180-22

Range Commanders Council
Data Systems Group

TRMC BIG DATA ANALYTICS IMPLEMENTATION GUIDE

ABERDEEN TEST CENTER
DUGWAY PROVING GROUND
ELECTRONIC PROVING GROUND
REAGAN TEST SITE
REDSTONE TEST CENTER
WHITE SANDS TEST CENTER
YUMA PROVING GROUND

NAVAL AIR WARFARE CENTER AIRCRAFT DIVISION PATUXENT RIVER
NAVAL AIR WARFARE CENTER WEAPONS DIVISION CHINA LAKE
NAVAL AIR WARFARE CENTER WEAPONS DIVISION POINT MUGU
NAVAL SURFACE WARFARE CENTER DAHLGREN DIVISION
NAVAL UNDERSEA WARFARE CENTER DIVISION KEYPORT
NAVAL UNDERSEA WARFARE CENTER DIVISION NEWPORT
PACIFIC MISSILE RANGE FACILITY

96th TEST WING
412th TEST WING
ARNOLD ENGINEERING DEVELOPMENT COMPLEX

SPACE LAUNCH DELTA 30
SPACE LAUNCH DELTA 45

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

This page intentionally left blank.

RCC 180-22

TRMC BIG DATA ANALYTICS IMPLEMENTATION GUIDE

October 2022

Prepared by

Data Sciences Group
Range Commanders Council

Published by

Secretariat
Range Commanders Council
US Army White Sands Missile Range
New Mexico 88002-5110

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

Preface

An enterprise approach for knowledge management and big data analytics will provide for more robust and reliable test data storage, greater discoverability and accessibility to this test data, and significantly improved ability to perform large-scale data analyses upon this test data. This results in improved test and evaluation of DoD weapon systems and hence better weapon systems for the warfighter. An effective enterprise approach requires a well-designed implementation that is agile, sustainable, and affordable. All test ranges will benefit from guidance on and evaluation tools for TRMC Big Data Analytics Architecture implementations.

For questions about this document, contact the Range Commanders Council Secretariat.

Secretariat, Range Commanders Council
ATTN: TEWS-TDR
1510 Headquarters Avenue
White Sands Missile Range, New Mexico 88002-5110
Telephone: (575) 678-1107, DSN 258-1107
E-mail: rcc-feedback@trmc.osd.mil

This page intentionally left blank.

Acronyms

BDAA	Big Data Analytics Architecture
CNCF	Cloud Native Computing Foundation
DevSecOps	development, security, and operation
FOSS	free and open-source software
MOSA	modular open system approach
OCI	Open Container Initiative
T&E BDA-KM	Test and Evaluation Big Data Analytics and Knowledge Management
TRMC	Test Resource Management Center

This page intentionally left blank.

1. Introduction

The purpose of this document is to provide high-level implementation guidance for Test and Evaluation Big Data Analytics and Knowledge Management (T&E BDA-KM) software solutions that are based upon the Test Resource Management Center (TRMC) Knowledge Management and Big Data Analytics Architecture (BDAA) Framework as documented in the TRMC's *Knowledge Management and Big Data Analytics Architecture Framework*.[1] For simplicity of reference, the above publication will be referred to in this document as the "TRMC BDAA".

As an important point of clarification, this implementation guidance only addresses T&E BDA-KM software solutions based upon the TRMC BDAA. Furthermore, this implementation guidance only identifies high-level key characteristics, and more specifically the top five key characteristics, of these software solutions.

The rationale for only looking at the top five key characteristics is to provide a quick direction check for candidate T&E BDA-KM solutions rather than provide an in-depth analysis.

The organization of this implementation guidance is as follows.

a. Review the published DoD guidance pertinent to a viable T&E BDA-KM software implementation.

b. Identify the top five key characteristics of a viable T&E BDA-KM software implementation derived from DoD guidance.

c. Provide an implementation scorecard derived from the top five key characteristics of a viable T&E BDA-KM software implementation.

This implementation guidance was conducted by the Range Commanders Council Data Sciences Group and is a follow-on to *TRMC Big Data Analytics Architecture Assessment*.[2]

2. Department of Defense Guidance

This section reviews the published DoD guidance that is pertinent to a viable T&E BDA-KM software implementation. This review is by intent limited to DoD-level publications because a viable T&E BDA-KM software implementation must be applicable to all Services. While the set of reviewed publications is not intended to be comprehensive, it is intended to be representative.

2.1 DoD Cloud Strategy

The *DoD Cloud Strategy*[3] states the following.

---

[1] Test Resource Management Center. *Knowledge Management and Big Data Analytics Architecture Framework*. Version 13. 31 January 2019. Retrieved 27 June 2022. Available at https://www.trmc.osd.mil/wiki/display/JMETC/Big+Data+Knowledge+Management?preview=%2F55968739%2F55968745%2FBigDataArchitecture-v13-2019-01-31-DistA.pdf.

[2] Range Commanders Council. *TRMC Big Data Analytics Architecture Assessment*. SR-21-003. September 2021. May be superseded by update. Retrieved 31 October 2022. Available at https://www.trmc.osd.mil/wiki/x/7wTkBw.

[3] Department of Defense. *DoD Cloud Strategy*. December 2018. Retrieved 27 June 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

> The Department of Defense (DoD) has entered the modern age of warfighting where the battlefield exists as much in the digital world as it does in the physical. Data and our ability to process data at the ready are differentiators to ensure mission success. Cloud is a fundamental component of the global infrastructure that will empower the warfighter with data and is critical to maintaining our military's technological advantage. (*DoD Cloud Strategy*, Forward)
>
> DoD must enable decision makers to use modern data analytics, such as AI and machine learning (ML), at the speed of relevance to make time-critical decisions rapidly in the field to support lethality and enhanced operational efficiency. (*DoD Cloud Strategy*, p. 5)

The *DoD Cloud Strategy* identifies guiding principles for the implementation of an "extensible and secure cloud environment that spans the homeland to the global tactical edge, as well as the ability to rapidly access computing and storage capacity to address warfighting challenges at the speed of relevance." (*DoD Cloud Strategy*, p. 7) The principle that is particularly relevant to a viable T&E BDA-KM implementation is the following.

> Leverage Commercial Industry Best Practices: Maximizing competition to ensure that DoD is getting the best technology and value; and Leverage industry open standards and best practices to avoid lock-in and provide maximum flexibility for future cloud advances. (*DoD Cloud Strategy*, p. 8)

The Joint Enterprise Defense Contract was a key component of the *DoD Cloud Strategy*, and although it was cancelled and replaced with the Joint Warfighter Cloud Capability[4], the guiding principles of the *DoD Cloud Strategy* are still pertinent.[5]

2.2 DoD Artificial Intelligence Strategy

The *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*[6] states the following.

> AI [Artificial Intelligence] is rapidly changing a wide range of businesses and industries. It is also poised to change the character of the future battlefield and the pace of threats we must face. We will harness the potential of AI to transform all functions of the Department positively, thereby supporting and protecting U.S. service members, safeguarding U.S. citizens, defending allies and partners, and improving the affordability, effectiveness, and speed of our operations. The women and men in the U.S. armed forces remain our enduring source of strength; we will use AI-enabled information, tools, and systems to empower, not replace, those who serve. (*Summary of the 2018…*, p. 4)

[4] Department of Defense. *Future of the Joint Enterprise Defense Infrastructure Cloud Contract*. 6 July 2021. Retrieved 27 June 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.
[5] Department of Defense. *Interim Guidance for Implementation of the Department of Defense Cloud Strategy*. 16 April 2020. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.
[6] Department of Defense. *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*. 12 February 2019. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

The focus area that is particularly relevant to a viable T&E BDA-KM implementation is the following.

> *Engaging with the open-source community*. The open-source community is a vibrant global incubator of talented individuals and transformative ideas. We will contribute our data, challenges, research, and technologies to this community and engage with the open-source ecosystem as a vehicle for attracting talent, identifying and advancing new AI technologies that can transform defense, and broadening our accessible technology base." (*Summary of the 2018…*, p. 12)

2.3 DoD Digital Modernization Strategy

The *DoD Digital Modernization Strategy*[7] states the following.

> The DoD's Big Data Platform (BDP) is a secure, extensible, scalable, agile, and open infrastructure platform designed to provide a distributed computing solution. The BDP is based on DISA-developed open source technologies and is capable of ingesting, storing, and visualizing multiple petabytes of data. It acts as a common platform that can support a variety of cyberspace missions and organizations across DoD. The BDP provides the ability to perform aggregation, correlation, historical trending, and forensic analysis against structured and unstructured data from a variety of systems and sensors supporting both NIPRNET and SIPRNET environments. The DODIN Operations and DCO missions require a capability that can translate enterprise scale data into simple, dynamic visualizations that depict event relationships and answer leaders' information requirements. (*DoD Digital*…, p. 42)
>
> DoD's approach to Big Data Analytics is based upon a Big Data Platform (BDP), a secure, extensible, scalable, agile, and open infrastructure platform designed to provide a distributed computing solution. BDP instances are capable of ingesting, storing, and visualizing multiple petabytes of data from a variety of sources (e.g., mobile devices, aerial (remote) sensing, software logs, cameras, microphones, radio-frequency identification (RFID) readers, and wireless sensor networks). (*DoD Digital*…, p. 44)
>
> Big Data Analytics applications enable analysis of volumes of structured data, plus other forms of data (semi-structured and unstructured) that are often left untapped by conventional business intelligence (BI) and analytics programs. Data mining tools sift through data sets in search of patterns and relationships enabling aggregation, correlation, historical trending, and forensic analysis against the data. Machine learning algorithms can analyze large data sets and perform deep learning, a more advanced offshoot of machine learning to support predictive analytics that forecast behavior and other future developments. (*DoD Digital*…, p. 45)

[7] Department of Defense. *DoD Digital Modernization Strategy*. 5 June 2019. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

2.4 DoD Data Strategy

The *DoD Data Strategy*[8] states the following.

> Warfighters at all echelons require tested, secure, seamless access to data across networks, supporting infrastructure, and weapon systems out to the tactical edge. The advanced capabilities provided by DoD's Digital Modernization program depend upon enterprise data management policies, standards, and practices. Sensors and platforms across all domains must be designed, procured, and exercised with open data standards as a key requirement. Survival on the modern battlefield will depend upon leveraging and making connections among data from diverse sources, using analytic tools for superior situational awareness, and coordinating information for disaggregated-precision effects. (*DoD Data Strategy*, p. 1)

The capabilities that are particularly relevant to a viable T&E BDA-KM implementation are the following.

> This strategy emphasizes access to data and the capability to adjust requirements in stride with changes in technology and data sources. DoD architecture, enabled by enterprise cloud and other open-architecture capabilities, must allow pivoting on data more rapidly than adversaries are able to adapt. The ability to develop and deploy lightweight applications rapidly and continuously in support of user needs revolutionizes how DoD uses data and leads to a strategic advantage. An agile architectural approach enables incremental value to be delivered by balancing emergent design and intentional architecture. This agile approach allows the architecture of data and systems (even a large solution) to evolve over time, while simultaneously supporting the needs of current users. (*DoD Data Strategy*, p. 5)

> DoD employs a family of standards that include not only commonly recognized approaches for the management and utilization of data assets, but also proven and successful methods for representing and sharing data. Given the diversity of DoD systems, these standards should be applied at the earliest practical point in the data lifecycle and industry standards for an open data architecture should be used wherever practical. Standards are not an end unto themselves, but rather, they provide value when enabling data and information to be readily and securely utilized and exchanged. Additionally, physical encoding of the data interchange specifications will allow operations in congested and contested environments. Additionally, the DoD CDO will work with the Director, Operational Test and Evaluation (DOT&E), to ensure that data-related material capabilities are tested and evaluated so the effectiveness and suitability of the technologies are known. (*DoD Data Strategy*, p. 5)

The goal that is particularly relevant to a viable T&E BDA-KM implementation is the following.

[8] Department of Defense. *DoD Data Strategy*. 2020. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

> The goal of making data accessible enables authorized users to obtain the data they need when they need it, including having data automatically pushed to interested and authorized users. Data accessibility must comply with Public Law (P.L.) 115-435, the Foundations for Evidence-Based Policymaking Act of 2018. DoD is making data, including warfighting, intelligence, and business data, accessible to authorized users. Accessibility requires that protective mechanisms (e.g., security controls) are in place for credentialed users to ensure that access is permitted in accordance with laws, regulations, and policies.
>
> DoD will know it has made progress on making data accessible when:
>
> **Objective 1**: Data is accessible through documented standard Application Programming Interfaces (APIs).
> **Objective 2**: Common platforms and services create, retrieve, share, utilize, and manage data.
> **Objective 3**: Data access and sharing is controlled through reusable APIs. (*DoD Data Strategy*, p. 7)

2.5 DoD Enterprise DevSecOps Strategy Guide

The *DoD Enterprise DevSecOps Strategy Guide*[9] states the following.

> The DoD CIO and the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD A&S) recognize the urgent need to rethink our software development practices and culture by leveraging the commercial sector for new approaches and best practices. DevSecOps is such a best practice as it enables the delivery of resilient software capability at the speed of relevance, a central theme of software modernization across the DoD. DevSecOps is a proven approach widely adopted by commercial industry and successfully implemented across multiple DoD pathfinders. DevSecOps is a core tenant of software modernization, technology transformation, and advancing an organization's software development ecosystem to be more resilient, while ensuring cybersecurity and metrics/feedback are paramount. (*DoD Enterprise*…, p. 1)

The principles that are particularly relevant to a viable T&E BDA-KM implementation are the following.

> *Relentless pursuit of Agile principles and culture within a software factory construct* (Italics added). The Agile Manifesto[10] captures core competencies that define functional relationships that every DevSecOps team should value:
>
> - Individuals and Interactions over Processes and Tools
> - Working Software over Comprehensive Documentation
> - Customer Collaboration over Contract Negotiations

[9] Department of Defense. *DoD Enterprise DevSecOps Strategy Guide*. Version 2.1. September 2021. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.
[10] Beck, K. et al. *Manifesto for Agile Software Development*. Retrieved 14 July 2022. Available at https://agilemanifesto.org/.

- Responding to Change over Following a Plan

The first core competency emphasizes the value and importance that individuals work together, but it should not be interpreted that processes and tools are irrelevant. The same holds true for the other core competencies; documentation is still needed, but not at the cost of working software; Agile teams still create sprint plans, etc. (*DoD Enterprise*…, pp. 11-12)

Adoption of Cloud-smart and data-smart architectural motifs throughout (Italics added). There is an optimistic vision that portrays the Cloud as offering endless computing capacity, guaranteed availability, and lower operational costs. The reality is that an improperly architected application remains as brittle in a Cloud environment as it did operating in a Regional Data Center. If not re-architected, it may actually be more unreliable and more expensive to operate. The shift to Cloud must be accompanied by the adoption of new architectural design patterns and an overpowering preference to build atop existing enterprise services instead of reinventing duplicative capabilities. Further, data generation, transport, and consumption shows no signs of abating. Software architectures must consciously acknowledge this with smarter application programing interfaces (API) designs, caching strategies, and data tagging/labeling. (*DoD Enterprise*…, p. 13)

2.6 DoD Enterprise DevSecOps Playbook

The *DoD Enterprise DevSecOps Playbook*[11] states the following.

A modular open system approach (MOSA) is an acquisition and design strategy consisting of a technical architecture that adopts open standards and supports a modular, loosely coupled and highly cohesive system structure.1 U.S. Code Title 10 Section 2446a, and DoD Instruction 5000.02 require MOSA. A modern software architecture predicated upon microservices and software containers meet MOSA requirements.

A container is a lightweight, standalone, executable package of software that includes everything needed to run a business service except the OS; code, runtime, system tools, system libraries and settings. Containers run in isolated processes from one another, so several containers can run in the same host OS without conflicting with one another. All containers must be Open Container Initiative compliant.2 The DoD DevSecOps Strategy requires a CNCF Certified Kubernetes cluster for container orchestration.

A microservice architecture is an approach to application development where discrete, modular business services are bundled inside of a software container. These business services are then loosely coupled and rapidly composed using lightweight protocols. The primary functional benefit of this approach when executed properly is that each service can advance independently from the other services. Numerous non-functional benefits also exist, including more

[11] Department of Defense. *DevSecOps Playbook*. Version 2.1. September 2021. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

> agility in scaling to demand, multiple upgrade options that don't impact the user population, more precise cyber hardening at a per-service level, and inherent support for failure and recovery. (*DevSecOPS Playbook*, p. 6)

2.7 DoD Enterprise DevSecOps Reference Design

The *DoD Enterprise DevSecOps Reference Design: CNCF Kubernetes*[12] states the following.

> This DoD Enterprise DevSecOps Reference Design is specifically for Cloud Native Computing Foundation (CNCF)[13] Certified Kubernetes[14] implementations. This enables a Cloud agnostic, elastic instantiation of a DevSecOps software factory anywhere: Cloud, On Premise, Embedded System, Edge Computing. (*DoD…CNCF Kubernetes*, p. 1.)
>
> Kubernetes is a container orchestrator that manages the scheduling and execution of Open Container Initiative (OCI)[15] compliant containers across multiple nodes… OCI is an open governance structure for creating open industry standards around both container formats and runtimes. The container is the standard unit of deployment in this reference design. Containers enable software production automation in this reference design, and they also allow operations and security process orchestration. (*DoD…CNCF Kubernetes*, p. 6)

The benefits of adopting Kubernetes that are particularly relevant to a viable T&E BDA-KM implementation are the following.

> **Multimodal Environment**: Code runs equally well in a multitude of compute environments, benefitting from the K8s [Kubernetes] API abstraction.
>
> **Resiliency**: Self-healing of unstable or crashed containers.
>
> **Adaptability**: Containerized microservices create highly-composable ecosystems.
>
> **Scalability**: Application elasticity to appropriately scale and match service demand.
>
> The adoption of K8s [Kubernetes] and OCI compliant containers are concrete steps towards true microservice reuse, providing the Department with a compelling ability to pursue higher orders of code reuse across an array of programs. (*DoD…CNCF Kubernetes*, p. 7)

The assumptions that are particularly relevant to a viable T&E BDA-KM implementation are the following.

---

[12] Department of Defense. *DoD Enterprise DevSecOps Reference Design: CNCF Kubernetes*. Version 2.1. September 2021. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.
[13] Cloud Native Computing Foundation website, https://www.cncf.io/.
[14] Cloud Native Computing Foundation website, https://www.cncf.io/certification/software-conformance/.
[15] Open Container Initiative website, https://opencontainers.org/.

> No specific Kubernetes implementation is assumed, but the selected Kubernetes implementation must have submitted conformance testing results for review and certification by the CNCF.
>
> Vendor lock-in is avoided by mandating a Certified Kubernetes implementation; however, product lock-in into the Kubernetes API and its overall ecosystem is openly recognized.
>
> Adoption of hardened containers as a form of immutable infrastructure results in standardization of common infrastructure components that achieve consistent and predictable results. (*DoD…CNCF Kubernetes*, p. 4)

2.8 DoD Software Development and Open Source Software

The *Software Development and Open Source Software*[16] states the following.

> The Department must follow an "Adopt, Buy, Create" approach to software, preferentially adopting existing government or OSS [Open Source Software] solutions before buying proprietary offerings, and only creating new non-commercial software when no off-the-shelf solutions are adequate.
>
> OSS meets the definition of "commercial computer software" and therefore, shall be given equal consideration with proprietary commercial offerings, in accordance with Section 2377 of Title 10, U.S.C.
>
> In accordance with FAR 13.104, refusal to consider all OSS based solely on software being open source may be contrary to statutory and regulatory preferences for commercial products, and would unnecessarily restrict competition. OSS should be considered to the maximum extent practical. (*Software Development and*…, p. 3)

The advantages of open-source software that are particularly relevant to a viable T&E BDA-KM implementation are the following.

> The continuous and broad peer-review enabled by publicly available source code supports software reliability and security efforts through the identification and elimination of defects that might otherwise go unrecognized by a more limited core development team.
>
> The unrestricted ability to modify software source code enables the Department to respond more rapidly to changing situations, missions, and future threats.
>
> Reliance on a particular software developer or vendor ("vendor lock-in") due to proprietary restrictions may be reduced by the use of OSS, which can be operated and maintained by multiple vendors, thus making it easier to replace and upgrade components as technology and mission needs change. At some level, lock-in may be likely, based on product, architecture, or platform constraints, in spite of using OSS.

[16] Department of Defense. *Software Development and Open Source Software*. 24 January 2022. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

> Since OSS typically does not have a per-seat licensing cost, it can provide a cost advantage in situations where many copies of the software may be required and can mitigate risk of cost growth in licensing for situations where the total number of users may not be known in advance. (*Software Development and*…, pp. 3-4)

2.9 DoD Software Modernization Strategy

The *DoD Software Modernization Strategy*[17] states the following.

> The vision for software modernization is simple - deliver resilient software capability at the speed of relevance. Resilience implies software that is high-quality and secure, able to withstand and recover in the face of challenging conditions. Speed of relevance implies the accelerated delivery needed to maintain a competitive advantage. The approach is practical - unify efforts across DoD and partner with industry-leading software institutions to produce a portfolio of best-in-class software capabilities enabled by DoD processes. (*DoD Software Modernization*…, p. 2)

The goals that are particularly relevant to a viable T&E BDA-KM implementation are the following.

> *Accelerate the DoD Enterprise Cloud Environment* (Italics added). The DoD Enterprise Cloud Environment is the foundation for software modernization. The multi-cloud, multi-vendor approach still holds true. The requirement for cloud across all classification domains, from enterprise to tactical edge, is still valid. The need to transition from disparate cloud efforts to a structured, integrated, and cost-effective cloud portfolio remains the Department's intent. Working with commercial cloud service providers continues to be critical as the Department technically evolves. DoD and commercial cloud service providers must work together to quickly and securely deploy cloud services and ensure transparency of cybersecurity activities to maintain the protection of DoD data. (*DoD Software Modernization*…, p. 6)
>
> *Establish Department-wide Software Factory Ecosystem* (Italics added). As mentioned earlier, software increasingly defines military capabilities; therefore, DoD must scale its ability to produce secure and resilient software at speed to maintain a competitive advantage. This strategy recognizes that the modern approaches and tools, as well as the technical talent needed to do this, are not without cost. The Department must pursue an enterprise-wide approach, establishing a software factory ecosystem that takes advantage of investments already made by the Military Services (e.g., Air Force Platform One, Navy Overmatch Software Armory, Marine Corps Business Operations Support Services, and Army Coding Resources and Transformation Ecosystem) and scales their success to enable cross-Program/cross-Service use as espoused in the 2019 Defense Innovation Board Software Acquisition and Practices Report… DoD must establish requirements for a reasonable number of

[17] Department of Defense. *DoD Software Modernization Strategy*. Version 1.0. November 2021. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

> approved enterprise providers to efficiently scale software factories, minimize unnecessary platform duplication, and advance DevSecOps. (*DoD Software Modernization*…, p. 7)

3. Top 5 Key Characteristics for Software

This section identifies the top five high-level key characteristics of a viable T&E BDA-KM software implementation based upon the federal guidance identified in the previous section.

The rationale for only looking at the top five key characteristics is to provide a quick direction check for candidate T&E BDA-KM solutions rather than provide an in-depth analysis.

3.1 Utilize Modular Open Systems Approach

Utilize a MOSA that support the rapid evolution and independent replacement of specific services and components. More specifically, utilize a microservices software architecture approach of loosely coupled services that are independent of each other and interact with each other through open interface standards. This enables the ability to share data and analytics across multiple platforms.

3.2 Minimize Proprietary Components

Utilize free and open-source software (FOSS) components wherever possible, where FOSS is defined as public access to source code and unrestrictive licenses. If no other option is available, then use proprietary software. This maximizes flexibility and minimizes the maintenance and support burden.

3.3 Leverage Scalable Services

Design and implement service architectures that scale quickly. Use distributed compute technologies (e.g., Apache Hadoop[18], Apache Spark[19]) for "scale-out" performance (i.e., increasing performance by using additional CPUs rather than using faster CPUs) and big data optimized storage technologies (e.g., Apache Parquet[20], Apache Cassandra[21]) for "scale-out" storage capacity.

3.4 Prioritize Portable Solutions

Design and implement solutions that can be easily deployed on-premises, off-premises, and a mix of both. These solutions should be deployable on vendor-agnostic hardware and commodity hardware in order to maximize affordability. Prioritize the utilization of OCI compliant containers and CNCF certified Kubernetes to orchestrate and manage the containers.

[18] The Apache Software Foundation. Apache Hadoop website: https://hadoop.apache.org/. Retrieved 18 July 2022.
[19] The Apache Software Foundation. Apache Spark website: https://spark.apache.org/. Retrieved 18 July 2022.
[20] The Apache Software Foundation. Apache Parquet website: https://parquet.apache.org/. Retrieved 18 July 2022.
[21] The Apache Software Foundation. Apache Cassandra website: https://cassandra.apache.org. Retrieved 18 July 2022.

3.5 Employ Agile and DevSecOps Development

Design and implement solutions that adhere to Agile software development principles and methods in addition to DevSecOps software development principles and methods.

4. Implementation Scorecard for Software

This section provides an implementation scorecard based upon the key characteristics of a viable T&E BDA-KM software implementation.

TRMC Big Data Analytics Software Implementation Scorecard					
Top 5 Key Characteristics			**Rating**		
			Outstanding	**Satisfactory**	**Unsatisfactory**
1	**Utilize Modular Open Systems Approach**				
	1.1	Use services and components that are loosely-coupled (i.e., independent of each other)			
	1.2	Use open and standardized interfaces between services and components			
2	**Minimize Proprietary Components**				
	2.1	Prioritize usage of FOSS			
	2.2	Minimize dependencies on any one product or company			
3	**Leverage Scalable Services**				
	3.1	Use distributed compute technologies (e.g., Apache Hadoop, Apache Spark) for “scale-out” performance			
	3.2	Use big-data optimized storage technologies (e.g., Apache Parquet, NoSQL databases)			
4	**Prioritize Portable Solutions**				
	4.1	Deployable on-premises, off-premises, and a mix of both			
	4.2	Deployable on vendor agnostic commodity hardware			
	4.3	Use Open Container Initiative (OCI) compliant containers			
	4.4	Use Cloud Native Computing Foundation (CNCF) Kubernetes container management			
5	**Employ Agile and DevSecOps Development**				
	5.1	Adhere to Agile software development principles and methods			
	5.2	Adhere to DevSecOps software development principles and methods			
OVERALL RATING					

This page intentionally left blank.

APPENDIX A

Citations

Beck, K. et al. *Manifesto for Agile Software Development*. Retrieved 14 July 2022. Available at https://agilemanifesto.org/.

Cloud Native Computing Foundation website, https://www.cncf.io/.

———.https://www.cncf.io/certification/software-conformance/.

Department of Defense. *DevSecOps Playbook*. Version 2.1. September 2021. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *DoD Cloud Strategy*. December 2018. Retrieved 27 June 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *DoD Data Strategy*. 2020. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *DoD Digital Modernization Strategy*. 5 June 2019. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *DoD Enterprise DevSecOps Reference Design: CNCF Kubernetes*. Version 2.1. September 2021. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *DoD Enterprise DevSecOps Strategy Guide*. Version 2.1. September 2021. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *DoD Software Modernization Strategy*. Version 1.0. November 2021. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *Future of the Joint Enterprise Defense Infrastructure Cloud Contract*. 6 July 2021. Retrieved 27 June 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *Interim Guidance for Implementation of the Department of Defense Cloud Strategy*. 16 April 2020. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *Software Development and Open Source Software*. 24 January 2022. Retrieved 14 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

———. *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*. 12 February 2019. Retrieved 12 July 2022. Available at https://search.usa.gov/search?affiliate=defense_gov.

Open Container Initiative website, https://opencontainers.org/.

Range Commanders Council. *TRMC Big Data Analytics Architecture Assessment*. SR-21-003. September 2021. May be superseded by update. Retrieved 31 October 2022. Available at https://www.trmc.osd.mil/wiki/x/7wTkBw.

Test Resource Management Center. *Knowledge Management and Big Data Analytics Architecture Framework*. Version 13. 31 January 2019. Retrieved 27 June 2022. Available at https://www.trmc.osd.mil/wiki/display/JMETC/Big+Data+Knowledge+Management?preview=%2F55968739%2F55968745%2FBigDataArchitecture-v13-2019-01-31-DistA.pdf.

The Apache Software Foundation. Apache Cassandra website: https://cassandra.apache.org. Retrieved 18 July 2022.

———. Apache Hadoop website: https://hadoop.apache.org/. Retrieved 18 July 2022.

———. Apache Parquet website: https://parquet.apache.org/. Retrieved 18 July 2022.

———. Apache Spark website: https://spark.apache.org/. Retrieved 18 July 2022.

*** * * END OF DOCUMENT * * ***