.) to create new content. Data analytics are generally used in tandem with Generative AI to inform what kind of content needs to be created***, for example***, the creation of dynamic and realistic virtual marksmanship scenarios tailored to the soldier's training needs. These scenarios could adapt in real time to the soldiers' actions, presenting increasingly challenging targets or varying environmental conditions to ensure a comprehensive training experience.

To illustrate how these different AI functionalities can be applied within military training content, we will follow one example throughout and detail it in **Figure 1,** which provides the basis for additional examples in subsequent sections as we discuss additional methods of categorizing types of AI.

Technology

- **Machine Learning (ML) "**is the general term for when computers learn from data" (Brown, 2021) and at its core includes the use of algorithms and statistical models to "train" a system to perform a specific task that it wasn't explicitly programmed to do by recognizing patterns in data and making predictions as new data arrives.
- **Deep Learning (DL)** is a subset of Machine Learning, so technically DL is ML, but ML isn't necessarily DL. DL is the use of algorithms that can draw conclusions and analyze data "with a logical structure similar to how a human would conclude" (Wolfewicz,2023). This is done through the use of multiple layers of algorithms—called Artificial Neural Networks (ANN)—which were modeled on the neural network of a human brain. This means that the system can understand complex patterns in data and ingest unstructured data.
- **Natural Language Processing (NLP)** "deals with how computers understand, process, and manipulate human languages" according to the Network of the National Library of Medicine (NNLM). While not technically a subset of machine learning, NLP often uses both ML and DL to perform tasks associated with the interpretation of meaning, translation of languages, recognition of patterns, and classification (to name just a few.)
- **Computer Vision** is the simulation of human 'vision' in a machine, ultimately enabling the said machine to "interpret and analyze the visual world" (2024), via visual inputs from eye tracking to posture and environmental placement.
- **Robotics** involves designing and using machines (robots) to perform tasks or interact with the environment that typically would require human intervention. Now, whilst robotics is not solely AI technology; the application of other AI technologies within robotics such as machine learning, computer vision, and NLP have fundamentally revolutionized the possibilities in the application(s) of robots as they become increasingly capable of replacing typically human tasks.
- **Expert Systems** in its meaning is very simple; a machine (or system) that simulates the judgment, behavior, and or knowledge of an 'expert' in a particular field. The premise hasn't changed much since it was born by Dr. Edward Feigenbaum in the 1970s to support chemists in determining molecular structures (Dennis, 2024), however with developments in things like NLP, DL, ML, and more; what took Dr. Feigenbaum 10 years to develop can now be achieved in the time it takes an expert to input the rules into the system.

There are *many* other ways of classifying AI such as; by learning type (supervised/unsupervised/semi-supervised/reinforcement); by model of interaction (human in the loop/on the loop/out of the loop); automated vs. autonomous; and by so on—possessing a baseline understanding of AI as it relates to capability, functionality type, functionality, and technology provides the fundamental understanding necessary to look at *how* AI can be used in application to harness AI in meaningful ways—even if those ways are minimal at first.

AI IN PRACTICE

This section is less academic and more practical, building on the knowledge from the above sections, and provides specific examples of how AI can support and enhance military training, focusing on practical, real-world applications that underscore the strategic integration of AI technologies.

For these examples we will use a soldier; Private First Class (PFC) Smith. PFC Smith is completing their basic marksmanship course at Ft. Liberty, and it is assumed that there is a virtual marksmanship trainer that PFC Smith will be using.

- **Targeted Training:** PFC Smith uses a virtual marksmanship trainer with an EXCON station which is bidirectionally connected to the Army Training Information System (ATIS). This station uses a trained Artificial Neural Network (ANN) to analyze his past scores, comparing them with data from thousands of other students. The AI identifies that PFC Smith is below average in a particular task but excels in others. The system flags these areas for targeted training.
  - ***Impact of Capability:*** The AI system identifies specific weaknesses and strengths, enabling personalized training that addresses individual needs.
- **Class Assessment:** A deep learning algorithm continuously assesses PFC Smith's entire class using historical student data. This algorithm determines whether PFC Smith's deficiency in a task is common among his peers (indicating a potential training delivery issue) or if he is an outlier.
  - ***Impact of Capability:*** Identifying training deficiencies helps improve the training curriculum, while pinpointing individual outliers allows for targeted remediation, leading to better overall training outcomes.
- **Virtual Tutoring:** If PFC Smith is identified as an outlier, his virtual training is tailored to his needs by a virtual tutor—an expert model that uses past data to suggest the best course of action to remedy the deficiency. For instance, if similar students historically responded better to a different teaching method, the virtual tutor implements this approach.
  - ***Impact of Capability:*** Customized remediation for PFC Smith enhances training effectiveness for all students and supports instructors by providing targeted insights and strategies for each trainee.
- **Instructor Feedback:** Instructors receive AI-generated feedback on their training methods. The AI identifies patterns, such as lower student performance on rainy days, or subtle habits of expert instructors that contribute to better student outcomes.
  - ***Impact of Capability:*** The AI system helps instructors improve their training techniques by highlighting effective practices and areas for improvement that might not be obvious.
- **Wargaming:** Another example is a wargame designed to inform future force structure and procurement decisions. A genetic algorithm (a type of AI algorithm), monitored by a trained ANN, excels at finding unorthodox ways to solve problems. The Ukraine war shows that asymmetrical warfare in unexpected forms can play a massive role, and demonstrates that the battlefield of today wasn't necessarily envisioned, let alone the battlefield of tomorrow. Letting an AI algorithm 'evolve' tens of thousands of times without human interaction with the sole goal of defeating an opponent takes hours or days—compared with running wargames a handful of times a year. It has the added benefit of rooting out issues in current modeling and simulation tools, as the incredibly odd strategies these algorithms devise will find any weaknesses in the tools. For example, if you provide a genetic algorithm with a set amount of resources and an objective, it will randomly choose how to utilize the resources and will fail miserably at first. However, with hundreds or thousands of 'runs', each a 'genetic branch,' an ANN will grade each result according to rules and 'kill' nearly all of the branches. The top few will be permuted dozens of times, and all run again. After many iterations of this process (referred to as generations), an efficient solution will be generated. A human then reviews it and modifies the weights (or adds nodes/rules) on the ANN.
  - ***Impact of Capability:*** The use of AI in wargaming effectively creates a turbocharged wargame where AI will attempt thousands of ideas, many so unorthodox no human would even consider it. This accelerates the development of innovative strategies and improves the robustness of simulation tools, preparing the military for unexpected future conflicts, and identifying highly innovative approaches
- **Pilot Training**: As a final example, consider the habits of a highly experienced fighter pilot versus one fresh out of flight school. The experienced pilot unconsciously performs tasks such as checking their instruments in a particular fashion, which inexperienced pilots do not. Both groups look at key instruments roughly the same number of times; however, experienced pilots watch different instruments longer and have very calculated gaze

patterns. This particular difference between skill levels has been observed and documented. However, there are many undiscovered techniques used by experts in every field that aren't taught because they're not consciously recognized by experts. AI excels at finding patterns such as this and, when given data between an inexperienced and experienced individual, can identify the deltas. This facilitates better training and reduces time until mastery. (Mohan, 2020) (Babu, Jeevitha, Shree, Prabhakar, Saluja, Pashilkar, Biswasb, 2019)

- ***Impact of Capability:*** AI identifies undiscovered techniques used by experts, which in turn can be defined to facilitate better training programs and reduce the time required for new pilots to achieve mastery.

DATA STRATEGY VS. AI STRATEGY – THE ICEBERG MODEL

An AI strategy and a data strategy are not one and the same. In definition, an AI Strategy is "simply a plan for integrating AI into an organization so that it aligns with and supports the broader goals of the business[organization]] (Finio, 2024), and a Data Strategy is a "plan that defines the technology, processes, people, and rules required to manage an organization's information assets[data]" (Amazon, n.d.) (In recent years, there has been a growing understanding that AI requires data, leading to a greater appreciation for the importance of a clearly defined data strategy. However, the absence of a cohesive AI strategy alongside a data strategy often fails, commonly referred to as falling into the "valley of death."

While a data strategy focuses on the collection, management, and analysis of data, an AI strategy defines how the data is utilized. It includes setting goals, allocating budgets, securing resources, and defining the necessary steps to implement AI solutions effectively. Moreover, it involves identifying key stakeholders who will be crucial in navigating the complexities and challenges of deploying AI technologies. And thus, having both an AI strategy, therefore, is vital for ensuring that AI initiatives not only cross the "valley of death" but also achieve sustainable and impactful results.

AI Strategy

Developing an AI strategy can seem daunting, but with time, resources, expertise, and cooperation one can set themselves up for success. Having a strategy in place significantly increases your chances of achieving meaningful outcomes with AI. This paper aims to describe the core components necessary for developing an AI strategy, explain their importance, and address some of the common challenges and hurdles others have experienced.

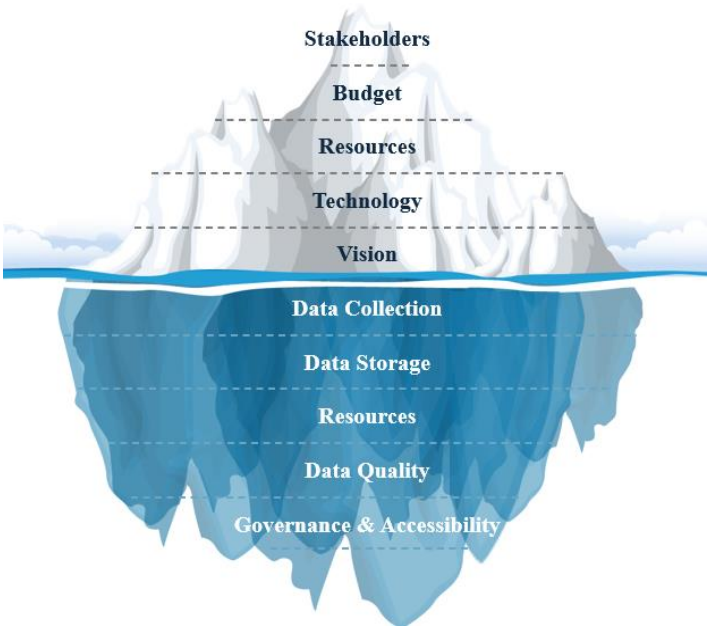

Figure 1. AI Strategy Iceberg Model

The **AI Strategy Iceberg Model (Figure 1):** If we were to use the metaphor of a successfully implemented AI System as an Iceberg, the visible portion above the water--and working its way up from the waterline in order of importance–would represent the AI Strategy that enabled the successful initiative. The part below the surface that fundamentally supports the full weight of the AI strategy is the data and Data Strategy, working down from the waterline in order of importance. Just as the bulk of an iceberg lies hidden beneath the water, the success of an AI strategy heavily relies on the quality and structure of the underlying data. Without a robust data strategy, the AI strategy above would lack the necessary support and foundation to thrive.

It's important to note that there are numerous sub-layers and complexities within each component, but focusing on the main elements:

Vision: The most important component of any AI Strategy, vision, defines what one is trying to achieve and why one is trying to achieve it which in turn will drive all other components of the strategy. AI is a tool to achieve a particular vision in the same way you might use Excel to more efficiently create a budget.

- **Common Challenges** include failing to clearly define objectives and the driving motivation behind them. This lack of clarity can lead to over engineering or scope creep. Using AI without a defined goal will not result in quantifiable value.
- **Questions to Answer:**
  - What, specifically, are we trying to achieve with this AI implementation?
  - Who will use the AI system, and who will benefit from this capability/improvement?
  - Why is it beneficial to pursue this AI project? And what specific improvements or changes are expected?

Technology: Choosing the right AI technology is pivotal to the success of your AI initiative. The type of AI technology you select should be *driven* by the vision and specific objectives defined in the vision. As discussed above, different AI technologies offer different capabilities—and in turn require different things. Understanding what their capabilities aree will help inform decisions about which technology to leverage, and thus what resources are required.

- **Common Challenges** include; selecting technology that doesn't align with the vision, failing to consider the ease of integration with existing infrastructure, "reinventing the wheel", underestimating the complexity of certain AI technologies, or not understanding the limitations and strengths of different AI approaches. It is crucial to research and evaluate the technologies that best fit the vision to ensure the desired outcomes can be achieved effectively.
- **Questions to Answer**
  - What type of AI technology best aligns with our vision and objectives (e.g., ML, NLP, Computer Vision)?
  - What specific functionalities are required to achieve our goals (e.g., predictive analytics, automation, image recognition, conversational agents)?
  - How will this AI technology integrate with our existing systems and workflows? And what changes or upgrades will be necessary to support the integration?
  - What are the future-proofing considerations to ensure long-term viability?
  - Are there limitations to what technology we can use based on the intended user (e.g., can we use open-source capabilities or do we need an isolated instantiation for cybersecurity reasons)

Resources: Having a clear understanding of the resources available is crucial for any AI initiative. It is imperative that *what* the goal is (vision) and what *type* of AI (Technology) is identified to permit an analysis of resources required. This includes personnel, infrastructure, and time and organizational support required. It is essential to assess and secure the necessary resources before embarking on an AI journey.

- **Common Challenges** include failing to understand what resources are needed, underestimating the need for skilled personnel, lacking adequate hardware and software, failing to secure sufficient time and executive buy-in. Without these resources, AI projects are likely to falter or fail to deliver the desired outcomes.
- **Questions to Answer**
  - Do we have the necessary AI and data science expertise in-house, or do we need to hire or train staff?
  - What roles and skills are required for our AI projects (e.g., data scientists, machine learning engineers, project managers)? understanding what type of AI you will use will also help determine the answer to this question
  - Do we have the hardware needed to support AI, such as powerful servers and GPUs, or will we need to invest in new equipment?
  - Are our current software tools and platforms sufficient, or do we need to acquire new AI-specific technologies?
  - How much time can we realistically allocate to developing and implementing AI projects?
  - Is there buy-in and support from top leadership for our AI initiative(s)?

Budget: Understanding and setting a budget is a key component of any successful AI Strategy. Budget considerations will not only directly impact the choice of technology, the scope of the effort, and the resources you can allocate—but it will offer the longevity necessary to sustain the project across the Valley of Death.

- **Common Challenges** include underestimating the costs associated with AI development, not accounting for ongoing maintenance and operational costs, failing to secure sufficient funding upfront, and not budgeting for unexpected costs. Without a well-defined budget, AI projects can quickly become financially unviable and without a contingency plan, the first unexpected cost may derail the project.

- **Questions to Answer**
  - What is the estimated initial cost for developing and implementing the AI project?
  - How will these costs be funded (internal budget, external funding, partnerships)?
  - What are the ongoing costs for maintaining and updating the AI systems—or the hardware/software required for the AI system?
  - Are there subscription or licensing fees for software and platforms?
  - How will we handle unexpected costs or resource needs?
  - What contingency plans are in place if the project exceeds the budget?
  - How will the budget be allocated across different aspects of the project (personnel, technology, training, etc.)?
  - How will the budget be allocated across the different phases of the project (will it be frontloaded, or will it be a steady budget over the life of the project.)

Stakeholders: The 'Tip of the Iceberg' of any successful AI Strategy. These are the people who stand to make or break any AI initiative and extend beyond the people directly involved in AI development. It includes custodians of the existing infrastructure in which the AI system will reside. It also includes higher level management. Identifying and engaging the correct stakeholders ensures that the project has a path to success. Without clearly established stakeholder roles, projects can get held up in 'analysis paralysis', be blockaded by processes that were never written with an AI use case in mind, or lose momentum and risk being shelved during tough times.

- **Common Challenges** include failing to identify *all* relevant stakeholders, not securing strong executive sponsorship, and lacking a clear ownership structure. Without these, AI projects will struggle with alignment, support, and accountability and minimize chances of success.
- **Questions to Answer:**
  - Who are the key stakeholders within the organization that will be involved in or impacted by the AI project?
  - Who are the external stakeholders, such as partners, vendors, or regulatory bodies, that need to be considered?
  - What roles will each stakeholder play in the AI initiative (e.g., project sponsor, project manager, technical lead, end-users)?
  - Who will be responsible for driving the project and ensuring its success?
  - Who are the executive sponsors or champions who will advocate for the AI project at the highest levels of the organization? And how will they be involved in decision-making and support?
  - Who will own the AI project and be accountable for its outcomes?
  - How will ownership be maintained throughout the project lifecycle, especially during leadership changes?
  - How will differing priorities and perspectives be managed to ensure alignment and cooperation?

Data Strategy

Data strategy constitutes the hidden part of the iceberg. The data strategy is the solid foundation that facilitates any AI endeavor. Without a well-crafted data strategy, AI efforts are entirely useless. A well-defined data strategy involves managing the entire lifecycle of data, from collection, storage, processing, and utilization which ensures that data is accessible, high-quality, and secure. This enables, and underpins, the success of the AI System.

Just as the development of an AI Strategy consists of a multitude of sub-layers, complexities, and theories of thought, so too does the development of a successful Data Strategy. However, when considering the Iceberg Model, it's important to recognize that a Data Strategy in *support* of an AI Strategy may differ from a standalone Data Strategy. For example, the vision/goals of a standalone Data Strategy would be duplicative in the Iceberg Model, as they are captured within the broader AI strategy. Let's explore the key components of a Data Strategy as part of the Iceberg Model:

Data Collection: Data is the lifeblood of any AI system so effective data collection starts with understanding the AI Strategy to know what data is needed and what it will be used for. Ensuring consistent and comprehensive data collection from the relevant sources is fundamental to building a solid AI system, and requires systematically gathering relevant, accurate, and timely data from various sources.

- **Common Challenges** include misunderstanding of what data is needed, or an inconsistency in data formats (leading to difficulties in consolidation). Another *very* common mistake is the collection of incomplete data or

only selected parts of data based on the needs of today—which limits additional future use cases. Additionally, many AI initiatives end before they ever get off the ground due to inaccessibility of data or 'data silos' within organizations, limiting the overall effectiveness or scalability of AI initiatives. Without addressing these issues, the collected data may lack the necessary completeness and consistency for effective/intended use.

- **Questions to Answer:**
  - What data do we need to collect to support the AI Strategy?
  - How will we ensure data collection is consistent across different sources?
  - What tools and processes will we use for data collection?
  - How will we address data silos within the organization?

Data Storage: Once data collection has been defined where it will be stored, cost, accessibility and security need to be considered. It is important to balance the amount of data collected with cost considerations and make tradeoffs as appropriate for individual visions.

- **Common Challenges** often revolve around scalability issues where initial developments may not have considered future use cases, which results in a solution that cannot scale to meet operational needs. The sheer cost of data storage can also be prohibitive, leading to hasty decisions about which data should be stored or discarded. Ensuring data is accessible while maintaining security adds another layer of complexity. Additionally, defining when and how to archive or delete data requires careful planning to ensure compliance with regulations and organizational policies.
- **Questions to Answer:**
  - How much data will we store? And do we understand which data needs to be stored?
  - What storage solutions will we use? (e.g. cloud, on-premises, hybrid)
  - What are the security and/or regulatory requirements for this data? And what measures will we take to comply with data protection regulations?
  - Are there mechanisms to scale our storage solution should the use case require it?
  - How will we manage the costs associated with data storage?
  - What criteria will we use to decide which data to store?
  - When and how will we decide that data is no longer needed and should be deleted or archived? And what processes will we use for securely deleting or archiving data?

Resources: Just as it is critical to define the resources and budget as components of the AI Strategy, Understanding and securing the necessary resources is crucial for the successful implementation of a Data Strategy that supports AI initiatives. This includes considering budget, stakeholders, available technology, architecture, and ongoing support. Ensuring these resources are aligned with the broader AI strategy will help in effectively managing data and maximizing the impact of AI projects.

- **Common challenges** include underestimating the budget required for comprehensive data management, assuming that all resources required to fulfill the Data Strategy are captured within the AI Strategy, and failing to identify and engage key stakeholders—especially when they are different from those involved with the AI strategy, and not having access to the appropriate technology and architecture. Additionally, securing ongoing support for data initiatives can be difficult without clear alignment with the broader AI strategy. Without addressing these issues, data projects can face delays, increased costs, inefficiency, and internal conflict.
- **Questions to Answer:**
  - What budget/resources is/are allocated in support of fulfilling the Data Strategy? And is this separate from the AI Strategy budget/resources?
  - Who are the key stakeholders involved and how do they differ from those involved with the AI Strategy?
  - What technology is available to support our Data Strategy needs? And will we reuse existing resources, or will we need to purchase additional software/hardware?
  - How will we integrate new technology with existing systems?
  - What data architecture will we use to support scalability, security, and accessibility?
  - How will this architecture integrate with our AI systems and other processes?

Data Quality: High-quality data is essential to avoid duplicates, errors, and outdated information. Ensuring that AI models are trained on accurate and reliable data leads to better performance and more meaningful outcomes. This involves not just cleaning the data but also organizing it into a clear hierarchy, so it can be easily queried and used effectively.

- **Common Challenges** include the presence of duplicates and errors within datasets, which can significantly skew AI model results; outdated data, which can render insights irrelevant or incorrect; and non-uniform, unstructured data, which makes querying and analysis difficult. Data cleaning processes are time-consuming, complex, and often cost-prohibitive. While retroactively cleaning data may not always be feasible, ensuring that all newly input data is of high quality will not only support the production of reliable and accurate outcomes for current needs but also ensure that data collected today will be usable for future applications.
- **Questions to Answer:**
  - What processes will we implement to ensure data quality?
  - How will we identify and correct duplicates and errors in our data?
  - How frequently will we update our data to ensure it remains current?
  - What metrics will we use to measure data quality?

Data Governance & Accessibility: Data accessibility ensures that data is available to those who need it while protecting it from unauthorized access. Clearly defining how the data will be governed and managed will ultimately dictate how data is accessed, ensuring that proper protocols are in place for both security *and* usability.

- **Common Challenges** include the challenge of balancing the need for security with the ease of access, as overly restrictive security measures can hinder data access, while lenient measures put data integrity at risk. Defining and managing user permissions is complex, as it requires determining appropriate access levels for different stakeholders, however often there is a lack of clear governance and ownership of data management. This ultimately results in a lack of a shared understanding of standards/controls, or established mechanisms to support efficient changes as requirements evolve.
- **Questions to Answer:**
  - Who needs access to the data, and how will we provide it?
  - How will we ensure data is accessible while maintaining security?
  - What systems and processes will we use to manage user permissions?
  - What mechanisms do we need to change/alter user permissions? And do all stakeholders have a full, mutual understanding of these mechanisms?
  - Who governs the data, and what are their responsibilities?

Living Documents

In developing a successful AI initiative, the AI and Data Strategies must not only work in tandem, with the Data Strategy providing the necessary foundation for AI efforts, but both strategies should also be considered living documents, to evolve as new technologies emerge, objectives shift, and organizational needs change.

For instance, if we were to define an AI Strategy (and supporting Data Strategy) to enhance educational methodologies at military training institutions, the 'Vision' could be that instructors receive automatically generated reports grouping students based on the learning methodologies they respond to best, to maximize the amount of far transfer achieved from a 2-week training course.

The AI Strategy would include (in addition to other elements as outlined in the Iceberg Model); **Technology** (Data aggregation tools to collect and integrate data from various sources, NLP to analyze written student feedback, statistical analytics tools to identify key factors influencing student outcomes, and data visualization tools like Power BI to highlight trends and patterns in student performance across different courses and teaching methodologies.) **Resources and Stakeholders** (Identifying the necessary resources, engaging key stakeholders, and determining if faculty needs training to interpret the generated reports and adjust teaching methods accordingly)—but it should also include mechanisms for **adaptability** – as the curriculum evolves, how new methodologies will be accounted for within the AI models for the AI Strategy to remain effective.

In the ever-evolving world of military training, the AI Strategy and Data Strategy are like dynamic duo superheroes—they need to be clear, well-defined, and ready to adapt at a moment's notice. These strategies should be living documents, capable of evolving with changing educational needs and technological advancements. This flexibility ensures that military training institutions can continuously improve their training programs, and stay relevant and effective in a rapidly changing landscape. It's all about having a clear vision and a defined path, but also being prepared to pivot as needed to keep up with the times. In short, your AI and Data Strategies should be as robust and dynamic as the training programs they support.

FUTURE ROADMAP AND RECOMMENDATIONS

As we chart the path forward, it's crucial to not only consider the current landscape of AI but also anticipate future advancements. This forward-thinking approach will ensure that our AI strategy remains relevant, adaptable, and innovative. Here are some key trends and actionable recommendations to guide your AI journey in military training:

- **Embracing Future Trends in AI Development**
  - Rapid Advancements: Stay ahead by keeping an eye on the latest AI developments, and have a mechanism in place that will allow you to integrate new solutions quickly. Whether it's new machine learning algorithms or cutting-edge technologies, staying informed and poised to try new applications will result in new capabilities without risk to existing capabilities.
  - AI-Enhanced Simulations: Leverage products which integrate AI to create more realistic and dynamic simulations. This will make training scenarios more effective and immersive, providing a significant boost to training outcomes with minimal cost.
  - Swarming Technologies: Explore the potential of AI-powered swarming technologies. These systems can coordinate multiple autonomous units, enhancing operational effectiveness in complex military scenarios.
- **Incorporating Advancements into Your Strategy**
  - Start Small and Scale: Begin with small, manageable AI projects. This approach allows for incremental learning and minimizes risk, making it easier to scale successful initiatives.
  - Collaborate for Innovation: Partner with academic institutions and research organizations. Leveraging their expertise and data can drive innovation and result in more robust AI models tailored to your needs.
  - Focus on Incremental Improvements: Implement AI-driven efficiencies that reduce costs and accelerate time-to-field solutions. These small enhancements can have a significant cumulative impact.
- **Actionable Recommendations for Integrating AI into Military Training**
  - Identify Clear Use Cases: Pinpoint specific areas where AI can add value. Clear use cases will guide your efforts and ensure that AI applications are targeted and effective.
  - Invest in Data Infrastructure: Build a robust data infrastructure to support your AI initiatives. High-quality, well-organized data is the foundation of successful AI projects.
  - Develop a Skilled Team: Assemble a team with the right mix of AI expertise and domain knowledge, and don’t be afraid to consider partners and solution providers as part of that team. Solution providers specialize in areas that you may not, leverage their knowledge and grow together.
  - Implement Strong Governance: Establish governance frameworks to ensure ethical and responsible AI use. Address issues like bias, transparency, and accountability proactively.
  - Foster a Culture of Innovation: Create an environment that encourages experimentation and rewards innovation. This will help uncover new AI applications and drive continuous improvement.

FINAL THOUGHTS

The monumental impact AI is poised to have on M&S in military training cannot be overstated. And the potential for exponentially rapid releases of new and improved capabilities, driven by AI, is unlike anything we've seen before. However, it is exactly this awe at the ‘art of the possible’ that has led to less of a ‘valley’ and more of a ‘seismic chasm’ of death. We must remember, that it's not about chasing the next big thing but about making practical, impactful changes. Start small, focus on high-impact, quick wins, and build from there.

AI and Data Strategies must evolve hand in hand, adapting to new technologies and shifting goals, and by implementing the recommendations and strategy developments explored here, from starting with manageable projects to fostering a culture of innovation, we can bridge the gap between AI's potential and its realized benefits. This strategic, measured approach will ensure that our AI initiatives are not only effective but also *sustainable*, positioning us to lead in the AI-driven future of military training.

The time for action is now—harnessing the power of AI to transform our training programs, achieve tangible improvements, and maintain our edge against adversaries doesn’t happen overnight, but it has to start somewhere. Somewhere small, somewhere meaningful, and somewhere that results in true benefits to our warfighters.

REFERENCES

Arm Technology. (n.d.). Computer Vision. Glossary. Retrieved 2024, from https://www.arm.com/glossary/computer-vision.

Amazon. (n.d.). Glossary . What is data strategy? - data strategy explained. https://aws.amazon.com/what-is/data-strategy/

Babu, M. (2020, September). Application of Eye Gaze Tracking for analysis of Pilot's visual scan pattern: Simulator and inflight studies. Research Gate . https://www.researchgate.net/publication/345166499_Application_of_Eye_Gaze_Tracking_for_Analysis_of_Pilot's_Visual_Scan_Pattern_Simulator_and_Inflight_Studies

Babu, M. D., JeevithaShree, D. V., Prabhakar, G., Saluja, K. P. S., Pashilkar, A., & Biswas, P. (2019). Estimating Pilots' Cognitive Load From Ocular Parameters Through Simulation and In-Flight Studies. Journal of eye movement research, 12(3), 10.16910/jemr.12.3.3. https://doi.org/10.16910/jemr.12.3.3

Brown, Sara (2021, April 21). Machine learning explained. MIT Sloan School of Management. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Dennis, M. Aaron (2024, April 25). Edward Albert Feigenbaum. Encyclopedia Britannica. https://www.britannica.com/biography/Edward-Albert-Feigenbaum

Finio, M. (2024, May 28). How to build a successful AI strategy. IBM Blog. https://www.ibm.com/blog/artificial-intelligence-strategy/

Institute of Data. (2023, October 17). Exploring the differences between Narrow AI, General AI, and Superintelligent AI. https://www.institutedata.com/blog/exploring-the-differences-between-narrow-ai-general-ai-and-superintelligent-ai

Khan, Hanif. (2021). Types of AI | Different Types of Artificial Intelligence Systems. fossguru.com/types-of-ai-different-types-of-artificial-intelligence-systems.

Knight, W. (2016, November 14). Understanding the four types of AI, from reactive robots to self-aware beings. The Conversation. https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

LeewayHertz. (n.d.). Generative AI tech stack: A comprehensive guide to generative AI technologies. LeewayHertz. https://www.leewayhertz.com/generative-ai-tech-stack/

Network of the National Library of Medicine. (n.d.). Natural language processing. NNLM. https://www.nnlm.gov/guides/data-glossary/natural-language-processing

Wolfewicz, A. (2023, February 15). Deep Learning vs. machine learning – what's the difference? Levity. https://levity.ai/blog/difference-machine-learning-deep-learning