DRAFT — UNOFFICIAL — NOT FOR OPERATIONAL USE

DATA LITERACY TECHNICAL REFERENCE

ADRP-TR

USAREUR-AF

CHAPTER 1

Data Literacy Technical Reference

HEADQUARTERS
UNITED STATES ARMY EUROPE AND AFRICA
(USAREUR-AF)
Wiesbaden, Germany

DRAFT — NOT FOR OFFICIAL USE. FOR TRAINING PLANNING PURPOSES ONLY.

24 MARCH 2026

DRAFT — UNOFFICIAL — NOT FOR OPERATIONAL USE

DATA LITERACY — TECHNICAL REFERENCE

HEADQUARTERS UNITED STATES ARMY EUROPE AND AFRICA Unit 29351 APO AE 09014 Wiesbaden, Germany

2026

PROPONENT: Headquarters, United States Army Europe and Africa (USAREUR-AF), C2DAO

SUMMARY OF CHANGE Data Literacy Technical Reference

This edition incorporates the following updates from the initial publication:

- **Governance update.** Adds the Army data stewardship 4-tier hierarchy (MADOs, Data Stewards, Functional Data Managers, C2DAOs) per Army CIO Data Stewardship Policy (April 2024) at paragraph 9-7.
- **Architecture update.** Adds reference to Unified Data Reference Architecture (UDRA) v1.1 (February 2025) and data mesh principles at paragraph 3-7c and Appendix C.
- **Policy update.** Updates Governing References to reflect that the Army Data Plan (2022) has been superseded in part by Army CIO guidance (2024) and UDRA v1.1 (2025).
- **VAULTIS-A update.** Updates data quality framework from VAUTI (5 dimensions, AR 25-1 2019) to VAULTIS-A (8 dimensions, DDOF Playbook v2.2, December 2025). Adds Linked, Secure, and Auditable dimensions. Establishes 85% minimum weighted average as DDOF Phase 3 quality gate. Full supersession chain documented at paragraph 4-8.
- **Decision dominance.** Adds paragraph 1-2e framing decision dominance as an enduring Army operational objective.
- **NATO references.** Adds AJP-3.2 (Allied Joint Doctrine for Land Operations) to interoperability section and references.
- **Architecture resources.** Expands reference to USAREUR-AF technical architecture and implementation resources available through C2DAO.
- **UDRA data quality dimensions.** Adds the seven UDRA v1.1 measurement dimensions (Accuracy, Completeness, Conformity, Consistency, Uniqueness, Integrity, Timeliness) at paragraph 4-9, linking them to the VAULTIS-A framework.
- **Data mesh architecture.** Adds Chapter 11 covering UDRA data mesh concepts: data products, data domains, computational governance, and the six UDRA services.

- **Army Data Plan strategic objectives.** Adds Chapter 12 with the eleven strategic objectives (SO-01 through SO-11) and five strategic enablers, highlighting SE-05 (Talent) as the mandate for MSS training.

This publication establishes doctrine for data literacy across United States Army Europe and Africa (USAREUR-AF) and the broader Army. It provides the foundational reference for all personnel who produce, manage, analyze, or consume data in support of military operations. It is platform-agnostic. Concepts apply regardless of system or tool.

The Army Data Plan (2022) established the foundational framework; current governance and architecture are codified in subsequent Army CIO guidance (2024) and the Unified Data Reference Architecture (UDRA) v1.1 (2025). This publication incorporates those authorities.

USAREUR-AF is the Army Service Component Command (ASCC) to United States European Command (USEUCOM) and United States Africa Command (USAFRICOM), responsible for theater land operations across the European and African AOR. USAREUR-AF supports NATO Article 5 collective defense commitments and is integrated into Joint All-Domain Command and Control (JADC2). Subordinate commands include III Corps, V Corps (Forward), 21st Theater Sustainment Command (21 TSC), 7th Army Training Command (7ATC), 10th AAMDC, 56th MDC-E, and SETAF-AF. Data literacy is foundational to USAREUR-AF's ability to operate effectively across this multi-nation, multi-domain theater.

PREFACE

This publication applies to all Army commands and organizations under USAREUR-AF. It addresses the full spectrum of data literacy — from the Soldier reading a readiness dashboard to the staff officer building an analytical pipeline to the commander interpreting data products before making a decision.

Data literacy is not an IT competency. It is a warfighting competency. Commanders and staff who cannot critically evaluate data will make worse decisions. Units that produce poor-quality data will degrade the situational awareness of every echelon above them. Data literacy is therefore a readiness issue.

The proponent of this publication is Headquarters, USAREUR-AF, C2DAO. Send comments and recommended changes to the proponent.

CHAPTER 1 — FOUNDATIONS OF DATA LITERACY

BLUF: Data literacy is the ability to read, understand, evaluate, and communicate data. It is a core competency for every Soldier and leader in a data-driven Army. Units that develop data literacy outperform those that do not — in speed of decision, quality of analysis, and reduction of operational error.

1-1. What Data Literacy Is

1-1a. Data literacy is the capacity to read, work with, analyze, and communicate using data. It is not the same as programming skill or data science. A data-literate Soldier does not need to write code. A data-literate Soldier needs to ask the right questions of data, evaluate whether data is trustworthy, understand what a chart shows and what it does not, and communicate findings clearly to decision-makers.

1-1b. Data literacy exists on a continuum. At the base level, it means understanding what a table or report represents. At advanced levels, it means constructing analytical pipelines, applying statistical methods, and building decision support tools. This publication addresses the full continuum.

1-1c. Data literacy is distinct from the following related competencies:

Competency	Definition	Relationship to Data Literacy
Information Technology (IT)	Managing systems, networks, and infrastructure	Enables data movement; not the same as analysis
Data Science	Statistical modeling, machine learning, advanced analytics	Upper end of the data literacy continuum
Intelligence Analysis	Interpretation of adversary intent and capability	Consumes data products; requires data literacy
Data Engineering	Building pipelines, ETL, databases	Technical implementation; supports data literacy

NOTE

A unit can have excellent IT support and zero data literacy. The systems are available but no one knows how to use the data they generate. This is a common failure mode.

1-2. The Cognitive Hierarchy — From Data to Understanding

1-2a. Army doctrine (ADP 6-0, Mission Command; ADP 3-13, Information) defines a four-level cognitive hierarchy that is foundational to all data work:

Level	Definition (ADP 6-0)	Operational Example
Data	Unprocessed signals communicated between nodes	Raw GCSS-A equipment status codes, MEDPROS records, SIGINT intercepts
Information	Data organized and processed to provide context and meaning	A dashboard showing "847 NMC vehicles by type across V Corps"
Knowledge	Information analyzed to provide meaning and value	Assessment that NMC rates spiked 40% after rotation due to Class IX shortfall
Understanding	Knowledge synthesized with judgment to comprehend the situation	G4's determination that readiness will degrade below C2 within 60 days without intervention

1-2b. Data platforms automate the lower tiers (data → information) so humans can focus on the higher tiers (knowledge → understanding). This hierarchy explains why data literacy is a warfighting competency, not an IT skill: the goal is not to manage data — it is to accelerate the commander's progression from raw data to sound understanding.

1-2c. Per ADP 3-13 (November 2023), **information is a dynamic of combat power** — at the same level as firepower, mobility, and survivability. ADP 3-13 defines five information activities: Enable, Protect, Inform, Influence, Attack. Data platforms primarily serve the Enable function (establishing infrastructure to collect, process, store, and disseminate data) and the Protect function (safeguarding data and networks through access controls and classification enforcement).

1-3. Why Data Literacy Matters for Military Operations

1-3a. In USAREUR-AF, allied sensors, data, and AI-powered machine learning tools are already integrated into real-time decision-making. Data literacy is what allows formations to participate in — and benefit from — that capability.

1-2b. Modern military operations generate enormous volumes of data across every warfighting function. Personnel systems track readiness by individual Soldier. Logistics systems track supply on hand, in transit, and requested. ISR (Intelligence, Surveillance, and Reconnaissance) platforms generate imagery, signals, and pattern of life data. Communications systems record network performance and anomalies. Each of these data streams has value. That value is realized only when personnel can interpret and act on the data.

1-2b. Conversely, bad data — or data that is not understood — creates operational risk. A readiness report based on stale data misleads the commander about actual combat power. A supply forecast built on incorrect inputs fails to position materiel where needed. An ISR product misread due to scale or projection error sends forces to the wrong location.

1-2c. Data literacy reduces these risks. It does not eliminate uncertainty — that is impossible in war — but it ensures that the uncertainty is understood rather than hidden inside a spreadsheet or dashboard.

1-2d. Speed matters. A formation that can rapidly assemble, analyze, and visualize relevant data and brief findings to the commander in decision cycle time has a competitive advantage over one that cannot. Data literacy is therefore a component of tempo. Units that are not data-literate operate slower and with less confidence than those that are.

1-2e. Decision dominance. The enduring Army operational objective is decision dominance: the ability to make better decisions, faster, than the adversary across the full competition continuum. Decision dominance is not achieved by having more data — it is achieved by having better-understood, better-managed, and more rapidly analyzed data than the adversary. In the USAREUR-AF AOR, where the threat employs sophisticated information operations and electronic warfare to degrade Allied situational awareness, decision dominance requires formations that can maintain data quality under pressure and extract decision-relevant insight faster than threat forces can deny or degrade the data. Data literacy is the human component of this competitive advantage.

WARNING

Speed of analysis is not the same as accuracy of analysis. Commanders must resist the temptation to accept fast-delivered data products without validating the underlying data quality. A rapid but wrong answer is worse than a slower correct one.

1-3. The Data Spectrum: Raw Data to Decision

1-3a. Not all data is immediately actionable. Data must be transformed into information, and information must be understood in context before it becomes knowledge that supports a decision. This progression is the data spectrum.

Stage	Definition	Military Example
Raw Data	Unprocessed facts, signals, observations	Vehicle fuel levels recorded by sensor
Information	Data processed into meaningful form	Fuel consumption rate by vehicle class
Knowledge	Information contextualized and interpreted	Forecast that Class III resupply is needed in 48 hours
Decision	Action taken based on knowledge	Commander authorizes logistics package to forward

1-3b. Personnel at every echelon operate somewhere on this spectrum. Soldiers and junior NCOs frequently deal in raw data — they record it, enter it, and transmit it. Staff officers and senior NCOs often work at the information and knowledge levels — they process, analyze, and brief. Commanders operate primarily at the knowledge and decision levels — they interpret finished analysis and decide.

1-3c. The chain is only as strong as its weakest link. If a Soldier enters erroneous raw data, the information derived from it is wrong, the knowledge built on that information is wrong, and the decision made on that knowledge is wrong. Data literacy obligations flow to every echelon.

> **NOTE**
>
> The data spectrum is not linear in all cases. Sometimes raw data directly enables a decision (e.g., a sensor alarm). Context determines where on the spectrum a given piece of data falls.

1-4. Why Data-Literate Forces Outperform

1-4a. Data-literate formations demonstrate measurable advantages in several areas:

Speed of Decision Cycle. A data-literate staff assembles the SITREP (Situation Report), LOGSTAT (Logistics Status Report), and key indicators faster. They know which systems hold which data, how to query it, and how to present it. This shortens the commander's decision cycle.

Reduced Operational Error. Data-literate personnel catch errors before they propagate. They notice when a report looks inconsistent with the prior report. They ask why the number changed, rather than accepting it uncritically.

Better After-Action Analysis. Formations that produce and retain quality operational data learn faster. After-action reviews (AARs) that draw on actual data — rather than memory — are more accurate and produce better lessons learned.

Improved Resource Allocation. Logistics, personnel, and maintenance decisions made on clean, current data are more precise. Readiness is maximized and waste is reduced.

Credibility with Higher Echelon. Staff products built on verifiable, well-sourced data carry more weight with higher headquarters than products based on anecdote or unvalidated reporting.

1-4b. The inverse is equally true. Data-illiterate formations are slower, more prone to surprise, and produce lower-quality staff products. They are also more susceptible to deception — an adversary who understands your data gaps can exploit them.

1-5. The Data-Driven Army: Doctrinal Context

1-5a. Army doctrine has long recognized the importance of information in military operations. The Military Decision-Making Process (MDMP), the Intelligence Cycle, and logistics doctrine all assume that commanders and staff will have access to timely, accurate information. Data literacy is the mechanism by which that assumption is fulfilled in the modern information environment.

1-5b. MDMP integration. The MDMP (defined in ADP 5-0) is the Army's primary tool for operational planning. Every step of the MDMP relies on data. Mission analysis requires data on enemy, terrain, weather, civil considerations, and friendly forces. Course of action (COA) development requires data on

logistics supportability and force availability. COA analysis requires data to wargame outcomes. Synchronization requires data to coordinate timing, sequencing, and resource allocation. Chapter 8 of this publication addresses MDMP data requirements in detail.

1-5c. The Intelligence Cycle. The Intelligence Cycle (Direction, Collection, Processing, Exploitation, Dissemination — DCPED) is fundamentally a data cycle. Data is collected, processed into intelligence products, and disseminated to decision-makers. Data literacy at every step of the cycle — from the collector who records what they observe to the analyst who synthesizes reporting into finished intelligence — determines the quality of the output.

1-5d. Logistics. Modern logistics is data-intensive. Authorized stockage levels, demand forecasting, distribution planning, and asset visibility all depend on data. Units that accurately report their Class I through Class IX status enable the supply chain to function. Units that do not — whether because they lack data literacy or because they are careless about reporting — impose costs on the entire formation.

1-5e. Personnel. Readiness reporting, personnel accounting, casualty tracking, and medical status are all data-dependent. The Personnel Status (PERSTAT) report, the Unit Status Report (USR), and casualty reporting all require personnel at the lowest levels to enter data accurately and on time.

1-5f. Nesting Under ADP 3-13, Information.

1-5g. This publication implements ADP 3-13, Information, by establishing the data literacy foundation required for information advantage in the USAREUR-AF AOR. ADP 3-13 establishes that information is combat power — that commanders who can generate, protect, and exploit information more effectively than the adversary gain a decisive advantage. Data literacy is the prerequisite for that capability. Formations that cannot accurately collect, manage, analyze, and communicate data cannot achieve the information advantage that ADP 3-13 describes. Every chapter of this publication supports the broader information operations framework of ADP 3-13 by building the competencies that make information advantage possible.

NOTE

ADP 3-13 addresses information at the conceptual level. This publication addresses the foundational data skills — collection, quality, analysis, security, and integration into planning — that operationalize ADP 3-13 in USAREUR-AF formations.

1-5g. NATO and EUCOM AOR Context.

1-5h. USAREUR-AF operates in the USEUCOM area of responsibility (AOR), which encompasses NATO Allied nations from the Atlantic to the Black Sea. Data operations in this theater must account for coalition interoperability. Data systems, data products, and data sharing practices must align with NATO Architecture Framework version 4 (NAFv4) and support interoperability with Allied Joint Publications (AJP-01, AJP-3, AJP-5). Data literacy requirements in this publication apply to multinational operations and exercises as well as unilateral Army operations.

1-5i. Data interoperability is a central concern for coalition effectiveness. The barrier is trust: Allied nations must have confidence in where their data goes, how it is handled, and what security controls govern it. Data literacy — including rigorous handling discipline, accurate classification, and documented data lineage — is how units build that trust at the working level. Formations that cannot demonstrate disciplined data practices cannot be trusted with coalition data and cannot participate fully in NATO integrated operations.

1-5j. The real-time command-and-control system USAREUR-AF has built in Europe — integrating allied sensors, data, and AI-powered tools across nations — represents a level of transformation the CG has compared to the introduction of radio-enabled command in World War II. That system must now scale globally. USAREUR-AF is the proving ground. Formations in this theater that master data-driven operations are building the template the Army will replicate across every AOR.

1-5h. Responsibilities at Each Echelon.

Echelon	Primary Data Responsibility
Soldier / Junior Enlisted	Accurate data entry and reporting; following data handling procedures
NCO (Sergeant through Sergeant Major)	Verifying subordinate data quality; enforcing standards; briefing data products to leaders
Company Grade Officer (2LT through CPT)	Integrating data from subordinate elements; producing staff products; evaluating data quality
Field Grade Officer (MAJ through COL)	Setting data requirements; evaluating analytical products; making data-informed decisions
General Officer / Senior Leader	Establishing data culture; demanding data quality; resourcing data capabilities

CHAPTER 2 — DATA CONCEPTS AND STRUCTURES

BLUF: Data exists in multiple types, formats, and storage architectures. Understanding the differences determines whether you can use the data, how you can analyze it, and where its limitations lie. This chapter provides the foundational vocabulary.

2-1. Data Types

2-1a. Data is not monolithic. The type of data determines how it can be stored, queried, and analyzed. Three primary data types are relevant to Army operations.

Structured Data. Structured data is organized in a defined schema — rows and columns with consistent field types. Every record conforms to the same format. Most administrative and logistics data is structured. Examples: personnel roster entries, maintenance work orders, supply requisition records, LOGSTAT entries.

Semi-Structured Data. Semi-structured data has some organizational elements but does not conform to a rigid schema. It often uses tags, keys, or nested structures to organize content. Examples: JSON-formatted sensor telemetry, XML-based message traffic, email with structured headers and unstructured body text.

Unstructured Data. Unstructured data has no predefined format or schema. Most human- generated content is unstructured. Examples: after-action review narratives, imagery, intercepted communications transcripts, debriefing notes, voice recordings.

2-1b. Military operational data uses all three types. A complete operational picture integrates structured logistics data, semi-structured sensor feeds, and unstructured reporting from human sources. Data literacy requires the ability to recognize which type you are working with and apply appropriate analytical methods.

Data Type	Army Example	Key Characteristic
Structured	Personnel database, maintenance tracker	Fixed schema; easy to query
Semi-Structured	Sensor telemetry in JSON, USMTF messages	Flexible schema; requires parsing
Unstructured	ISR imagery, AAR text, HUMINT reports	No schema; requires specialized processing

NOTE

Converting unstructured data into structured or semi-structured form (a process called extraction or parsing) is a data engineering task. Consuming the output of that process is a data literacy task. Know which one you are being asked to do.

2-2. Data Formats

2-2a. Within each data type, data is stored and transmitted in a variety of formats. Format determines which tools can read the data and how it must be processed.

Tabular. Tabular data organizes information in rows (records) and columns (fields). This is the most common format for operational reporting. Comma-Separated Values (CSV) files and database tables are tabular. Most readiness reports, personnel records, and logistics reports are tabular.

Time-Series. Time-series data records measurements at successive points in time. Vehicle sensor readings, network traffic metrics, and daily readiness percentages are time-series. Time-series data requires analysis methods that account for temporal ordering — the sequence of the observations matters.

Geospatial. Geospatial data encodes location information — coordinates, boundaries, routes, and overlays. Geospatial data is addressed in detail in Chapter 6.

Text/Document. Text data is human-readable prose or semi-structured documents. SITREP narratives, orders, and intelligence products are text data. Searching and analyzing text requires different methods than tabular analysis.

Imagery. Imagery data represents visual information — satellite imagery, aerial reconnaissance, ground-level photography. Imagery requires specialized exploitation tools and trained analysts. Data literacy for imagery focuses on understanding resolution, currency, coverage, and the difference between what is visible and what is confirmed.

2-2b. A single operational product may use multiple formats. A force protection report might include a tabular incident log, geospatial incident mapping, and a text narrative. Data literacy requires the ability to work across formats.

2-3. Databases, Data Lakes, and Data Warehouses

2-3a. Data is stored in different architectures depending on its purpose, volume, and usage pattern. Understanding these architectures prevents common errors — particularly using data from a system optimized for one purpose to answer questions it was not designed to support.

Database (Relational). A relational database stores structured data in tables with defined relationships between them. It is optimized for reading and writing individual records quickly and reliably — what data professionals call Online Transaction Processing (OLTP). Personnel and logistics systems use relational databases. Data is current, well-structured, and normalized (stored without redundancy). Best for: lookups, individual record retrieval, current status.

Data Warehouse. A data warehouse stores historical, integrated data from multiple source systems. It is optimized for analytical queries across large volumes — Online Analytical Processing (OLAP). A data warehouse might aggregate readiness data from multiple systems and time periods to support trend analysis. Best for: trend analysis, cross-system reporting, historical lookups.

Data Lake. A data lake stores large volumes of raw data in its native format — structured, semi-structured, and unstructured — until it is needed for analysis. Data lakes are flexible but require strong governance to prevent them from becoming unusable repositories of mixed- quality data (sometimes called a "data swamp"). Best for: storing diverse data types, enabling future analysis, feeding machine learning pipelines.

Architecture	Optimized For	Data State	Risk
Relational Database	Current record transactions	Structured, normalized	Stale if not refreshed; limited history
Data Warehouse	Historical analytical queries	Structured, integrated	Latency; may lag source systems
Data Lake	Flexible storage of all types	Raw, any format	Governance failure; unusable swamps

CAUTION

Do not assume that data from a source system is the same as data from a warehouse built from that system. Warehouses introduce latency and transformation logic. Always know the recency and transformation history of analytical data before presenting it as current.

2-4. Schemas, Fields, Records, and Relationships

2-4a. These terms describe the structure of data. They are not technical jargon — they describe how data is organized so that humans and machines can find and interpret it.

Schema. A schema is the blueprint for how data is organized. It defines what fields exist, what type of data each field holds (text, number, date, etc.), and what rules apply (e.g., this field cannot be blank). Every database table has a schema. Knowing the schema tells you what information is available and what is not.

Field. A field is a single unit of data — one column in a table. A personnel record might have fields for Last Name, First Name, Rank, Unit, MOS (Military Occupational Specialty), and Date of Last Physical.

Record. A record is a single complete entry in a dataset — one row in a table. A maintenance record for one vehicle at one point in time is one record.

Relationship. A relationship links two tables through a shared field (called a key). A maintenance table and a vehicle registry table share a vehicle identification number. Joining them through that key creates a complete picture — what vehicle it is and what happened to it. Understanding relationships is essential for combining data from multiple sources, which is a common analytical task.

NOTE

When data from two sources is joined, the quality of the resulting dataset depends on the quality of both inputs. A high-quality maintenance record joined to an inaccurate vehicle registry produces a combined record that is unreliable. Quality problems multiply when data is combined.

2-5. Metadata

2-5a. Metadata is data about data. It describes the content, context, quality, and provenance of a dataset without being the dataset itself. Metadata answers the questions: Where did this data come from? When was it collected? Who collected it? What does each field mean? What are its known limitations?

2-5b. Metadata matters operationally because data without context cannot be trusted. A table of incident locations is useful if the metadata tells you the collection period, reporting unit, and coordinate system. The same table without that metadata could be misinterpreted, misplotted, or used to draw invalid conclusions.

2-5c. Common metadata elements:

Metadata Element	What It Tells You
Source	Which system or organization produced the data
Collection date/time	When the data was captured
Last updated	When the record was most recently modified
Collection method	How the data was gathered (sensor, manual entry, automated feed)
Data owner	Who is responsible for the data's accuracy
Classification	The handling requirements for the data
Field definitions	What each column means; units of measure
Coverage	What geographic area, time period, or population the data represents

2-5d. Data products — charts, maps, reports — must include relevant metadata. A decision briefed to a commander must identify the data source and its recency. A map product must identify its coordinate system and the currency of the underlying data.

WARNING

A data product without metadata provenance cannot be defended. If the commander asks where the number came from and no one knows, the product has zero credibility. Always document the source.

2-6. Data Volume, Velocity, and Variety — The 3 Vs

2-6a. The 3 Vs describe the challenge of managing modern data at scale.

Volume. Volume is the quantity of data generated. Modern military operations generate data at a scale that exceeds human ability to review manually. ISR platforms generate petabytes of imagery. Logistics systems record millions of transactions. Network sensors generate continuous telemetry. Volume demands automated processing — no individual can manually review all data generated by a modern formation.

Velocity. Velocity is the speed at which new data arrives. Some data arrives in near-real- time (sensor feeds, network traffic); some arrives in batch (daily reports, weekly USR submissions). Analytical systems must be designed to handle the velocity at which data arrives. A system designed for daily batch processing will fail if required to handle real-time sensor streams.

Variety. Variety is the diversity of data types and formats. A theater-level operations center integrates tabular logistics data, geospatial overlays, imagery, text reports, voice communications metadata, and network telemetry simultaneously. Each data type requires different storage, processing, and analytical approaches.

2-6b. For the individual analyst, the 3 Vs frame the question: "Can I work with this data manually, or do I need automated tools?" When volume, velocity, or variety exceeds manual capacity, automation becomes necessary — not a luxury.

2-7. Master Data, Transactional Data, and Analytical Data

2-7a. Data within an organization serves different functions. Understanding the function determines how the data should be managed and used.

Master Data. Master data is the core reference data that all other data refers to. Personnel records (who is in the unit), equipment registries (what equipment the unit has), and organizational hierarchy data (what unit belongs to what formation) are master data. Master data changes infrequently and must be highly accurate. Errors in master data propagate throughout every system that references it.

Transactional Data. Transactional data records events and activities — what happened, when, and to whom. Maintenance work orders, supply requisitions, personnel actions, and incident reports are transactional. Transactional data is created constantly and reflects the operational tempo of the formation.

Analytical Data. Analytical data is derived from master and transactional data to support decision-making. Readiness percentages, trend lines, forecast models, and summary statistics are analytical data. Analytical data is only as good as the master and transactional data from which it is derived.

Data Category	Example	Management Priority
Master Data	Vehicle registry, personnel roster	High accuracy; controlled updates; single authoritative source
Transactional Data	Work orders, supply requests, incident logs	Complete and timely recording; chronological integrity
Analytical Data	Readiness percentages, trend analysis	Correct derivation logic; traceable to source

CHAPTER 3 — THE DATA LIFECYCLE

BLUF: Every dataset follows a lifecycle from creation to disposal. Understanding where data is in its lifecycle determines how it should be handled, who is responsible for it, and how much it can be trusted. Gaps in the lifecycle create data quality failures and operational risk.

3-1. Ingestion — How Data Enters a System

3-1a. Ingestion is the process by which data enters an analytical or storage system. Data can be ingested in several ways:

Manual Entry. A Soldier or analyst types data into a form, spreadsheet, or database. Manual entry is the most common ingestion method for operational reporting and the most common source of data quality errors. Human error, inconsistent formatting, and intentional rounding or estimation all degrade manually entered data.

Automated Feed. A system-to-system connection transfers data without human intervention. Sensor platforms, network monitoring tools, and integrated logistics systems use automated feeds. Automated feeds are faster and less prone to human error, but can fail silently — data stops flowing and no one notices until the absence of data is discovered.

Bulk Upload. A file (CSV, spreadsheet, XML, etc.) is loaded into a system at once. Bulk uploads are common when migrating data between systems or loading historical records.

API (Application Programming Interface) Integration. Systems exchange data through defined programmatic interfaces. APIs enable near-real-time integration between systems without file transfers.

3-1b. Each ingestion method has failure modes. Know the method for each data source you use and validate that the data arrived correctly.

> **CAUTION**
>
> Silent feed failures are operationally dangerous. If an automated data feed stops, the system may display old data without indicating that it is stale. Establish data freshness checks — timestamp validation — for all automated feeds used in operational decision-making.

3-2. Storage and Organization

3-2a. Once ingested, data must be stored in a way that enables retrieval, analysis, and security controls. Storage decisions affect every downstream use of the data.

3-2b. Structured data is stored in relational database tables, flat files, or data warehouse schemas. Semi-structured data is stored in document databases or object stores. Unstructured data (imagery, documents, audio) is stored in file systems or object stores.

3-2c. Organization is as important as storage. Data that is stored but cannot be found is functionally unavailable. Good data organization includes:

- Consistent naming conventions for files, tables, and fields
- Logical folder or schema structures that reflect operational function
- Indexing of commonly searched fields to enable fast retrieval
- Version control or audit trails for datasets that are updated over time

3-2d. Physical and logical security controls must be applied at storage. Classification- appropriate storage systems must be used for classified data. Access controls must be configured to enforce need-to-know. Chapter 7 addresses data security in detail.

3-3. Transformation and Enrichment

3-3a. Raw data rarely arrives in a form ready for analysis. Transformation processes convert raw data into usable analytical data. Common transformations include:

Cleaning. Correcting errors, standardizing formats, and removing duplicates. A personnel roster with rank abbreviated inconsistently ("SGT," "Sgt," "Sergeant") requires cleaning before it can be reliably searched or filtered.

Normalization. Converting data to a standard format or unit of measure. Converting all distances to kilometers, all weights to kilograms, or all times to UTC (Coordinated Universal Time) enables valid comparison across records.

Joining. Combining data from multiple tables or sources using shared keys. Joining a maintenance record to a vehicle registry enriches the maintenance record with vehicle attributes.

Aggregation. Summarizing individual records into summary statistics. Rolling individual Soldier readiness records into a company readiness percentage is aggregation.

Enrichment. Adding context from an additional data source. Adding weather data to an incident dataset enriches the incidents with environmental context.

3-3b. Every transformation changes the data. Document what transformations were applied and in what sequence. This documentation is part of data lineage (see paragraph 3-7).

WARNING

Transformation errors compound. A cleaning step that incorrectly standardizes a value produces incorrect results in every downstream join, aggregation, and analysis. Validate transformation outputs, not just raw inputs.

3-4. Analysis and Reporting

3-4a. Analysis is the process of applying methods to transformed data to extract meaning. Analysis ranges from simple descriptive summaries (what happened, how many, when) to predictive modeling (what is likely to happen next). Chapter 5 addresses analysis methods in detail.

3-4b. Reporting is the presentation of analytical outputs to consumers. A report translates analysis into a product that decision-makers can use. Effective reports lead with the conclusion (BLUF), support it with data, and are calibrated to the audience's level of technical sophistication.

3-4c. Analysis and reporting are not the same as data entry or data management. Personnel assigned to reporting functions must understand both the analytical methods used and the operational context of what they are reporting.

3-5. Distribution and Consumption

3-5a. After analysis, data products must reach their consumers. Distribution includes all mechanisms by which data products are delivered: dashboards, automated reports, briefings, message traffic, and file transfers.

3-5b. Distribution must enforce classification and access controls. A data product cannot be distributed on an unclassified system if the underlying data is classified. Distribution pathways must be established before data products are produced — not after.

3-5c. Consumers of data products have responsibilities. They must:

- Understand the classification and handling requirements of products they receive
- Not further distribute products beyond their authorized recipients
- Report data quality concerns back to the producer
- Not modify data products and re-release them without explicit authorization

NOTE

A data product is not the same as data. A finished analytical product — a chart, map, or summary table — may have different classification and distribution rules than the raw data it was built from. Both the product and the underlying data must be handled according to their respective classification markings.

3-6. Archiving and Retention

3-6a. Data has a defined lifespan. Retention requirements vary by data type, operational context, and regulatory requirement. Retaining data longer than required creates storage burden and security risk. Disposing of data before required retention periods expire creates legal and operational risk.

3-6b. Operational data generated during exercises and operations is typically subject to Records Management requirements under Army Regulation (AR) 25-400-2, The Army Records Information Management System (ARIMS). Data stewards must know the retention schedule for data they manage.

3-6c. Archived data must remain readable. Data stored in obsolete formats may become inaccessible when the systems that read those formats are decommissioned. Long-term archives should use open, well-documented formats.

3-6d. Disposition — the final deletion or transfer of data at the end of its retention period — must be documented. Records of what was disposed of, when, and how must be maintained.

3-7. USAREUR-AF 5-Layer Data Stack Architecture

3-7a. USAREUR-AF manages operational data across a structured five-layer architecture. This model defines how data systems are organized from physical infrastructure through to operational decisions. All personnel working with USAREUR-AF data systems should understand which layer they are operating in and what responsibilities that layer carries.

Layer	Name	Description	Military Operational Example
Layer 1	Infrastructure	Physical and virtual storage, compute resources, and connectivity that all data systems rely on	NIPRNET/SIPRNET transport, cloud compute nodes in theater, forward-deployed storage
Layer 2	Integration	Data ingestion, pipeline processing, and ETL (Extract-Transform-Load) that moves data between systems	Automated feed from GCSS-Army logistics system into the theater analytical environment; bulk upload of daily readiness reports
Layer 3	Semantic	Ontology, data meaning, governance, and cross-system definitions — the layer where data acquires consistent meaning across systems and organizations	Ontology object types for unit readiness, equipment status, and route conditions; NATO-aligned data dictionaries ensuring a "vehicle" means the same thing in every subordinate system
Layer 4	Analytics	Analysis, reporting, and dashboards that transform integrated, semantically consistent data into decision-relevant products	Theater readiness dashboard; ISR collection gap analysis; sustainment forecast for V Corps (Forward)
Layer 5	Activation	Applications, automated workflows, and command decisions that act on analytical outputs	Automated resupply triggers based on Class III consumption thresholds; commander decision to request OPCON forces based on readiness threshold crossing

3-7b. Each layer depends on the integrity of the layers below it. A dashboard at Layer 4 that draws on poorly integrated data at Layer 2 will produce inaccurate products regardless of the quality of the visualization. Commanders and staff who identify data quality problems must diagnose which layer is the source — the fix at a Layer 4 product (the dashboard) may actually require a correction at Layer 2 (the pipeline) or Layer 1 (the source system feed).

3-7c. The 5-Layer Data Stack is the USAREUR-AF standard architecture for operational data systems. It is implemented consistent with the Army's Unified Data Reference Architecture (UDRA) v1.1 (February 2025), which provides the Army's current reference architecture based on data mesh principles: distributed data ownership, domain-aligned data products, and federated governance. In practice, this means data ownership and management responsibility rests with the functional domain (S1, S2, S4, etc.) closest to the data, rather than centralizing all data management in a single IT element. For detailed technical implementation guidance, see the USAREUR-AF Cross-Domain Architecture (CDA) Portal (listed in the References section of this publication).

NOTE

The UDRA v1.1 and supporting enterprise architecture documentation are available through C2DAO. Contact your unit data steward or C2DAO for implementation guidance, ontology design references, and NATO doctrine crosswalks.

3-8. Data Provenance and Lineage

3-8a. Data provenance is the origin of a dataset — where the data came from, how it was collected, and who created it. Data lineage is the record of how data has moved and changed from its origin to its current state. Together, provenance and lineage establish the trustworthiness of a data product.

3-8b. A data product without documented provenance and lineage cannot be independently verified. If the analysis is questioned — and in operational decision-making, analysis will be questioned — there is no way to trace the numbers back to their source to confirm or correct them.

3-8c. Every data product presented to a commander must be able to answer: Where did this come from? When was it collected? What was done to it between collection and now? Who is responsible for its accuracy?

3-8d. Lineage documentation does not require a sophisticated tool. A simple notation in a report or briefing slide ("Source: GCSS-Army, pulled 01MAR2026 at 0600Z; filtered to units in EUCOM AOR; aggregated by brigade") satisfies the minimum standard.

CHAPTER 4 — DATA QUALITY

BLUF: Data quality is not a technical concern — it is an operational concern. Poor-quality data produces poor decisions. Every person in the data chain is responsible for data quality at their level. Catching quality problems early is always cheaper than correcting them after a decision has been made.

4-1. Dimensions of Data Quality

4-1a. Data quality is not a single attribute. It is a composite of multiple dimensions, each of which must be assessed independently. A dataset can be accurate but incomplete. It can be complete but untimely. It can be valid but inconsistent with other sources.

Dimension	Definition	Example Failure
Accuracy	Data correctly represents reality	Vehicle reported as FMC (Fully Mission Capable) when it is actually NMC (Not Mission Capable)
Completeness	All required data is present	Maintenance records missing operator information
Consistency	Data agrees across systems and time periods	USR shows 85% readiness; LOGSTAT shows data inconsistent with that figure
Timeliness	Data is current enough for its intended use	Supply levels reported 72 hours ago used as current in a fast-moving operation
Validity	Data conforms to defined rules and formats	Date field contains text string; grid coordinate in wrong format
Uniqueness	Each real-world entity appears only once	Same vehicle entered twice with slightly different serial numbers

4-1b. Assessing data quality requires evaluating all six dimensions. A high score on five dimensions and a failure on one can still render data unusable — a completely filled, accurate, consistent, and valid dataset that is 30 days old is useless for real-time readiness reporting.

4-2. Common Data Quality Failures and Their Operational Impact

4-2a. Data quality failures have predictable causes and predictable operational consequences. The following are the most common failures encountered in Army operational data.

Manual Entry Errors. Transcription errors, wrong units of measure, transposed digits, and incorrect date formats are common in manually entered data. In logistics, a manual entry error in a fuel quantity can distort the consumption rate calculation, triggering an unnecessary or inadequate resupply request.

Stale Data. Data that was accurate when collected but is no longer current. In a dynamic operational environment, personnel status, equipment readiness, and supply levels change continuously. A readiness report that is 48 hours old in a high-tempo operation may be operationally meaningless.

Inconsistent Reporting Standards. When different units apply different interpretations to the same reporting requirement, aggregate data becomes inconsistent. If one battalion counts vehicles in maintenance as Non-Mission Capable (NMC) and another counts them as Partially Mission Capable (PMC), the brigade-level readiness calculation is distorted.

Duplicate Records. The same entity (person, vehicle, equipment item) entered multiple times creates overcounts. Duplicate records in a personnel database inflate headcount. Duplicates in an equipment database misrepresent available assets.

Missing Values. Fields left blank where data is required create gaps in analysis. If incident reports do not consistently include location data, geospatial analysis of incident patterns becomes unreliable.

Data Type Violations. Data entered in the wrong format — text in a numeric field, a date in the wrong format, a coordinate in the wrong system — can cause processing errors or silent calculation failures.

4-3. Data Validation vs. Data Verification

4-3a. These terms are related but distinct. Confusing them leads to incomplete quality assurance.

Data Validation confirms that data conforms to defined rules and formats. Validation asks: Is this value in the correct format? Is it within an acceptable range? Is the required field populated? Validation can be automated — a system can check that every date is a valid date, that every grid coordinate is a valid MGRS (Military Grid Reference System) string, that every readiness percentage is between 0 and 100.

Data Verification confirms that data accurately represents reality. Verification asks: Is this value true? Does the record match what actually exists? Verification requires comparing data to an authoritative source — physically checking the equipment, cross- referencing with another system, or confirming with the reporting unit. Verification cannot be fully automated.

4-3b. Both are required. A dataset can be valid (correctly formatted) but wrong. A vehicle serial number can be a valid format but belong to a different vehicle. Validation catches format errors; verification catches substantive errors.

4-4. Profiling Your Data — What to Look for Before You Trust It

4-4a. Data profiling is the process of examining a dataset to understand its content, structure, and quality before using it for analysis. Every analyst should profile any dataset before presenting findings based on it.

4-4b. Basic profiling steps:

1. **Count records.** How many records are in the dataset? Does that number make sense given the expected population?
2. **Check date ranges.** What is the earliest and latest date in the dataset? Does the time range match what was expected?
3. **Identify missing values.** For each field, what percentage of records have a value? High missing rates indicate a collection problem.
4. **Check for duplicates.** Are there records that appear multiple times? Deduplication may be required.
5. **Review distributions.** For numeric fields, what is the minimum, maximum, mean, and median? Are there values that appear to be errors (e.g., a weight of 0 or a date of 1900)?
6. **Check categorical consistency.** For fields with defined value lists (rank, unit, equipment type), are all values in the expected list? Unexpected values indicate coding errors or inconsistent standards.
7. **Cross-check totals.** If the dataset represents a subset of a known universe, do the totals add up to the expected total?

NOTE

Data profiling is not a one-time activity. Datasets that are refreshed regularly must be profiled each time they are refreshed. A dataset that was high quality last week may have developed quality issues in the latest update.

4-5. The Cost of Bad Data in Operational Decision-Making

4-5a. The cost of bad data manifests in several forms:

Operational Error. A decision based on bad data produces a wrong outcome. Forces dispatched to the wrong location based on an erroneous grid. Resupply that does not arrive because the request was based on incorrect consumption data. Medical evacuation delayed because casualty data was entered incorrectly.

Lost Credibility. A commander or staff officer who presents bad data to higher headquarters loses credibility. Data products that are later found to be based on poor-quality data undermine trust in the producing element. This credibility loss persists beyond the individual incident.

Rework. Analysis built on bad data must be redone when the errors are discovered. Rework consumes time and resources and delays decisions. The cost of rework almost always exceeds the cost of quality control at the point of data entry.

Compounding Error. Bad data shared across systems or used as input to further analysis compounds. An error in a source system propagates to every product built from that system. The further downstream the error is discovered, the more costly the correction.

4-5b. Quantifying the cost of bad data is difficult in a military context, but the operational consequence is clear: bad data degrades the commander's ability to make informed decisions, which degrades the unit's ability to execute the mission (MSN).

4-6. Data Quality Responsibilities at Each Echelon

4-6a. Data quality is a shared responsibility. No single echelon can ensure quality alone.

Echelon	Quality Responsibility
Individual Soldier	Enter data accurately and on time; follow reporting standards; report anomalies
NCO	Verify subordinate entries; enforce reporting standards; conduct spot-checks
Staff Section	Validate data quality before forwarding or using for analysis; document sources
Data Steward	Monitor data quality metrics; identify systemic problems; coordinate corrections
Commander	Establish quality standards; resource quality controls; hold units accountable

4-6b. Data quality failures at the Soldier level are usually individual errors. Data quality failures at the NCO and staff levels are usually systemic — they indicate that standards are not being enforced or that the reporting process is poorly designed. Commanders must diagnose which type of failure they are dealing with before taking corrective action.

4-7. Automated Quality Checks — What They Can and Cannot Catch

4-7a. Automated quality checks run against data without human intervention to flag potential quality problems. They are fast, consistent, and scalable. They are also limited.

What automation can catch: - Format errors (date fields with invalid dates, coordinates in wrong format) - Range violations (values outside defined min/max bounds) - Null checks (required fields that are blank) - Duplicate record detection - Referential integrity violations (a record references a key that does not exist in the reference table) - Statistical outliers (values more than N standard deviations from the mean)

What automation cannot catch: - Plausible-but-wrong values (a vehicle FMC/NMC status recorded as the wrong code but in valid format) - Systematic bias in reporting (an entire unit consistently underreporting NMC status) - Missing records (automation cannot detect data that was never entered) - Substantive verification (whether the data reflects reality)

4-7b. Automated checks are a first line of defense, not a complete solution. They must be supplemented by human review and periodic verification against authoritative sources.

4-8. The VAULTIS-A Standard

The USAREUR-AF operational quality gate for all data products is **VAULTIS-A** — eight dimensions of data quality established by the DDOF Playbook v2.2 (T2COM C2DAO, December 2025). VAULTIS-A extends the DoD Data Strategy's VAULTIS framework (7 dimensions) by adding Auditable as the eighth dimension.

Lineage: VAUTI (AR 25-1, 2019, 5 dimensions) → VAULTIS (DoD Data Strategy, 2020, 7 dimensions) → **VAULTIS-A** (DDOF Playbook v2.2, 2025, 8 dimensions). VAULTIS-A is the current authoritative standard.

VAULTIS-A Dimension	Target	Definition	How to Evaluate
V — Visible	100%	Data is discoverable in the catalog or data product	Can you find this dataset without already knowing where it is?
A — Accessible	99%	Authorized users can retrieve it with acceptable latency	Can all users who need it actually access it within SLA?
U — Understandable	100%	Complete metadata and user guide; meaning is clear	Does it have a data dictionary, column descriptions, and context?
L — Linked	100%	Traced to authoritative sources and consuming products	Can you trace this data from source to consumption?
T — Trusted	95%	Accuracy validated, sponsor sign-off obtained	Is there a documented source, refresh schedule, quality check, and sponsor?
I — Interoperable	90%	Compatible with approved platforms; uses standard identifiers	Does it use standard identifiers shared with other data sources?
S — Secure	100%	Compliant with classification and access control policy	Are markings correct and access controls enforced?
A — Auditable	100%	Full provenance trail and access logs maintained	Can you trace every transformation and who accessed this data?

Minimum Gate Score: 85% weighted average across all eight dimensions to pass DDOF Phase 3. Products below threshold are returned for remediation with documented deficiencies.

> **NOTE**
>
> A data product that fails any VAULTIS-A dimension is not operationally ready. Commanders should require VAULTIS-A compliance certification before incorporating any new data source into decision-making.

WARNING

Earlier references to "VAUTI" (5 dimensions) reflect the superseded AR 25-1 (2019) standard. VAULTIS-A adds three critical dimensions — Linked, Secure, and Auditable — that address data traceability, security compliance, and provenance. Do not evaluate data products against the old 5-dimension standard.

4-9. UDRA Data Quality Dimensions

4-9a. The Unified Data Reference Architecture (UDRA) v1.1 defines seven measurement dimensions for data quality. These dimensions provide the operational criteria for assessing whether data meets the standards described in the VAULTIS-A framework (paragraph 4-8). VAULTIS-A defines what makes data fit for use; the UDRA dimensions define how to measure it.

4-9b. The seven UDRA data quality dimensions are:

Dimension	Definition	Measurement
Accuracy	Data correctly reflects true values	Comparison to authoritative source
Completeness	Data contains expected information at specified time	Null/missing field analysis
Conformity	Data follows agreed policies, standards, and procedures	Schema validation
Consistency	Values uniformly represented within and across data sets	Cross-dataset comparison
Uniqueness	One-to-one alignment between observed event and record	Duplicate detection
Integrity	Pedigree, provenance, and lineage known and aligned with business rules	Lineage audit
Timeliness	Time between event occurrence and data availability	Latency measurement

4-9c. Applying the dimensions. Data stewards and analysts use these seven dimensions as the measurement layer when conducting quality assessments. Each VAULTIS-A gate (paragraph 4-8) maps to one or more of these measurement dimensions. For example, the VAULTIS-A "Trusted" dimension (95% target) is assessed by measuring Accuracy, Completeness, and Integrity against the authoritative source. The "Auditable" dimension (100% target) is assessed through Integrity and Timeliness checks on lineage records.

4-9d. Units conducting DDOF Phase 3 quality gate assessments (paragraph 4-8) should apply these seven dimensions as the measurement protocol. Document findings against each dimension and report aggregate scores to the C2DAO for theater-wide quality tracking.

NOTE

These seven dimensions provide the measurement criteria underneath the VAULTIS-A quality framework. VAULTIS-A defines what makes data fit for use; these dimensions define how to measure it. Source: UDRA v1.1, Table 8.

CHAPTER 5 — DATA ANALYSIS FUNDAMENTALS

BLUF: Analysis is the process of extracting meaning from data. It must be driven by a question, not by the desire to produce a chart. Analysis at the wrong type for the right question wastes time and misleads decision-makers. Every analyst must match the analytical method to the operational question.

5-1. Descriptive, Diagnostic, Predictive, and Prescriptive Analytics

5-1a. Analytics exists on a spectrum of complexity and operational utility. Each type answers a different question.

Analytics Type	Question Answered	Army Example
Descriptive	What happened?	Unit readiness percentage last 30 days
Diagnostic	Why did it happen?	Root cause of readiness decline
Predictive	What will happen?	Forecast readiness 30 days out based on maintenance pipeline
Prescriptive	What should we do?	Recommended parts prioritization to restore readiness

5-1b. Most Army staff work is descriptive. Most operational value is in diagnostic and predictive. Most units have not yet developed the capability for systematic prescriptive analytics. Advancing up this spectrum requires both data quality investment and analytical skill development.

5-1c. Higher-order analytics depends on lower-order foundation. Predictive analysis is only as reliable as the descriptive data that feeds it. If readiness data is inconsistently reported, a readiness forecast built on that data will be unreliable. Fix the descriptive layer before attempting advanced analytics.

5-2. Summary Statistics — What They Tell You and What They Hide

5-2a. Summary statistics condense a large number of values into a few representative numbers. They are useful for communication but can obscure important detail.

Mean (Average). The sum of all values divided by the count. The mean is sensitive to outliers. A single extreme value significantly pulls the mean toward it. If nine vehicles have a readiness rate of 80% and one has 10%, the mean readiness is 73% — which may not reflect the operational reality of having one severely degraded vehicle.

Median. The middle value when all values are sorted. The median is resistant to outliers. In the above example, the median is 80%, which better represents the majority of vehicles. Use median when outliers may distort the mean.

Mode. The most frequently occurring value. Useful for categorical data (most common equipment type, most common MOS, most common failure code).

Range. The difference between the maximum and minimum values. Range indicates variability but is highly sensitive to outliers.

Standard Deviation. A measure of how spread out values are around the mean. A low standard deviation means values cluster near the mean (consistent performance). A high standard deviation means values are widely spread (inconsistent performance). High variability in readiness across subordinate units is operationally significant even if the mean looks acceptable.

5-2b. No single statistic tells the full story. Always present at least mean and variability together. Always ask whether outliers are meaningful or noise.

5-3. Distributions and Outliers — Reading a Dataset

5-3a. A distribution describes how values in a dataset are spread across their possible range. Understanding the distribution of a dataset is essential before applying analytical methods or drawing conclusions.

Normal Distribution. A symmetric, bell-shaped distribution where most values cluster near the mean and fewer appear at extremes. Many natural and performance metrics follow approximately normal distributions. Statistical tests and common analytical methods assume normality — applying them to non-normal data produces incorrect results.

Skewed Distribution. A distribution where most values cluster at one end with a tail at the other. Maintenance downtime is often right-skewed — most vehicles require short repair times, but a few require very long repair times. Reporting the mean of a skewed distribution overstates the typical experience.

Bimodal Distribution. A distribution with two peaks, suggesting two distinct sub-populations. A bimodal readiness distribution might indicate two distinct groups of equipment with different maintenance histories.

5-3b. Outliers are values that differ significantly from the bulk of the data. Outliers may indicate genuine exceptional events, errors, or fraud. Before treating an outlier as noise and removing it, investigate its cause. A vehicle with 5% readiness in a fleet averaging 85% is either severely degraded (operationally significant) or a data entry error (quality problem). Either way, it requires attention.

CAUTION

Do not remove outliers from datasets without investigation and documentation. Removing data points because they are inconvenient is analytical misconduct. If an outlier is confirmed as an error, document the correction. If it is a genuine extreme value, it must be included and explained.

5-4. Correlation vs. Causation — Critical for Operational Decisions

5-4a. Correlation means that two variables tend to move together — when one changes, the other tends to change in a predictable direction. Causation means that one variable directly causes a change in the other.

5-4b. Correlation does not establish causation. This is one of the most important and most frequently violated principles in analytical work. Two variables can be highly correlated for reasons that have nothing to do with a causal relationship.

5-4c. Example: Vehicle maintenance hours and vehicle readiness are correlated — more maintenance hours are associated with higher readiness. This seems to suggest that investing more maintenance hours increases readiness. This is partially true but incomplete. Causation runs in both directions: more maintenance hours produce higher readiness, but lower readiness also generates more maintenance hours. A new maintenance initiative that increases maintenance hours does not guarantee a proportional readiness increase if the underlying problem is parts availability or technician skill, not labor hours.

5-4d. Before recommending a course of action (COA) based on a correlation, establish the mechanism by which one variable affects the other. If no mechanism can be identified, the correlation may be spurious.

WARNING

Briefing a commander that one factor "causes" an outcome when only correlation has been demonstrated is analytical error. Use precise language: "These two factors are correlated" rather than "This factor causes that outcome" unless causation has been established through controlled analysis or subject matter expertise.

5-5. Visualization Principles — Good Charts vs. Misleading Ones

5-5a. Visualization translates data into a visual form that communicates patterns, trends, and relationships. Good visualization accelerates understanding. Bad visualization misleads.

Principles of effective visualization:

Match chart type to data type. Bar charts compare discrete categories. Line charts show change over time. Scatter plots show relationships between two variables. Maps show geographic distribution. Using the wrong chart type for the data distorts the message.

Start axes at zero. A bar chart or line chart with a y-axis that does not start at zero exaggerates differences between values. A 2% difference appears as a dramatic cliff when the y-axis spans only 1% around the data range.

Label everything. Every axis, every unit, every data source. A chart without labeled axes is an incomplete data product.

Avoid chartjunk. Three-dimensional bar charts, decorative graphics, and excessive color add visual noise without informational content. Simplicity communicates faster.

Show uncertainty. A forecast should show confidence intervals, not just a single projected value. Presenting a single predicted value without uncertainty bounds implies a precision that the analysis does not support.

Use color deliberately. Color should encode information, not decorate. Red conventionally indicates warning or degradation; green indicates health. Do not reverse this convention.

5-5b. Misleading visualization patterns to avoid:

Pattern	How It Misleads
Truncated y-axis	Exaggerates differences between values
Cherry-picked time window	Shows a favorable trend by excluding unfavorable periods
Dual y-axes	Can imply correlation between unrelated variables
Pie charts with many slices	Makes small differences impossible to compare
3D charts	Distorts relative bar heights; adds no information
Unlabeled axes	Makes the chart uninterpretable and indefensible

---

5-6. Hypothesis-Driven Analysis — Start with a Question, Not a Chart

5-6a. Analysis must begin with a question, not with data. A common failure mode is "data exploration" that produces many charts but no conclusions. Charts produced without a guiding question do not answer anything.

5-6b. Hypothesis-driven analysis follows a structured process:

1. **State the question.** What does the commander or staff need to know? Example: "Why has vehicle readiness declined in the past 30 days?"
2. **Form a hypothesis.** A testable explanation for what might answer the question. Example: "Readiness declined because of increased NMC due to a specific component failure."
3. **Identify the data needed to test it.** What data, if examined, would confirm or refute the hypothesis? Example: Maintenance work orders categorized by failure type over the past 60 days.
4. **Collect and analyze the data.** Apply appropriate methods.
5. **Evaluate the hypothesis.** Does the data support or refute it? Was the hypothesis wrong? Form a new hypothesis if needed.
6. **Communicate findings.** Brief the conclusion with supporting evidence.

5-6c. Hypothesis-driven analysis prevents the common failure of "fishing" for a positive result by trying many analyses until one looks interesting. If you try enough hypotheses, statistical chance will produce an apparent result. Stating the hypothesis before looking at the data imposes analytical discipline.

5-7. Communicating Findings to Commanders

5-7a. Analysis is only valuable if it reaches the commander in a form they can act on. The most sophisticated analysis is worthless if it cannot be communicated clearly.

5-7b. Apply the BLUF (Bottom Line Up Front) principle. Lead with the conclusion. Support it with the key data. State the recommended COA if analysis supports one. The commander can ask for more detail if needed. Do not bury the conclusion in the middle of a lengthy brief.

5-7c. The "So What" discipline. Every data finding must be accompanied by a "So What" — the operational significance of the finding. "Vehicle readiness is 72%" is a finding. "At 72% readiness, the battalion cannot execute its current MSN with all vehicles organic to the formation without augmentation" is a finding with operational significance.

5-7d. Calibrate to the audience. A data-literate commander can absorb more technical detail. A commander who is not data-literate requires plain language, visual aids, and an explicit statement of the confidence level of the analysis. Never assume the commander shares the analyst's technical background.

5-7e. Acknowledge uncertainty. Every analytical product has limitations. State them. "This forecast assumes continuation of current maintenance rates and parts availability. It does not account for potential surges in operational tempo." Failing to state limitations is a form of misrepresentation.

CHAPTER 6 — GEOSPATIAL DATA AND GEOINT LITERACY

BLUF: Geospatial data represents the physical world in digital form. It is foundational to military operations — virtually all MSN planning involves geography. Data-literate personnel understand how to read geospatial data, what its limitations are, and how to integrate it with other data types.

6-1. Geospatial Data Basics — Coordinates, Projections, and Layers

6-1a. All geospatial data represents locations on the surface of the Earth. The two foundational concepts are the coordinate system and the projection.

Coordinate Systems. A coordinate system defines how locations are expressed as numbers. The most common systems in Army use are:

- **Geographic Coordinates (Latitude/Longitude).** Express location as angles from the equator (latitude) and prime meridian (longitude). Expressed in degrees, minutes, and seconds or decimal degrees.
- **MGRS (Military Grid Reference System).** The standard military coordinate system, derived from the Universal Transverse Mercator (UTM) projection. Expresses location as a grid zone designation, 100km grid square identifier, and easting/northing coordinates. Standard for operational overlays and grid references.
- **UTM (Universal Transverse Mercator).** Divides the Earth into 60 longitudinal zones and expresses locations as easting and northing measurements in meters within each zone.

6-1b. **Projections.** The Earth is a sphere; a map is flat. A projection is the mathematical transformation used to represent the curved surface of the Earth on a flat plane. All projections introduce distortion — of area, shape, distance, or direction. No projection preserves all four properties simultaneously.

6-1c. **Layers.** Geospatial data is organized in layers, each representing a different type of feature. Terrain is one layer. Roads are another. Incident locations are another. Friendly and enemy positions are additional layers. Layers are overlaid to create a composite picture.

6-2. Common Geospatial File Formats

6-2a. Geospatial data is stored and exchanged in standardized formats.

Format	Description	Common Use
Shapefile (.shp)	Multi-file format for vector data (points, lines, polygons)	Unit boundaries, route overlays, facility locations
GeoJSON (.geojson)	JSON-based vector format; human-readable	Web-based mapping; data exchange
KML/KMZ (.kml, .kmz)	Keyhole Markup Language; used in Google Earth and similar tools	Simple overlays; sharing with non-specialized users
GeoTIFF (.tif)	Raster format with embedded geographic coordinates	Satellite imagery, terrain elevation data
DTED (Digital Terrain Elevation Data)	Elevation data in standardized grid format	Terrain analysis, LOS (Line of Sight) calculations

6-2b. Coordinate system and projection information must be included with any geospatial dataset. Data without this information cannot be reliably displayed or combined with other datasets — it has no reference point on the Earth's surface.

6-3. Reading Geospatial Visualizations

6-3a. A geospatial visualization (map) is a data product. Like any data product, it must be evaluated critically.

Elements every map must have: - Title (what the map shows) - Scale (relationship between map distance and real-world distance) - North arrow (orientation) - Legend (what each symbol and color represents) - Coordinate grid or graticule - Source and date of underlying data - Classification marking

6-3b. Map scale affects what can and cannot be seen. A small-scale map (covering a large area) cannot show individual building positions. A large-scale map (covering a small area) can show individual structures but not theater-level context. Always verify that the map scale is appropriate for the operational question being answered.

6-3c. Symbol density can mislead. A map with many incident symbols in one area may suggest a hotspot — or may simply reflect an area with better reporting, not a higher actual incident rate. Always consider whether the pattern on the map reflects reality or reflects data collection patterns.

6-4. Integrating Geospatial and Attribute Data

6-4a. Geospatial data becomes most powerful when integrated with attribute data — the non-spatial characteristics of a feature. A point on a map representing an installation becomes far more useful when linked to attribute data containing the installation's capacity, current occupancy, status, and unit assignment.

6-4b. Integration requires a common key. The geospatial record and the attribute record must share an identifier that allows them to be joined. If the identifier is not consistent between the two datasets — a common problem when systems use different naming conventions for the same location — the join will produce errors or omissions.

6-4c. Always validate the result of a geospatial-attribute join. Spot-check several records to confirm that the joined data makes sense — that the attribute data assigned to each geographic feature actually describes that feature.

6-5. Limitations of Geospatial Data

6-5a. Geospatial data is powerful but has specific limitations that affect its operational use.

Resolution. Imagery and raster data have finite resolution — there is a minimum feature size that can be distinguished. A satellite image with 30-meter resolution cannot show a feature that is 5 meters across. Know the resolution of your geospatial data and do not draw conclusions that require finer detail than the resolution supports.

Currency. Geospatial data has a collection date. The world changes; the data does not update automatically. A road network captured six months ago may not reflect current conditions — bridges may have been destroyed, new routes opened, or areas flooded. Always check the collection date of geospatial data, especially in areas of active military operations.

Accuracy. Position accuracy varies by collection method. GPS-surveyed points may be accurate to meters. Hand-digitized features from paper maps may have positional errors of hundreds of meters. Know the accuracy of the data and assess whether that accuracy is sufficient for the MSN.

Coverage Gaps. A geospatial dataset may not cover the entire area of operations (AOR). Gaps may result from collection limitations, cloud cover over imagery, or deliberate exclusions. Never assume that absence of data on a map means absence of a feature in the real world.

CHAPTER 7 — DATA SECURITY, CLASSIFICATION, AND HANDLING

BLUF: Data security is not optional and it is not the sole responsibility of the S6 or information system security officer. Every person who touches data bears responsibility for handling it correctly. Failure to handle data properly creates operational risk — to the mission, to Soldiers, and to national security.

7-1. Classification Levels and Their Impact on Data Handling

7-1a. The United States Government classifies information at three levels: Confidential, Secret, and Top Secret. Each level has defined handling requirements. Additional special access programs and compartments — Sensitive Compartmented Information (SCI) and Special Access Programs (SAP) — impose additional requirements beyond the base classification level.

7-1b. Classification determines:

- Which systems the data may be stored on or transmitted through
- Who may access the data
- How the data must be physically secured
- How the data must be destroyed when no longer needed
- How documents containing the data must be marked

7-1c. Data that is not classified may still be controlled. Controlled Unclassified Information (CUI) is unclassified information that requires protection under law, regulation, or government-wide policy. CUI categories include Law Enforcement Sensitive (LES) and Privacy Act-protected information. (Note: For Official Use Only (FOUO) was a legacy marking now retired and superseded by CUI.) CUI must be handled according to applicable CUI policy even though it is unclassified.

7-2. Need-to-Know vs. Need-to-Share

7-2a. Need-to-know is the traditional standard for access to classified information. An individual has a need-to-know if access to the information is required to perform their official duties. Clearance level alone does not grant access — a person with a Top Secret clearance does not automatically have access to all Top Secret information.

7-2b. Need-to-share is a complementary principle that recognizes the operational risk of information hoarding. Failure to share relevant information across organizational boundaries has historically contributed to intelligence failures and operational disasters.

7-2c. Commanders and staff must balance both principles. Applying only need-to-know creates information silos that degrade combined arms integration. Applying only need-to-share risks exposing sensitive information to parties who should not have it.

7-2d. The resolution: share at the classification level and with the audience appropriate to the information. Produce products at the lowest classification level that satisfies the operational requirement. Share laterally and vertically to the extent authorized.

7-3. Marking Requirements for Data Products

7-3a. All data products — regardless of format — must be marked according to their classification. A spreadsheet, a map, a chart, a database query result, and a briefing slide are all data products and all require appropriate marking.

7-3b. Classification markings must appear:

- At the top and bottom of every page of a document
- On every slide of a briefing
- On charts and maps as a notation within or adjacent to the product
- In the file name or metadata of electronic files, in accordance with local policy

7-3c. Derivative classification — creating a new product based on classified source material — requires that the derivative product be marked at the highest classification level of any source used in its creation, unless a declassification authority has authorized otherwise.

WARNING

A briefing that aggregates Secret-level data from multiple sources is itself a Secret product, even if no individual piece of data in the briefing is marked Secret in isolation. Aggregation can raise classification level. When in doubt, consult your unit security officer before distributing.

7-4. Access Controls and Data Compartmentalization

7-4a. Access controls limit who can read, modify, or delete data. They are implemented technically (usernames, passwords, role-based permissions) and procedurally (access logs, approval processes).

7-4b. Role-based access control (RBAC) grants access based on an individual's role in the organization rather than granting access individually. An S2 staff officer has access to intelligence products appropriate to their role; a logistics officer does not. RBAC reduces the risk of inadvertent access and simplifies management when personnel rotate.

7-4c. Compartmentalization separates data into discrete compartments, each with its own access controls. An individual must be separately authorized for each compartment they access, even if they hold the appropriate clearance level.

7-4d. Access logs — records of who accessed which data and when — are essential for accountability and incident response. If a data breach or unauthorized access is suspected, access logs are the primary investigative tool. Systems that do not maintain access logs cannot support accountability.

7-5. Risks of Improper Data Handling

7-5a. Improper data handling creates risk at multiple levels.

Spillage. Classified information processed or stored on a system with a lower classification level (e.g., classified data transmitted on an unclassified network) is a security incident requiring immediate reporting, investigation, and remediation. Spillage can compromise operations and sources.

Aggregation Risk. Individually unclassified pieces of information, when combined, can reveal classified information. The individual pieces — a unit's location, its readiness percentage, its scheduled departure date — may each be unclassified. Together, they reveal the unit's operational plan, which may be classified.

Insider Threat. Personnel with authorized access who misuse that access — intentionally or inadvertently — are an insider threat. Data handling procedures, access controls, and audit logs are all controls against insider threat.

Data Exfiltration. The unauthorized transfer of data from a secure environment to an insecure one. Exfiltration can occur via removable media, email, personal devices, or unauthorized cloud services. Procedures restricting removable media and personal devices in secure spaces mitigate this risk.

7-6. NATO Classification Markings and Coalition Data Interoperability

7-6a. USAREUR-AF operates in a coalition environment. Data products shared with or received from NATO and partner nation forces carry NATO classification markings that differ from U.S. national classification markings. All USAREUR-AF personnel handling multinational data must understand the equivalency and handling requirements for both systems.

7-6b. NATO classification levels are UNCLASSIFIED, NATO RESTRICTED, NATO CONFIDENTIAL, NATO SECRET, and COSMIC TOP SECRET. The general equivalency to U.S. classifications is as follows: NATO RESTRICTED approximates CUI; NATO CONFIDENTIAL approximates U.S. CONFIDENTIAL; NATO SECRET approximates U.S. SECRET; COSMIC TOP SECRET approximates U.S. TOP SECRET. However, these are not exact equivalencies. When in doubt about the appropriate handling of NATO-marked material, consult the unit security officer.

7-6c. Data sharing with NATO and Allied partners must comply with applicable Technical Arrangements, Memoranda of Understanding (MOU), and NATO security policy. Data products that aggregate U.S.-origin and NATO-origin information may carry multiple marking requirements. The more restrictive handling requirement governs the combined product.

7-6d. USAREUR-AF data systems must comply with NATO Architecture Framework version 4 (NAFv4) to ensure interoperability with Allied data environments. NATO NAFv4 alignment is particularly relevant for data products, system interfaces, and ontology definitions shared across the coalition. For NAFv4 implementation guidance, contact C2DAO.

7-6e. Coalition data sharing requirements in the EUCOM AOR are further governed by NATO Allied Joint Publications. AJP-3 (Allied Joint Doctrine for the Conduct of Operations) and AJP-3.2 (Allied Joint Doctrine for Land Operations) establish data sharing obligations for combined land operations; AJP-5 (Allied Joint Doctrine for the Planning of Operations) governs information requirements in multinational planning. USAREUR-AF data products generated in support of combined operations must satisfy the data interoperability requirements established in these publications. NATO doctrine crosswalks are available through C2DAO.

7-7. Responsibilities of Data Owners, Stewards, and Consumers

7-7a. Data security responsibility is distributed across three roles:

Data Owner. The organizational element responsible for a dataset. The data owner authorizes access, establishes classification, and is accountable for the data's security and accuracy. Data owners are typically commanders or senior leaders, not technicians.

Data Steward. The individual responsible for day-to-day management of the data on behalf of the data owner. The data steward manages access requests, enforces handling procedures, monitors for quality problems, and coordinates with system owners.

Data Consumer. Any individual who accesses and uses data. Consumers are responsible for using data only within their authorized access, handling it according to its classification, and reporting security incidents or quality problems.

7-7b. These roles are not mutually exclusive. A staff officer may be a data steward for their section's data while simultaneously being a consumer of data from other sections. Clear role assignment prevents accountability gaps.

7-8. Zero Trust Architecture

The Army Unified Network Plan 2.0 (March 2025) mandates Zero Trust Architecture (ZTA) across all Army systems. ZTA operates on the principle that no user, device, or system is trusted by default — even inside the network perimeter. Every access request is verified against user attributes, device posture, and data sensitivity. In practice, ZTA means: - Users must authenticate and have their attributes verified for each access request - Access grants are time-limited and context-specific, not permanent - All access is logged and auditable - The combination of CBAC markings and attribute-based policies implements ZTA at the data layer

---

CHAPTER 8 — DATA IN THE MDMP AND DECISION-MAKING PROCESS

Data in Multi-Domain Operations. Army doctrine grounds data requirements in Multi-Domain Operations (MDO). MDO integrates capabilities across land, air, sea, space, and cyberspace — and data is the connective tissue. The Army Data Plan (2022) established the foundational framework; current governance and architecture are codified in subsequent Army CIO guidance (2024) and the Unified Data Reference Architecture (UDRA) v1.1 (2025). USAREUR-AF operates in a contested, degraded, and operationally limited (CDO/CDOL) environment where adversaries actively target data infrastructure, communications, and command systems. In this environment, data literacy is not a staff function — it is a warfighting requirement. Units that cannot ingest, understand, and act on data faster than their adversaries will be at a decision disadvantage.

BLUF: The Military Decision-Making Process (MDMP) is a structured analytical process that is fundamentally data-dependent. Data quality problems that exist before MDMP begins degrade every subsequent step. Data literacy at the staff level directly enables better planning and better decisions.

8-1. Where Data Fits in the MDMP

8-1a. The MDMP (defined in ADP 5-0) is the Army's primary planning process. It consists of seven steps: Receipt of Mission, Mission Analysis, Course of Action (COA) Development, COA Analysis (Wargame), COA Comparison, COA Approval, and Orders Production. Data informs and shapes every step.

8-1b. Data is not optional at any step of the MDMP. A plan built on incomplete, inaccurate, or misunderstood data is a flawed plan. The quality of the plan is bounded by the quality of the data on which it is based.

8-2. Data Requirements at Each MDMP Step

MDMP Step	Key Data Requirements
Receipt of Mission	Higher headquarters order (EXORD, OPORD, WARNO), commander's guidance, timeline
Mission Analysis	Enemy (OPFOR) data, terrain and weather analysis, civil considerations, friendly force status, logistics status, time-space analysis
COA Development	Unit capabilities data, task organization options, logistics supportability data, terrain overlays

MDMP Step	Key Data Requirements
COA Analysis (Wargame)	Enemy COA data, resource consumption rates, timing and sequencing constraints
COA Comparison	Decision support criteria, weighted evaluation data, risk assessment data
COA Approval	Commander's intent, selected COA, branch and sequel triggers
Orders Production	Coordinating instructions, task organization, fire support plan, logistics plan

8-3. Information vs. Data Requirements — IR, PIR, FFIR

8-3a. The MDMP produces formal data and information requirements that drive collection planning.

Priority Intelligence Requirements (PIRs). Questions that, if answered, would significantly affect the commander's decision-making. PIRs are enemy-focused. They drive ISR collection and intelligence analysis. Each PIR should have a defined decision point — when does the commander need the answer?

Information Requirements (IRs). The broader set of questions the commander needs answered. PIRs are a subset of IRs focused on enemy and adversary activity.

Friendly Force Information Requirements (FFIRs). Questions about friendly force status, resources, and capabilities that the commander needs answered to make decisions. A readiness threshold that triggers a request for reinforcement is an FFIR trigger.

8-3b. PIRs, IRs, and FFIRs are data requirements expressed in operational terms. They define what data the staff must collect, analyze, and report. Translating these requirements into specific data queries — what systems hold the relevant data, what fields to query, how to analyze and present the results — is a data literacy task.

8-4. Data in IPB — Intelligence Preparation of the Battlefield

8-4a. IPB (Intelligence Preparation of the Battlefield), described in ATP 2-01.3, is the systematic process of analyzing the enemy, terrain, weather, and civil considerations in a specific AOR to support military planning. IPB is a data-intensive process.

8-4b. Terrain analysis in IPB uses geospatial data — elevation models, vegetation layers, hydrography, road networks, and urban area data — to assess trafficability, observation and fields of fire, cover and concealment, obstacles, and key terrain. The quality of geospatial data directly determines the accuracy of terrain analysis.

8-4c. Weather analysis uses meteorological data — precipitation, temperature, wind, visibility — to assess the operational impact of weather on friendly and enemy capabilities. Weather forecast accuracy degrades with range. IPB must account for forecast uncertainty.

8-4d. Civil considerations analysis integrates population data, infrastructure data, and economic data to assess the operational impact of civil factors. Civil data is often the least reliable input to IPB — it may be outdated, incomplete, or deliberately denied.

8-4e. Enemy analysis integrates all-source intelligence to assess enemy capabilities, intentions, and most likely and most dangerous COAs. The quality of enemy intelligence directly constrains the reliability of enemy COA analysis. Acknowledge intelligence gaps explicitly rather than filling them with assumptions.

8-5. Data Visualization in COA Development and Comparison

8-5a. COA development produces graphical and written products — operational overlays, scheme of maneuver narratives, fire support plans. These products are data visualizations.

8-5b. Operational overlays must accurately represent the intended scheme of maneuver, timing, and boundaries. Graphical errors — incorrect unit symbols, inaccurate boundaries, missing phase lines — create ambiguity that degrades mission execution.

8-5c. COA comparison uses a decision matrix — a structured comparison of COAs against weighted evaluation criteria. The quality of the COA comparison depends on the quality of the data used to score each criterion. Subjective scores without data backing are difficult to defend.

8-6. Decision Support Matrices and Data Backing

8-6a. A Decision Support Matrix (DSM) links decision points to specific information triggers. When a specific condition is met — an enemy action, a resource threshold, a time event — the matrix identifies the decision the commander must make and the COA branches available.

8-6b. Decision triggers in the DSM are data thresholds. "If enemy force exceeds two battalion equivalents at checkpoint 7 by H+4" is a data condition. The staff must monitor the relevant indicators to detect when the trigger condition is met.

8-6c. Effective DSM design requires clarity about what data will be monitored, how frequently it will be assessed, who is responsible for monitoring, and what the threshold is. Vague triggers — "if the situation deteriorates" — are unmonitorable and useless.

8-7. Data Literacy, Information Advantage, and ADP 3-13

8-7a. ADP 3-13, Information, establishes that information is combat power. The ability to generate, protect, and exploit information more effectively than the adversary constitutes information advantage — a decisive operational asset in the USAREUR-AF AOR and across the USEUCOM theater. Data literacy is the prerequisite for information advantage.

8-7b. The MDMP produces information requirements (PIRs, FFIRs, IRs) precisely because information — derived from quality data — shapes every planning step. A formation that cannot reliably collect, manage, and analyze operational data cannot satisfy its own information requirements, cannot produce accurate staff products, and cannot generate the quality analytical outputs that information advantage demands. Data quality failures at the staff level are therefore not administrative deficiencies — they are direct degradation to the command's ability to achieve information advantage.

8-7c. In the USAREUR-AF AOR, information advantage has particular operational weight. USAREUR-AF supports NATO Article 5 commitments and joint operations under USEUCOM. The adversary in this theater invests heavily in information operations, deception, and electronic warfare specifically to degrade allied situational awareness. USAREUR-AF formations that produce high-quality, well-sourced, timely data products are more resistant to adversary deception and better positioned to identify and exploit adversary data gaps.

8-7d. Data literacy in the MDMP is therefore not merely procedural competence — it directly contributes to the information advantage that ADP 3-13 defines as central to modern operations. Commanders must treat staff data literacy as an enabler of the broader information operations framework.

8-8. After-Action Data Collection and Lessons Learned Pipelines

8-8a. The After-Action Review (AAR) process generates operational lessons. AARs that rely solely on human memory are less reliable than those supplemented by operational data. What actually happened — as recorded in logs, communications, and reporting — may differ from what participants remember.

8-8b. Establish data collection practices before and during operations to support post-operation analysis. Define what data will be recorded, by whom, and in what format. This preparation enables higher-quality AARs and richer lessons learned products.

8-8c. Lessons learned that are not systematically captured and disseminated are lost. A formal lessons learned pipeline — collection, analysis, validation, dissemination, and incorporation — preserves institutional knowledge and prevents repeating avoidable mistakes.

CHAPTER 9 — DATA ROLES AND RESPONSIBILITIES

BLUF: Data management requires clearly defined roles. Ambiguity in data roles creates accountability gaps, quality failures, and security vulnerabilities. Every organization must assign and enforce data roles explicitly.

9-1. Data Producer vs. Data Consumer vs. Data Steward

9-1a. Three fundamental roles exist in any data ecosystem.

Data Producer. An individual, system, or organization that creates or collects data. A Soldier entering a maintenance work order is a data producer. A sensor generating telemetry is a data producer. The quality of a data ecosystem is limited by the quality of its producers.

Data Consumer. An individual, system, or organization that uses data to derive information or make decisions. A staff officer querying a logistics database is a data consumer. A commander interpreting a readiness dashboard is a data consumer. Consumers are responsible for understanding the limitations of the data they consume.

Data Steward. An individual responsible for managing and maintaining data on behalf of its owner. Stewards enforce quality standards, manage access, document lineage, and coordinate between producers and consumers.

9-1b. The same individual may occupy multiple roles simultaneously. A battalion S4 may be a data producer (entering logistics data into the system), a data consumer (analyzing readiness trends), and a data steward (managing access to the battalion's logistics data).

9-2. Staff Responsibilities for Data

9-2a. Each staff section has specific data responsibilities tied to its functional area.

Staff Section	Primary Data Domain	Key Data Products
S1 (Personnel)	Personnel status, strength accounting, casualty data	PERSTAT, USR (personnel portion), casualty reports
S2 (Intelligence)	Enemy data, terrain analysis, weather, civil considerations	IPB products, intelligence summaries, SALUTE reports
S3 (Operations)	Operations status, task organization, synchronization	OPORD, FRAGORD, SITREP (operational portion), decision support matrix

Staff Section	Primary Data Domain	Key Data Products
S4 (Logistics)	Supply status, maintenance status, transportation	LOGSTAT, OPSTAT, Class I-IX status reports
S6 (Communications)	Network status, system availability, data transport	Network status reports, communications plan

9-2b. Staff sections must not operate as data silos. Cross-staff data integration — combining personnel, logistics, intelligence, and operations data — produces the holistic operational picture that commanders need. Data literacy across all staff sections enables this integration.

9-3. Data Owner vs. System Owner vs. Data Custodian

9-3a. These three roles are often confused but have distinct responsibilities.

Data Owner. The organizational element or leader accountable for a specific dataset. The data owner sets access policy, authorizes use, and is accountable for the data's accuracy and security. In Army contexts, the data owner is typically the commander or senior functional leader for the relevant functional area.

System Owner. The organizational element responsible for managing the information system on which data resides. The system owner maintains the system's availability, security, and performance. The system owner controls the platform; the data owner controls the data on the platform.

Data Custodian. The individual or organization with physical or technical custody of the data — managing storage, backup, and transmission. The S6 or information management officer often serves as data custodian for systems they operate.

9-3b. When a data quality problem is discovered, responsibility is determined by the role. A quality problem caused by incorrect data entry is the data owner's problem. A quality problem caused by a system failure is the system owner's problem. A quality problem caused by improper storage or transmission is the custodian's problem.

9-4. The Data Literacy Continuum — The USAREUR-AF Training Framework

9-4a. Data literacy is not binary. It exists on a continuum. The USAREUR-AF training framework organizes this continuum into five tiers, designated TM-10 through TM-50. The first three tiers apply to all personnel; the TM-40 and TM-50 tiers are specialist tracks for designated data roles. All tiers are cumulative — each builds on the one below.

Tier	Title	Description	Target Population
TM-10	Operator	Navigates Maven Smart System; consumes data products; follows data handling procedures; accurately enters data; observes security requirements	All Soldiers and Civilians — every MOS and staff function
TM-20	Builder	Builds basic Workshop applications; runs light transforms; creates and shares data products without specialist tools; identifies data quality problems	All personnel — builds on TM-10
TM-30	Advanced Builder	Designs pipelines and Ontology objects; builds governed dashboards and complex transforms; profiles data quality; mentors TM-10/20 users	Data-adjacent personnel (17/25-series, S6/G6/G2)
TM-40 Series (WFF)	Warfighting Function Specialist	Role-specific MSS integration within an assigned warfighting function. Six tracks: Intelligence (40A), Fires (40B), Movement & Maneuver (40C), Sustainment (40D), Protection (40E), Mission Command (40F). Applies MSS tools to functional domain workflows.	WFF functional staff (INT, FIRES, M2, SUST, PROT, MC); requires TM-30 prerequisite
TM-40 Series (Technical)	Technical Specialist	Role-specific mastery within a designated technical specialty. Eight tracks: ORSA (40G), AI Engineer (40H), ML Engineer (40M), Program Manager (40J), Knowledge Manager (40K), Software Engineer (40L), UX Designer (40N), Platform Engineer (40O). Produces command-level data products independently.	Designated specialist billets (ORSA, AI Eng, ML Eng, PM, KM, SWE, UX, PE); requires TM-30 prerequisite
TM-50 Series	Advanced Specialist	Advanced practitioner capability within a specialist track (TM-50G through TM-50O); leads, mentors, and develops new capability; research-grade analytical and engineering proficiency. No TM-50 WFF tracks.	TM-40G–O graduates in senior or lead roles; requires TM-40 in same track

9-4b. The goal is not for every Soldier to reach TM-50. The goal is for every Soldier to complete TM-10, for all personnel to achieve TM-20, for data-adjacent personnel to reach TM-30, for WFF functional staff to complete the appropriate TM-40 WFF track (TM-40A–F, prereq TM-30), and for the formation to have technical specialist coverage across the TM-40 specialist tracks (TM-40G–O, prereq TM-30) appropriate to its mission. TM-50 capability (TM-50G–O) is required in senior data roles and for personnel responsible for developing and sustaining formation data capability. There are no TM-50 WFF tracks — the TM-50 series applies only to the eight technical specialist tracks (G through O).

9-5. Command Responsibilities for Data Culture

9-5a. Data culture — the values, behaviors, and practices around data in an organization — is set at the command level. Commanders who demand data quality get it. Commanders who accept poor-quality data products enable a culture of sloppy data practices.

9-5b. Commanders establish data culture through:

Demanding Data Quality. When a briefing contains unsourced or obviously questionable data, the commander asks where the number came from and what the source quality is. This demand, applied consistently, teaches the staff that data provenance matters.

Resourcing Data Capabilities. Data literacy training, analytical tools, and dedicated data roles require resources — time, people, and equipment. Commanders must explicitly resource these capabilities or accept degraded analytical capacity.

Holding Units Accountable. Subordinate units that consistently report inaccurate, incomplete, or late data must be corrected. Data quality problems at subordinate units that are not corrected by the commander become the commander's quality problems.

Modeling Data Literacy. Commanders who ask good data questions, who understand the difference between correlation and causation, and who demand sourced analysis create a culture where data literacy is valued at every level.

9-6. Training Requirements and Qualification Criteria

9-6a. Data literacy training must be systematic and progressive — not a one-time event. The USAREUR-AF training framework uses the TM series to define minimum standards by role and echelon.

9-6b. Minimum training standards by echelon and role:

Population	Minimum Standard	Training Path
All Soldiers and Civilians	TM-10 (Operator) — Navigate Maven, consume data products, follow data handling and security procedures	Unit-level training; TM-10 self-paced course
All personnel	TM-20 (Builder) — Build basic data products; identify and report quality issues	Unit training; TM-20 course (builds on TM-10)
Data-adjacent billets (17/25-series, S6/G6/G2)	TM-30 (Advanced Builder) — Design pipelines and Ontology objects; govern data products	TM-30 course (builds on TM-20); mentored practicum
WFF functional staff (INT, FIRES, M2, SUST, PROT, MC)	TM-40 Series (WFF Track) — MSS integration within assigned warfighting function	TM-40A through TM-40F per WFF assignment; prerequisite TM-30
ORSA, AI Eng, ML Eng, PM, KM, SWE billets	TM-40 Series (Technical Specialist) — Role-specific track at the assigned specialist level	TM-40G through TM-40O per billet type; prerequisite TM-30
Senior data leads, capability developers, training cadre	TM-50 Series (Advanced Specialist) — Research-grade proficiency; leads and develops capability in specialist track	TM-50G through TM-50O per track; prerequisite TM-40G–O in same track

Population	Minimum Standard	Training Path
Commander/Senior Leader	Commander-level (direction, evaluation, resourcing) — Not a technical track; governed by this publication and the companion senior leader guide	Data Literacy for Senior Leaders; unit immersion; C2DAO advisory

9-6c. Data literacy qualification is a readiness consideration. A unit with systematically low data literacy across the staff cannot produce the quality of analytical products required for complex operations. Commanders must assess and develop data literacy as part of overall readiness. TM-10 and TM-20 completion should be tracked as unit-level training metrics comparable to weapons qualification.

9-7. Army Data Governance Structure — The 4-Tier Stewardship Hierarchy

9-7a. The Army CIO Memo (April 2024) establishes a four-tier data stewardship hierarchy governing how data policy is set, enforced, and executed across the Army enterprise. All USAREUR-AF personnel with data management responsibilities must understand where their role fits within this structure.

Tier	Role	Authority	Scope
1	**Mission Area Data Officers (MADOs)**	Set enterprise data policy within their mission area	Four appointed officials covering Warfighter, Intelligence, Business, and Enterprise IT mission areas
2	**Data Stewards**	Establish access policies, retention requirements, and data generation standards for specific data types	Appointed by MADOs; authority is domain-specific
3	**Functional Data Managers**	Execute day-to-day data management for specific programs and systems	Typically PEOs and PMs; operational implementation
4	**Command Chief Data and Analytics Officers (C2DAOs)**	Consume and enforce enterprise data policy at the command level	Command-level; do NOT create Army enterprise data policy

9-7b. USAREUR-AF's Chief Data and Analytics Officer (CDAO) function operates at Tier 4. USAREUR-AF consumes and enforces enterprise data policy established by MADOs and Data Stewards. USAREUR-AF does not create Army enterprise data policy. Within the USAREUR-AF AOR, the C2DAO establishes command-level data standards that implement, but do not supersede, Army enterprise policy.

9-7c. This hierarchy is not a reporting chain — it is a policy chain. Data policy flows downward from MADOs through Data Stewards to Functional Data Managers and C2DAOs. Operational data problems that require policy changes above the command level must be escalated through the appropriate MADO or Data Steward, not resolved at command level through workaround procedures.

9-7d. **Governance must be upstream.** The most common failure mode in enterprise data programs is treating governance as a documentation exercise — something that happens after systems are built and data is flowing. Effective data governance requires participation in requirements definition, funding decisions, and acquisition processes *before* systems are fielded. A data architecture built without governance participation will be technically coherent but operationally inconsistent: different systems will define the same entity differently, access will be granted inconsistently, and quality will degrade at every handoff. Architecture without governance is documentation. Governance without architecture is policy without mechanism. Both are required. This principle applies at every tier of the Army data stewardship hierarchy and at the command level within the USAREUR-AF AOR. Data personnel must have a seat at the table during system acquisition and requirements development, not just during post-fielding integration.

NOTE

For current USAREUR-AF data governance contacts and policy memoranda, contact C2DAO or your unit data steward.

CHAPTER 10 — PRINCIPLES OF DATA-DRIVEN OPERATIONS

BLUF: Eight principles govern data-driven military operations. These principles are not aspirational — they are doctrinal standards. Adherence to these principles is the difference between a formation that uses data effectively and one that is served by data that it cannot use.

10-1. Principle 1: Data Is Ammunition — Treat It Accordingly

Data has operational value equivalent to physical resources. It enables decisions that determine outcomes. Losing, corrupting, or mishandling data degrades operational effectiveness in ways analogous to losing ammunition or equipment. Data must be protected, inventoried, and accounted for. Data produced carelessly or handled sloppily is as dangerous as ammunition improperly stored — it creates risk for the entire formation. Commanders must instill in their formations the understanding that data quality and security are readiness issues, not administrative concerns.

10-2. Principle 2: Quality Over Quantity — Bad Data Is Worse Than No Data

More data is not inherently better. A dataset with high volume and low quality degrades analysis. A small, clean, well-documented dataset is more operationally useful than a large, dirty, poorly documented one. Bad data is worse than no data because it produces false confidence — it gives analysts and decision-makers something to cite, when that something is wrong. Formations must resist the temptation to measure data capability by volume. The standard is quality: accuracy, completeness, consistency, timeliness, validity, and uniqueness.

10-3. Principle 3: Context Transforms Data Into Information

Raw data without context is not information. The number "72" is meaningless. "72% vehicle readiness as of 011200ZMAR2026, compared to a brigade standard of 85%, representing a 13- percentage-point decline from last month" is information. Context includes: what the data represents, when it was collected, what the reference standard is, and how it compares to prior periods or other elements. Every data product must supply context. Analysts are responsible for providing context; consumers are responsible for demanding it.

10-4. Principle 4: Analysis Precedes Visualization

Charts and maps are the output of analysis, not the substitute for it. A formation that builds charts before defining the analytical question has reversed the process. Visualization communicates analytical conclusions — it does not generate them. The question drives the analysis; the analysis drives the visualization. Formations that produce many charts but answer no specific questions are engaged in data theater, not data-driven operations. Every visualization must have a question it is designed to answer and an audience it is designed to serve.

10-5. Principle 5: Automation Enables Scale; Humans Enable Judgment

Automated data processing enables scale that no manual process can match. A formation operating in a data-rich environment cannot analyze its data manually — the volume is too great and the velocity too high. Automation must be embraced. But automation cannot replace human judgment. Automated systems process what they are designed to process and fail in ways they are not designed to recognize. Humans identify context, apply experience, and make decisions that automated systems cannot. The right architecture is automated processing with human review and judgment at decision points. Ceding judgment to automation is an abdication of command responsibility.

10-6. Principle 6: Lineage Enables Trust

A data product that cannot be traced back to its source cannot be defended. Lineage — the documented record of where data came from and what was done to it — is the basis of trust in a data product. When a commander asks "How do you know this?" the analyst must be able to answer: "This data came from this system, on this date, processed in this way, compared to this baseline." Formations that cannot answer this question are not data-driven; they are data-performing. Establishing lineage is not bureaucratic overhead — it is the mechanism by which analytical credibility is built and maintained.

10-7. Principle 7: Security by Design, Not as Afterthought

Data security cannot be retrofitted. When a data system, product, or pipeline is designed, security requirements must be built in from the beginning — not added after the fact. Classification decisions must be made at data creation, not after the product is already distributed. Access controls must be configured before data is stored, not after a breach is discovered. Data handling procedures must be established before an operation begins, not during it. Security by design requires that every individual involved in data production and management understand the security requirements applicable to their data before they begin working with it.

10-8. Principle 8: Commanders Own Data Culture in Their Formation

Data culture — the shared values and behaviors around data — is a command responsibility. No data officer, no system, and no training program can establish data culture without command support and modeling. The commander sets the standard by demanding quality, rewarding accurate reporting even when it carries bad news, and sanctioning sloppy data practices. A commander who accepts a readiness briefing built on unverified data teaches the staff that unverified data is acceptable. A commander who asks for the source, the sample size, and the confidence level teaches the staff that data rigor is expected. Data culture is a command climate issue, and the commander owns it.

CHAPTER 11 — UNIFIED DATA REFERENCE ARCHITECTURE — DATA MESH CONCEPTS

BLUF: The Unified Data Reference Architecture (UDRA) v1.1 establishes data mesh as the Army's architectural paradigm for data management. Data mesh replaces centralized data architectures with a model of distributed ownership, domain-aligned data products, and computationally enforced governance. Every formation that produces or consumes data operates within this architecture.

11-1. Data Products

11-1a. A data product is a logically pre-packaged unit of data and associated metadata produced to satisfy a consumer's mission or business demand. A data product is not raw data. It is curated, documented, and governed data that is ready for consumption without additional preparation by the consumer.

11-1b. Data products are self-describing: they include metadata that defines their content, schema, quality, lineage, refresh schedule, and responsible owner. A consumer can discover a data product in the catalog and understand what it contains, where it came from, how current it is, and who is accountable for its accuracy — without contacting the producer.

11-1c. Data products are computationally governed: quality standards, access controls, classification markings, and retention policies are enforced through automated mechanisms, not manual review alone. This ensures consistent governance at scale across the enterprise.

11-1d. Military application. In the USAREUR-AF context, a data product might be a theater readiness dataset published by G3 for consumption by V Corps and subordinate commands, a logistics consumption forecast published by 21 TSC for sustainment planning, or a geospatial threat layer published by G2 for operational planning. Each is owned by the producing domain, governed by automated policy, and consumed by authorized users across the theater.

11-2. Data Domains

11-2a. A data domain is an organization with specific functional expertise that produces data products. In the Army context, data domains align to staff functions and subordinate commands. The G2 section is a data domain for intelligence data. The G4 section is a data domain for logistics data. Each domain owns and is accountable for the data it produces.

11-2b. Domain ownership means the functional element closest to the data — the element with subject matter expertise — is responsible for its quality, accuracy, and availability. Data management responsibility does not centralize in a single IT element or data center. The producing domain manages its own data products in accordance with enterprise standards.

11-2c. This principle directly supports the Army data stewardship hierarchy (paragraph 9-7). Functional Data Managers within each domain are accountable for data quality; the C2DAO provides enterprise governance and standards enforcement, not centralized data ownership.

11-3. Computational Governance

11-3a. Computational governance is the automated enforcement of governance policies across data products and domains. It replaces manual, document-based governance with machine-enforceable rules. Two key concepts:

- **Standards as code.** Data quality standards, schema requirements, and naming conventions are encoded as executable rules that run automatically against data products. A data product that violates a schema standard is flagged or blocked before it reaches consumers.
- **Policies as code.** Access control policies, classification rules, retention schedules, and handling procedures are implemented as automated controls. Policy enforcement does not depend on individual compliance alone — it is built into the platform.

11-3b. Computational governance enables scale. A theater-level data environment with hundreds of data products cannot be governed through manual review and checklists alone. Automated governance ensures consistent enforcement across all domains and products while freeing data stewards to focus on judgment-intensive decisions.

11-4. UDRA Service Architecture

11-4a. The UDRA defines six core services that compose the Army data architecture. All data systems and platforms must provide or integrate with these services:

#	Service	Purpose
1	**Production**	Creation and curation of data products by source domains
2	**Orchestration**	Coordination and scheduling of data pipelines, transforms, and workflows across domains
3	**Consumption**	Discovery, access, and use of data products by authorized consumers
4	**Access Management**	Authentication, authorization, and enforcement of need-to-know / need-to-share policies

#	Service	Purpose
5	**API Brokerage**	Standardized interfaces enabling programmatic data exchange between systems and domains
6	**Computational Governance**	Automated enforcement of quality standards, policies, and compliance requirements

11-4b. These six services operate across all layers of the USAREUR-AF 5-Layer Data Stack (paragraph 3-7). The Production and Orchestration services map primarily to Layers 2-3 (Integration and Semantic). Consumption and API Brokerage map primarily to Layers 4-5 (Analytics and Activation). Access Management and Computational Governance are cross-cutting services that operate at every layer.

NOTE

Source: Unified Data Reference Architecture (UDRA) v1.1, ASA(ALT)/DASA(DES) and Army CIO, 6 September 2024. The UDRA is the authoritative Army architecture reference for data systems design. Implementation guidance specific to USAREUR-AF is available through C2DAO.

CHAPTER 12 — ARMY DATA PLAN — STRATEGIC OBJECTIVES

BLUF: The Army Data Plan (2022) established eleven strategic objectives and five strategic enablers that define the Army's data management vision. Strategic Enabler 05 (Talent) is the mandate for MSS training: the Army cannot achieve its data objectives without a trained workforce. Commanders must understand these objectives to nest their unit data programs within the Army's strategic direction.

12-1. Strategic Objectives

12-1a. The Army Data Plan defines eleven strategic objectives (SO) that guide Army-wide data management, governance, and analytics. These objectives are not aspirational — they are directed outcomes that all Army organizations must support.

SO	Title	Description
SO-01	Data Governance	Establish enterprise data governance with clear roles, responsibilities, and accountability at every echelon
SO-02	Data Architecture	Implement a unified data architecture that enables interoperability, discoverability, and reuse across the enterprise
SO-03	Data Standards	Adopt and enforce common data standards, metadata schemas, and naming conventions Army-wide
SO-04	Data Quality	Ensure data accuracy, completeness, consistency, timeliness, and fitness for use through measurable quality controls
SO-05	Data Access	Enable authorized users to discover and access data when and where needed, consistent with security requirements
SO-06	Data Sharing	Promote responsible data sharing across organizations, systems, and classification domains to maximize operational value
SO-07	Data Analytics	Scale analytical capabilities to transform data into actionable insights that support decision-making at all echelons
SO-08	AI/ML Integration	Integrate artificial intelligence and machine learning into Army data systems to enhance speed and quality of analysis
SO-09	Data Security	Protect data throughout its lifecycle through classification, access controls, and compliance with security policy

SO	Title	Description
SO-10	Data Infrastructure	Modernize data infrastructure (cloud, edge, tactical) to support data operations across all operating environments
SO-11	Performance Measurement	Establish metrics and accountability mechanisms to track progress against data management objectives

12-2. Strategic Enablers

12-2a. The Army Data Plan identifies five strategic enablers (SE) that underpin the eleven strategic objectives. Without these enablers, the strategic objectives cannot be achieved.

SE	Title	Description
SE-01	Leadership	Commander and senior leader commitment to data-driven operations as a warfighting imperative
SE-02	Policy	Authoritative policy framework governing data management, access, quality, and security
SE-03	Resources	Funding, equipment, and platform investment to sustain data operations at scale
SE-04	Partnerships	Collaboration across Army, Joint, interagency, and coalition partners for data interoperability
SE-05	**Talent**	**Trained workforce with data literacy, analytical skills, and technical expertise to execute data operations**

12-2b. **SE-05 (Talent) is the mandate for MSS training.** The Army Data Plan explicitly recognizes that the Army cannot achieve any of its data strategic objectives without a workforce trained in data literacy, data management, and data analytics. The MSS Training Program (TM-10 through TM-50) is USAREUR-AF's implementation of SE-05. Every MSS course directly supports one or more of the eleven strategic objectives by developing the human capital required to execute them.

12-2c. Nesting. Commanders aligning their unit data programs with the Army Data Plan should map their training and operational data activities to the relevant strategic objectives and enablers. The C2DAO can assist with this alignment.

NOTE

Source: Army Data Plan (2022), Office of the Chief Information Officer (OCIO). Superseded in part by Army CIO Data Stewardship Policy (April 2024) and UDRA v1.1 (February 2025) for governance and architecture specifics. The strategic objectives and enablers remain the foundational Army-wide framework.

APPENDIX A

DATA LITERACY SELF-ASSESSMENT

PURPOSE: This self-assessment enables individual Soldiers and leaders to identify their current data literacy level and determine a path for development. Answer each question honestly based on current capability, not aspirational capability.

SECTION 1 — DATA AWARENESS (LEVEL 1)

#	Question	Yes	Partially	No
1.1	I understand why accurate data entry matters for my unit's mission.			
1.2	I follow my unit's data handling and security procedures.			
1.3	I know how to identify the classification of a data product and handle it accordingly.			
1.4	I report data entry errors or system problems to my chain of command.			
1.5	I understand the difference between data, information, and a decision.			

Score: 5 Yes = Level 1 qualified. Less than 3 Yes = start here before proceeding.

SECTION 2 — DATA USER (LEVEL 2)

#	Question	Yes	Partially	No
2.1	I can read and interpret a tabular data report (e.g., a readiness table or logistics status report).			
2.2	I can identify when a chart is misleading (e.g., truncated axis, missing labels).			

#	Question	Yes	Partially	No
2.3	I ask about the source and date of data before using it to make a recommendation.			
2.4	I can identify obvious data quality problems (e.g., missing values, inconsistent entries).			
2.5	I understand the difference between the mean and median and know when each is more appropriate.			
2.6	I can read a geospatial map product and identify its key elements (scale, legend, source).			

Score: 5-6 Yes = Level 2 qualified. 3-4 = approaching. Less than 3 = continue Level 1 development.

SECTION 3 — DATA ANALYST (LEVEL 3)

#	Question	Yes	Partially	No
3.1	I can construct a data-driven staff product with documented sources and clear conclusions.			
3.2	I can profile a dataset to assess its quality before using it.			
3.3	I understand the difference between correlation and causation and apply that distinction in my work.			
3.4	I can build a basic visualization (chart or table) that accurately represents the underlying data.			
3.5	I can apply descriptive statistics (mean, median, standard deviation, range) appropriately.			
3.6	I can identify and document data lineage for a product I produce.			
3.7	I can articulate the "so what" of a data finding in operational terms.			

Score: 6-7 Yes = Level 3 qualified. 4-5 = developing. Less than 4 = focus Level 2 skill gaps first.

SECTION 4 — DATA PRACTITIONER (LEVEL 4)

#	Question	Yes	Partially	No
4.1	I can design and build a repeatable data processing pipeline (automated or scripted).			
4.2	I can apply statistical methods (regression, distribution testing, forecasting) to operational data.			
4.3	I can evaluate the design of a data collection system and identify structural quality risks.			
4.4	I can integrate data from multiple source systems using join operations.			
4.5	I have trained others in data analytical methods or tools.			
4.6	I can design a data product for a specific audience and analytical requirement.			

Score: 5-6 Yes = Level 4 qualified. 3-4 = developing. Less than 3 = focus Level 3 skill gaps first.

DEVELOPMENT PATH RECOMMENDATIONS

Current Level	Recommended Next Steps
Below Level 1	Unit data handling training; Army digital literacy training
Level 1	Study readiness and logistics reports for your unit; practice reading data products with intent
Level 2	Take a structured data analysis course; practice building briefing products with sourced data
Level 3	Learn scripting for data processing; pursue functional area training; develop pipelines for your section
Level 4	Pursue advanced statistical training; contribute to unit data strategy; develop training for subordinates

APPENDIX B

DATA QUALITY CHECKLIST

PURPOSE: Use this checklist before using any dataset for operational decision-making. Not every item applies to every dataset. Apply judgment about which checks are relevant to the data type and operational context.

SECTION 1 — PROVENANCE AND METADATA

#	Check	Status
B-1.1	Source system identified: I know which system produced this data.	☐ Pass / ☐ Fail / ☐ N/A
B-1.2	Collection date known: I know when this data was collected or last updated.	☐ Pass / ☐ Fail / ☐ N/A
B-1.3	Data is current enough for the intended use: the age of the data does not compromise its validity for this purpose.	☐ Pass / ☐ Fail / ☐ N/A
B-1.4	Collection method known: I understand how the data was gathered (automated, manual, etc.).	☐ Pass / ☐ Fail / ☐ N/A
B-1.5	Data owner identified: I know who is responsible for the accuracy of this data.	☐ Pass / ☐ Fail / ☐ N/A
B-1.6	Field definitions documented: I understand what each field/column means and its unit of measure.	☐ Pass / ☐ Fail / ☐ N/A

SECTION 2 — COMPLETENESS

#	Check	Status
B-2.1	Record count is as expected: the number of records in the dataset matches the expected population.	☐ Pass / ☐ Fail / ☐ N/A
B-2.2	Required fields are populated: fields that should always have a value are not blank.	☐ Pass / ☐ Fail / ☐ N/A

#	Check	Status
B-2.3	Missing value rate is acceptable: the percentage of blank values in each field is within tolerable limits.	☐ Pass / ☐ Fail / ☐ N/A
B-2.4	Time range coverage: the data covers the intended time period without gaps.	☐ Pass / ☐ Fail / ☐ N/A
B-2.5	Geographic coverage: the data covers the intended AOR without gaps.	☐ Pass / ☐ Fail / ☐ N/A

SECTION 3 — ACCURACY AND VALIDITY

#	Check	Status
B-3.1	Values are in expected formats: dates are valid dates, numbers are numbers, coordinates are in correct format.	☐ Pass / ☐ Fail / ☐ N/A
B-3.2	Values are within expected ranges: no readiness percentages above 100 or below 0; no negative quantities where negative values are impossible.	☐ Pass / ☐ Fail / ☐ N/A
B-3.3	Categorical values are from the approved list: status codes, unit designations, equipment types are from defined sets.	☐ Pass / ☐ Fail / ☐ N/A
B-3.4	Spot-check verification: a random sample of records has been verified against source documents or authoritative systems.	☐ Pass / ☐ Fail / ☐ N/A
B-3.5	Outliers investigated: extreme values have been examined and are either corrected or documented as genuine.	☐ Pass / ☐ Fail / ☐ N/A

SECTION 4 — CONSISTENCY

#	Check	Status
B-4.1	Cross-system consistency: data from this system is consistent with related data from other systems.	☐ Pass / ☐ Fail / ☐ N/A
B-4.2	Longitudinal consistency: current data is consistent with prior periods (unexplained discontinuities are flagged).	☐ Pass / ☐ Fail / ☐ N/A
B-4.3	Aggregation consistency: summary totals match the sum of component records.	☐ Pass / ☐ Fail / ☐ N/A

#	Check	Status
B-4.4	Reporting standard applied uniformly: all reporting units applied the same definitions and standards.	☐ Pass / ☐ Fail / ☐ N/A

SECTION 5 — UNIQUENESS

#	Check	Status
B-5.1	Duplicate records checked: the dataset has been checked for duplicate entries and duplicates resolved.	☐ Pass / ☐ Fail / ☐ N/A
B-5.2	Unique identifiers are unique: primary key fields (serial numbers, SSN, EIN, etc.) appear only once per entity.	☐ Pass / ☐ Fail / ☐ N/A

SECTION 6 — FITNESS FOR PURPOSE

#	Check	Status
B-6.1	This dataset covers the population relevant to my analytical question.	☐ Pass / ☐ Fail / ☐ N/A
B-6.2	This dataset is at the correct level of granularity for my question (not too aggregated, not excessively detailed).	☐ Pass / ☐ Fail / ☐ N/A
B-6.3	The data was not collected for a different purpose that might introduce systematic bias.	☐ Pass / ☐ Fail / ☐ N/A
B-6.4	I can document this dataset's quality assessment if asked to defend my analysis.	☐ Pass / ☐ Fail / ☐ N/A

CHECKLIST SCORING GUIDANCE

Result	Action
All Pass	Proceed with analysis; document the assessment.

Result	Action
One or more Fail in Sections 1-2	Resolve provenance/completeness issues before proceeding; brief the commander on known gaps.
One or more Fail in Sections 3-4	Investigate discrepancies; determine if they are correctable; proceed only if errors do not affect the analytical conclusion; document residual limitations.
Multiple Fail across sections	Do not use this dataset for operational decision-making until quality issues are resolved or an alternative source is identified.

APPENDIX C

REFERENCES AND RELATED PUBLICATIONS

ARMY DOCTRINE AND REGULATIONS

- **AR 25-1, Army Information Technology (Jul 2019)** — Statutory framework for Army data management, data governance, and IT policy. Established the original VAUTI (5-dimension) data quality principles. NOTE: VAUTI superseded by VAULTIS (DoD Data Strategy 2020, 7 dimensions) and extended to VAULTIS-A (DDOF Playbook v2.2, 8 dimensions). See paragraph 4-8.
- **ADP 3-13, Information** — Establishes doctrine for information as combat power and the foundation for information advantage. This publication implements ADP 3-13 at the data literacy level.
- **ADP 5-0, The Operations Process** — Defines the Military Decision-Making Process (MDMP) and the role of information in planning and execution.
- **ATP 2-01.3, Intelligence Preparation of the Battlefield (IPB)** — Defines IPB methodology and its data requirements.
- **AR 25-400-2, The Army Records Information Management System (ARIMS)** — Governs operational data retention requirements.

NATO PUBLICATIONS

- **NATO Architecture Framework version 4 (NAFv4)** — The NATO standard for architectural descriptions of systems, capabilities, and data. USAREUR-AF data systems must align to NAFv4 to ensure coalition interoperability within the EUCOM AOR.
- **AJP-01, Allied Joint Doctrine** — Overarching Allied Joint Publication governing NATO joint operations.
- **AJP-3, Allied Joint Doctrine for the Conduct of Operations** — Establishes data sharing requirements for combined operations. Applicable to all USAREUR-AF combined exercise and operational data products.
- **AJP-3.2, Allied Joint Doctrine for Land Operations** — Establishes data sharing and interoperability requirements for combined land operations in the EUCOM AOR. Primary reference for USAREUR-AF/NATO land force data integration.
- **AJP-5, Allied Joint Doctrine for the Planning of Operations** — NATO planning doctrine, including multinational information requirements for combined planning.

USAREUR-AF RESOURCES

- **C2DAO** — The authoritative point of contact for USAREUR-AF operational data architecture, ontology design references, technical implementation guidance, NATO NAFv4 alignment resources, and the USAREUR-AF 5-Layer Data Stack implementation documentation. Contact your unit data steward or C2DAO directly for current resources.

DOD AND ARMY STRATEGIC REFERENCES

The following are strategic guidance documents — not doctrine — that inform MSS training design and operational context.

- **Army CIO Data Stewardship Policy (April 2, 2024)** — Establishes the data stewardship hierarchy (MADO, Data Steward, Functional Data Manager, C2DAO) and data chain of responsibility. Current authoritative governance reference.
- **Army Data Plan (2022)** — Established the foundational Army-wide framework for data management, governance, and analytics in support of Multi-Domain Operations. Superseded in part by subsequent Army CIO guidance (2024) and UDRA v1.1 (2025); remains a foundational reference.
- **Unified Data Reference Architecture (UDRA) v1.1 (February 2025)** — Provides the Army's current reference architecture for data systems based on data mesh principles: distributed data ownership, domain-aligned data products, and federated governance.
- **DoD Data Strategy (2020)** — Establishes the VAULTIS framework (Visible, Accessible, Understandable, Linked, Trustworthy, Interoperable, Secure) as the DoD standard for data quality. Supersedes the 5-dimension VAUTI model from AR 25-1.
- **DDOF Playbook v2.2 (December 2025)** — T2COM C2DAO / HQDA CIO/G-6 / SAIS-ADD implementing document. Extends VAULTIS to VAULTIS-A (8 dimensions, adds Auditable). Establishes the 6-phase data product lifecycle and 85% minimum quality gate. Authoritative standard for all MSS data products.
- **DoD Data, Analytics & AI Adoption Strategy (November 2023)** — Establishes the AI Hierarchy of Needs and the DoD framework for scaling data, analytics, and AI adoption across the enterprise.
- **NATO Data Strategy for the Alliance (Feb 2025)** — Alliance-wide data governance mandate establishing common data governance principles across NATO nations. Directly applicable to USAREUR-AF data literacy requirements in the EUCOM AOR.
- **CALL 25-10, Commander and Staff Guide to Data Literacy (April 2025)** — CALL handbook providing an accessible introduction to data literacy for commanders and staff. Covers data interpretation, analytical pitfalls, and data tool use within military contexts. Companion resource to this publication's senior leader edition.

- **Brito, Gary M. "Data Literacy: How We Prepare for the Future." *Military Review* Online Exclusive, January 2025.** — TRADOC CG's article establishing data literacy as an Army-wide readiness imperative.
- **TRADOC OCKO Data Literacy Training Portal** — TRADOC's central hub for data literacy training across the enterprise, maintained by the Command Chief Data and Analytics Office (C2DAO). Includes Data Literacy 101, Data Immersion Course for Knowledge Managers, and the KM Qualification Course. Available at tradoc.army.mil/ocko/training-portal/data-literacy/.

Access requires a valid CAC and USAREUR-AF network account. Contact your unit S6 or C2DAO for access provisioning.

GLOSSARY

Accuracy — The degree to which data correctly represents the real-world entity or event it describes.

Aggregation — The process of combining individual records into summary values (e.g., summing or averaging).

Aggregation Risk — The risk that combining individually unclassified pieces of information reveals a classified fact.

After-Action Review (AAR) — A structured review of an event or operation to identify what happened, why, and what should be sustained or improved.

Analytical Data — Data derived from master and transactional data to support decision-making (e.g., readiness percentages, trend lines).

API (Application Programming Interface) — A defined interface that allows two software systems to exchange data programmatically.

AOR (Area of Operations) — A defined geographic area within which a commander has the authority to conduct military operations.

Archive — Long-term storage of data that is no longer actively used but must be retained for legal, regulatory, or historical purposes.

Army Data Plan (2022) — The foundational Army-wide strategic framework for data management, governance, and analytics. Defines eleven strategic objectives (SO-01 through SO-11) and five strategic enablers (SE-01 through SE-05). SE-05 (Talent) is the mandate for MSS training. Superseded in part by Army CIO Data Stewardship Policy (2024) and UDRA v1.1 (2025) for governance and architecture specifics; strategic objectives remain authoritative. (See Chapter 12.)

Attribute Data — Non-spatial descriptive information associated with a geographic feature.

Automation — The use of technology to perform data processing tasks without human intervention.

BLUF (Bottom Line Up Front) — An Army writing and briefing principle directing that the most important conclusion or recommendation be stated first.

Causation — A relationship in which one variable directly produces a change in another.

Classification — The formal designation of information as Confidential, Secret, or Top Secret based on the harm its unauthorized disclosure would cause.

C2DAO (Command Chief Data and Analytics Officer) — The command-level official responsible for consuming and enforcing Army enterprise data policy within a command's AOR. The USAREUR-AF C2DAO operates at Tier 4 of the Army data stewardship hierarchy. C2DAOs do not create Army enterprise data policy; they implement it. (Authority: Army CIO Data Stewardship Policy, April 2024.)

COA (Course of Action) — A possible plan for accomplishing a mission, developed and compared during the MDMP.

Computational Governance — The automated enforcement of governance policies across data products and domains. Implements standards as code (executable quality rules) and policies as code (automated access control, classification, and retention enforcement). Replaces manual, document-based governance with machine-enforceable controls. One of the six UDRA services. (See paragraph 11-3.)

Completeness — The degree to which all required data is present in a dataset.

Consistency — The degree to which data agrees across systems, sources, and time periods.

Controlled Unclassified Information (CUI) — Unclassified information requiring safeguarding under law, regulation, or government-wide policy.

Coordinate System — A reference system for expressing geographic location as numerical values.

Correlation — A statistical relationship in which two variables tend to change together.

CSV (Comma-Separated Values) — A plain text file format for tabular data in which fields are separated by commas.

Data — Raw facts, observations, or measurements that have not yet been interpreted or contextualized.

Data Consumer — An individual, system, or organization that uses data to derive information or support decisions.

Data Culture — The shared values, behaviors, and practices around data within an organization.

Data Domain — An organization with specific functional expertise that produces data products. In the Army context, data domains align to staff functions (G1, G2, G3, G4, etc.) and subordinate commands. The domain owns and is accountable for the data it produces. (See paragraph 11-2.)

Data Custodian — An individual or organization with physical or technical custody of data.

Data Engineering — The technical discipline of building systems and pipelines to collect, store, transform, and move data.

Data Mesh — An architectural paradigm for data management based on distributed ownership, domain-aligned data products, federated governance, and self-serve data infrastructure. Adopted by the Army through the Unified Data Reference Architecture (UDRA) v1.1. Replaces centralized data architectures. (See Chapter 11.)

Data Lake — A storage repository that holds large volumes of raw data in its native format until needed for analysis.

Data Lifecycle — The stages through which data passes from creation to disposal: ingestion, storage, transformation, analysis, distribution, archiving, and disposition.

Data Lineage — The documented record of a dataset's origin and all transformations applied to it.

Data Literacy — The ability to read, understand, evaluate, and communicate using data.

Data Owner — The organizational element or leader accountable for a specific dataset's accuracy, security, and authorized use.

Data Producer — An individual, system, or organization that creates or collects data.

Data Product — A logically pre-packaged unit of data and associated metadata produced to satisfy a consumer's mission or business demand. Data products are self-describing (include metadata for content, schema, quality, lineage, and ownership) and computationally governed (quality and access policies enforced through automated mechanisms). (See paragraph 11-1.)

Data Profiling — The examination of a dataset to assess its content, structure, and quality.

Data Provenance — The documented origin of a dataset, including its source and collection method.

Data Quality — The composite of accuracy, completeness, consistency, timeliness, validity, and uniqueness of a dataset.

Data Science — The application of statistical and computational methods to extract knowledge from data.

Data Steward — An individual responsible for managing data on behalf of its owner.

Data Swamp — A data lake that has become unusable due to inadequate governance, poor organization, or low-quality data accumulation.

Data Validation — The process of confirming that data conforms to defined rules, formats, and ranges.

Data Verification — The process of confirming that data accurately represents the real-world entity it describes.

Data Warehouse — A repository optimized for historical analytical queries, integrating data from multiple source systems.

Decision Support Matrix (DSM) — A planning tool that links specific information triggers to commander decisions and available COA branches.

Decision Dominance — An enduring Army operational objective: the ability to make better decisions, faster, than the adversary across the full competition continuum. Data literacy is a prerequisite for decision dominance. Formations that cannot manage and analyze data effectively operate at a decision-cycle disadvantage. (See paragraph 1-2e.)

Descriptive Analytics — Analysis that summarizes what has happened.

Diagnostic Analytics — Analysis that identifies why something happened.

Distribution — The pattern of how values in a dataset are spread across their possible range.

DTED (Digital Terrain Elevation Data) — A standardized format for terrain elevation data.

Duplicate — A record that appears more than once in a dataset, representing the same real-world entity.

Enrichment — The process of adding contextual information from an additional data source to an existing dataset.

EXORD (Execute Order) — A command directive that authorizes and initiates execution of an operation.

FFIR (Friendly Force Information Requirement) — Information about friendly force status that the commander requires to make decisions.

Field — A single data attribute within a record; one column in a tabular dataset.

FMC (Fully Mission Capable) — Equipment status indicating an item is able to perform all of its designated missions.

FRAGORD (Fragmentary Order) — A fragmentary order that amends an existing operation order.

GeoJSON — A JSON-based geospatial data format for representing vector geographic features.

GEOINT (Geospatial Intelligence) — Intelligence derived from the exploitation and analysis of imagery and geospatial information.

GeoTIFF — A raster image format with embedded geographic coordinate information.

GPS (Global Positioning System) — A satellite-based navigation system providing position, navigation, and timing data.

HUMINT (Human Intelligence) — Intelligence derived from human sources.

Information — Data that has been processed and interpreted to convey meaning.

Ingestion — The process by which data enters a system.

IPB (Intelligence Preparation of the Battlefield) — The systematic process of analyzing the enemy, terrain, weather, and civil considerations in an AOR to support military planning.

IR (Information Requirement) — A question the commander needs answered to make a decision.

ISR (Intelligence, Surveillance, and Reconnaissance) — An integrated capability that synchronizes collection, processing, and exploitation of information to support operations.

Join — A database operation that combines records from two tables based on a shared key field.

Key — A field used to uniquely identify a record or to link records between two tables.

KML/KMZ (Keyhole Markup Language) — A geospatial file format used to display geographic data.

Knowledge — Information that has been contextualized and interpreted to support decision-making.

LOGSTAT (Logistics Status) — A standardized report of a unit's logistics status.

MADO (Mission Area Data Officer) — One of four Army-appointed officials responsible for setting enterprise data policy within a defined mission area (Warfighter, Intelligence, Business, or Enterprise IT). MADOs operate at Tier 1 of the Army data stewardship hierarchy and establish policy that binds all subordinate tiers. (Authority: Army CIO Data Stewardship Policy, April 2024.)

Master Data — Core reference data that other data refers to (e.g., personnel records, equipment registries).

MDMP (Military Decision-Making Process) — The Army's structured process for planning military operations.

Median — The middle value in a sorted dataset; resistant to the influence of outliers.

Metadata — Data that describes the content, context, quality, and provenance of a dataset.

MGRS (Military Grid Reference System) — The standard military coordinate system derived from UTM.

MOS (Military Occupational Specialty) — The Army's system for designating a Soldier's primary job.

MSN (Mission) — The task assigned to a unit or individual.

Need-to-Know — The standard requiring that access to classified information be limited to those whose duties require it.

Need-to-Share — The complementary principle recognizing the operational risk of withholding relevant information from those who need it.

NMC (Not Mission Capable) — Equipment status indicating an item cannot perform any of its designated missions.

Normalization — The process of converting data to a standard format or unit of measure.

OLAP (Online Analytical Processing) — A database architecture optimized for complex analytical queries across large datasets.

OLTP (Online Transaction Processing) — A database architecture optimized for rapid reading and writing of individual records.

OPDATA — Operational data; data generated in the context of military operations.

OPFOR (Opposing Force) — A simulated enemy force in training; may also refer to actual adversary forces.

OPORD (Operations Order) — A directive that provides specific instructions to subordinate units for the conduct of an operation.

Outlier — A value in a dataset that differs significantly from the bulk of the data.

Predictive Analytics — Analysis that forecasts what is likely to happen.

Prescriptive Analytics — Analysis that recommends a course of action.

PIR (Priority Intelligence Requirement) — An intelligence requirement that takes priority in collection and analysis because the commander needs the answer to make a critical decision.

PMC (Partially Mission Capable) — Equipment status indicating an item can perform some but not all of its designated missions.

Projection — A mathematical transformation for representing the curved surface of the Earth on a flat plane.

RBAC (Role-Based Access Control) — An access control model that grants permissions based on an individual's organizational role.

Record — A single complete entry in a dataset; one row in a tabular dataset.

Relational Database — A database that stores structured data in tables with defined relationships between them.

Retention — The period during which data must be preserved before it may be disposed of.

SAP (Special Access Program) — A classified program imposing access controls beyond standard classification levels.

Schema — The blueprint defining the structure, field types, and rules for a dataset.

SCI (Sensitive Compartmented Information) — Classified information requiring special access controls beyond standard classification.

Semi-Structured Data — Data with some organizational elements but without a rigid schema.

Shapefile — A vector geospatial format consisting of multiple component files.

SITREP (Situation Report) — A standardized report of current operational situation.

Skewed Distribution — A distribution where most values cluster at one end with a tail at the other.

Spillage — A security incident in which classified information is processed on or transmitted through a system not authorized for that classification level.

Standard Deviation — A statistical measure of the dispersion of values around the mean.

Structured Data — Data organized in a defined schema with rows and columns.

Timeliness — The degree to which data is sufficiently current for its intended use.

Transactional Data — Data recording events and activities within a system (e.g., work orders, requisitions).

Transformation — A process applied to data to convert it from one form to another.

Uniqueness — The degree to which each real-world entity appears exactly once in a dataset.

UDRA (Unified Data Reference Architecture) v1.1 — The Army's current reference architecture for data systems (February 2025), issued by the Army Chief Information Officer. Based on data mesh principles: distributed data ownership, domain-aligned data products, and federated governance. The authoritative Army architecture reference; supersedes prior centralized data architecture guidance. Implemented at USAREUR-AF level through C2DAO.

Unstructured Data — Data with no predefined format or schema.

USR (Unit Status Report) — A standardized report of a unit's overall readiness.

UTC (Coordinated Universal Time) — The primary time standard used for military and scientific timekeeping.

UTM (Universal Transverse Mercator) — A coordinate system dividing the Earth into longitudinal zones.

Validity — The degree to which data values conform to defined rules, formats, and ranges.

VAUTI (Visible, Accessible, Understandable, Trustable, Interoperable) — The original five-dimension data quality framework from AR 25-1 (2019). **Superseded** by VAULTIS (DoD Data Strategy 2020, 7 dimensions) and VAULTIS-A (DDOF Playbook v2.2, 8 dimensions). See VAULTIS-A.

VAULTIS (Visible, Accessible, Understandable, Linked, Trustworthy, Interoperable, Secure) — The seven-dimension data quality framework established by the DoD Data Strategy (2020). Supersedes VAUTI. Extended to VAULTIS-A by the DDOF Playbook v2.2. See VAULTIS-A.

VAULTIS-A (Visible, Accessible, Understandable, Linked, Trusted, Interoperable, Secure, Auditable) — The eight-dimension data quality framework established by the DDOF Playbook v2.2 (T2COM C2DAO, December 2025). Extends DoD VAULTIS by adding Auditable. All USAREUR-AF data products must achieve 85% minimum weighted average across all eight dimensions to pass DDOF Phase 3 quality gate. This is the current authoritative standard. (See paragraph 4-8.)

Visualization — The representation of data in a visual form, such as a chart, graph, or map.

WARNO (Warning Order) — A preliminary notice of an order or action that is to follow.

5-Layer Data Stack — The USAREUR-AF standard architecture for operational data systems, organizing data capabilities across Infrastructure, Integration, Semantic, Analytics, and Activation layers.

Cross-Domain Architecture (CDA) — The USAREUR-AF enterprise architecture framework governing data systems design, ontology standards, and NAFv4 alignment. Implementation guidance available through C2DAO.

Information Advantage — As defined in ADP 3-13, the operational condition achieved when a force can generate, protect, and exploit information more effectively than the adversary. Data literacy is a prerequisite for information advantage.

JADC2 (Joint All-Domain Command and Control) — The DoD concept for connecting sensors, shooters, and decision-makers across all domains to enable faster, better-informed command decisions.

NAFv4 (NATO Architecture Framework version 4) — The NATO standard for describing systems, capabilities, and data architectures to ensure coalition interoperability.

USAREUR-AF (United States Army Europe and Africa) — The Army Service Component Command (ASCC) to USEUCOM, responsible for theater land operations across the European and African portions of the USEUCOM AOR.

USEUCOM (United States European Command) — The Combatant Command responsible for military operations across Europe, portions of Asia, the Arctic, and the Atlantic Ocean.

Data Literacy Technical Reference Headquarters, United States Army Europe and Africa (USAREUR-AF) Wiesbaden, Germany 2026

By order of the Commanding General, United States Army Europe and Africa.

DISTRIBUTION RESTRICTION: Distribution authorized to U.S. Government agencies and their contractors only. Other requests must be referred to Headquarters, USAREUR-AF, C2DAO, Wiesbaden, Germany.