

Archetype Core designs and delivers governed data platforms for regulated environments. The company combines Databricks and Unity Catalog architecture with production data engineering to modernize pipelines, govern access, tune performance and cost, and build audit-ready foundations for analytics and AI. Archetype Core is led by a data architect with nearly 10 years of software and data engineering experience, including nearly six years building and operating federal data systems on USCIS programs across identity, verification, and analytics.

DATABRICKS PARTNER | DATABRICKS CERTIFIED DATA ENGINEER PROFESSIONAL | SMALL BUSINESS, WOMAN-OWNED

Auditability built in. Lineage, quality gates, and audit records are part of the architecture from the start.

Senior technical delivery. Architecture leadership continues through implementation and production readiness.

Databricks specialization. Professional certification backed by working Databricks and Unity Catalog reference systems.

CORE CAPABILITIES

Unity Catalog and Governance Architecture

Metastore, catalog and schema architecture, privilege and ownership models, workspace bindings, governed tags and ABAC policy design for row filtering and column masking, with table-level filters where per-table logic is required, lineage and audit through system tables, external locations and storage credentials, and Hive metastore to Unity Catalog migration.

Lakehouse, Medallion, and Performance

Delta Lake and medallion design across bronze, silver, and gold layers, PySpark and SQL, Lakeflow batch and streaming pipelines, incremental and idempotent MERGE patterns, file layout and compaction, partitioning and clustering, query tuning, and warehouse sizing and cost control.

Data Engineering and Modernization

CDC, ETL/ELT, multi-source ingestion, data modeling, dbt models, migration, schema evolution, orchestration, data quality, and observability, delivered on AWS with S3, DMS, Glue, EMR, Redshift, EKS, OpenSearch, and Terraform.

Governed AI and Retrieval

Audit-ready RAG, source-grounded retrieval with citations, quality gates, confidence signals, model and retrieval metadata, and query-level audit records.

HOW WE ENGAGE

Archetype Core supports prime and subcontract teams on data modernization, Databricks and Unity Catalog implementation, data architecture, governance, assessments, and targeted prototypes. Senior architecture support can also be embedded within an established delivery team.

COMPANY DATA

POINT OF CONTACT

Mariana Lungu, Founder and Data Architect
mariana@archetypecore.com
571-510-0194

LEGAL NAME

Lungu Holdings, LLC DBA Archetype Core

UEI / CAGE

WBQJP6JPZFA3 / 20R49

ADDRESS

277 S Washington St, Ste 210 PMB 1137
Alexandria, VA 22314-3672

SAM.GOV / SIZE

Active and registered
Small business, woman-owned

NAICS

541512 (primary)
541511, 541519, 518210

PARTNER / CERTIFICATIONS

Databricks Partner
Databricks Certified
Data Engineer Professional
AWS Certified Solutions
Architect - Professional
CompTIA Security+
Microsoft Certified:
Azure Fundamentals (AZ-900)
Power Platform Fundamentals (PL-900)

FEDERAL BACKGROUND

Previous DHS Public Trust
Favorably adjudicated December 2024

FOUNDER DELIVERY RECORD

Verification analytics platform. Designed the data architecture for DMS and Glue ingestion into S3, dbt medallion models in Redshift, EKS batch processing, encryption and audit logging, and Terraform deployment, serving downstream Databricks workloads and Tableau dashboards.

Six-source identity pipeline. Multi-source CDC into Oracle, Spark on ephemeral EMR, and OpenSearch; 99%+ data accuracy and replication reduced from hours to minutes.

Architecture reach. Defined integration, access control, partitioning, and monitoring patterns used by 10 or more engineers across multiple teams and supporting 3 or more downstream applications and analytics consumers.

Performance and cost. Improved audit-query performance by 50% through partitioning and targeted Oracle indexing, with query tuning across Oracle and Redshift, and reduced storage cost by 30% on a 1 TB+ structured and semi-structured migration using AWS DMS.

Delivered by the founder on USCIS programs under prior primes. Shown as an individual delivery record, not corporate past performance.

REFERENCE SYSTEMS

Governed Databricks pipeline and Unity Catalog

Airflow on MWAA orchestrating PySpark into Databricks Delta Lake under a dedicated Unity Catalog catalog and schema, with bronze and gold Delta tables, idempotent Delta MERGE, run-level identifiers, per-record lineage, Great Expectations quality gates, dead-letter handling, a PostgreSQL audit trail, Terraform, and SHA-256 hashing of inputs and prompts.

Audit-ready RAG system

9,116 indexed policy chunks, 44 automated tests, citation-grounded answers, confidence tiers, model and retrieval metadata, and a 16-field query-level audit record.

Reference implementations built and operated by Archetype Core. Not client deployments.