Certificate in Master Data Migration

Data Security and Risk Management

Access Control – Concept: Mechanisms that restrict who can view or use resources. Related terms: authentication, authorization, role-based access control. Explanation: Access control enforces policies that determine permissible actions for users, devices, or processes. Example: A data migration tool grants read-only rights to analysts but edit rights to data stewards. Practical application: Implementing ACLs on migration repositories to prevent unauthorized changes. Challenges: Balancing security with usability, especially when multiple teams need differing levels of access during migration phases.

Authentication – Concept: Process of verifying identity. Related terms: multi-factor authentication, credential, single sign-on. Explanation: Authentication confirms that a user or system is who it claims to be, typically using passwords, tokens, or biometrics. Example: A migration engineer logs into the source database using a password plus a time-based one-time code. Practical application: Enforcing MFA for all privileged accounts that control master data loads. Challenges: Managing credential sprawl and user resistance to additional steps.

Authorization – Concept: Granting permission to perform actions. Related terms: access control, role, privilege. Explanation: After authentication, authorization determines which resources a user may access and what operations they may perform. Example: A data steward is authorized to approve data quality rules but not to delete source tables. Practical application: Configuring role-based policies that align with migration governance. Challenges: Keeping authorization matrices up-to-date as roles evolve during migration.

Audit Trail – Concept: Record of system activities. Related terms: logging, compliance, forensic analysis. Explanation: An audit trail captures who did what, when, and where, providing traceability for data migration actions. Example: Every load job writes entries showing the user, timestamp, and number of records processed. Practical application: Using audit logs to satisfy regulatory requirements such as GDPR. Challenges: Ensuring logs are tamper-proof and retained for the required period without overwhelming storage.

Availability – Concept: Ensuring data and services are accessible when needed. Related terms: uptime, redundancy, disaster recovery. Explanation: Availability measures the proportion of time systems are operational, critical for uninterrupted migration pipelines. Example: Deploying load-balanced migration services across two data centers to avoid single-point failures. Practical application: Monitoring service health to trigger failover during a migration window. Challenges: Balancing high availability with cost and complexity, especially in cloud environments.

Backup – Concept: Copy of data for restoration. Related terms: snapshot, recovery point objective, archival. Explanation: Backups protect against data loss by preserving a version of source or target datasets before migration. Example: Taking a full dump of the legacy master data before initiating incremental loads. Practical application: Scheduling nightly backups of migration staging tables. Challenges: Managing backup

size, retention policies, and ensuring backup integrity.

Baseline – Concept: Reference point for performance or security. Related terms: benchmark, control, comparison. Explanation: A baseline establishes the normal state of a system against which deviations are measured. Example: Recording average data transfer rates before migration to detect anomalies. Practical application: Using baseline metrics to identify potential security breaches during migration. Challenges: Defining accurate baselines in dynamic environments.

Business Continuity – Concept: Ability to maintain essential functions during disruption. Related terms: disaster recovery, resilience, contingency planning. Explanation: Business continuity planning ensures that critical data migration activities can continue despite incidents. Example: Having a secondary migration environment ready in case the primary site loses power. Practical application: Drafting run-books that outline steps for switching to backup infrastructure. Challenges: Coordinating across departments and keeping plans current with evolving migration scopes.

Cloud Security – Concept: Protecting data and services hosted in cloud platforms. Related terms: shared responsibility model, encryption-in-transit, identity-as-a-service. Explanation: Cloud security addresses threats specific to virtualized resources, such as misconfigured storage buckets. Example: Enforcing server-side encryption for data stored in an AWS S3 bucket used as a migration landing zone. Practical application: Applying cloud-native IAM policies to restrict access to migration workloads. Challenges: Navigating differing security controls across multiple cloud providers.

Confidentiality – Concept: Preventing unauthorized disclosure of information. Related terms: privacy, data classification, encryption. Explanation: Confidentiality safeguards sensitive master data from exposure during migration. Example: Masking personally identifiable information (PII) before loading it into a test environment. Practical application: Using column-level encryption for credit-card numbers during transit. Challenges: Balancing confidentiality with the need for data visibility for quality checks.

Data Classification – Concept: Categorizing data based on sensitivity and value. Related terms: confidentiality, data labeling, risk assessment. Explanation: Classification informs the security controls applied to each data set throughout migration. Example: Tagging customer address records as “high sensitivity” and applying stricter access controls. Practical application: Automating classification rules in the migration tool to enforce encryption for designated categories. Challenges: Maintaining consistent classification across legacy systems with divergent schemas.

Data Encryption – Concept: Transforming data into unreadable form without a key. Related terms: symmetric encryption, asymmetric encryption, key management. Explanation: Encryption protects data at rest and in transit, ensuring confidentiality even if intercepted. Example: Using AES-256 to encrypt CSV files before transferring them to the target data warehouse. Practical application: Enabling TLS for all migration APIs. Challenges: Managing encryption keys securely and rotating them without disrupting migration jobs.

Data Governance – Concept: Framework for managing data availability, usability, integrity, and security. Related terms: data stewardship, policy, compliance. Explanation: Governance defines roles, responsibilities, and processes that guide master data migration. Example: A data governance board approves the migration

schedule and validates quality metrics. Practical application: Embedding governance checkpoints into the migration workflow. Challenges: Aligning governance with agile migration timelines and cross-functional teams.

Data Masking – Concept: Obscuring sensitive data while preserving format. Related terms: de-identification, tokenization, pseudonymization. Explanation: Masking replaces real values with fictional ones to protect privacy in non-production environments. Example: Replacing real social security numbers with random but correctly formatted numbers for testing. Practical application: Applying dynamic data masking on source databases during extraction. Challenges: Ensuring masked data remains realistic enough for functional testing.

Data Migration – Concept: Moving data from one system to another. Related terms: extraction, transformation, load (ETL), cut-over. Explanation: Migration transfers master data assets while preserving integrity, often involving complex mappings. Example: Migrating product master data from an on-prem ERP to a cloud-based MDM platform. Practical application: Using automated pipelines to orchestrate incremental loads. Challenges: Managing data volume, downtime constraints, and unforeseen schema mismatches.

Data Privacy – Concept: Protecting personal information from misuse. Related terms: GDPR, CCPA, consent, anonymization. Explanation: Privacy regulations dictate how personal data must be handled during migration. Example: Obtaining consent before moving customer emails to a new marketing system. Practical application: Conducting privacy impact assessments before initiating migration. Challenges: Interpreting diverse legal requirements across jurisdictions.

Data Retention – Concept: Policies governing how long data is kept. Related terms: archival, disposal, compliance. Explanation: Retention rules determine the lifespan of migrated data and influence backup strategies. Example: Retaining transaction logs for seven years to satisfy financial regulations. Practical application: Configuring automated purge jobs after the retention period expires. Challenges: Balancing legal obligations with storage cost constraints.

Data Stewardship – Concept: Accountability for data quality and governance. Related terms: data owner, data custodian, stewardship. Explanation: Stewards oversee the accuracy, consistency, and security of master data throughout migration. Example: A product data steward validates that attribute values conform to the target taxonomy before load. Practical application: Assigning stewardship responsibilities in the migration project charter. Challenges: Securing sufficient resources and authority for stewards to enforce standards.

Data Transfer – Concept: Moving data between locations. Related terms: bandwidth, protocol, latency. Explanation: Transfer mechanisms affect migration speed and security. Example: Using SFTP with SSH keys to move large CSV files from a legacy server to a cloud bucket. Practical application: Scheduling transfers during off-peak hours to reduce network impact. Challenges: Managing network throttling and ensuring transfer integrity.

Data Validation – Concept: Checking data accuracy and completeness. Related terms: data quality, reconciliation, profiling. Explanation: Validation ensures that migrated data matches source expectations

and business rules. Example: Comparing row counts and checksum values between source and target tables after load. Practical application: Automating validation scripts as part of the migration pipeline. Challenges: Detecting subtle data anomalies that may not be caught by simple counts.

Data Warehouse – Concept: Central repository for analytical reporting. Related terms: OLAP, dimensional modeling, ETL. Explanation: Migration often involves moving master data into a data warehouse for downstream analytics. Example: Loading customer master records into a star schema for segmentation analysis. Practical application: Designing staging areas that enforce security before data reaches the warehouse. Challenges: Aligning warehouse security with enterprise policies while preserving performance.

De-identification – Concept: Removing or obscuring personal identifiers. Related terms: anonymization, pseudonymization, data masking. Explanation: De-identification reduces privacy risk by eliminating direct identifiers. Example: Stripping email addresses from a dataset used for statistical modeling. Practical application: Applying de-identification scripts during the extraction phase. Challenges: Maintaining data utility after removal of identifying fields.

Disaster Recovery – Concept: Restoring systems after a catastrophic event. Related terms: business continuity, RPO, RTO. Explanation: DR plans specify how migration environments are rebuilt after failure. Example: Restoring a failed migration server from a snapshot within the defined recovery time objective. Practical application: Conducting regular DR drills that simulate loss of the primary migration hub. Challenges: Ensuring DR sites have up-to-date configurations and data.

Encryption Key Management – Concept: Lifecycle handling of cryptographic keys. Related terms: key vault, rotation, escrow. Explanation: Effective key management is essential to maintain encryption security during migration. Example: Storing AES keys in a cloud-based key management service and rotating them monthly. Practical application: Integrating key retrieval APIs into the migration workflow. Challenges: Preventing unauthorized key access while avoiding operational bottlenecks.

Ethical Hacking – Concept: Authorized testing of security controls. Related terms: penetration testing, red team, vulnerability assessment. Explanation: Ethical hackers assess migration infrastructure for weaknesses before go-live. Example: Simulating a phishing attack to test employee awareness of migration-related credential handling. Practical application: Scheduling a pre-migration penetration test to uncover misconfigurations. Challenges: Coordinating testing without disrupting migration timelines.

GDPR – Concept: European Union regulation on data protection. Related terms: data subject rights, lawful basis, DPO. Explanation: GDPR imposes strict rules on processing personal data, influencing migration practices. Example: Documenting the legal basis for transferring EU customer records to a new CRM. Practical application: Conducting data mapping exercises to identify GDPR-covered fields before migration. Challenges: Interpreting ambiguous provisions and ensuring cross-border compliance.

Identity Management – Concept: Administration of user identities and access rights. Related terms: IAM, single sign-on, provisioning. Explanation: Identity management systems control who can initiate migration jobs and access data stores. Example: Using Azure AD to provision migration service accounts with limited permissions. Practical application: Automating role assignments based on group membership. Challenges:

Synchronizing identities across on-prem and cloud domains.

Incident Response – Concept: Structured approach to handling security events. Related terms: playbook, containment, forensic analysis. Explanation: Incident response defines steps to mitigate breaches that may arise during migration. Example: Isolating a compromised migration server and revoking its credentials. Practical application: Maintaining an incident response run-book that includes migration-specific contact lists. Challenges: Detecting incidents quickly in high-throughput migration pipelines.

Integrity – Concept: Assurance that data remains accurate and unaltered. Related terms: checksum, hash, tamper-evidence. Explanation: Integrity checks verify that master data has not been corrupted during migration. Example: Generating SHA-256 hashes for source files and confirming them after transfer. Practical application: Embedding hash verification into ETL jobs. Challenges: Managing hash collisions and ensuring performance does not degrade.

Least Privilege – Concept: Granting only the minimum permissions required. Related terms: role-based access control, separation of duties, privilege escalation. Explanation: Applying least privilege limits exposure if an account is compromised. Example: A migration script runs under a service account that can only insert into target tables, not delete. Practical application: Auditing permissions quarterly to enforce least-privilege principles. Challenges: Identifying the exact set of rights needed for complex migration tasks.

Logging – Concept: Recording system events for monitoring and analysis. Related terms: audit trail, log aggregation, SIEM. Explanation: Logs provide visibility into migration activities and support troubleshooting. Example: Capturing start and end timestamps for each data load job. Practical application: Forwarding logs to a centralized SIEM for real-time alerting. Challenges: Filtering noise from large volumes of log data.

Malware – Concept: Malicious software designed to disrupt or steal data. Related terms: ransomware, virus, trojan. Explanation: Malware can infiltrate migration environments, compromising data security. Example: A ransomware attack encrypts staging files, halting the migration pipeline. Practical application: Deploying endpoint protection on all migration servers. Challenges: Keeping defenses updated against evolving threats.

Multi-factor Authentication – Concept: Using two or more verification methods. Related terms: MFA, OTP, hardware token. Explanation: MFA strengthens authentication by requiring additional factors beyond passwords. Example: Requiring a push notification approval on a mobile device after entering a password. Practical application: Enforcing MFA for all accounts that can modify migration configurations. Challenges: Managing device enrollment and handling lost factors.

Network Segmentation – Concept: Dividing a network into isolated zones. Related terms: VLAN, DMZ, micro-segmentation. Explanation: Segmentation limits lateral movement of attackers within migration infrastructure. Example: Placing the migration staging area in a separate subnet with restricted inbound traffic. Practical application: Using firewalls to enforce strict rules between segmentation zones. Challenges: Configuring and maintaining segmentation policies without hindering data flow.

Non-repudiation – Concept: Preventing denial of actions performed. Related terms: digital signature, audit

log, accountability. Explanation: Non-repudiation provides proof that a specific user executed a migration operation. Example: A digitally signed commit log shows that the data engineer approved a load batch. Practical application: Implementing PKI-based signatures on migration scripts. Challenges: Managing certificate lifecycles and ensuring widespread adoption.

Penetration Testing – Concept: Simulated attack to uncover vulnerabilities. Related terms: ethical hacking, red team, exploit. Explanation: Pen tests assess the robustness of migration platforms before production use. Example: Attempting SQL injection on a migration API endpoint to test input validation. Practical application: Scheduling a test after major configuration changes. Challenges: Coordinating with migration teams to avoid service disruption.

Personal Identifiable Information – Concept: Data that can identify an individual. Related terms: PII, sensitive data, privacy. Explanation: PII requires special handling to comply with privacy laws during migration. Example: Email addresses, phone numbers, and national IDs are considered PII. Practical application: Applying encryption and masking to PII fields before loading into a test environment. Challenges: Identifying all PII elements in legacy schemas.

Risk Assessment – Concept: Systematic identification of threats and vulnerabilities. Related terms: risk register, threat modeling, impact analysis. Explanation: Assessment quantifies the likelihood and impact of security events affecting migration. Example: Evaluating the risk of data leakage when using public cloud storage for interim files. Practical application: Documenting findings in a risk register and assigning mitigation actions. Challenges: Keeping assessments current as migration scope evolves.

Risk Management – Concept: Process of controlling identified risks. Related terms: mitigation, acceptance, transfer. Explanation: Risk management implements controls to reduce or monitor threats throughout migration. Example: Deploying encryption to mitigate the risk of data interception. Practical application: Reviewing risk treatment plans during each migration phase gate. Challenges: Prioritizing limited resources among many identified risks.

Role-based Access Control – Concept: Permissions assigned to roles rather than individuals. Related terms: RBAC, least privilege, role hierarchy. Explanation: RBAC simplifies management by grouping users with similar responsibilities. Example: Assigning the “Data Migration Operator” role permission to execute load jobs but not alter configurations. Practical application: Mapping organizational roles to RBAC groups in the migration IAM system. Challenges: Ensuring roles accurately reflect evolving job functions.

Security Architecture – Concept: Design of security controls within an IT system. Related terms: defense in depth, security zones, framework. Explanation: Architecture defines how authentication, encryption, monitoring, and other controls interoperate for migration. Example: Designing a layered security model where data is encrypted, access is mediated by an API gateway, and logs feed a SIEM. Practical application: Reviewing architecture diagrams during migration planning. Challenges: Integrating legacy components that lack modern security features.

Security Incident – Concept: Event that compromises confidentiality, integrity, or availability. Related terms: breach, alert, response. Explanation: Incidents may arise from misconfigurations, insider actions, or external

attacks during migration. Example: Detecting unauthorized access to a staging database through anomalous login patterns. Practical application: Triggering automated containment scripts when an incident is identified. Challenges: Distinguishing false positives from genuine threats in high-volume environments.

Security Policy – Concept: Formal statement of security objectives and rules. Related terms: governance, compliance, standard. Explanation: Policies guide behavior and technical controls for migration activities. Example: A policy that mandates encryption of all data in transit during migration. Practical application: Disseminating the policy to all migration team members and enforcing compliance via audits. Challenges: Keeping policies relevant to emerging technologies used in migration.

Security Token – Concept: Digital credential used for authentication. Related terms: JWT, OAuth, SAML. Explanation: Tokens enable stateless, scalable authentication for migration services. Example: A JWT issued by an identity provider that the migration API validates on each request. Practical application: Configuring short-lived tokens to reduce exposure. Challenges: Safeguarding token secrets and handling token revocation.

Sensitive Data – Concept: Information that, if disclosed, could cause harm. Related terms: PII, PHI, confidential. Explanation: Sensitive data includes financial, health, or proprietary information requiring heightened protection. Example: Credit card numbers stored in a payment master table. Practical application: Applying field-level encryption for such columns during migration. Challenges: Accurately inventorying all sensitive elements across heterogeneous sources.

Threat Modeling – Concept: Structured analysis of potential attack vectors. Related terms: STRIDE, attack surface, risk assessment. Explanation: Modeling helps anticipate how adversaries might compromise migration processes. Example: Identifying that an unsecured FTP server used for interim file transfers is a high-risk component. Practical application: Documenting mitigations such as switching to SFTP with key authentication. Challenges: Keeping the model updated as architecture changes.

Vulnerability Management – Concept: Ongoing process of identifying, prioritizing, and remediating vulnerabilities. Related terms: patching, CVE, remediation. Explanation: Effective vulnerability management reduces the attack surface of migration environments. Example: Scanning migration hosts for known OpenSSL vulnerabilities and applying patches. Practical application: Integrating vulnerability scans into CI/CD pipelines for migration tools. Challenges: Balancing rapid migration schedules with the need for timely patching.

Zero-Trust Architecture – Concept: Security model that assumes no implicit trust. Related terms: micro-segmentation, continuous verification, least privilege. Explanation: Zero-trust requires verification for every access attempt, even within the same network. Example: Requiring authentication for each API call made by migration scripts, regardless of network location. Practical application: Deploying identity-aware proxies that enforce policy on every request. Challenges: Overhauling legacy systems that rely on perimeter-based security assumptions.