
ENKIDU

MIOVSKI GROUP

undefined

Glossary

Core terms used throughout this reference architecture. Each names a structural concept rather than a product or a sector.

Master data The shared, slowly-changing "nouns" an organization operates on — parties, assets, locations, products, accounts — as distinct from the transactional events that reference them.

Master entity The specific noun a given deployment masters first. This document uses a generic party as its running example; section 1.3 shows how the same model instantiates per sector.

Golden record The single authoritative, composite record for a master entity — assembled attribute-by-attribute from the best available source. A source of trusted data, not merely a single source of truth.

Contributing system Any operational system that holds its own version of the master entity and feeds the hub, classified abstractly (systems of entry, of record, of engagement, of scheduling/operations) rather than by product category.

Coexistence The MDM implementation style adopted here: the hub masters and cleanses the entity and stays bidirectionally synchronized with the contributing systems, which keep operating as their owners' systems of entry.

Bronze / Silver / Gold The three durable, separately-governed zones of the ingestion pipeline: raw and immutable (Bronze); working and cleansing (Silver); trusted and finished (Gold).

Survivorship The rules that select the best value per attribute when contributing systems disagree.

Identity resolution Recognizing that several source records describe the same real-world master entity.

Provenance / Chain of custody Respectively: where a value came from (which source contributed the winning value), and the ordered, recorded sequence of automated and human decisions that produced it.

HITL Human-in-the-loop — a steward decides where automation should not.

XREF The cross-reference mapping the golden ID to every source-system key.

Master Data Warehouse (DWH) Part 2 — the analytics-facing consolidation layer that conforms contributing-system data to MDM golden records and reference data.

Conformed dimension A golden entity (and its reference data) used consistently across the warehouse so that facts share one entity view.

Holding Zone Where DWH records for an unrecognized entity are parked until MDM produces a golden record.

Abstract

Every organization of scale runs multiple systems that each hold their own version of the same core entities — the same customer, citizen, patient, soldier, asset, or supplier appearing in several

systems under different identifiers, spellings, and status values. That fragmentation produces duplicate records, unreliable reporting, compliance exposure, and integration logic that grows combinatorially. The problem is universal; only the entity's name changes by sector.

This document describes the target-state architecture for an Enterprise Master Data Platform in two parts, deliberately written to be sector-independent. Part 1 — Master Data Management (MDM) maintains authoritative master data for a chosen entity so that operational systems interoperate against one agreed golden record. Part 2 — Master Data Warehouse (DWH) ingests data from the same contributing systems, consolidates it against those golden records and reference data, and serves it denormalized to analytics and AI. Both use the same Bronze -> Silver -> Gold staging discipline.

The architecture rests on four load-bearing decisions — a multi-stage ingestion pipeline, explicit named subsystems, a provenance-and-chain-of-custody capability, and an automated-plus-human-in-the-loop curation engine — and on a chosen coexistence implementation style that balances authority with acceptable change to source systems. Its trust posture — immutable raw data, attribute-level provenance, and a defensible chain of custody for every decision — is the same discipline MIOVSKI GROUP builds into the Origi trust framework, applied here to master data.

You cannot integrate what you cannot identify.

This platform makes the organization's core entities identifiable, trusted, and shareable — whatever the sector calls them.

1. Context and Problem Statement

1.1 The universal problem

An organization runs several systems that each hold their own version of the same core entities. The same entity appears in each with different identifiers, spellings, addresses, and status values — full name in one system, an initial-and-surname in another, a stale address in a third. That fragmentation produces duplicate records, unreliable reporting, compliance exposure, and integration logic that grows combinatorially as every system is wired to every other.

Master data is the shared, slowly-changing "nouns" the business operates on — distinct from the transactional events that reference them. The MDM hub exists to produce and maintain the authoritative representation of each master entity — the golden record — and to broker it so that source systems integrate against one trusted record instead of against each other.

1.2 A sector-neutral model

To keep the architecture independent of any one domain, contributing systems are described by their role, not by product category. Any real deployment maps its actual systems onto these roles:

System of entry

Where the entity is first created and owned. Authoritative for existence and core identity.

System of record

Where the entity's canonical operational data lives. Authoritative for specific attribute sets.

System of engagement

Where the entity is interacted with day-to-day. Contributes contact, activity, and status data.

System of scheduling / operations

Where the entity is deployed, rostered, or operated. Contributes operational identity and state.

Downstream consumer

Analytics, reporting, AI/ML, and operational integrations. Reads the golden record; does not master it.

The running example is a generic party entity. The architecture applies unchanged whether that party is a customer, a citizen, a patient, a member, a supplier, or an individual under a defence or public-safety mandate — and equally to non-party master entities such as assets, locations, products, or accounts.

1.3 The same model across sectors

The single insight that makes the platform reusable is that the structure is constant and only the instantiation changes. A deployment names its entity, assigns each contributing system to a role, and sets survivorship precedence — nothing in the architecture itself moves.

Commercial / enterprise

Entity: customer, supplier, product. Roles: CRM (engagement), ERP (record), billing (record), e-commerce (entry).

Public sector / government

Entity: citizen, case, benefit recipient. Roles: registration (entry), case management (record), service portals (engagement).

Healthcare

Entity: patient, provider, encounter. Roles: registration/ADT (entry), clinical record (record), scheduling (operations).

Defence / public safety

Entity: person, unit, platform, asset. Roles: personnel system (entry), operational C2 (operations), logistics (record).

Workforce / HR

Entity: worker, position, org unit. Roles: HRIS (entry), payroll (record), scheduling (operations), directory (engagement).

Because the model is entity-agnostic, an organization can master its highest-pain entity first and extend to further entities later without re-architecting.

1.4 Scope

In scope (Phase 1): one master entity and its core attributes, its hierarchy (parent/child, grouping, reporting lines), and the identity cross-reference across the contributing systems.

Contributing and consuming systems: the systems assigned to the roles above, plus downstream

analytics and AI.

Out of scope (later phases): additional master-data domains, full transactional integration, and decommissioning of legacy interfaces.

2. Architectural Principles and Decisions

The architecture is shaped by four critical decisions, plus a choice of MDM implementation style. These are the load-bearing commitments; everything downstream follows from them.

Decision 1 · Multi-stage ingestion

Data moves through three named zones — a Bronze landing zone for raw as-received data, a Silver working zone where deduplication, cleansing, and conflict resolution happen, and a Gold zone holding the finished golden records. Each zone is a durable, separately-governed boundary, not just a processing step.

Decision 2 · Explicit subsystems

Every functional capability of the hub is named, bounded, and assigned a responsibility — governance, modeling, quality, matching/survivorship, golden record, integration, hierarchy, stewardship, metadata/reference, and security/lineage/audit.

Decision 3 · Provenance and chain of custody

A dedicated Security, Lineage & Auditability capability records, for every golden attribute, where the value came from (provenance) and the ordered sequence of automated and human decisions that produced it (chain of custody) — sufficient to reconstruct and defend any record.

Decision 4 · Automated plus human-in-the-loop curation

The Curation Engine resolves identity and conflicts automatically where confidence is high, and routes ambiguous, low-confidence, or sensitive decisions to a human steward through an explicit review workflow. The split is governed by thresholds and policy, not left implicit.

2.1 Chosen implementation style — Coexistence

Of the four classic MDM styles, this design adopts Coexistence. The hub masters the entity record and stays bidirectionally synchronized with the sources, which keep operating as their owners' systems of entry. The hub does not force every system to surrender data entry (the Centralized style, a far larger change), but it does hold and publish the authoritative cross-system view — more control than a read-only Registry or an analytics-only Consolidation can give.

Registry — rejected

Thin index pointing back to sources; golden record assembled virtually on read. Too little control; can't hold a durable, cleansed golden record.

Consolidation — partial

Copies pulled centrally, mainly for reporting; sources still own everything. Good for analytics, but doesn't feed trusted data back to operations.

Coexistence — selected

Hub masters and cleanses data and syncs bidirectionally with operating sources. Balances authority

with acceptable change to source systems.

Centralized — later-phase option

Hub becomes the sole system of entry for the entity. Highest consistency, but the largest organizational disruption.

2.2 Supporting principles

Program, not project

Governance and stewardship are first-class and ongoing; the platform is necessary but not sufficient.

Raw data is immutable

Bronze is never edited in place — it is the replayable system of record for everything the hub received.

Every decision is attributable

No value reaches Gold without a recorded source and a recorded rationale (rule or steward).

Source of trusted data, not just single source of truth

The golden record is what everyone agrees to rely on, assembled attribute-by-attribute from the best source.

3. High-Level Reference Architecture

The architecture is read top-to-bottom as a flow and left-to-right by zone. Source systems feed an integration layer; data lands in Bronze, is curated in Silver, and is mastered in Gold; golden records are published back out to operational and analytical consumers. Two capabilities run as cross-cutting planes spanning every layer: Governance & Stewardship above, and Security, Lineage & Auditability below.

The two planes are deliberately wrappers rather than steps: governance decisions and audit records apply at every stage, not at a single point. A value standardized in Silver carries the same provenance discipline as a record published from Gold.

4. Subsystem Catalog

Each subsystem has one clear responsibility. Together they cover the full lifecycle of a master entity — created once, defined consistently, identified uniquely, governed centrally, shared across systems, maintained as it changes, and eventually retired.

SUB-01 · Data Integration & Connectors

Captures data from each source (CDC, API, batch/file, events), feeds Bronze, and syndicates golden records back out. The plumbing of the hub.

SUB-02 · Data Model & Entity Schema

Defines the master entity, its attributes, relationships, and hierarchies so every system interprets the data identically.

SUB-03 · Data Quality Engine

Profiles, parses, standardizes, validates, and enriches incoming data against the six quality dimensions — accuracy, completeness, consistency, timeliness, validity, uniqueness.

SUB-04 · Matching & Deduplication

Identity resolution: recognizes that two differently-keyed records describe the same real-world entity via exact, fuzzy, probabilistic, and AI-assisted matching.

SUB-05 · Survivorship & Conflict Resolution

Applies survivorship rules to merge the best attribute from each source into the golden record, and resolves conflicting values — automatically or by escalation.

SUB-06 · Stewardship Workflow (HITL)

Human review, exception handling, and approval flows for matches and changes the system is not confident enough to apply on its own.

SUB-07 · Golden Record & XREF

Assembles and stores the authoritative record, plus the cross-reference that maps the golden ID to every source-system key.

SUB-08 · Hierarchy & Relationship Mgmt

Maintains roll-ups and groupings — parent/child chains, organizational units, reporting lines — so logic and reporting respect how entities relate.

SUB-09 · Metadata & Reference Data Mgmt

Holds definitions, business meaning, ownership, and quality rules; governs the controlled code-lists (country, currency, status) master data depends on.

SUB-10 · Security, Lineage & Auditability

Access control plus the provenance and chain-of-custody record for every value and every decision. The cross-cutting trust capability.

SUB-11 · Data Governance & Stewardship Console

The operational surface for owners, stewards, and the council to set definitions, decision rights, and policies and to action exceptions.

SUB-12 · Publication & Syndication

Pushes trusted golden records to operational consumers (sync-back) and to analytics, via events and request/response APIs.

Mapping: these subsystems land on the ArchiMate Application layer; the people and policies that drive them (SUB-11) sit on the Business layer; hosting, storage, queues, and access controls (part of SUB-10) sit on the Technology layer.

5. The Ingestion Pipeline — Bronze · Silver · Gold

Ingestion is a three-zone medallion pipeline. Each zone is a durable boundary with its own purpose, access policy, and retention rules. Data is promoted forward only when it meets the criteria of the

next zone; nothing is deleted from Bronze, so the whole pipeline is replayable.

5.1 Bronze — Landing Zone

Bronze receives data exactly as the source sent it and never modifies it. Each record is stamped with ingestion metadata — source system, ingestion timestamp, batch identifier, and a checksum — and persisted to a durable store. This makes Bronze the replayable system of record: if a downstream rule changes or a bug is found, Silver can be rebuilt from Bronze without re-extracting from the sources. Reference and code-list values are resolved against the metadata catalog here so later stages share one vocabulary.

5.2 Silver — Working Zone

Silver is where all the cleansing and resolution happens, in order: profile and standardize (parse names, normalize address and identifier formats), validate against rules and enrich from trusted reference data, match and deduplicate to cluster records that represent the same entity, and then apply survivorship and conflict resolution to choose the best value for each attribute. Anything the system cannot resolve confidently is escalated to the stewardship workflow. Silver holds work in progress — candidate clusters, proposed merges, and unresolved conflicts — separately from the finished records in Gold.

5.3 Gold — Trusted Zone

Gold holds only finished golden records. A record is promoted here once it meets quality thresholds and all conflicts are resolved and recorded. Promotion assigns or confirms a stable golden ID, updates the cross-reference that maps that ID to every source key, refreshes hierarchy roll-ups, and publishes the record to consumers. Under the chosen coexistence style, corrected values are also synced back to the relevant source systems.

ZONE PROPERTIES

Contents	Bronze: raw, as-received · Silver: cleansed, in-progress clusters · Gold: finished golden records.
Mutability	Bronze: immutable, append-only · Silver: working / re-derivable · Gold: versioned, controlled.
Edited by	Bronze: no one · Silver: engine + stewards · Gold: promotion process only.
Typical access	Bronze: pipeline & audit only · Silver: stewards & engine · Gold: consumers (read), sync-back.
Promotion gate	Bronze: parseable & registered · Silver: thresholds met, conflicts resolved · Gold: (terminal).

6. Curation Engine — Automated and Human-in-the-Loop

The Curation Engine is the part of Silver that decides identity and resolves conflicts. Its defining feature is a deliberate split between an automated lane that acts on high-confidence, policy-clear decisions and a human-in-the-loop (HITL) lane that routes everything ambiguous, low-confidence, or sensitive to a steward. The split is driven by explicit confidence thresholds and governance policy — never left implicit.

6.1 The automated lane

When a match score sits above the auto-merge threshold and no governed-sensitive attribute is in conflict, the engine acts on its own: it merges the cluster, applies survivorship rules to select the best value per attribute, records the rule that fired and its rationale, and moves the result toward Gold. This is the high-volume path — most records should clear automatically.

6.2 The human-in-the-loop lane

Decisions that are uncertain, contested, or sensitive are routed to a steward through the stewardship workflow. The steward sees the candidate cluster, the competing source values, the match scores, and the lineage so far, and decides — to merge, reject, un-merge a previous merge, or override a survivorship pick. Every steward action is captured with the steward's identity and a reason, and feeds back as tuning signal so the automated lane improves over time.

6.3 Where a human must intervene

The decision points below are routed to HITL by policy, regardless of score, because the cost of an automated mistake is high or the judgement is inherently human.

Mid-band match — HITL

Score falls between the auto-merge and no-match thresholds — genuinely ambiguous identity.

Authoritative conflict — HITL

Two trusted sources disagree on a key attribute and survivorship can't break the tie.

Sensitive attribute change — HITL

Changes to governed/regulated fields (national ID, protected data) require explicit sign-off.

Potential bad merge / split — HITL

A proposed merge would combine two real entities, or a suspected over-merge needs to be un-merged.

Validation exception — HITL

A record fails a rule that cannot be auto-corrected (impossible date, broken hierarchy loop).

New / unmapped attribute — HITL

A source sends a field with no model mapping; a steward decides how it is handled.

Clear high-confidence match — AUTO

No human needed — applied automatically with recorded rationale.

7. Security, Lineage and Auditability

This capability answers three questions for any value in any golden record: who is allowed to see or change it (security), where did it come from (provenance / lineage), and what sequence of decisions produced it (chain of custody). It runs as a cross-cutting plane, capturing evidence at every step of the pipeline rather than at a single checkpoint.

7.1 Security and access control

Role- and attribute-based access over zones and fields — stewards see Silver work-in-progress; consumers see only published Gold values they are entitled to. Encryption and masking of sensitive attributes in transit, at rest, and in steward views by entitlement. Least-privilege by zone: Bronze is restricted to the pipeline and auditors; nobody edits it.

7.2 Lineage and provenance

For every golden attribute the hub records which source contributed the winning value, which competing values were considered, and which survivorship rule or steward decision selected it. This is attribute-level provenance, not just record-level — one attribute may come from a system of engagement while another comes from a system of record, and each is traceable independently.

7.3 Audit and chain of custody

Every transformation and decision — automated or human — is written to an append-only audit log: the standardization that reformatted a value, the match that clustered two records, the survivorship rule that fired, the steward who approved a merge and why. The log is sufficient to reconstruct how any record reached its current state and to defend that record to an auditor or regulator.

Why this matters for compliance: privacy regimes — GDPR, PIPEDA, the Privacy Act, CCPA, HIPAA, and their sectoral equivalents — require knowing exactly where an entity's data lives in order to find, correct, or delete it. Attribute-level provenance plus the cross-reference make that a query, not an investigation. This is the same provenance-and-chain-of-custody discipline the Origi trust framework applies to AI systems, here applied to master data.

8. Golden Record and Publication

The golden record is the authoritative representation of a master entity, assembled attribute-by-attribute from the best available source. It is not a copy of any one system; it is the agreed composite plus the cross-reference that links it back to every source key.

Golden ID and cross-reference

Each entity receives a stable golden ID that never changes even as source records come and go. The XREF maps that golden ID to each contributing system's local key — this is what lets a consuming system ask for the trusted record for a given entity.

Hierarchy and relationships

The hub maintains parent/child chains and organizational roll-ups so that reporting and logic respect how entities relate. A relationship change is resolved once, in the hub, and published everywhere — rather than drifting between systems.

Publication and sync-back

Trusted records are published two ways: as events for systems that want to react to changes, and via request/response APIs for systems that look records up on demand. Consistent with the coexistence style, corrected values flow back to the source systems so the whole estate converges on the golden record rather than diverging from it.

9. Governance and Stewardship

MDM succeeds or fails on governance, not tooling. The platform is necessary but not sufficient; the agreements about what an entity is, who owns the definition, and who arbitrates conflicts are what make the golden record trustworthy. These roles operate the stewardship console and set the policies that drive the curation engine's automated/HITL split.

Data Owner

Business-side authority for the entity domain; approves definitions, survivorship policy, and source-of-truth precedence per attribute.

Data Steward

Day-to-day guardian of quality; works the HITL queue — merges, un-merges, resolves conflicts, fixes exceptions.

Data Custodian

IT/technical side; operates and secures the hub, the stores, and the integration plumbing.

Governance Council

Cross-functional body; sets standards, arbitrates cross-domain disputes, and signs off on threshold and policy changes.

10. Cross-Cutting Concerns

Scalability and replay

Because Bronze is immutable, Silver and Gold can be rebuilt by replaying — essential when match rules or survivorship policy evolve.

Idempotency

Re-ingesting the same source record produces the same result; matching is deterministic given the same inputs and ruleset version.

Latency modes

Event/CDC paths support near-real-time updates for operational sync-back; batch paths handle bulk loads and reconciliation.

Versioning

Golden records, rulesets, and the entity model are versioned so any past state — and the rules that produced it — can be reconstructed.

Data quality monitoring

The six dimensions — accuracy, completeness, consistency, timeliness, validity, uniqueness — are measured continuously, not just at load.

Phased rollout

Start with the entity domain where bad data hurts most; expand to further domains and tighten toward a more centralized style only if the organization commits to it.

11. Master Data Warehouse — Overview and Relationship to MDM

The Master Data Warehouse (DWH) is the analytics-facing consolidation layer. It pulls data from the same contributing systems as Part 1, but its job is different: where MDM resolves identity and produces authoritative master and reference data, the DWH integrates the broader operational and transactional data and conforms it to those golden entities, then reshapes it for query and modelling.

The relationship is a clean division of authority. MDM (Part 1) is the dimension authority — it decides what an entity is and supplies the golden record, the cross-reference, and the reference code-lists. The DWH (Part 2) is where data meets those conformed dimensions — it never re-resolves identity itself; it looks identity up from MDM. This avoids two systems making conflicting identity decisions.

The same three zones, different semantics: both parts stage data through Bronze -> Silver -> Gold, but the Silver and Gold contracts differ. In the DWH, Silver holds the consolidated, normalized integrated model (joined to MDM master and reference data); Gold holds the denormalized, analytics- and AI-ready dimensional and wide-table form.

12. The DWH Ingestion Pipeline — Bronze · Silver · Gold

The pipeline is a three-zone medallion flow, like Part 1, but tuned for analytics consolidation rather than identity resolution. Bronze lands the raw extracts; Silver consolidates and normalizes them against MDM's master and reference data; Gold denormalizes the result for consumption.

12.1 Bronze — Landing Zone

Bronze receives raw extracts from each contributing system exactly as sent, stamps them with ingestion metadata, and persists them immutably. As in Part 1, this makes the pipeline replayable: if the Silver consolidation logic or a Gold model changes, the warehouse can be rebuilt from Bronze without re-extracting from sources.

12.2 Silver — Consolidation Zone (normalized)

Silver is where consolidation happens against MDM data. Records are normalized and cleansed, then their source business keys are resolved to golden IDs through the MDM cross-reference, reference codes are joined to their standardized values, and golden-record attributes are attached as conformed dimensions. The result is a normalized integrated model (third-normal-form) in which every entity reference points at a golden ID rather than a source-local key. This is the step that detects an unrecognized entity and triggers the Holding Zone path.

12.3 Gold — Serving Zone (denormalized)

Gold reshapes the integrated model into the denormalized forms analytics and AI consume efficiently — dimensional/star schemas, wide tables, and pre-computed aggregates — and publishes them to data marts, BI tools, and the feature store. Because identity was already conformed in Silver, Gold can denormalize freely without risking inconsistent entity references.

DWH ZONE PROPERTIES

Contents Bronze: raw source extracts · Silver: consolidated integrated model · Gold: dimensional / wide tables.

Shape Bronze: as-received · Silver: normalized (3NF) · Gold: denormalized.

MDM dependency Bronze: none · Silver: golden ID, XREF, reference data · Gold: inherited from

Silver.

Identity references Bronze: source-local keys · Silver: resolved to golden IDs · Gold: golden IDs throughout.

Consumer Bronze: pipeline & audit · Silver: pipeline (internal) · Gold: BI, AI/ML, ad-hoc.

13. Pipeline Component Catalog

Every component of the DWH ingestion pipeline is named with one responsibility. Together they implement the fully-automated flow and the holding-zone exception path.

DWH-01 · Source Extractors / Connectors

Pull data from each contributing system via CDC, batch, API, or file, and deliver it to Bronze. One adapter per source contract.

DWH-02 · Bronze Raw Store

Durable, immutable landing for as-received extracts plus their ingestion metadata; the replay point for the whole pipeline.

DWH-03 · Normalization & Cleansing

Conforms data types and formats, applies structural fixes, and removes within-source duplicates before integration.

DWH-04 · Key Resolution Service

Maps each source business key to its golden ID via the MDM cross-reference. The gate that detects entities MDM does not yet know.

DWH-05 · Reference Data Join

Resolves coded values to their standardized descriptions and hierarchies using MDM-governed reference data.

DWH-06 · Master Data Enrichment

Attaches golden-record attributes to incoming data as conformed dimensions, so every system's data shares one entity view.

DWH-07 · Integration / Normalized Model

Assembles the consolidated third-normal-form integrated model in Silver, with all references pointing at golden IDs.

DWH-08 · Holding Zone Manager

Parks records for unresolved entities, submits the core entity to the MDM pipeline, listens for resolution, and reprocesses held records.

DWH-09 · Denormalization / Dimensional Modeler

Transforms the integrated model into star schemas and wide tables for Gold.

DWH-10 · Aggregation & Pre-computation

Materializes summaries, metrics, and roll-ups so common analytics queries are fast.

DWH-11 • Serving Layer / Data Marts

Exposes Gold to BI, ad-hoc analytics, and the AI/ML feature store through query interfaces.

DWH-12 • Orchestration & Scheduling

Drives the end-to-end automated flow — dependencies, triggers, retries, and the holding-zone callback. The engine that makes the pipeline hands-off.

DWH-13 • Data Quality & Monitoring

Checks freshness, completeness, and reconciliation across zones; raises alerts and reconciliation reports.

DWH-14 • Lineage & Audit (shared)

Extends the Part 1 lineage capability across DWH transforms, so a Gold metric traces back through Silver and Bronze to its sources.

14. Automated Flow and the Holding Zone

The DWH pipeline runs completely automated, end to end, driven by the orchestration engine — no human step is required for a normal load. There is exactly one exception, and even it is handled automatically: when incoming data references an entity that has no golden record yet, that data cannot be conformed, so it is set aside rather than loaded with a broken or guessed identity.

14.1 The exception path

When the Key Resolution Service (DWH-04) finds no golden ID for an entity, the Holding Zone Manager (DWH-08) does three things, in order: it parks the full record in the Holding Zone; it passes the core entity data into the MDM ingestion pipeline (Part 1) so a golden record can be created; and it waits for MDM to signal completion. Once MDM publishes the new golden record and updates the cross-reference, the held records are automatically re-evaluated — now they resolve — and continue on to Gold.

14.2 Why park rather than guess

Loading a record whose entity has no golden ID would either create an orphaned reference or force the DWH to invent an identity — which is exactly the authority that belongs to MDM, not the warehouse. Parking keeps the warehouse internally consistent and preserves the single-authority principle: the DWH never resolves identity itself; it defers to MDM and waits.

14.3 Where humans do and don't appear

The DWH path is fully automated — including the parking, the hand-off, and the reprocessing. A human only enters the picture inside Part 1: if MDM's own curation engine cannot resolve the new entity automatically, a steward handles it there. From the warehouse's point of view that is invisible — it simply waits for the golden record, however MDM produced it.

HOLDING STATES

Parked Record held; core entity submitted to MDM; awaiting a golden record. (automatic)

Triggered MDM has published the golden record and updated the XREF. (automatic)

Reprocessing Held records re-evaluated; keys now resolve to the new golden ID. (to Gold)

Released Records conformed and promoted; Holding Zone entry cleared. (done)

Escalated MDM could not resolve after policy retries / timeout. (exception)

15. End-to-End: Part 1 and Part 2

Seen together, the two parts form one platform with one authority for identity. MDM resolves and publishes golden records and reference data; the DWH consolidates everything else against them and serves analytics and AI. The holding-zone loop is the seam that keeps them in step — the warehouse never races ahead of the master data it depends on.

15.1 Key takeaways

- Master data is the shared nouns of the business; MDM produces a golden record — a single source of trusted data — and is a program, not a project.
- The architecture is sector-independent: name the entity, assign each contributing system to a role, set survivorship precedence, and the structure applies unchanged from commercial to public-sector to defence deployments.
- The Bronze -> Silver -> Gold pipeline keeps raw data immutable and replayable, isolates cleansing work, and promotes only finished records.
- Curation is explicitly split: automate the confident, escalate the ambiguous and the sensitive to a human.
- Every value carries its provenance and every decision its chain of custody — the basis of both trust and compliance, and the same discipline Origi applies to AI systems.
- Coexistence balances authority with acceptable change; tighten toward centralized only by deliberate choice.
- Part 2 — the Master Data Warehouse consolidates contributing-system data against MDM's golden records: Silver normalized, Gold denormalized and analytics/AI-ready.
- The DWH pipeline is fully automated; the one exception — an entity with no golden record — is handled by parking the data, feeding the entity to MDM, and reprocessing on resolution, with no human step in the warehouse path.
- One authority for identity: the warehouse never resolves entities itself; MDM does, and the warehouse waits.

References

- [1] MIOVSKI GROUP / Enkidu Enterprises (2026). Trust Assurance Overview (Origis), v1.0.
- [2] MIOVSKI GROUP / Enkidu Enterprises (2026). Aegis Assure — AI Data Governance and Ethics Assessment, v3.0.
- [3] Enkidu Enterprises (2026). Tangram Overview, v1.0.
- [4] The Open Group. ArchiMate 3.2 Specification.
- [5] DAMA International. DAMA-DMBOK: Data Management Body of Knowledge (2nd ed.).

END OF DOCUMENT