**White Paper: Data Architecture & Governance**

Foundational Practices for Veteran-Centric Systems (VA Context)

## Executive Summary

In digital government and healthcare systems—particularly within the U.S. Department of Veterans Affairs (VA)—effective data architecture and governance are essential for ensuring data integrity, interoperability, and user-centric service delivery. This paper outlines the core domains of enterprise data management, explores how these are applied in the VA’s ecosystem, and references leading resources such as DAMA DMBOK2, Data.gov, and internal VA governance frameworks.

1. Core Areas of Data Architecture and Governance

## 1.1 Data Modeling

- Purpose: Define the logical structure and meaning of data to support consistent storage, querying, and transformation.

- Types:

- Conceptual: Captures business entities (e.g., Veteran, Claim) and their relationships.

- Logical: Translates business definitions into attribute-level models.

- Physical: Implements data structures in DBMS systems like Oracle or Postgres.

- Best Practices:

- Adopt canonical models to ensure consistency.

- Maintain centralized data dictionaries.

- Normalize for consistency; denormalize for performance where needed.

## 1.2 Metadata Management

- Definition: Governance of metadata—information about data such as structure, lineage, and ownership.

- Types:

- Technical: Schemas, transformations, lineage.

- Business: Glossaries, definitions, data owners.

- Operational: Update timestamps, refresh schedules, audit events.

- Tools:

- Alation, Collibra, Informatica EDC.

- AWS Glue Catalog for automatic schema discovery and tagging.

## 1.3 Master Data Management (MDM)

- Purpose: Provide a single, authoritative view of shared entities (Veterans, providers, locations).

- Capabilities:

- Entity resolution and survivorship.

- Golden record creation and publication.

- Stewardship and data quality enforcement.

- Government Domain Examples:

- Veteran identities, provider rosters, VA facilities.

- Architecture Models:

- Registry: Resolves to source, no persistence.

- Consolidation: Golden record is persisted.

- Coexistence/Transactional: Bidirectional sync.

2. VA-Specific Context: Veteran-Centric Data Strategy

## 2.1 Veteran-Centric Modeling

- Approach: Design systems and schemas that place the Veteran at the center.

- Use Cases:

- Link claims (BGS) with health data (Cerner, VistA) and identity (MPI, VA Profile).

- Identifiers:

- EDIPI, ICN, SSN, VA File Number.

- MPI as source-of-truth for identity resolution.

- Challenges:

- Conflicts in IDs.

- Maintaining longitudinal data continuity.

## 2.2 Interoperability Across Systems

- Exchange Patterns:

- MPI ↔ VA Profile ↔ Corp DB ↔ BGS

- VA ↔ DOD ↔ CMS ↔ States

- Standards:

- HL7 FHIR (health), NIEM (interagency), JSON/XML APIs.

- Integration Architecture:

- ESB and API Gateway layers used to decouple systems.

## 2.3 Data Governance at the VA

- Governance Council:

- Defines data policies and taxonomy.

- Assigns stewards and review boards.

- Artifacts:

- DART, DMPs, and DQRs document access, lineage, and quality.

- VA MDM Goals:

- Prevent data duplication.

- Secure interoperability with DOD and CMS.

- Comply with HIPAA and 38 CFR §1.500.

3. Tools and Best Practice Frameworks

## 3.1 DAMA DMBOK2

- Description: Data Management Association's 11-domain framework for enterprise data management.

- Domains:

1. Governance

2. Architecture

3. Modeling

4. Storage

5. Security

6. Integration

7. Content Management

8. Reference/MDM

9. BI/Data Warehousing

10. Metadata

11. Quality

## 3.2 Data.gov

- Purpose: Public data catalog for U.S. government.

- VA Usage:

- Schema standardization.

- Integration for public research and transparency.

## 3.3 VA Technical Reference Model (TRM)

- Function: Provides VA-approved technology stack.

- Usage:

- Ensures compliance.

- Reference for architects and developers.

4. Architecture Example: Veteran MDM System

## Scenario: Create MDM service to reconcile identity data across VA systems.

- Step 1: Ingest records via Kafka, file drop, or REST.

- Step 2: Match records using fuzzy matching algorithms.

- Step 3: Generate golden records and internal GUIDs.

- Step 4: Publish updates via message bus.

- Step 5: Track lineage and steward assignments.

## 5. Resources and Reference Materials

| Resource | Description |

|---------------------------|-------------------------------------------------------|

| DAMA DMBOK2 | Industry standard for data governance |

| Data.gov | Federal schema/catalog portal |

| VA TRM | Technical compliance and architecture registry |

| NIEM | Interagency data sharing framework |

| HL7 FHIR | Health data interoperability standard |

| Collibra, Alation, Informatica | Tools for metadata and master data governance |

## Conclusion

Veteran-centric delivery requires VA systems to prioritize secure, high-quality, and standardized data architecture. With MDM, governance, and metadata management, VA can improve accuracy, enhance interoperability, and streamline services.

## Next Steps:

- Finalize data domain definitions and stewards.

- Normalize identifiers and implement MPI integration.

- Deploy metadata discovery and stewardship tooling.