# Modeling and Simulation (M&S) Government Reference Architecture (GRA) for Synthetic Training

(Version 1.41 – 17 July 2023)

FOREWORD

Our ability to conduct training in an uncertain and complex future environment will rely on digital, synthetic, and encrypted capabilities to render realistic and relevant warfighting domains. These capabilities are critical to ensure the integration of operations across all domains and to achieve successful military endeavors with our Joint and coalition partners.

This Government Reference Architecture (GRA) provides a high-level focus on principles and practices for Synthetic Training Modeling and Simulation efforts, with alignment to other DoD reference architectures, modernization strategies, and the Data Concept Plan. It outlines the necessary capabilities to render a training environment that is relevant to future threats, with a particular emphasis on digital, synthetic, and encrypted capabilities to integrate operations across all domains.

Its structured approach prevents duplication of effort, enables unity of effort across all warfighting domains, and provides flexibility to scale and instantaneously provision realistic environments for distributed Airmen. Use this document as an essential set of guiding principles when establishing requirements and relaying them to their material leaders to implement.

JAMES C. SLIFE, Lt Gen, USAF
Deputy Chief of Staff, Operations

# Table of Contents

# Tables



# Figures

# STRATEGIC PURPOSE

This Government Reference Architecture (GRA) provides a framework that can be used by architects and developers to meet the intent of Department of Defense (DoD) guidance while shaping the USAF Synthetic Training Enterprise.

> ***Vision:*** *to establish and maintain a realistic, integrated environment that allows our forces to train operationally and tactically.*

As stated in the DAF Posture Statement 2022, "current platforms will not fully support tomorrow's demands" and that "we maintain air superiority in the future by introducing game-changing technology that includes digital engineering, open mission systems architecture, and agile software." This requires incorporating composability into the digital environment so that the United States Air Force (USAF) can rapidly adjust (**see Figure 1**)to new technology to meet the challenges provided by peer and near-peer adversaries. This **Enterprise Level Reference Architecture** is modifiable and extensible and provides developers the means to create solution architectures that support specific Synthetic Training requirements while also being compliant with Department level policy and guidance.

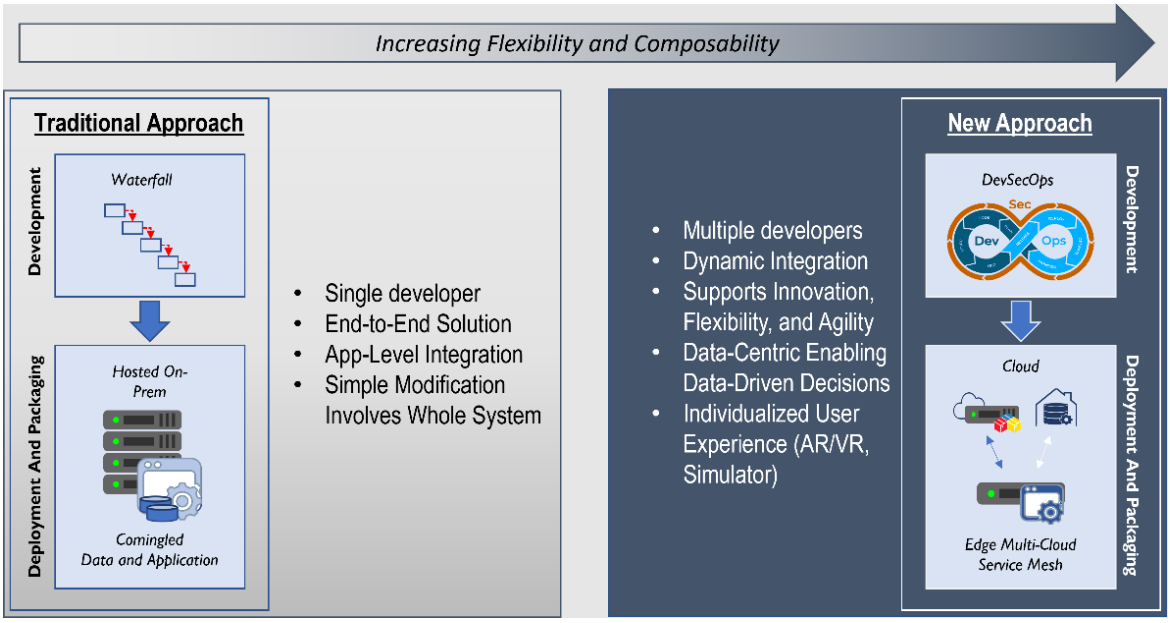


*Figure 1. AF Priorities Push Toward a New Approach*

**Figure 1** shows the traditional waterfall software development approach. This approach is a linear and sequential process that involves distinct phases: requirements, design, development, testing, deployment, and maintenance. Each phase relies on the completion of the previous phase before the next phase can begin. Oftentimes, changes made in a subsequent phase may require revisiting earlier phases, increasing the cost of the system development. In contrast, the new approach employs DevSecOps (Development, Security, and Operations) and the need for Data-Centric Enabling Data-Driven Decisions.

DevSecOps is an agile iterative approach to software development emphasizing collaboration and automation across different teams and stages within the development life cycle. DevSecOps integrates security and operations into the development process at the start rather than an afterthought. According to the "State of DevOps Report 2020" by Puppet and CircleCI, one of the key differences between traditional and modern approaches to software development is the focus on cloud deployment and packaging in DevOps and DevSecOps [Puppet and CircleCI. (2020)].

Current policy encourages and technology enables adoption of cloud computing solutions when appropriate in meeting program requirements. While program requirements will dictate specific solution architectures cloud-based solutions can reduce the attack surfaces; reduce overhead; and can provide leadership on-demand visibility of status of subjects including software updates, database versions, and cybersecurity assessments. Program Managers should weigh the priority of their full set of requirements and evaluate the cost/benefits of implementing a cloud-based solution. If appropriate, then the associated DevSecOps should place a greater emphasis on creating software that is optimized for cloud environments.

Another significant difference between the two approaches is the role of testing. In the traditional waterfall approach, testing is often done at the end of the development process, in a separate phase. In DevSecOps, testing is integrated throughout the development process, using automated testing tools and techniques that help to identify and fix defects early in the development process.

Overall, the DevSecOps approach provides increased flexibility, composability, collaboration, and responsiveness than the traditional waterfall approach. It enables teams to develop, test, and deploy software quickly and efficiently, while also improving security and reliability by incorporating these concerns into the development process from the beginning.

To achieve Data-Driven Decisions through a Data-Centric approach, M&S training data is essential. A common data set is required to create synthetic training environments and generate standardized outputs for analysis.

However, as with the traditional approach, Synthetic Training M&S currently lacks an enterprise-wide technical and procedural process to discover, retrieve, and transform available and appropriate training data. Currently M&S training data is contained in silos. To effectively correct this, an enterprise-wide technical and procedural process is necessary to support operational force training, exercises, and related activities effectively. This has begun as documented in the September 2022 M&S Operational Training Data Concept Plan.

By using Data-Centric Enabling Data-Driven Decisions, relevant and integrated training data products can be quickly produced. These products can support evolving live, synthetic, and blended training requirements, which are crucial for the M&S training enterprise to develop a comprehensive and coherent approach to provide on-demand data and data services.

This GRA is an initial step into composable thinking, which is an approach to problem-solving emphasizing the creation of modular, flexible, and interoperable solutions. It involves breaking down complex problems into smaller more manageable components that can be combined and recombined in different ways to create new solutions. This is a shift from traditional monolithic or integrated solutions to more loosely coupled and interoperative ones. It encourages the use of standard interfaces and protocols that allow different components to communicate with each other, regardless of the underlying technology or platform.

Composable thinking is particularly relevant in the context of digital transformation and the rise of cloud computing, microservices, and application program interfaces (APIs). By adopting a composable mindset, organizations can create solutions that are more agile, scalable, and adaptable to changing M&S training needs.

Some key principles of composable thinking include:

- **Componentization**: breaking down complex systems into smaller, modular components that can be easily combined and reused in different contexts.
- **Standardization**: using common interfaces, protocols, and data formats to ensure interoperability and compatibility across different components and systems.
- **Flexibility**: designing solutions that can be easily adapted and modified to meet changing business needs and market conditions.
- **Automation**: using tools and technologies to automate the creation, deployment, and management of composable solutions.

> *The greatest **danger** in times of turbulence is not the turbulence; it is to **act with yesterday's logic**.*
>
> *- Peter Drucker*

By embracing creative use of the principles and incorporating composability into planning and decision-making, organizations gain freedom of maneuver and flexibility in sustaining adaptive operations. In summary, composable means that computing, storage, and networking resources are abstracted from their physical locations and can be managed by software through a web-based interface [HPE-1]. Encouraging composability means …

1. Everything is changeable.
2. Embracing risk.
3. Think modular products.
4. Choosing flexibility.

This GRA is an abstract framework, which is a high-level conceptual model that provides the concept, principles, and guidelines for applying and orchestrating a composable M&S enterprise. The composable M&S enterprise model reapplies the core and advanced principles of Service-Oriented Architecture (SOA), Mesh App and Service Architecture (MASA), and the core principles of Data-Centric Security Architecture (DCSA). It lays out common concepts and definitions as the foundation for the development for M&S Synthetic Training efforts, facilitating frictionless rearrangement of capabilities and systems while creating new pathways to enhancing training objectives and achieving DoD strategic goals.

It is important to note that this **GRA is a living document**, and as such, it is influenced by new strategy and policies. As **new ideas, components, patterns, and views** are introduced or created, they will be adjudicated for incorporation into this document. This document does attempt to identify the necessary concepts, relationships, and components that will require further elaboration and/or implementation.

In addition, this GRA is NOT a solution architecture. Each individual program will have its own set of unique requirements. As such, as the name implies, this document is used as reference for developers to assist them in the creation of their individual solution architecture.

# 1 Introduction

A Reference Architecture (RA) is focused on a specific subject area, guides the instantiations of capability architectures and solutions architectures for the subject area. An RA provides common language and terminology, guides the application of technology, supports traceability of requirements to validate future architectures, and provides a method to implement common standards, patterns, and views. This RA is government owned and is at the Enterprise level to enable various communities to implement DoD and AF strategic guidance (See **Figure 2**) as part of their solution architecture.

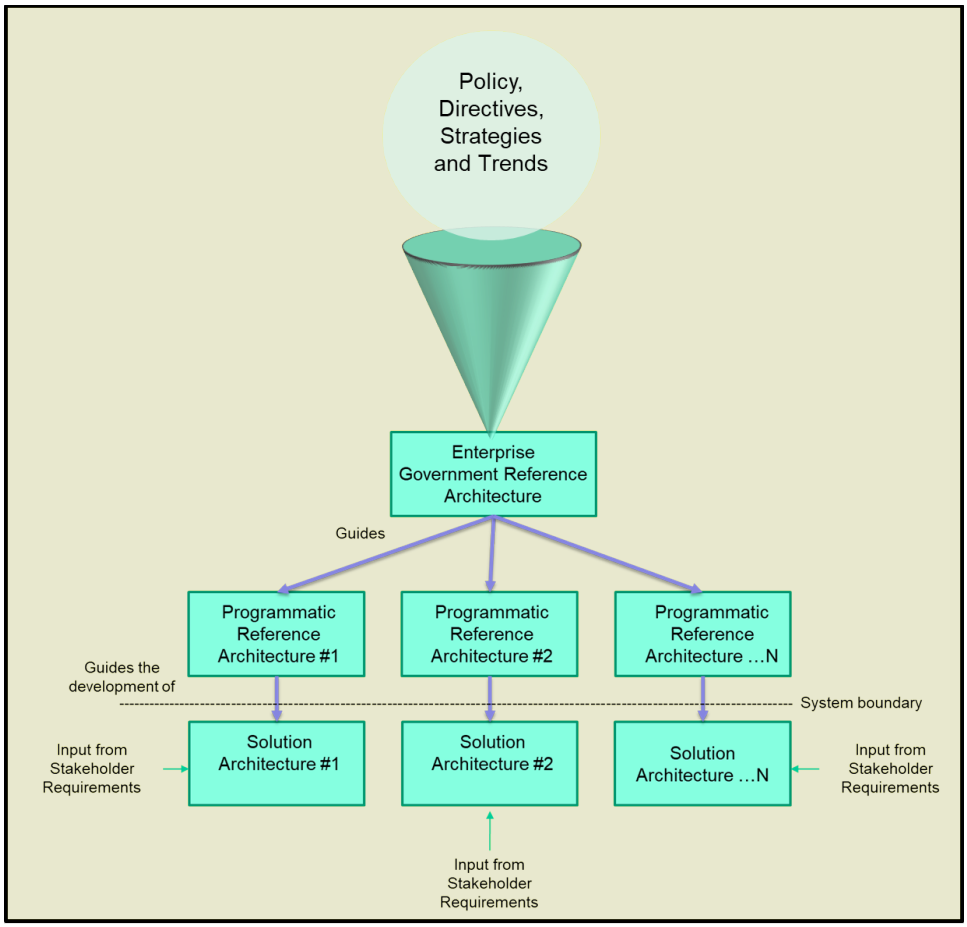


*Figure 2. Lineage from Strategy to Reference Architectures to Solutions Architectures*

According to the DoD Reference Architecture Description, a Solution Architecture is defined as a framework or structure that portrays the relationships among all the elements of something that answers a problem. It describes the fundamental organization of a system, embodied in its components, their relationships with each other and the environment, and the principles governing its design and evolution. Solution architecture instantiations are guided by all or part of a Reference Architecture.

This GRA is a description of the important concepts to support M&S Synthetic Training efforts toward composable architecture and to foster relationships within the Air Force. As stated earlier, it is not intended to be specific enough to govern the implementation of any individual software system implementation. Rather, it is a framework for guiding implementations in general, with the aim of standardizing or harmonizing certain key aspects of those implementations to support flexibility, reusability, composability, and technical interoperability.

It is intended to be used to support M&S Synthetic Training efforts by guiding their evolution; ensuring functional M&S Synthetic Training requirements for various M&S systems are being met.

This is the "Go-To Guidance" for a composable enterprise and, as such, does not address the following:

- Detailed specifications for operational systems
- Detailed specifications of information exchanges or services
- Recommendations or standards for integration infrastructure products

## 1.1 Purpose

The purpose of this GRA is to inform requirements owners and program managers along with other stakeholders on a modern and evolving framework for synthetic training environments. Specifically, this GRA intends to support increased data sharing, reduce lifecycle costs, and streamline operational complexity, while leveraging DoD and DAF infrastructure.

This GRA is intended to support the transition and transformation of existing simulation capabilities that support training to evolve towards the next generation of training infrastructure and architectures. This GRA incorporates tenets of the DoD Digital Modernization Strategy, the DoD Data Strategy, the DoD Cloud Strategy, the DoD Digital Engineering Strategy, the DoD Cyber Strategy, and the DoD Artificial Intelligence Strategy. It supports the Operational M&S Vision of "constructing a relevant training environment, which allows weapons systems and operators to interact in a dynamic, realistic manner," as articulated in the Operational Training Infrastructure (OTI) 2035 Flight Plan.

The cost of modernizing the training infrastructure from scratch would be overwhelming. This GRA aims to provide relief to this problem by providing practitioners with a set of documents that represent the Synthetic Training M&S efforts very best practices, experiences, and lessons learned when modernizing or transforming away from our current technical debt. Project managers leading Synthetic Training M&S efforts can start with this document rather than starting from nothing, dramatically accelerating their transformation efforts.

This GRA supports a migration approach to a Multi-Level Security (MLS)-protected environment to conduct widely accessible synthetic training with all DoD members and selective coalition partners dynamically. An MLS-protected environment is an interoperable, fully functional, optimized, and accurate representation of the training environment that meets the needs of the AF warfighter for Joint and coalition training. i.e., training as we intend to fight.

## 1.2 Scope

The focus of this GRA is on the underlying technology that supports Synthetic capabilities for Operational Training a. It aligns with other DoD reference architectures and modernization strategies and provides a

structured approach. The framework of this GRA can be adopted by M&S communities outside of operational training and modified for their own use.

## 1.3 Vision, Goals, and Strategy

The overall vision of this GRA is to enable accurate representation of an operationally and tactically relevant training environment, provide for interaction and effects of agile, resilient, and transparent data at the user's security level while securing that data when created, curated, and shared within the training environment to accelerate training with precision, speed, and scale. As pointed out in the DoD-IG Audit of Training Ranges Supporting Aviation Units, "the currently available live ranges are insufficient for aviation units to train as they will fight, thus the need for synthetic training environment that connects to Joint and coalition partners." [DoD-IG-2019-081, 2019] The capabilities identified in **Figure 3** are representative of an end-state synthetic training environment implementation, when applied, accommodates any variation of joint/coalition partners. The goals and capabilities of this vision are described in the subsequent sections.

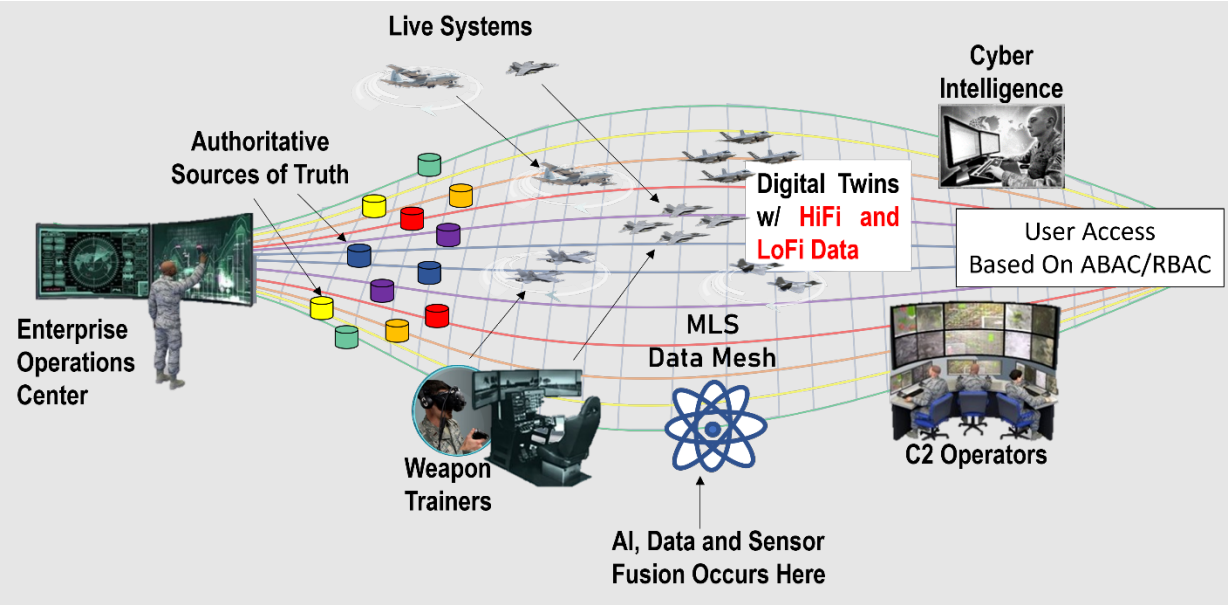


*Figure 3. End-State Synthetic Training Environment*

The high-level goals and objectives of this GRA are focused on meeting many of the goals identified in DoD and AF strategic guidance focused specifically on security within a composable M&S enterprise.

*Table 1. Goals and Objectives*

| Goal | Description |
|---|---|
| ***Data Sharing*** | Enable conceptual and semantic technical interoperability and sharing of data, information, or services |
| ***Leverage DoD and AF Infrastructure*** | Reliance on transformational enterprise services and assets for the background infrastructure so resources may be realigned toward improving quality of M&S solutions, providing Defense M&S assets that are accessible world-wide |

| Goal | Description |
|---|---|
| ***Reduced Lifecycle Costs*** | Consolidate redundant functionality, reduce integration errors and integration timeframe, and decouple functionality from obsolete and increasingly costly applications while leveraging existing investments |
| ***Streamline Operational Complexity*** | Apply DevSecOps principals to reduce preparation, operation, and maintenance cycle time through continuous improvement and continuous deployment. |
| ***Apply Zero Trust Principles*** | Trust nothing. Verify everything and everyone every time–- Zero trust principles are a set of guiding principles for workflow, system design and operations that can be used to improve the security posture of any classification or sensitivity level. [NIST SP 800-207] |
| ***Representative Simulation Environments*** | Synthetic representations of the operational environment, including the representations of entities, behaviors, environmental effects, and interactions with the operational environment, that are operationally valid and at the appropriate resolution for the intended use |
| ***Agility and Innovation*** | M&S solutions designed for modularity and composability, producing solutions based on a set of loosely coupled services facilitating rapid restructuring and reconfiguration and enabling use of existing capabilities in new and innovative ways |

## 1.4 High Level Operational Concept

DoD and the DAF have strategic goals to "Treat Data as a Strategic Asset," to realize the benefits of decision-making, information sharing, cloud migration, AI, and other DoD objectives. which are dependent upon DoD's data being visible, accessible, understandable, trusted, and interoperable as prescribed by DoD Instruction (DoDI) 8320.07. To achieve this goal, the current simulation architecture must migrate. [DoD Mod Strat, 2019]

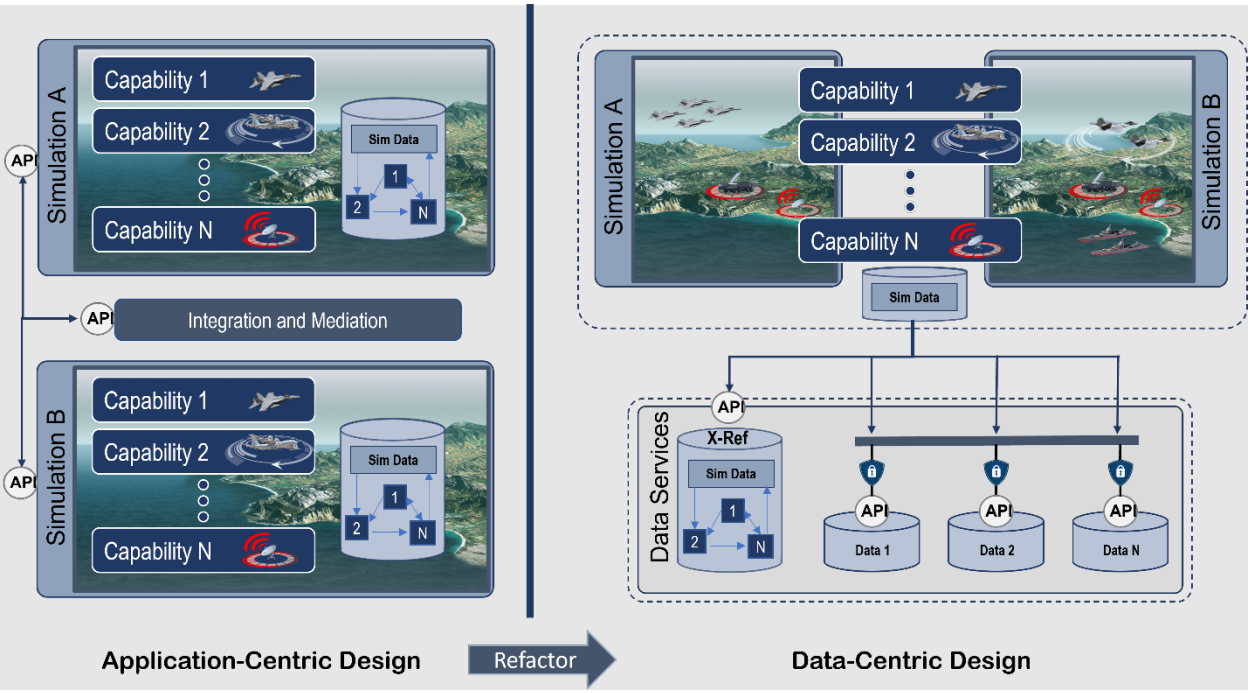


*Figure 4. Transition from Application-Centric to Data-Centric*

As shown on the **left side of Figure 4** (Current Capability as Application-Centric Design), simulators and simulations were deliberately designed for training within specific domains and modified over time to meet

growing requirements. Each simulation and simulator contain its own databases rather than using shared data. To adhere with current DoD regulations and to protect data, the network and training must operate at one classification level. This causes issues when a specific classification level platform, i.e., 5th Gen aircraft, needs to interact with a lower classification platform, i.e., 4th Gen aircraft, resulting in misrepresentation and ineffective training.

By shifting to a data-centric enterprise as shown on the **right side of Figure 4** (Desired Capability as Data-Centric Design), data-centricity will make the data available for all simulators and edge devices, facilitate analytics, and completes one step towards MLS. When a lower-level classification system requests data from or calculates data based on an interaction from a higher-level classification system, the MLS application or method used will respond with the correct level of data for that requestor. A data-centric enterprise will need to be designed and configured to accommodate any variation of joint/coalition partners to process these requests away from the requestor and return the correct security level results to the requestor with relatively low latency. A data-centric enterprise will also provide shorter integration times and more readily available training opportunities, which leads to an ability to conduct Joint training exercises and short-fuse mission rehearsals.

### 1.4.1 Decision Points, Components, and Capabilities

The desired capability aims to connect warfighters across Joint, deployed, and home-station training environments with accurate and updatable operating environments and capabilities that are secure from external and foreign surveillance. To overcome current limitations in distributed training limitations, the desired capability will leverage transformative technologies to perform computing and synchronizing functions as efficiently as possible. With technological advancements, the capability will enable resource pooling of scalable & ubiquitous computer processing power that is dynamically configurable to meet No-notice or Short Notice demands at the point of need, that support distributed training interactions with an optimized and measurable computing environment. This data will then be used by data scientist and Artificial Intelligence (AI) software for analysis and feedback. Status communication data can still be logged for After-Action Review replay and analysis.

The organization of shared and trusted data orchestrated effectively throughout the enterprise bakes in:

- *Security* by labeling and tagging data consistently;
- *Technical Interoperability* through common standards;
- *Scalability* using Platform as a Service (PaaS) architecture;
- *Representation* through consistent & improved models; and
- *Multi-Fidelity* through application modularity, by design and not as an external solution or add-on.

The operating concept is revolutionary within the existing AF M&S Synthetic Training efforts because data is the strategic key asset and is self-describing (i.e., meta-tagging with tag and value pairs) and does not rely on the application for interpretation and meaning.

Users will connect their simulators/training devices to the system ubiquitously (using any device, in any location) through a purpose-built authentication application (e.g., a Portal) using the appropriate login credentials. Based upon the person or systems credentials, their edge device will have access up to the classification they are allowed to interact with and their viewpoint into the shared environment will be shaped through the lens of their edge device's capabilities.

### 1.4.2 Capability Objectives

The capability objectives defined in the following table guide the design and development of a composable enterprise. They provide a basis for defining the scope and objectives of potential projects, as well as a framework for evaluating the effectiveness and suitability of final products.

*Table 2. Capabilities and Description*

| Capability | Description |
|---|---|
| ***Low Latency*** | The target latency of the new simulation environment should be 20 milliseconds (ms) or less. |
| ***Lightweight/Agile Simulator and Edge device access*** | The ability to access a simulation over the network running remotely from a machine with minimal computational capability. |
| ***Simulation container components*** | Distinct simulation functions are contained in separate software containers designed to facilitate reuse. |
| ***Micro-services and Macro-services*** | Functionality encapsulated in discrete services to access specific data stores. Services are loosely coupled. Services may be dependent on other services and organized in a layered hierarchy. |
| ***Application Programming Interfaces (API)*** | APIs are based on established industry or DoD standards to facilitate technical interoperability and composability. |
| ***Scalable*** | The ability to scale training from Tier 1 (large, strategic) to Tier 4 (small, tactical) of the simulation to handle a growing amount of work in a capable manner or its ability to be enlarged to accommodate that growth. For example, the ability to add users or increase the number of entities without human intervention, e.g., to add computer hardware. |
| ***Zero Trust*** | Minimize uncertainty by enforcing accurate, least privilege per-request access decisions in information systems and services. Employ continuous verification to minimize impact if a breach occurs (internally or externally) and automate context collection and response. |
| ***Identity Access Management*** | Allow access to models, simulations, services, and data repositories to only authorized users complying with applicable policy and guidance. |
| ***Multi- Level Security (MLS)*** | Controlled interface providing ability to manually and/or automatically access and/or transfer information between two or more different security classification levels. This is based on approved security classification guides (SCG). |
| ***Data Centric*** | Maintain data consistency, integrity, and relationships across large datasets. Ensure that data is secured and is a critical piece of the architecture. Focus is on data model, data stores, and security of data items. Data as a service. |

## 1.5 Linkages to Other Architectures.

- AF Data Services Reference Architecture [DoD DSRA, 2019]
- Defense Modeling and Simulation Reference Architecture [DMSRA, 2020]
- DoD Cyber Security Reference Architecture Integration [DoD CS RA, 2021]
- DoD Enterprise DevSecOps Reference Design [DoD ED RD, 2019]
- DoD ICAM Reference Design [DoD ICAM RD, 2020]
- DoD Zero Trust Reference Architecture [DoD ZT RA, 2021]
- NIST Special Publication 800-207 Zero Trust Architecture [NIST 800-207, 2020]

# 2 PRINCIPLES

This section documents the objectives to be addressed and satisfied by users of this GRA. These objectives are stated in the form of principles, the intent of which is to guide the choices made in developing the architecture.

The traditional approach to an enterprise has been using application-centric development and management of those resources within the applications. As stated earlier, this approach is proving to be not only inefficient in current operations but also an impediment in moving towards future growth strategies such as the introduction of new threat systems or in the application of Augmented Reality/Virtual Reality (AR/VR) systems. The current strategic direction for data focused solutions is only providing point solutions and are still disparate solutions with limited successes, and they impede, rather than enable, an adaptive training enterprise. The principles identified below are intended to guide the M&S enterprise by embracing adaptability in its design and enabling the enterprise to plan for the future of data-centric design:

## 2.1 Principle: Manage Architectural Technical Debt

Architectural technical debt is a metaphor used to describe sub-optimal architectural design and implementation choices that bring short-term benefits at the cost of the long-term gradual deterioration of the quality of existing software systems. All current edge devices and simulators will need to undergo an examination of their "architectural technical debt" as they are evaluated for modernization in keeping with the goals of the composable M&S enterprise.

The quality of the components within the M&S enterprise matter to the end-user. Better internal quality of the software enables developers to add models and components more quickly and cheaply as the operational environment changes. If the architecture is managed effectively with close observation regarding the goals of the composable M&S enterprise, developers can more easily add new features.

**Rationale**

Linking software design decisions with their financial costs is described by the metaphor "Technical Debt". It captures the extent to which design decisions that are expedient in the short-term can lead to increased system costs in future. Currently, the foundational component of our simulators is software, and our simulations are software systems. Over time, each new version or update to these software systems creates another system to build and extend. This approach restricts the developer by bearing the consequences of past design decisions. Typically, as systems grow and evolve, their architectures can degrade, increasing maintenance costs and reducing developer productivity. This raises the question as to when it might be appropriate to redesign ("refactor") a system, to reduce what has been called "architectural debt". Along with the current building blocks of software (i.e., languages, libraries, and platforms) which change greatly every few years, the M&S software systems must adapt as well.

Currently, most edge devices and simulators within the M&S enterprise have been working for many years, some for decades. Almost all have had modifications from their original design to work with new requirements; thus, an increase in their architectural technical debt has occurred. A potential modernization of these edge devices and simulators with guidance through the M&S GRA document will need to have their "architectural technical debt" evaluated to determine how best to proceed.

**Implications**

- Large-scale mature simulation applications can suffer from ad hoc changes to the codebase and expose security vulnerabilities. By managing the technical debt through redesign refactoring of components, security vulnerabilities are more easily identified and removed.
- While modernizing the composable M&S enterprise, management of architectural technical debt is everyone's responsibility. As such, each system should support a Technology Capability Plan identifying what must be fixed and when and it is clearly articulated to stakeholders.
- Managing the architectural technical debt of the M&S components should reduce the cost for future change, and enable adding features with less effort, time, and cost.

## 2.2 Principle: Innovate and Experiment with New Technology

Innovation and use of new technologies enables the Air Force to improve and enhance the training environment. By proactively encouraging the innovation process and adopting new technologies, we enable a secured and improved training environment. The importance is to start small in investment and expand when successful.

**Rationale**

- The innovation process allows M&S Synthetic Training efforts to stay current and keep pace with our peers and near peers.
- Innovation might help M&S Synthetic Training efforts improve in one specific aspect, but innovation in areas where it already excels can make it better still.
- Innovation helps keep focus on both the present and future expectations rather than past efforts that are now outdated.

**Implications**

- There are many benefits for enabling innovation for the M&S enterprise. Some of these benefits include:
    - Improved productivity – When new technologies are employed, they provide greater technical efficiency which increases user productivity.
    - Reduced costs – With improved productivity, reduced costs are a natural consequence. The less time employees spend on tasks that don't require innovation, the more they can do with their allotted work hours.
    - Better user experience – Increased efficiency benefits the M&S enterprise; it also helps users by providing better and faster M&S support.
    - Increased competitiveness – Innovation provides a competitive advantage as well. Specifically, it helps companies maintain their edge against competitors who may not be embracing change.

## 2.3 Principle: Evaluate Legacy Systems for Inclusion in the Composable M&S Enterprise

To migrate legacy systems (edge devices and simulators) from being device-centric or application-centric to having a data-centric architecture and extend their service use, a path forward migration plan must be implemented to achieve the listed principles of a modernized enterprise to evaluate whether those systems can be included in the composable M&S enterprise. Some systems will be able to undergo changes while

other systems cannot and may be placed into a sunset plan. Questions to ask in support of a sunset decision should include:

- Can the system be reconfigured to access a remote data repository?
- Can the system be reconfigured to discover new data repositories?
- Is the latency low enough while accessing the data repositories?

**Rationale**

Legacy systems currently house their own data files (terrain, ballistics, EW, weather, etc.). Some systems can be configured to work in this new architecture of composable M&S enterprise. As a test, a representative one or few of those systems can be migrated while others continue to support their current training missions. The legacy system migration plan can start with changes to the system to access data on a remote system. These remote data files, the data repository, will become the trusted data source.

If a system can be reconfigured to access the data repository, it will begin to fulfill the principle of Collaboration and Communication (Principle 2.5) through using Authoritative Sources of Truth (ASoT). A data principle or data steward will manage the data stored in these files to become the trusted data sets for all systems to communicate with in the new architecture. If the system can be reconfigured to accept the changes, the system will begin the process towards the principle of Scalability.

Once Scalability, Collaboration and Communication, and Composable M&S enterprise principles are met, if there is a change in data policy or standards, the migrated legacy system will be able to adapt to these changes relatively quickly to further work within the principle of Development Agility (Principle 2.6); being able to adapt for the training requirements.

As the legacy system interacts with the new data repositories, some new data may need to be discovered and interacted with such as sensor packages, weapons systems, etc. Instead of using the current practice of building internal custom code on a legacy system, the migrated system will be able to discover and communicate with the data repository by way of a standard (i.e., an API) comprised of language neutral datasets. If code is needed to achieve any new connections, instead of creating new code, systems will reuse system interfaces and information exchange implementations (such as API's) available to the enterprise to improve consistency and reliability of information by having all information users draw from the same source. This provides for the principle of Model Flexibility (Principle 2.7).

As legacy systems are evaluated to be migrated, some will be identified as having software that cannot be upgraded or changed. A plan to sunset these systems will need to be implemented. Other systems which can be migrated fully should follow the principles to this M&S GRA to implement these new communication methods.

**Implications**

This principle implies the following about the GRA:

- Adoption of a modular approach that allows organizations to implement a subset of the full architecture, achieving some initial benefit while retaining the option of adopting more of the architecture later.
- The GRA should encourage the sharing of information and functionality between systems in a way that minimizes the implementation dependencies between them.

- The GRA should encourage communication between systems using government data rights standards, NATO Standardization Agreements (STANAGs), and open standards rather than proprietary approaches.
- The GRA should provide mechanisms to separate the logic of information exchange (e.g., the routing and transforming of messages that flow between organizations) from the logic of technological solutions.

## 2.4 Principle: Promote Scalability for Small and Large Operations

"Scalability is the measure of a system's ability to increase or decrease in performance and cost in response to changes in application and system processing demands." [GART-1]. A composable enterprise should provide the ability to integrate software and hardware components for small operations with a few participants, as well as to enterprise-wide operations that reach across MAJCOM boundaries. In addition, M&S Synthetic Training efforts that are composable must have access to combine components in a creative way to further enable future paths.

**Rationale**

Current M&S Synthetic Training efforts have not been designed for growth, can be quite costly to maintain, and are unable to rival the speed of our potential adversaries' technological advances. M&S Synthetic Training efforts must ensure that it can accommodate and sustain technical growth. The current lifespan of technology components is at least three years, and the M&S enterprise needs to adapt for change in infrastructure and components.

In addition, the M&S enterprise must have confidence that M&S Synthetic Training capabilities grow and that there will be minimal rework and reinvestment across all efforts. This principle promotes an architecture that will satisfy the needs of an initial implementation and that will have the ability to scale systems in the future with the advantage of lower total cost of ownership for the enterprise.

**Implications**

This principle implies the following about the GRA:

- Adoption of a modular approach that allows organizations to implement a subset of the full architecture, achieving some initial benefit while retaining the option of adopting more of the architecture later.
- Encouragement of use of government data rights standards, NATO STANAGs, and open standards with a broad range of implementations available in the marketplace, from less expensive implementations with modest capabilities, to larger investments that support an increased volume of information sharing.
- Encourage the use of clear descriptions, straightforward discovery, and reuse of services across organizations to provide access to components for potential future paths.
- The components need to scale vertically (i.e., adding more power to a machine or resources to a training center) or horizontally (i.e., distributing processing activities across network).
- Predictive monitoring of resources' usages.

## 2.5 Principle: Data is Shared and Used for Collaboration and Communication

Data is a valuable resource; it has real, measurable value. Most Air Force assets are carefully selected, managed, and funded; data should not be an exception. Data assets are to provide maximum benefit to the enterprise. Currently, data is spread amongst various silos within the Air Force, and while necessary within a specific group can oftentimes cause high maintenance and management costs. The lack of an enterprise-wide technical and procedural process for discovering, retrieving, and transforming available and appropriate training data necessary to support operational force training, exercises, and related activities makes it difficult for developers and users of synthetic training applications to discover and gain access to the data required to meet their mission objectives.

To prepare M&S Synthetic Training efforts for collaborating and communicating across systems, the enterprise needs to work from a single source of truth (SSOT) to achieve data security, data integrity and data resilience. This ensures that the data-driven decisions are based on the same data across the enterprise. An SSOT eliminates duplicate data entries, reduces identifying correct data; and enables the ability to improve data intelligence across the enterprise.

**Rationale**

Shared data will result in improved decisions since the M&S Synthetic Training efforts will rely on fewer (ultimately one virtual) sources of more accurate and timely managed data for synthetic training. It reduces the costs associated with data management by eliminating the need to maintain multiple databases supporting the same datasets. Data is captured once and used many times. Open sharing of data and the release of data must be balanced against the need to restrict the availability of classified, proprietary, and sensitive data. This principle of data sharing will continually need to reference the principle of data security. Under no circumstances will the data sharing principle violate the principle of data security causing private or confidential data to be compromised. For further directive related guidance, please refer to M&S Operational Training Data Concept Plan (dated Sep 23, 2022).

**Implications**

This principle implies the following:

- To enable data sharing, the M&S Synthetic Training efforts must develop and abide by a common set of policies, procedures, and standards governing data management and access.
- Preserve investment in legacy systems by migrating legacy system data into a shared data environment.
- To define the shared environment for most or all M&S Synthetic Training efforts, standard data models, data elements, and other metadata elements will need to be identified. To ensure data security, data integrity and data resilience, developers must leverage data tagging and data segmentation approaches to manage releasability and accessibility of data.
- Shared data will become a virtual single source of truth.
- Maximize leveraging the DAF Data Fabric for Synthetic Training use where applicable and capable. More info can be found at DAF Data Platforms **[DAF Data Platforms]**

## 2.6 Principle: Development Agility

A government reference architecture for sharing information between and among various M&S Synthetic Training efforts should accommodate changes in policy, information flow, and organization system implementation without forcing investments or changes in unrelated systems or exchanges.

**Rationale**

While events that trigger information exchange remain constant, the policy responses and the flow of information following these events are in constant change. This principle promotes an architecture that allows information sharing practitioners to respond to—and even to thrive in—this dynamic environment.

Technologies within organizations change frequently as well. The days of purchasing a technological solution and leaving it untouched for years at a time are long past.

New capabilities available from vendors and improvements in internal operations both compel a more rapid rate of change. This principle promotes an architecture that separates system implementations from one another, reducing the impact of change to one on the others.

**Implications**

This principle implies the following about the GRA:

- The GRA should encourage the sharing of information and functionality between systems in a way that minimizes the implementation dependencies between them.
- The GRA should encourage the definition of system interfaces that reflect what the interfaces do, as opposed to how they work.
- The GRA should provide mechanisms to separate the logic of information exchange (e.g., the routing and transforming of messages that flow between organizations) from the logic of technological solutions.

## 2.7 Principle: Model Flexibility

To stay ahead of the planned need for machine learning via data in the modernized M&S GRA environment, there is a choice of building simple inflexible model or a complicated flexible model. A flexible model is a proactive approach while an inflexible model is reactive and will take time for changes and integration testing.

**Rationale**

Simulation models for USAF needs can be weather, ballistics, instructorless flight training, etc. A flexible model can identify actions needed before or after a given operation or control input or activity to either improve or mitigate the outcome or event. A flexible model can evaluate the impact of changes in policy prior to implementation.

A key characteristic of a model is to be specific enough to learn a target function based on its training data, yet flexible enough such that the learned target function provides acceptable performance on unseen observations [SPR-1].

**Implications**

- Simple models that do not adhere to flexibility design standards will be limited to only working for one set of data making it unusable for shared use.
- The flexibility of a model can be described as how much a mode's behavior is influenced by characteristics of the data; therefore, the flexibility or inflexibility of a model is a characteristic that should not be overlooked. It can mean the difference between a useful or useless tool (model) [SPR-1].

## 2.8 Principle: Gapless Security Protection

The data-centric security model re-thinks how to implement security access to resources and is determined by dynamic policy—including the observable state of every edge device and simulator, and application/service—and may include other behavioral and environmental attributes. The data-centric security models embrace DoD's Zero Trust (ZT) capabilities high-level goals: Gapless Security Protection, Optimize Data Management Operations and Dynamic Credentialing and Authorization. Confidence levels are built from multiple attributes of the subject being authenticated (identity, location, time, device security posture) and allow a much more thorough evaluation of access requests beyond credential verification.

**Rationale**

"To keep pace with today's dynamic and increasingly sophisticated cyber threat environment, the Federal Government must take decisive steps to modernize its approach to cybersecurity, including by increasing the Federal Government's visibility into threats, while protecting privacy and civil liberties. The Federal Government must adopt security best practices; advance toward Zero Trust..." [Executive Order 14028, Sec 3 (a)] "Zero Trust is the term for an evolving set of cybersecurity paradigms that move defenses from static, network-based perimeters to focus on users, assets, and resources. Zero Trust assumes there is no implicit trust granted to assets or user accounts based solely on their physical or network location." [NIST 800-207, 2020] Executive Order 14028 requires all federal agencies to "...develop a plan to implement Zero Trust Architecture, which shall incorporate, as appropriate, the migration steps that the National Institute of Standards and Technology (NIST) within the Department of Commerce has outlined in standards and guidance, describe any such steps that have already been completed, identify activities that will have the most immediate security impact, and include a schedule to implement them." [Executive Order 14028, Sec 3 (b)(ii)]

A transition to a 'zero trust' approach to security provides a defensible architecture for this new environment. Currently, network-centric, and device-centric security strategies inevitably leave gaps between protected systems using network perimeter fencing. "[H]ardened perimeter defense can no longer suffice as an effective means of enterprise security." [DoD ZT RA] As a result, the National Security Agency (NSA) is assisting DoD with integrating the Zero Trust framework within NSS and DoD, starting with the release of the "Advancing Zero Trust Maturity throughout the User Pillar" Cybersecurity Information Sheet (CSI) to help system operators' mature identity, credential, and access management (ICAM) capabilities. To achieve a mature Zero Trust framework, systems must integrate and harmonize the capabilities from the following seven pillars: user, device, data, application/workload, network/environment, visibility and analytics, and automation and orchestration. [NSA Press Release] Regarding the data pillar, the success of DoD missionsare increasingly dependent on structured tagged data within and external to originating systems. Advanced analytics also depend on these dependencies. This results in:

- Interoperability challenges between applications, organizations, and with external partners
- System inefficiencies and vulnerabilities
- Poor user experience
- Inability to fully leverage cloud computing, data analytics, machine learning, and artificial intelligence" [DoD ZT RA]

ZT embeds security principles throughout the architecture for the purpose of protecting data and service operations, while preventing, detecting, responding, and recovering from malicious cyber activities. Although straightforward in principle, the actual implementation and operationalization of ZT incorporates several areas which need to be smartly integrated and that include software defined networking, data tagging, behavioral analytics, access control, policy orchestration, encryption, automation, as well as end-to-end ICAM. [DoD ZT RA]

ZT focuses on protecting critical data and resources, not just the traditional network or perimeter security. [DoD ZT RA] Effective data-centric security reduces security gaps and potential security breaches, keeping sensitive data protected at its security level by being protected and encrypted everywhere data is created, shared, and stored. This is possible when a data protection solution provides the following:

- Persistent zero-trust protection at a granular data level enforcing security and releasability classifications. "Identify authoritative sources for user attributes and implement data tagging for all critical resources to support more granular access models. "[NSA CSI, pg 16]
- Cross-platform operability that enables the M&S enterprise to protect data (and make data available for authorized use) on edge devices and simulators. "User accesses are granular to the specific resource being requested, considering the user and their device, as well as the sensitivity of the application and specific data associated with the request. "[NSA CSI, pg 17]
- Just in Time (JIT) / Just Enough Access (JEA) – in addition to **granular** access rules, granting "...privileges to controlled resources only for predetermined periods of time on an as-needed basis." [NSA CSI, pg 14]

**Implications**

This principle implies the following about the GRA:

- ZT principles and practices are to be applied within each of the Department of Defense (DoD) Cloud Computing Security Requirements Guide (CC SRG) Impact Levels.
- However, ZT does not have an approved solution for classified boundary protection. Currently NSA has not approved ZT as a Cross Domain Solution (CDS). Today, ZT does not meet the threshold to be the gatekeeper for classified information. Until further notice, there are no ZT systems currently approved to be used as a CDS.
- Every edge device and simulator, application and data flow are authenticated and explicitly authorized using least privilege, multiple attributes, and dynamic cybersecurity policies. **ZT implements continuous multi-factor authentication, micro-segmentation, encryption, and robust auditing to Data.** [DoD ZT RA]
- All resources are consistently accessed in a secure manner using multiple attributes (dynamic and static) to derive confidence levels for contextual access to resources. Access to resources is conditional and access can dynamically change based on action and confidence levels resulting from those actions. [DoD ZT RA]

- While most systems will eventually implement Zero Trust security model, there are certain types of systems that may not be amenable to the specific constraints of Zero Trust designs, and as such will be allowed to continue to operate without the Zero Trust constraints.
- "All Joint Electronic Warfare operational scenarios will require MLS capabilities" [DoD AMBIT Study 2023], which is partially enabled by ZT. A data-centric security architecture can enhance the ability to plan, integrate, execute, and debrief in training, which contributes to combat readiness. Program managers should continue to apply the most appropriate security solutions necessary to ensure data transfers meet their needs.

## 2.9 Principle: Alignment with Best Practices and Experience with M&S community

As the USAF moves toward developing a digital ecosystem to support the various M&S Synthetic Training efforts, the desired end state is to create a cohesive and holistic M&S Community of Interest (COI) which operates in a productive environment under common policies, standards, processes, and best practices to modernize readiness which will increase lethality.

**Rationale**

To remain competitive, the AF must sharpen its ability to pivot, by recognizing the geopolitical environment and transform accordingly and rapidly to meet the challenges ahead. This will require us to adopt common industry best practices as well as United States' and international standards.

We must find every opportunity to adopt commercial best practices that will transform our existing software architectures, network architectures and application protocols of training systems, as well as to the organizational structures and cultures of developmental teams, acquisition offices, and simulation support personnel. We must adopt practices such as Design Thinking (empathize warfighters pain points, help further define requirements, ideate on pathways to articulate approaches to address requirements), iterate on concepts and approaches, liaise with science and technology (S&T) and research and development (R&D), measure and learn from pilot efforts, present concise requirements, and find the way forward to solution foundries, product development centers and commercial providers.

The AF will continue to develop and publish templates, best practices, technical assistance, and other materials to support the authorization of cloud computing products and services and increase the speed, effectiveness, and transparency of the processes, consistent with standards and guidelines established by the Director of the National Institute of Standards and Technology and relevant statutes.

The intent of this pivot is to transform the current M&S Synthetic Training efforts from the current traditional, monolithic, industrial era approach to a 21st century approach; leveraging a shared environment to reduce latency and increase scalability and flexibility to deliver to the warfighter ubiquitous and on-demand capability to assemble simulation with no or short notice. Alignment with Best Practices increases our ability to rapidly provision, recycle and remove the need to constantly configure physical infrastructure.

**Implications**

This principle implies the following:

- Leveraging the power of data.
- Providing a 21st century simulation capability responsive to the demands of modern combat.
- Model industry best practice on common design patterns.

- Adopt agile and DevSecOps practices.
- A new "M&S ecosystem" is required where M&S products can be more readily identified and accessed by many users to meet their specific requirements.
- This "as a Service" paradigm must support the integration of multiple simulated and real systems into a unified simulation environment whenever the need arises.

## 2.10 Summary of Principles

The principles presented in this GRA are intended to inform and guide the development of solution-focused RAs. It is important that they be considered collectively, understanding that no one principle takes precedence of the others.

*Table 3. Goals Mapped to Principles*

| Principles | Goals |
|---|---|
| ***Manage Architectural Technical Debt*** | • Streamline Operational Complexity<br>• Agility and Innovation<br>• Reduced Lifecycle Costs |
| ***Innovate and Experiment with New Technology*** | • Agility and Innovation<br>• Reduced Lifecycle Costs |
| ***Evaluate Legacy Systems for inclusion in the Composable M&S Enterprise*** | • Leverage DoD and AF Infrastructure<br>• Streamline Operational Complexity<br>• Reduced Lifecycle Costs |
| ***Promote Scalability for Small and Large Operations*** | • Data Sharing<br>• Leverage DoD and AF Infrastructure<br>• Streamline Operational Complexity<br>• Reduced Lifecycle Costs |
| ***Data is Shared and Used for Collaboration and Communication*** | • Data Sharing<br>• Leverage DoD and AF Infrastructure<br>• Streamline Operational Complexity<br>• Agility and Innovation<br>• Representative Simulation Environments |
| ***Development Agility*** | • Data Sharing<br>• Leverage DoD and AF Infrastructure<br>• Streamline Operational Complexity<br>• Agility and Innovation |
| ***Model Flexibility*** | • Agility and Innovation<br>• Representative Simulation Environments |
| ***Gapless Security Protection*** | • Apply Zero Trust Principles<br>• Data Sharing<br>• Streamline Operational Complexity<br>• Agility and Innovation |
| ***Alignment with Best Practices and Experience with M&S community*** | • Data Sharing<br>• Streamline Operational Complexity<br>• Agility and Innovation |

# 3 TECHNICAL POSITIONS

This GRA provides common, accepted, and experience-based guidance that supports the effort to identify, select, and apply the appropriate technical position for current efforts and guide the design/planning of emerging technical capabilities. Technical positions are based on subject area principles and assist in scoping real-world solutions. This forces organizations to identify relevant subject area standards and specifications as well as justify their choices and tradeoffs. Technical positions are followed and implemented as part of the solution to drive compliance.

It recognizes Air Force challenges and needs with applying new and emerging technology to support various M&S Synthetic Training efforts, as such, this GRA will assist Air Force organizations with the appropriate guidance to the solutions that can be considered when evaluating/determining innovation pathways. This GRA will leverage existing efforts such as SISO's eXtended Reality (XR) Interoperability Standards Standing Study Group, the Simulator Common Architecture Requirements and Standards (SCARS) Engineering Control Board (SECB) standards development; as well as the many DAF M&S Council (DAFMSC) Cross Functional Teams (CFT) formed to address issues and leverage knowledge of well-known experts. Additional information will be gathered from these development activities to provide the level of planning, implementation and execution needed to support USAF Synthetic Training M&S stakeholder system sustainment requirements and included in in a subsequent GRA Technical Positions document.

# 4 PATTERNS

The use of patterns supports the USAF in achieving desired outcomes without having to recreate the wheel. Patterns are guidelines that the USAF can refer to that have been tested or determined as acceptable or are considered best practices for others to follow/repeat. New pattern concepts are discovered from S&T and R&D efforts; others emerge from solution architectures.

The patterns will assist the functional groups in configuration and deployment of simulation resources to meet solutions in a manner consistent with principles stated above. These patterns allow M&S Synthetic Training efforts to rapidly reuse existing objects in the system in a manner consistent with overall design.

Patterns included in this GRA are Data Patterns, Integration Patterns and Migration Patterns. Choosing one pattern over another is determined by the cost associated and benefit gained from the achieved outcome. In most cases, the functional domain, its application for the end-user and time for implementation drives the decision. Some patterns are better suited than others for a particular use-case.

## 4.1 Data Patterns

Data patterns support the integration of applications and services within an organization, ensuring the availability of accurate and up-to-date data. Data integration involves combining data of various types and formats from any source in the organization into a central repository such as a data lake or data warehouse. This process aims to create a unified resource that can be used consistently for analysis or other purposes.

This GRA delves into four essential data integration patterns: Extract, Transform, Load (ETL); Extract, Load, Transform (ELT); Data Mesh, and Data Fabric. Data consolidation has conventionally relied on ETL methods. However, the emergence of cloud storage technologies has facilitated the adoption of more efficient ELT processes, which are considered more efficient. The key distinction between these processes lies in when and where data transformations occur (see **Table 4**).

It's crucial to consider several factors when selecting the appropriate pattern for your requirements. These factors include:

- Schema:
  - Harmonization of data from different sources
  - Facilitation of analytics
  - Development of data products, such as simulations
- User Experience:
  - Ensuring intuitive and user-friendly data integration
  - Seamless access and interaction with integrated data
  - User satisfaction and ease of use
- Processing Speed:
  - Efficient and timely data integration and transformation
  - Minimization of processing delays or bottlenecks
  - Real-time or near-real-time data processing capabilities
- Agility:
  - Flexibility to adapt to changing data sources and formats
  - Quick response to evolving business requirements
  - Scalability to handle growing data volumes and complexities

Each pattern has its own distinct advantages and considerations, and choosing the right one depends on an organizations' specific integration needs. Therefore, these factors should be carefully considered when selecting the appropriate data pattern to meet an organizations' integration needs.

### 4.1.1 Data Integration - Extract, Transform, Load (ETL)

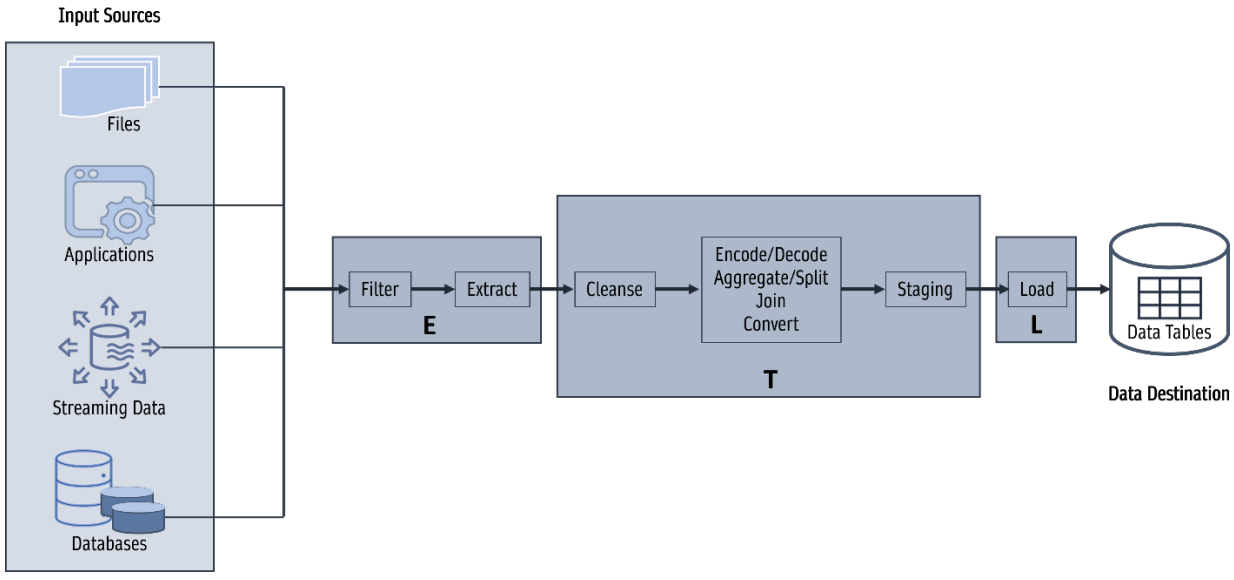


*Figure 5. Extract, Transform, Load – Logical Architecture*

To transfer data from one system to another, a data pipeline is required. This pipeline must comprehend the structure and meaning of the data while also defining its path through the systems. Data ingestion is a common and relatively straightforward type of data integration, involving regular integration of data from one system into another. Data integration may encompass additional processes like cleansing, metadata

tagging, enrichment, and other preparations to ensure the data is ready for its destination. This process is known as Extract, Transform, and Load (ETL).

In the initial step, data is gathered from identified sources using an extraction tool. The data then undergoes the data manipulation step in the ETL process. During this step, the input data is cleansed, mapped, and curated to align with the schema of the data warehouse. Finally, in the Load step, the staged data is loaded into the target data tables within the data warehouse. It's worth noting that data can be loaded in batches or streams, depending on whether real-time data updates are involved. The design of the ETL process can quickly become intricate.

ETL is particularly suitable for compute-intensive data workflows that require manipulation before entering a target system. This manipulation may involve tasks such as removing personally identifiable information or handling other controlled information.

### 4.1.2 Data Integration - Extract, Load, Transform (ELT)

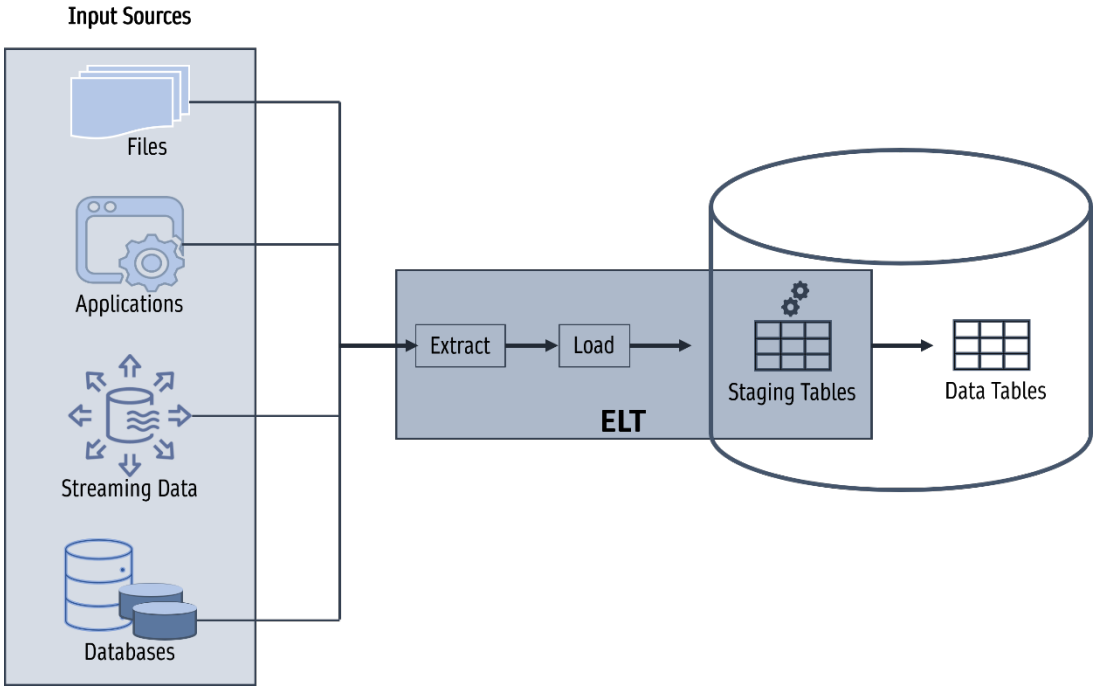


*Figure 6. Extract, Load, Transform – Logical Architecture*

Due to the increasing utilization of cloud-based data warehouses and the rising volume of unstructured data, the Extract, Transform, Load (ETL) process remains essential. However, in the context of ELT, the transformation step occurs after loading all input data into the data destination. As illustrated in Figure 6, the input data is initially extracted and loaded into staging data tables within the data destination. Following the execution of data transformations, the processed data is then transferred to the final data tables and stored for future use.

The majority of enterprise data consists of unstructured formats such as images, videos, PDF files, PowerPoint documents, and so on. This type of data poses challenges in terms of accessibility and processing. The manner in which unstructured data is processed is critical, and ELT currently surpasses ETL in this regard. ELT offers superior processing capabilities for semi-structured and unstructured data,

while ETL is typically utilized for structured data. Consequently, ELT will play a vital role in overcoming challenges and enhancing the interpretation of unstructured data.

### 4.1.3 Data Fabric

A Data Fabric pattern, **Figure 7**, emphasizes on building a *knowledge graph* of metadata that holds relationships between data sources and makes those data sources interoperable. Machine Learning and upcoming yet immature technologies such as semantic knowledge graphs and active metadata management are aimed at facilitating data fabric architecture. In addition, a Data Fabric pattern relies on Data Virtualization, an early version of the Data Fabric, and a concept that doesn't require ingestion of data beforehand but accesses the data via metadata stores dynamically with clever techniques like caching and push down query optimization. Industry sources state that a data fabric is approximately 5 to 10 years away from realization.

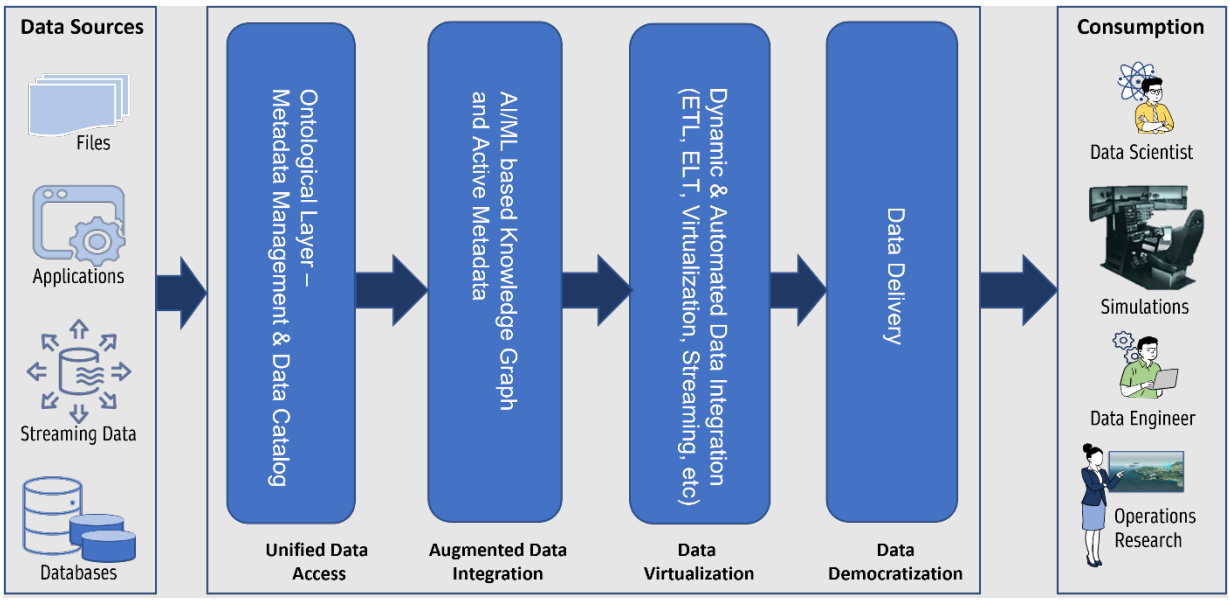


*Figure 7. Data Fabric – Logical Architecture*

### 4.1.4 Data Mesh

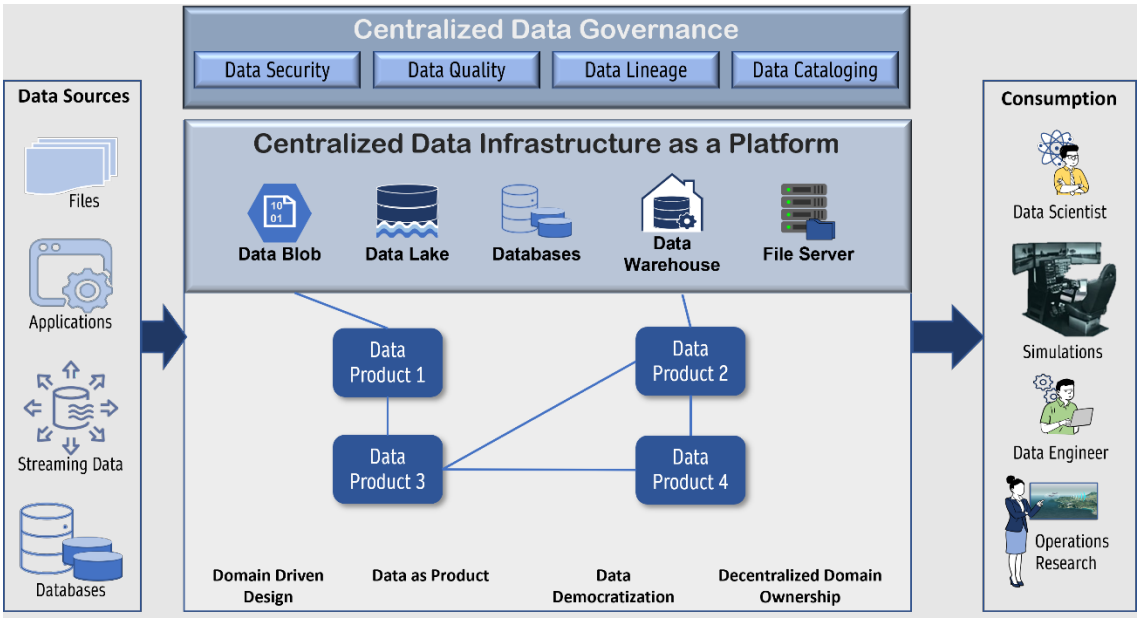


*Figure 8. Data Mesh Logical Architecture*

Data Mesh architecture, **Figure 8**, is based on Domain Driven Design and aims at delivering Data as a Product (DaaP). The idea is to give the ownership and onus to the domain teams to build and govern the data products and expose the service to serve the data product to other domains, a concept called Data as a Service (DaaS). Data products help solve latency problems due to scaling issues when large data sets are made interoperable inside a data fabric. All data doesn't need to sit in a single Data Lake but consists of its own set of data stores such as object storage (object stores), RDBMS, Data Warehouse, Data Lake, Data Lakehouse, etc. In other words, Data Mesh relies on the concept of Federated Data. Rather than looking at enterprise data as one huge data repository, data mesh considers it as a set of repositories of data products.

### 4.1.5 Comparison of Data Patterns

In Table 4 a comparison of each of the patterns is discussed based on potential components necessary within the simulation enterprise specific to data. This table provides a general comparison; however, the specific features and capabilities of each approach may vary depending on the implementation and tools used to develops the system.

*Table 4. Data Patterns Comparison Chart*

| Component | ETL | ELT | Data Mesh | Data Fabric |
|---|---|---|---|---|
| **Schema** | Defined before load | Defined before load | Defined within domains and data products | Defined based on target system |
| **Data** | Transformed before load | Loaded as raw data and transformed within the target system | Owned by autonomous data teams | Unified view of data across sources |
| **End User Experience** | Delayed data availability | Near real-time or real-time data availability | Improved data discovery and accessibility | Seamless access to integrated data |
| **Positive Features** | • Data quality and consistency,<br>• Supports Legacy Architectures<br>• Source system decoupling | • Scalability,<br>• Flexibility,<br>• Superior Data Tagging<br>• Real-time insights | • Decentralized ownership,<br>• Scalability,<br>• Reduced data silos | • Data integration,<br>• Scalability,<br>• Simplified management |
| **When is the schema applied** | Before loading data | Before and after loading data | Defined within data products | Defined based on target system |
| **Speed or processing** | Flexible Data loading process | Efficient processing of large volumes of data | Depends on domain teams and data product implementations | Scalable and elastic processing |
| **Use of SQL** | Compute intensive data workflows | Commonly used for transformations within the target system | Depends on the implementation | Depends on the target system and technologies used |

| Component | ETL | ELT | Data Mesh | Data Fabric |
|---|---|---|---|---|
| **Agility** | Sufficient for Traditional Data processing | High agility due to real-time or near real-time processing | Promotes autonomy and agility of data teams | Promotes flexibility and agility in data operations |

## 4.2 Integration Patterns

Integration patterns are intended to provide a minimum necessary standard for interface, component, and workflow implementations to ensure consistency and technical interoperability.

### 4.2.1 Mesh App and Service Architecture (MASA)

MASA is a non-proprietary architectural model introduced by Gartner and focuses on the dynamic connection of people, process, services, devices, and things. This pattern is multidimensional in which an application is an interconnected mesh of independent services and applications sharing functionality with other applications and external systems via APIs.

"MASA is a distributed architecture where the application comprises multiple fit-for-purpose application experiences provided by front-end UIs and multiple back-end services that provide core application functionality. MASA supports a heterogeneous and multi-grained back-end ecosystem where functionality may come from custom systems, data virtualization, large systems of record, microservices or enterprise applications.

In MASA, API mediation provides an abstraction layer that insulates the concerns of the app experience design away from the constraints and implementation details of the enterprise application functional logic. This abstraction allows the mesh applications to integrate capabilities and data from multiple enterprise applications and other systems and services, while supporting seamless user interactions across multiple channels that support a variety of user experiences." (Dayley & Skowron, 2021)

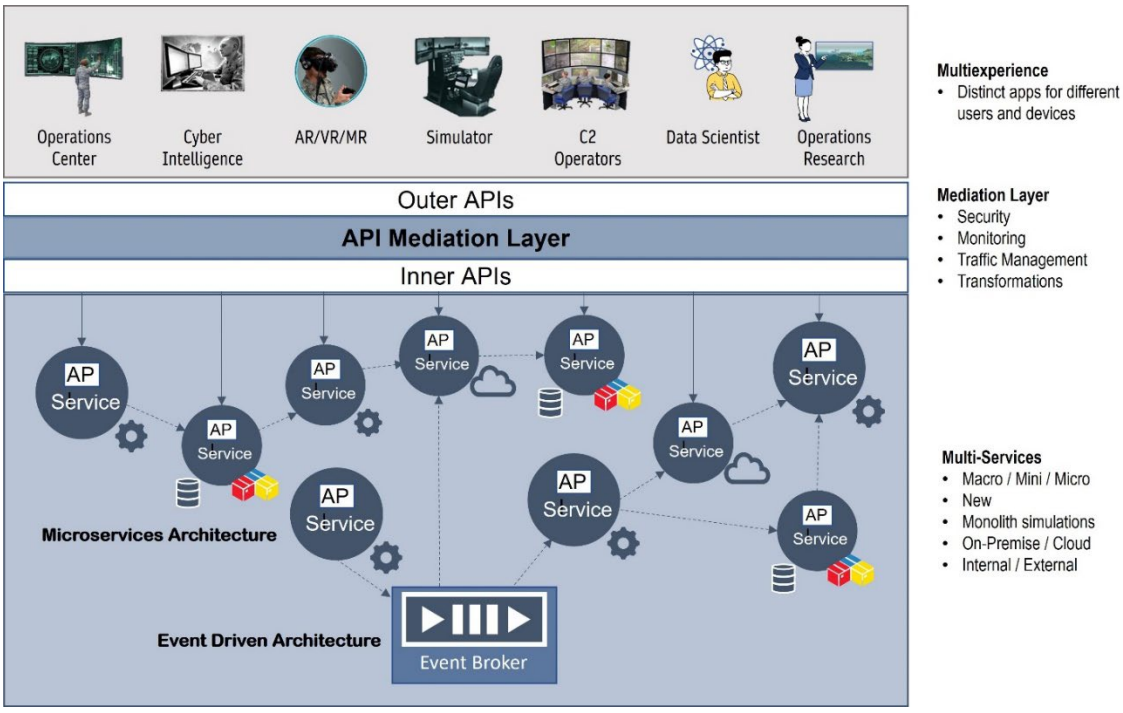


*Figure 9. Simulation Applications in a Mesh App and Service Architecture*

The MASA pattern shown in **Figure 9** is a demonstration of how this could apply to the simulation ecosystem. The user interface becomes part of the app mesh that delivers a continuous user experience, and all benefit from modifications or updates to services within the ecosystem. As monolith simulations migrate towards a digital environment, they can still function within this environment and all users would benefit. The benefits for adopting this pattern are:

- **Agility**: A decoupled component architecture enables development teams to adapt to changing requirements and replace aging components.
- **Scalability**: Scalability enables the enterprise to efficiently expand or collapse the training capacity for specific training features or capabilities in a fine-grained and dynamic manner. Independent services enable the simulation to scale specific capabilities independently of others, whereas large monolithic simulation systems require entire instances of the simulation system just to support increased demand for a narrow scope of functionality.
- **Cohesive UX**: MASA enables cohesive user experiences by providing access across multiple devices and channels, integrated with training needs.
- **Extend legacy systems**: This modular service architecture provides the ability to create new independent app experiences and custom integrations using legacy system capabilities, as well as add capabilities from other services to enhance legacy applications.
- **Integration**: Using a mediated integration approach provides the ability to connect varying technologies with enterprise simulation services, functionality and data using APIs.
- **Modernization**: The distributed nature of abstracted discrete components enables modernization of individual parts at a pace the USAF can tolerate.

- **Innovation**: The very structure of the component architecture enables development teams to optimize and innovate with new technologies, provide new user interfaces, and experiment with providing new capabilities.
- **Consistency**: Using an API mediation layer simplifies security and access policy enforcement, traffic management, monitoring, and logging by providing a single place to configure policies and monitor the APIs.
- **Faster delivery**: Decoupling the apps and services from each other using APIs enables independent release cadences. This provides the developers to release new capabilities faster, get feedback, adapt, and innovate for some M&S applications, while using a slower cadence for those capabilities or systems that rarely change.
- **Reduced technical debt**: This supports the use of open standards technologies and the ability to minimize proprietary implementations which result in technical debt. Additionally, a modular architecture allows development teams to iteratively reduce technical debt over time rather than a "big bang" approach, which would cause unnecessary delay in M&S Training & Readiness capabilities.

### 4.2.2 Edge Hybrid

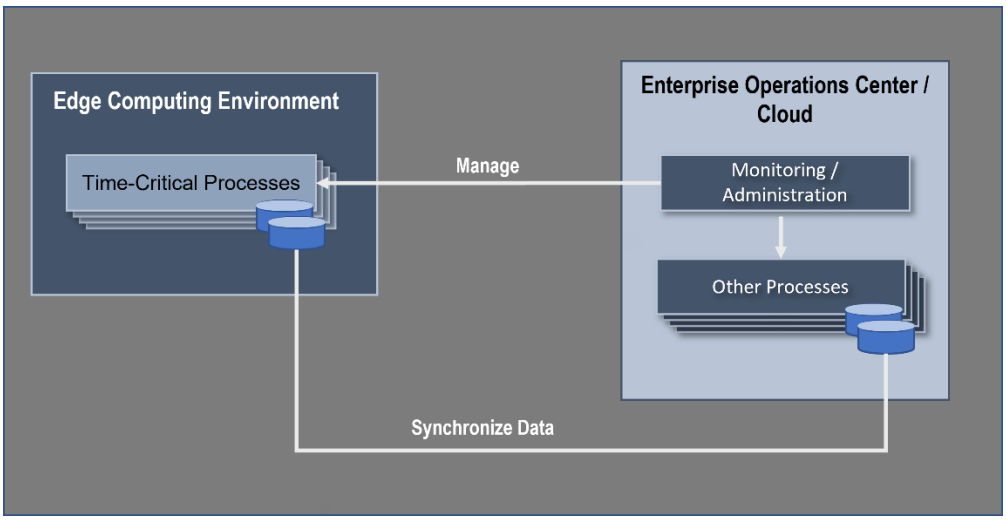


*Figure 10. Edge Hybrid Logical Architecture*

A hybrid cloud refers to an infrastructure for storage and computing that combines private cloud services, a public cloud, and/or on-premises infrastructure. These different resources are orchestrated to seamlessly work together. To effectively run workloads in the cloud, clients need fast and reliable network connectivity. In today's world, network connectivity is generally sufficient and not a hindrance to cloud adoption. However, there are situations where continuous connectivity cannot be relied upon. In an edge hybrid setup, the network link serves a noncritical role primarily for management purposes and asynchronous data synchronization or upload, rather than being essential for time-critical computing or transactions. The edge hybrid pattern, **Figure 10**, was designed and developed by a leading technology company and is used by commercial cloud providers to address challenges with running processes that had intermittent or unreliable network services, but still had a need to remain connected to a distributed architecture. This pattern ensures that the time-critical processes (e.g., sensor fusion) remain at the edge and the data is synchronized with the other distributed components. Additionally, when deploying

augmented reality (AR), a significant obstacle is achieving the necessary computational performance with extremely low latency to maintain a real-time experience. Placing demanding and time-critical processing and data streaming at the edge, near the AR device user, resolves these issues without imposing costly demands on the rest of the cloud architecture. This computational and low-latency performance also supports the simulation of complex electronic warfare (EW) interactions and provides an opportunity for model reuse in accordance with the recommendations of the DoD AMBIT report.

Compared to other cloud or non-cloud-based solutions, an Edge-Hybrid often emerges as the preferred choice for numerous organizations. It offers a range of advantages, including scalability, security, cost-effectiveness, control, and speed. [Fortinet]

The benefits of this approach are:

- **Ensure low latency:** Running significant processes that are time-critical at the edge ensures low-latency and self-sufficiency. This is especially true for the real-time processing needs for edge digital twins. Edge computing allows for real-time interaction with users because data is processed closer to where it is generated, allowing for real-time interaction.
- **Self-sufficiency:** Even in the event of network connectivity failure or temporary unavailability, the ability to carry out critical transactions remains intact. Simultaneously, leveraging the cloud for a substantial portion of the workload offers notable advantages.
- **Reuse existing infrastructure:** The existing infrastructure can be retained thus reducing cost expenditures while modifying software architecture to take advantage of new technology enhancements.
- **Incrementally refactor:** By maintaining the existing infrastructure, systems can be incrementally migrated to determine best services at the edge and communication and services across the enterprise.
- **Greater agility:** Edge-to-cloud platforms give organizations the flexibility to respond quickly to requests, capitalize on opportunities when they arise, and accelerate time-to-field for new products.
- **Performance flexibility:** A hybrid environment with edge computing capabilities brings computing resources closer to where they are needed by users, devices, etc., which can dramatically improve performance.

### 4.2.3 Hyperconverged Infrastructure (HCI)

"Hyper-convergence is a type of infrastructure approach that is largely software-defined with tightly integrated compute, storage, networking, and virtualization resources running on commodity hardware. This stands in contrast to a traditional converged infrastructure, where each of these resources is typically handled by a discrete hardware component that serves a singular purpose." [DoD Mod Strat, 2019]

Hyperconverged infrastructure (HCI) consolidates all local hardware and software components into a unified system (as depicted in **Figure 11**). This consolidation reduces complexity, simplifies management, and eliminates the need for separate hardware elements like routers and switches. HCI prioritizes user needs and significantly reduces the time required for deploying new services to mere minutes. This architectural approach aligns with the recommendation of the Defense Science Board (DSB) to identify specific simulations that can offer high-quality training comparable to major training centers while being operated from home stations. [DSB GEMS 2021] HCI provides a unified environment that seamlessly supports on-premises software that ensures data replication for time-sensitive access, addressing low-latency requirements.

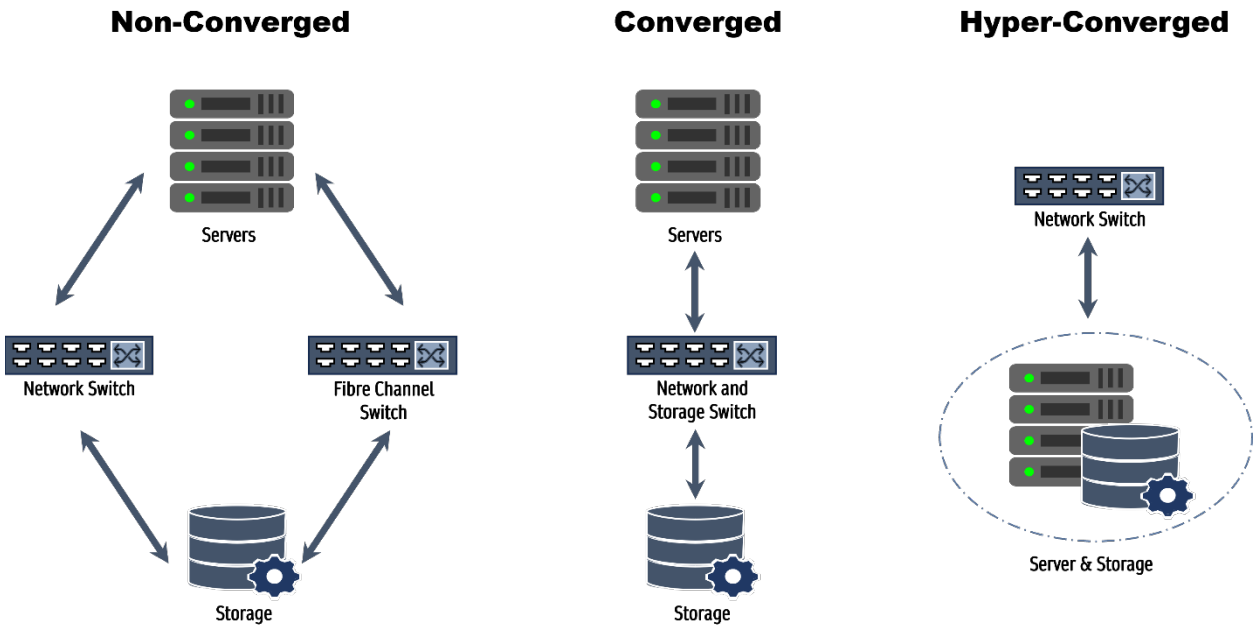


*Figure 11. Difference between non-converged, converged, and hyper-converged*

HCI offers the potential to reduce large portions of on-prem equipment (see figure 9) or eliminate technical debt and while eventually facilitating the implementation of other integration patterns in the future. Experts in the industry assert that HCI encompasses the benefits of cloud solutions while retaining the advantages and control of on-premises deployment. It simplifies complexity, enhances operational efficiency, enables rapid resource scalability, and optimizes IT infrastructures, [RIVER-1]

The benefits for adopting this pattern are:

- **Simplified Management:** HCI offers a centralized local management interface that simplifies the administration of the entire infrastructure stack. A unified management console decreases the need to manage separate components; hance troubleshooting is streamlined. [TECH-1]
- **Scalability:** HCI enables linear scalability, allowing local organizations the option to increase clusters to meet growing local requirements. This ensures local resources can be scaled on-demand without disrupting ongoing operations. [MINDSIGHT-1]
- **Improved Performance:** Hyperconverged systems leverage technologies at the local level to reduce data movements across wide-area networks, HCI minimizes latency. resulting in improved application performance. [INSIGHT-1]
- **Data Protection and Resiliency:** HCI often includes data redundancy and distributed storage, this provides a high level of resilience and minimize the risk of data loss. [MINDSIGHT-1]
- **Cost Efficiency:** HCI can lead to cost savings by simplifying management at the local level while offering more predictable scaling and enabling organizations to align infrastructure costs with their actual needs. [MINDSIGHT-1]
- **Ease of Migration:** Hyperconverged infrastructure simplifies the migration process by providing a standardized platform that abstracts the underlying hardware. This means applications and workloads can be easily moved between different nodes in the HCI cluster without needing to make

significant changes or reconfiguration. This flexibility facilitates seamless migration of workloads and minimizes downtime during the transition.

## 4.3 Migration Patterns

Migrating a simulation or application to the cloud or on-premises equipment (OPE) requires a review to determine the best fit from five alternative patterns:

- Rehost (Lift and Shift) - move simulation software to another platform (physical, virtual, or cloud) without modifying the code, features, or functions.
- Replatform (Lift and Reshape) - move to another platform and recode some software, but not the code structure, features, or functions.
- Refactor (Restructure and optimize) - modification of existing code to meet modern standards without changing external behavior to remove technical debt and improve nonfunctional attributes.
- Rearchitect (Re-Engineer) - create a new application architecture that enables improved performance and new capabilities.
- Rebuild (Rewrite, redesign) - complete recode from scratch while preserving the original scope and specifications.
- Replace (Repurchasing, "Drop and Shop") - sunset existing and purchase a new one considering new requirements and needs at the same time.

### 4.3.1 Strangler Pattern

The strangler pattern, **Figure 12**, enables development teams to migrate legacy monolithic systems and avoid the drawbacks of major code rewrites. This pattern was first described by Martin Fowler in 2004 (Fowler, 2004) and is analogous to the behavior of the Australian Strangler Fig vine. These vines seed in the upper branches of a host tree and slowly work their way down to root in the soil. As they continue to grow, the vines slowly take over thus strangling and killing the host tree.

The strangler pattern was initially developed to be event-driven, bi-directional communication between the legacy system and the new architecture. Matt Heusser (Heusser 2020) identified seven basic steps of the Strangler Pattern, which are depicted in Figure 12 and Figure 13. By adopting this pattern for legacy systems, developers can slowly deprecate a legacy system over time while incrementally adding new functionality. This approach leaves the legacy system in place meeting the end-user's needs and making valuable enhancements to the system using the new architecture. The strangler pattern enables the developers to focus on the pain points, solving high-value problems, and mitigating risks.

**Figure 13** demonstrates the iterative process needed to modify the legacy monolithic system into a modern data-centric software system. It begins with identifying specific service functions that can be refactored into subsystem modules; new subsystems can be added to the legacy system which align with modernized features. Later, a façade interface is created to abstract the calls and responses from the legacy. This interface now becomes the main contact point.

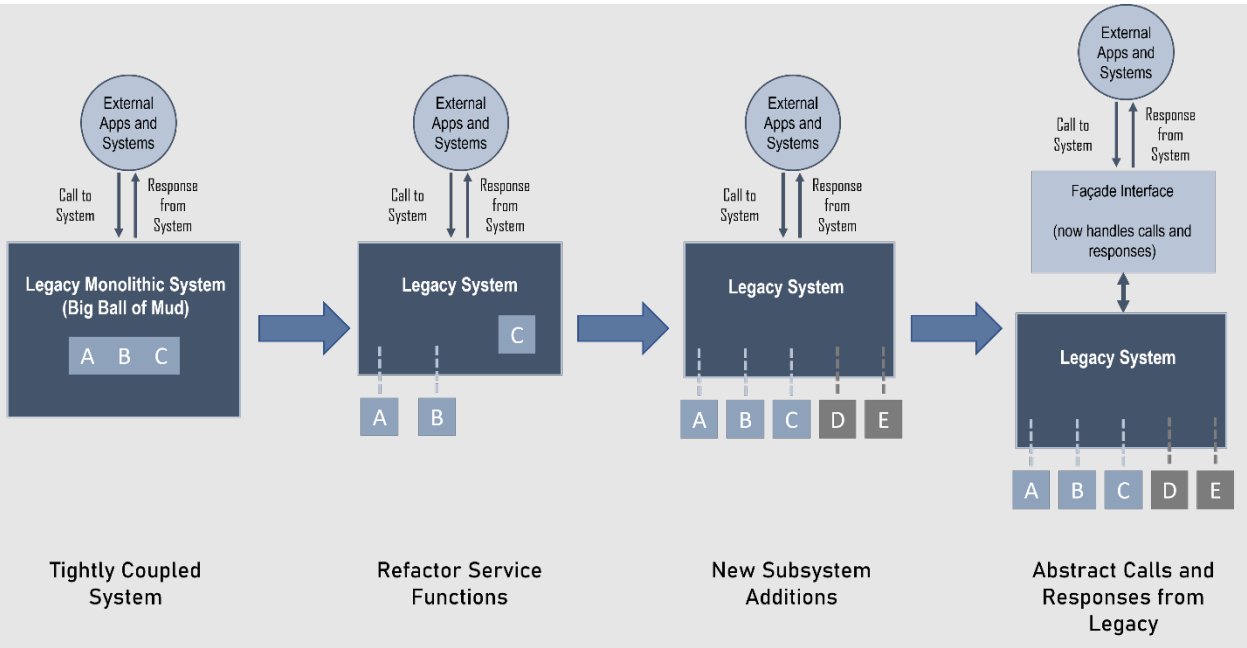


*Figure 12. Initial Steps to Refactor and Strangle Legacy*

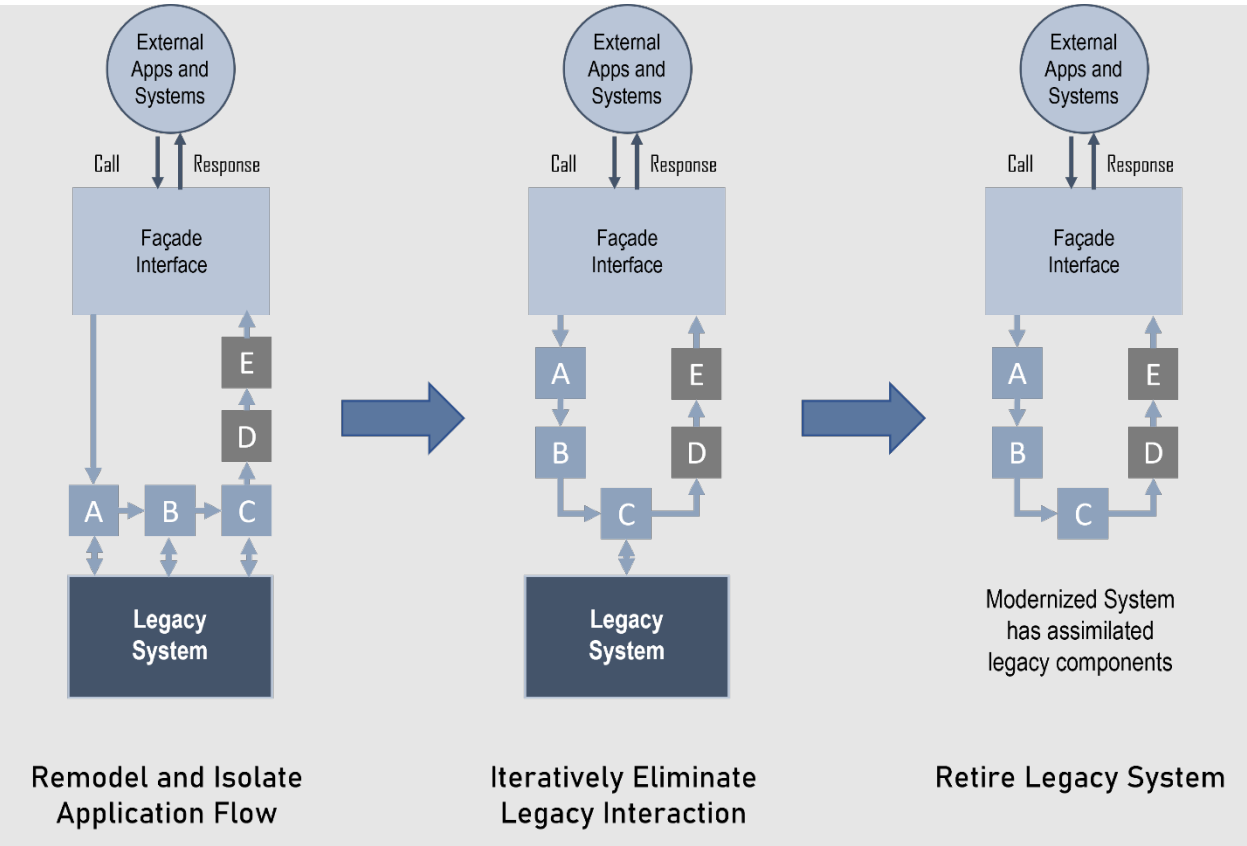


*Figure 13. Remodel and Iteratively Eliminate Legacy System*

As shown in **Figure 13**, the modifications continue with remodeling and isolating the application code while some services are still coupled to the legacy system. Continue with the modernization by iteratively eliminating subsystems interactions from their legacy system. Once all interactions have been assimilated into the modernized system, retire the legacy system. The façade remains as the main interface for all external applications and systems.

### 4.4 Additional Patterns

While this version of the M&S Synthetic Training GRA identifies the initial guidance to transition towards the ability to instantaneously provision realistic scenario-based training environments for any of our distributed warfighters, additional patterns will be incorporated in the next version of the GRA, since an RA is a living document.

## 5 Conclusion

As stated in the DAF Posture Statement 2022, "current platforms will not fully support tomorrow's demands" and that "we maintain air superiority in the future by introducing game-changing technology that includes digital engineering, open mission systems architecture and agile software." This resonates with the previous 2015 Air Force Future Operating Concept, which stated, "The current Air Force must design, plan and implement tangible decisions if it wishes to organize, train, equip, and provide future AF forces..." [AF FOC, 2015] Our ability to conduct training in uncertain and complex environments will rely on digital, synthetic, and encrypted capabilities to render realistic and relevant warfighting domains. This guidance is necessary to dynamically render realistic and relevant warfighting environments at scale across all domains to accomplish military endeavors successfully and swiftly with our Joint and coalition partners.

# 6 References

[AF 2035 FP Supp], Operational Training Infrastructure 2035 Flight Plan Supplement for Operational Modeling and Simulation, 17 June 2020

[AF FOC, 2015], https://www.af.mil/News/Article-Display/Article/617301/af-releases-future-operating-concept/

[DAF Posture, 2022], https://www.armed-services.senate.gov/download/joint-statement-061721

[DAF Data Platforms], https://usaf.dps.mil/sites/13057/CND/Shared%20Documents/PhaseIII/23/DAF-Data-Platforms.aspx

[Dayley & Skowron, 2021], Dayley, B., & Skowron, J. (2021, October 22). Using MASA to Deliver Agile Enterprise Application Architecture. https://www.gartner.com/document/4007276

[DMSRA, 2020], https://de-bok.org/asset/513f2b204cee1b4ae10168a4e0144dcc4bdcfea6

[DoD AMBIT Study 2023] DoD America's Mid-Band Initiative Team (AMBIT) Electronic Warfare (EW) Simulation Study, 22 Feb 2023.

[DoD CS RA, 2023] DoD Cybersecurity Reference Architecture v5.0, 30 Jan 2023, https://dodcio.defense.gov/Portals/0/Documents/Library/CS-Ref-Architecture.pdf

[DoD ICAM RD, 2010], DoD Enterprise ICAM Reference Design, June 2020, https://dodcio.defense.gov/Portals/0/Documents/Cyber/DoD_Enterprise_ICAM_Reference_Design.pdf

[DoD-IG-2019-081], Audit of Training Ranges Supporting Aviation Units in the U.S. Indo-Pacific Command, (2019, April 17) https://media.defense.gov/2019/May/08/2002129129/-1/-1/1/DODIG-2019-081.PDF

[DoD Mod Strat, 2019], DoD Digital Modernization Strategy, 12 Jul 2019, https://media.defense.gov/2019/Jul/12/2002156622/-1/-1/1/DOD-DIGITAL-MODERNIZATION-STRATEGY-2019.PDF

[DoD Reference Architecture Description, June 2010] https://dodcio.defense.gov/Portals/0/Documents/Ref_Archi_Description_Final_v1_18Jun10.pdf

[DoD ZT RA] Zero Trust Reference Architecture, v2.0 Jul 2022, https://dodcio.defense.gov/Portals/0/Documents/Library/(U)ZT_RA_v2.0(U)_Sep22.pdf

[Executive Order (EO) 14028 on Improving the Nation's Cybersecurity, 12 May 2021, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

[Fortinet] Hybrid Cloud vs. Multi-cloud: What's the Difference? https://www.fortinet.com/resources/cyberglossary/what-is-hybrid-cloud

[GART-1], https://www.gartner.com/en/information-technology/glossary/scalability

[INSIGHT-1] “6 Pros (and 3 Cons) of Hyperconverged Infrastructure” https://www.insightsforprofessionals.com/it/data-center/hyperconverged-infrastructure-pros-and-cons

[Fowler, 2004], Fowler, M. (2004, June 29). Strangler Fig Application. Retrieved from Martin Fowler: https://martinfowler.com/bliki/StranglerFigApplication.html

[Heusser, 2020], Heusser, M. (2020, June 29). What is the strangler pattern and how does it work? Retrieved from TechTarget Network: https://www.techtarget.com/searchapparchitecture/tip/A-detailed-intro-to-the-strangler-pattern

[HPE-1], https://www.hpe.com/us/en/what-is/composable-infrastructure.html

[MINDSIGHT-1] “5 Benefits of Hyperconvergence (HCI): An Infrastructure Report” https://gomindsight.com/insights/blog/5-benefits-of-hyper-convergence/

[NIST 800-207, 2020], https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

[NSA Press Release], NSA Press Release "NSA Releases Recommendations for Maturing Identity, Credential, and Access Management in Zero Trust,” 14 March 2023, https://www.nsa.gov/Press-Room/Press-Releases-Statements/Press-Release-View/Article/3328152/nsa-releases-recommendations-for-maturing-identity-credential-and-access-manage/

[NSA CSI], National Security Agency | Cybersecurity Information Sheet on Advancing Zero Trust Maturity Throughout the User Pillar, version 1.1, April 2023, https://media.defense.gov/2023/Mar/14/2003178390/-1/-1/0/CSI_Zero_Trust_User_Pillar_v1.1.PDF

[OMB M-22-09], Moving the U.S. Government Toward Zero Trust Cybersecurity Principles,26 Jan 2022, https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf

[Puppet and CircleCI. (2020)], State of DevOps Report 2020. Retrieved from https://puppet.com/resources/report/state-of-devops-report/.

[RIVER-1], Red River, “An Introduction To Hyper-Converged Infrastructure”, https://www.redriver.com/wp-content/uploads/2020/03/Intro_to_HCI_Whitepaper-datacenter.pdf

[SPR-1], https://spr.com/data-science-back-basics-flexible/

[TECH-1] “Why hyper-converged infrastructure simplifies IT management” https://www.techtarget.com/searchdatacenter/feature/Why-hyper-converged-infrastructure-simplifies-IT-management

# Appendix A: Vocabulary

Vocabulary provides context dependent ontology of semantic classification and meaning of the acronyms, terms and definitions of architecture elements used within the subject area. It enables a common understanding of terms and consistency of definitions used across the subject area. It includes acronyms, a taxonomy of terms, and definitions that are used in the GRA and relevant to solution architectures.

**Abstraction**: The process of selecting the essential aspects of a Simuland to be represented in a model or simulation while excluding those aspects that are not relevant to the purpose of the model or simulation. The set of elements produced by this process. (https://ac.cto.mil/de-ms-glossary/)

**Architectures**: The structure of components in a program/system, their interrelationships, and the principles and guidelines governing their design and evolution over time. (DoDI 5000.70)

**Assumption**: A supposition on the current situation or a presupposition on the future course of events, either or both assumed to be true in the absence of positive proof, necessary to enable the commander in the process of planning to complete an estimate of the situation and decide on the course of action. (JP 1-02)

**Attribute**: A property or characteristic of an entity, object, or event (e.g., color, weight, sex, time of occurrence). (https://ac.cto.mil/de-ms-glossary/)

**Attribute Based Access Control (ABAC):** An access control method where subject requests to perform operations on objects are granted or denied based on assigned attributes of the subject, assigned attributes of the object, environment conditions, and a set of policies that are specified in terms of those attributes and conditions. (Hu, Vincent et. Al., "Guide to Attribute Based Access Control (ABAC) Definition and Considerations", *NIST Special Publication 800-162,* January 2014).

**Augmented reality**: Technology that allows you to be add information to the visual environment around you. Examples would include recognizing and adding information to objects or generating digital holograms in physical space. This contrasts with virtual reality which is a completely immersed experience. (https://ac.cto.mil/de-ms-glossary/)

**Authoritative data**: A recognized or official data production source with a designated mission statement, source, or product that publishes reliable and accurate data for subsequent use by customers. An authoritative data source may be the functional combination of multiple separate data sources. (DoDI 8320.05)

**Battlespace:** The physical environment in which the simulated warfare will take place and the forces that will conduct the simulated warfare. All elements that support the front-line forces (i.e., logistics, intelligence) are included. (https://ac.cto.mil/de-ms-glossary/)

**Coalition**: An arrangement between two or more nations for common action. Source: JP 1-02 (2016)

**Common use:** Services, materiel, or facilities provided by a DoD agency or a Military Department on a common basis for two or more DoD agencies, elements, or other organizations as directed. (DoDI 5000.61)

**Commonality:** A quality that applies to materiel or systems: a. possessing like and interchangeable characteristics enabling each to be utilized, or operated and maintained, by personnel trained on the others

without additional specialized training; b. having interchangeable repair parts and/or components; and c. applying to consumable items interchangeably equivalent without adjustment. (JP 1-02)

**Common-use M&S:** M&S applications, services, or materials provided by a DoD component to two or more DoD components. (https://ac.cto.mil/de-ms-glossary/)

**Community of Interest (COI):** A collaborative group of users (working at the appropriate security level or levels) who exchange information in pursuit of their shared goals, interests, missions, or business processes, and must have a shared vocabulary for the information exchanged. The group exchanges information within and between systems. (Committee on National Security Systems (CNSS) Glossary, CNSSI 4009)

**Composable:** In a composable infrastructure, compute, storage, and networking resources are abstracted from their physical locations and can be managed by software through a web-based interface. Composable architecture is the process of scaling storage, networks, databases, and compute functionality in a more agile fashion. In composable architecture, APIs are used to enable flexibility and ecosystem management. Building a composable architecture helps you scale — and compose — your IT footprint faster. ("What is Composable Infrastructure?", Hewlett Packard Enterprise, www.hpe.com)

**Composable M&S Enterprise:** The composable M&S enterprise is created from interchangeable building blocks, which allow it to rearrange and reorient as needed in response to internal or external factors. These building blocks can be reused across the enterprise among different M&S groups, increasing the efficiency of organizations in implementing novel applications in response to changes.

**Composability:** The capability to select and assemble reusable modeling and simulation components in various combinations into simulation systems to meet user requirements. (https://ac.cto.mil/de-ms-glossary/)

**Conceptual model**: The description of what the model or simulation will represent, the assumptions limiting those representations, and other capabilities needed to satisfy the user's requirements. A collection of assumptions, algorithms, relationships, and data that describe a developer's concept about the simulation. (https://ac.cto.mil/de-ms-glossary/)

**Containers:** Cloud, desktop and IT services build on open-source technologies to facilitate and accelerate application capabilities. They are executable units of software in which application code is packaged along with libraries and dependencies in standard architectures that allow execution in most platforms. They are operating system virtualizations that leverage isolated processes that control CPU, memory, and IO utilization. They contain only software, libraries, and dependencies, not the underlying hardware nor operating systems. The absence of the guest operating system and its complexity, allows containers to be extremely lightweight, fast, portable applications that can scan horizontally in the cloud. They fit modern application development (Ci/CD) patterns, like DevSecOps, DevOps, serverless and microservices. (dla.mil)

**Constraint:** An externally imposed limitation on a process, model, or dataset. An equation (equality or inequality) that must be satisfied for a possible solution to be determined feasible. (https://ac.cto.mil/de-ms-glossary/)

**Cross domain solution:** A form of controlled interface that provides the ability to manually and/or automatically access and/or transfer information between different security domains. (NIST SP 800-37 Rev 1)

**Data:** A representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means. (DoDD 8320.02)

**Data as a Product (DaaP):** The result of applying product thinking into datasets, making sure they have a series of capabilities including discoverability, security, explorability, understandability, trustworthiness, etc. (towardsdatascience.com/data-as-a-product-vs-data-products)

**Data as a Service (DaaS):** A strategy made to simplify the world of storing and processing large amounts of data. (towardsdatascience.com/daas-data-as-a-service)

**Data asset:** Any entity that is comprised of data. For example, a database is a data asset that is comprised of data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application. For example, a service that returns individual records from a database would be a data asset. Similarly, a website that returns data in response to specific queries (e.g., www.weather.com), would be a data asset. A human, system, or application may create a data asset. (DoDD 8320.02)

**Data-centric:** An environment where data is the primary and permanent asset separated from systems/applications making data available to a broad range of tools and analytics within and across security domains of enrichment and discovery. (ref. IC Data Management Lexicon, January 2020)

**Data dictionary:** A specialized type of database containing metadata that is managed by a data dictionary system. A repository of information describing the characteristics of data used to design, monitor, document, protect, and control data in information systems and databases. (https://ac.cto.mil/de-ms-glossary/)

**Data logger:** A device that accepts data outputs from a simulation or federation and stores them for processing and replay. (https://ac.cto.mil/de-ms-glossary/)

**Data model:** In a database, the user's logical view of the data in contrast to the physically stored data, or storage structure. A description of the organization of data in a manner that reflects the information structure of an enterprise. (SISO-REF-002-1999)

**DevSecOps:** (Development, Security, and Operations) automates the integration of security at every phase of the software development lifecycle. DevSecOps is a culture and an evolution where security is given the utmost priority in the software development life cycle. (DevSecOps, IBM Cloud Education, 30 July 2020)

**Digital twin:** Virtual model designed to accurately represent a physical object. Sensors produce data that when aggregated reproduce all aspects of a physical object's performance (energy output, temperature, geospatial orientation, relative orientation in space, weather conditions). The digital twin is used to run a simulation to train a user, experiment to create improvements, and test the physical objects performance inside a digital environment. Digital twin simulations continuously adjust as new sensor data is consumed leading to both higher fidelity and credibility overtime. Fully functional digital twins provide comprehensive and predictive analytics that are not available via conventional analytics (sbir.gov).

**Distributed simulation:** A simulation that has multiple modules, which can be run on multiple processors. The processors can be co-located in the same room or located in remote sites. (https://ac.cto.mil/de-ms-glossary/)

**DoD Community**: A DoD activity area, enabled by M&S, that has an established executive-level management structure. Examples of such activities that meet these criteria include acquisition, training, and analysis. (DoDD 5000.59)

**DoD Community M&S Strategic Plan:** A high-level DoD plan created and published by a DoD Community to facilitate and achieve the DoD M&S Strategic Vision, goals, and objectives for that DoD Community. (DoDI 5000.70)

**Ecosystem**: Digital ecosystem utilizes computerization and tools in a shared environment to facilitate a groups achievement of common goals. The utilitarian use of common technologies saves effort, resources, and benefits all. The sum is greater than its parts. By condensing processes and data into a smart responsive environment the group makes huge strides in advancement not otherwise available. The ceiling of possibilities has risen exponentially using containers, machine learning and serverless platforms. Each ecosystem is a unique "stack" of software pieces that serve a purpose on their own and together allowing the group to move faster and maximize. (dodcio.defense.gov)

**Edge device:** Any piece of hardware that controls data flow at the boundary between two networks. (https://www.techtarget.com/searchnetworking/definition/edge-device)

**Endpoint**: Endpoints are often defined as end-user devices, such as mobile devices, laptops, and desktop PC's, although hardware such as servers in data centers are also considered endpoints. This can be any device that transmits a network packet. Devices such as zero clients, virtualized systems, and infrastructure equipment (i.e., routers and switches) are considered endpoints. (Endpoint Security Vendors, The Cyber Research Databank, www.cyberdb.co)

**Enterprise:** An arbitrarily defined functional and administrative entity that exists to perform a specific, integrated set of missions, and achieve associated goals and objectives, encompassing all the primary functions necessary to perform those missions. (DoDI 5000.70)

**Enterprise model**: Information model(s) that presents an integrated top-level representation of processes, information flows, and data. (https://ac.cto.mil/de-ms-glossary/)

**Exercise**: A military maneuver or simulated wartime operation normally involving planning, preparation, execution, and after-action review. Also see simulation exercise. (https://ac.cto.mil/de-ms-glossary/)

**Experimentation**: A process using simulation to identify, develop, assess, and recommend changes to doctrine, organizational structure, training, materiel, leadership and education, people, and facilities required to achieve advances in operational capabilities. (https://ac.cto.mil/de-ms-glossary/)

**Extensibility:** The ability of a model, simulation, or data structure to accommodate additional values or iterations of data over time without impacting the initial design. (https://ac.cto.mil/de-ms-glossary/)

**Extract, Load, and Transform**: The processes a data pipeline uses to replicate data from a source system into a target system such as a cloud data warehouse. (www.stitchdata.com/resources/what-is-elt)

**Extract, Transform, and Load**: The general procedure of copying data from one or more sources into a destination system which represents the data differently from the source(s) or in a different context than the source(s). (https://www. https://www.leadingedge-group.com/glossary/etl)

**Fidelity**: The degree to which a model or simulation represents the state and behavior of a real-world object or the perception of a real-world object, feature, condition, or chosen standard in a measurable or perceivable manner; a measure of the realism of a model or simulation. (https://ac.cto.mil/de-ms-glossary/)

**Infrastructure:** The supporting hardware, software, communication, and information security services that a system requires to operate but can be shared by multiple systems for scalability. (https://ac.cto.mil/de-ms-glossary/)

**Instantiation**: The creation of a real instance or realization of an abstraction or template such as a class of objects or a computer process. Refers to the creation of an object (or an "instance" of a given class) in an object-oriented programming (OOP) language. (Zola, Andrew, "Instantiation", TechTarget)

**Latency:** (1) The time delay between action and result. (2) The time delay between any two simulators, from submitting a message from the sending simulation to receiving this message by the recipient simulation. (3) The time interval required for a simulation to begin its response to a stimulus after it has been presented with a stimulus or stimuli (e.g., input of data, occurrence of an event). (4) The time interval required by a simulation to respond to a stimulus more than the time interval required for the corresponding real world or standard event. (https://ac.cto.mil/de-ms-glossary/)

**Legacy model (legacy simulation, legacy M&S):** Any model or simulation that was developed either in the past or for a different purpose. (https://ac.cto.mil/de-ms-glossary/)

**Metadata**: Searchable information describing the characteristics of data; data or information about data; or descriptive information about an object's data, data activities, systems, and holdings. For example, metadata for a model or simulation will include keywords and a description of the capabilities along with developer and user information. (DoDD 8320.02)

**Multi-Level Security (MLS):** Concept of processing information with different classifications and categories that simultaneously permits access by users with different security clearances and denies access to users who lack authorization. (CNSSI 4009)

**Modeling and simulation (M&S):** (1) The discipline that comprises the development and/or use of models and simulations. (DoDD 5000.59, DoDI 5000.61). (2) The use of models, including emulators, prototypes, simulators, and stimulators, either statically or over time, to develop data as a basis for making managerial or technical decisions. (https://ac.cto.mil/de-ms-glossary/)

**M&S assets:** M&S tools, data, and services, including models and simulations, and data assets. (https://ac.cto.mil/de-ms-glossary/)

**Modeling and simulation (M&S) data, M&S data:** Data used to develop models or simulations, and/or used as input to models and simulations, and/or data produced by models and simulations. (https://ac.cto.mil/de-ms-glossary/)

**Modeling and Simulation (M&S) Community of Interest:** A collaborative group of users that must exchange modeling and simulation information in pursuit of its shared goals, interests, missions, or business

processes and therefore must have shared vocabulary for the information it exchanges. (https://ac.cto.mil/de-ms-glossary/)

**Modeling and simulation (M&S) services**: An activity that enhances the ability to effectively and efficiently use M&S to accomplish a mission. (DoDD 5000.59). Examples include M&S standards development and promulgation, technical interoperability, Verification, Validation, and Accreditation (VV&A) process development, and workforce development. (DoDI 5000.70 addition)

**OpenXR:** An application programming interface (API) specification for the delivery of 3D, Augmented Reality and Virtual Reality software. (www.metaverse.io)

**Operational data**: Data and information created from data, used in, or in support of, a combined/joint operation, by the Combined/Joint Force Commander, their components, and operating forces that support any appropriate operational use (e.g., planning, analysis, and assessment of friendly and enemy activity, etc.) to increase speed of informed decisions and operational effectiveness.

**Operational M&S**: The policies, processes, and resources that comprise the oversight and/or use of operationally relevant models and simulations for the warfighter to organize, train, equip, maintain operational readiness, and prepare current and future forces across the full spectrum of military operations.

**Operational Training**: Mission essential task training in support of operational forces readiness. Distinguishes itself from initial skills training due to its focus on employment of weapon systems and skills in an operational setting. Spans all domains and Tier 1 (large, strategic) to Tier 4 (small, tactical) events across the breadth of multiple security levels and releasability restrictions.

**Operational Training Infrastructure**: The framework and resources essential to accomplishing Air Force Operational Training objectives. It includes such elements as embedded training capability, training systems, airspace, ranges and off-range lands, scoring & feedback systems, targets, pods/instrumentation/weapon system interface devices, aggressors/contract air, threat environment generators, networks, synthetic environments, operational training centers, workforce, and cybersecurity.

**Platform as a Service (PaaS):** The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment. (NIST SP 800-145)

**Realistic Training**: The deliberate practice of individual and collective tasks to enable tactical and technical proficiency that support mission accomplishment in a training environment that approximates the operational environment in both sufficient complexity and substance. ("Enhanced Realistic Training White Paper: Delivering Training Capabilities for Operations in a Complex World", United States Army Combined Arms Center, 26 January 2016)

**Reference Architecture**: An authoritative source of information about a specific subject area that guides and constrains the instantiations of multiple architectures and solutions. (DoD Reference Architecture Description, June 2010)

**Representation**: Models of the entity or phenomenon associated and its effects. Representations using algorithms and data that have been developed or approved by a source having accurate technical knowledge are often considered authoritative. (https://ac.cto.mil/de-ms-glossary/)

**Requirement:** An established need justifying the timely allocation of resources to achieve a capability to accomplish approved military objectives, missions, or task. (DAFPD 16-10)

**Reuse**: The practice of using again, in whole or part, existing M&S tools, data, or services. (DoDD 5000.59)

**Role Based Access Control (RBAC)**: Access control based on user roles (i.e., a collection of access authorizations a user receives based on an explicit or implicit assumption of a given role). Role permissions may be inherited through a role hierarchy and typically reflect the permissions needed to perform defined functions within an organization. A given role may apply to a single individual or to several individuals. (NIST SP 800-53 Rev 4)

**Scalability**: The ability of a simulation to maintain time and spatial consistency as the number of entities and accompanying interactions increase. (SISO-REF-002-1999)

**Simulation**: A method for implementing a model over time. (DoDD 5000.59, DoDI 5000.61, DoDI 5000.70)

**Simulation environment**: The operational hardware, software including databases, communications, and infrastructure in which a simulation operates. (https://ac.cto.mil/de-ms-glossary/)

**Simulator**: A device, computer program, or system that performs simulation. (IEEE 610.3-1989)

**Stakeholders**: M&S stakeholders are individuals and organizations who are developers and/or users of DAF-related modeling & simulation capabilities. This includes regular and reserve Air Force and Space Force personnel, DAF civil servants, support contractors, and the formal organizations and informal professional communities to which these individuals belong. (DAFPD 16-10)

**Standard**: A rule, principle, or measurement established by authority, custom, or general consent as a representation or example. (DAFPD 16-10)

**Synthetic Environment**: The integrated set of data elements that define the environment within which a given simulation application operates. The data elements include information about the initial and subsequent states of the terrain including cultural features and atmospheric and oceanographic environments throughout an exercise. The data elements include databases of externally observable information about instantiable entities and are adequately correlated for the type of exercise to be performed. More generally, it is the combination of simulators and constructive environments. (IEEE Std 1278.1-2012)

**Technical Debt:** Consists of design or implementation constructs that are expedient in the short term but set up a technical context that can make future change **costlier** or **impossible**. Technical debt may result from having code issues related to architecture, structure, duplication, test coverage, comments and documentation, potential bugs, complexity, coding practices, and style which may accrue at the level of overall system design or system architecture, even in systems with great code quality. DoDI 5000.87, dated 2 Oct 2020 (www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500087p.pdf)

**Technical interoperability:** The ability of a model or simulation to provide services to and accept services from other models and simulations, and to use these exchanged services to operate effectively together. (SISO-REF-002-1999)

**Virtual:** An entity or data that is derived from a modeled or simulated representation of the actual or anticipated system. (https://ac.cto.mil/de-ms-glossary/)

**Virtual reality**: An environment represented by models and simulations made possible by technology that allows you to be fully immersed into a virtual world. Screens are attached directly to one's head via a headset so that head and eye-tracking makes it possible to look around in the virtual environment. This contrasts with Augmented Reality, which layers the spatial computing environment on top of the physical world around you. This environment is interactive, allowing the participant to look and navigate about the environment, enhancing the immersion effect. Also known as virtual environment and virtual world. (https://ac.cto.mil/de-ms-glossary/)

**Virtual simulation**: A simulation involving real people operating simulated systems. (https://ac.cto.mil/de-ms-glossary/)

**Warfare simulation**: A model of warfare or any part of warfare. (https://ac.cto.mil/de-ms-glossary/)

**Zero Trust**: Zero Trust is the term for an evolving set of cybersecurity paradigms that move defenses from static, network-based perimeters to focus on users, assets, and resources. Zero Trust assumes there is no implicit trust granted to assets or user accounts based solely on their physical or network location (i.e., local area networks vs. the Internet) or based on asset ownership (enterprise or personally owned. [NIST SP 800-207]

## Appendix B: Acronyms

ABAC – Attribute-Based Access Control

AFAMS – Air Force Agency for Modeling and Simulation

AI – Artificial Intelligence

API – Application Programming Interface

AR – Augmented Reality

CCMD – Combatant Command

CDS – Cross Doman Solution

COP – Common Operating Picture

DAF – Department of the Air Force

DCA – Data Centric Architecture

DM – Data Management

DoD – Department of Defense

DCSA – Data-Centric Security Architecture

DSRA – Data Services Reference Architecture

ELT – Extract, Load, and Transform

ETL – Extract, Transform, and Load

EW – Electronic Warfare

GRA – Government Reference Architecture

M&S – Modeling and Simulation

MASA – Mesh App and Service Architecture

ML – Machine Learning

MLS – Multi-Level Security

OSA – Open Systems Architecture

RA – Reference Architecture

RBAC – Role Based Access Control

SOA – Service-Oriented Architecture

SysML - Systems Modeling Language

STANAG - Standardization Agreement

USAF – United States Air Force

VR – Virtual Reality

XR – eXtended Reality