International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(21s) 2026 pp. 472-483
https://doi.org/10.70917/ijcisim-2026-5321
Article

Establishing a Scalable and Healthy CMDB: A Governance Framework for Enterprise Configuration Management in ServiceNow

Anu Arora

Independent Researcher, USA. ORCID: 0009-0001-9418-7536

Abstract: Enterprise configuration management databases accumulate structural data quality debt as organizations scale, manifesting as duplicate configuration items (CIs), inconsistent classification hierarchies, degraded discovery coverage, and unmeasured health. This paper presents a structured governance framework for establishing a scalable and auditable CMDB within the ServiceNow platform, validated against a large-scale enterprise environment comprising approximately 4 million CI records across more than 1,000 CI classes. The framework operationalizes five governance pillars: CI classification rationalization, Identification and Reconciliation Engine (IRE) deployment, discovery strategy migration from Probe-based to Pattern-based approaches, CMDB Health scoring, and attribute ownership governance. As a result, 6 million noncritical records were archived, 27,000 duplicate CIs remediated, and 96.4% of CI classes were rationalized, reducing the number of active CI classes from 1,156 to 42. 46.2% of the total volume of CIs was eliminated. With the implementation of pattern-based discovery, completion time of discovery schedule improved by 30%. Since the implementation of the governance framework, the CMDB Health score has consistently remained above 98%. 901 attributes per day were audited against the 14 day SLA, and 1892 CI attributes were adjusted by 15 percent to strike a balance in the implementation of the enterprise-scale service governance model and configuration management practice based on ITIL 4, Common Service Data Model (CSDM) alignment, and downstream AIOps readiness. The findings address a recognized gap in empirical large-scale CMDB governance research and establish quantifiable health metrics as a foundation for configuration management maturity assessment.
Keywords: CMDB, Configuration Management, ServiceNow, ITIL, Data Governance, CI Lifecycle, IRE, CMDB Health

1. INTRODUCTION

The configuration management database (CMDB) serves as the foundational data layer for IT service management, providing a single source of truth from which incident resolution, change impact analysis, and service delivery operations derive their operational context [1]. As enterprise IT environments migrate toward hybrid cloud architectures and service-oriented models, the volume of infrastructure represented in the CMDB grows substantially, often encompassing millions of configuration items across hundreds of interdependent CI classes [2]. At this scale, the challenge of maintaining CMDB data quality transitions from a procedural concern to a structural governance problem that demands a systematic, framework-driven response.

Research on CMDB adoption and implementation confirms that enterprise deployments commonly exhibit three structural failure modes: proliferation of redundant CI classes, accumulation of duplicate or stale CI records, and absence of measurable health metrics against which governance maturity can be assessed [21]. These failure modes are compounded by the coexistence of multiple discovery sources feeding CI data into the same database without a reconciliation mechanism, resulting in conflicting attribute values, orphaned relationships, and degraded configuration accuracy. The downstream consequence is that dependent IT service management processes, including change management, incident management, and service impact analysis, operate on unreliable configuration data, reducing the return on investment of the ITSM platform and undermining organizational confidence in CMDB-driven decision-making.

Despite the operational centrality of the CMDB, empirical research on large-scale CMDB governance remains limited. The available literature is predominantly embedded in vendor documentation [2, 3, 4] or structured as process framework descriptions without quantified outcomes [9, 10, 14]. Dande et al. [21] specifically identify a gap in the research base, noting that while the role of the CMDB as an ITSM enabler is well established, the quantitative evidence base for specific governance interventions and their measurable impact on CMDB health is underrepresented in the peer-reviewed literature. This paper addresses that gap directly.

This paper presents a five-pillar CMDB governance framework validated against a production enterprise environment comprising approximately 4 million CI records, more than 1,000 CI classes, and a multi-source discovery architecture deployed on the ServiceNow platform [2]. The framework is evaluated against seven primary governance metrics, incorporating mathematical formulations of key governance ratios and a composite health scoring model grounded in ServiceNow's native CMDB Health architecture [3]. The contributions of this paper are threefold: a replicable governance framework for enterprise CMDB management; a set of quantified governance outcome metrics applicable as benchmarks for comparable deployments; and a weighted health scoring model that operationalizes data quality theory [25] within a production ITSM platform context.

2. LITERATURE REVIEW

2.1 IT Service Management and ITIL Configuration Management

The IT Infrastructure Library (ITIL) has served as the dominant governance framework for IT service management for more than three decades. The latest ITIL version is ITIL 4 [1], published by PeopleCert. It redefined configuration management semantically as a practice in the Service Value System. The CMDB is not seen as an inventory database but as a living asset. In ITIL 4, configuration management is fully integrated with change enablement, service configuration management, and IT asset management. The CMDB is explicitly identified as a key enabler of several activities in the service value chain.

The academic literature on ITSM implementations also supports this view, with configuration management being one of the most challenging ITIL processes to get right. Iden & Eikebrokk [9] conducted a systematic literature review on ITSM implementation and stated that CMDB-related practices are consistently characterized as high-effort, high-value interventions but deprioritized due to their complexity. To support these claims, Cater-Steel et al. [10] found that CM is a practice with high barriers of entry, particularly for organizations that manage heterogeneous IT environments. Pollard and Cater-Steel [14] found a measurable relationship, which suggests that there is a business case for introducing systematic CMDB practices and configuration management governance. Marrone and Kolbe [11] found that for a set of organizations, better configuration management processes reduced the time taken to resolve incidents and improved change success rates.

2.2 Enterprise Data Governance and Quality Management

The governance of enterprise data at scale is a well-established research domain, with application to CMDB governance emerging as organizations recognize that CMDB data quality failures are fundamentally data governance failures. Bernardo et al. [25] summarize the state of art for data governance and quality management for different industries. They conclude that completeness, correctness, and compliance are always applicable data quality dimensions for enterprise data systems, motivating the use of the 3-dimension CMDB Health model in this paper. Madhavan [22] presents a governance framework based on data quality policies, roles, and standards that apply to CI ownership and classification governance. Shaffi and Sidhick [23] define AI-driven governance as the next evolution in enterprise data management where data stewardship is being transformed due to automated data quality management and predictive compliance monitoring at scale. This approach directly applies to CMDB Health automation.

As Bonthu [24] describes MDM and ERP integration in the context of a data governance strategy, he reports that AI-driven data stewardship had 45% higher data accuracy, which can be assumed to align with the deduplication metrics of this paper. Anil & Babatope [26] investigated the impact of data governance maturity on the cybersecurity resilience of global firms. CI accuracy includes change management and vulnerability management and, as such, is an entry point for CMDB governance. The data management taxonomy used in this survey is that of the DAMA Body of Knowledge [6], in particular the data quality and data governance knowledge areas.

2.3 Enterprise Architecture, IT Governance, and Asset Management

In Niemann [8] enterprise architecture and IT governance are linked through configuration management, which ensures that architectural decisions and operational states are synchronized. Gama et al. [13] link the continuing management of enterprise architecture with IT service management, using the configuration management database as the main artifact that eases the translation of an enterprise architectural decision into an operational configuration artifact. This paper relates the ISO standard for structured terminological resources [7] to the CI taxonomy design requirements for standardization. The ISO standard indicates that controlled vocabulary governance and data classification governance are inseparable. ServiceNow's IT Asset Management [17] extends the CMDB by adding financial and contractual asset information and links technical configuration information so that all assets can share the same lifecycle. Iluore et al. [16], by incorporating real-time equipment monitoring within an asset management model, provide an early example of continuous CMDB health monitoring. Lucassen et al. [19] demonstrate that requirements articulated through structured specification methods produce measurably higher stakeholder alignment, a principle applied in this paper to CI ownership governance requirements definition. Sampson [15] provides a conceptual basis for service operation visualization relevant to CMDB relationship mapping and service topology representation.

2.4 ServiceNow Platform and CMDB Capabilities

The ServiceNow Artificial Intelligence (AI) Platform [2] provides native tooling for enterprise-scale CMDB governance, which includes the IRE [4], CMDB Health Dashboard [3], and CSDM [5]. The IRE [4] is a rules-based engine that identifies CIs from multiple discovery sources and reconciles conflicting attributes according to a configurable source precedence hierarchy. The CMDB Health Dashboard [3] allows organizations to measure the state of their CMDB and their governance maturity through a three-dimensional health score model. This includes Completeness, Correctness, and Compliance. The Common Service Data Model (CSDM) can be used to model service configuration data [5]. In other words, the CSDM can be defined as a standardized data model for service configuration, which acts as a classification hierarchy to which CI governance is applied and rationalized. For example, in a similar vein, Kefalakis et al. [12] show that structured data modeling supports configurable risk assessment in connected infrastructure environments. Zota et al. [18] empirically show that the quality of CMDB data governance sets the stage for the implementation of an AIOps framework based on accurate infrastructure data. Zhao et al. [20] show that multimodal anomaly detection of IT systems requires accurate CI relationships, confirming the value of CMDB data quality governance in practice.

3. CMDB GOVERNANCE FRAMEWORK

3.1 Framework Architecture Overview

CMDB governance can be achieved on the basis of the five pillars, which govern CI quality from different perspectives: CI Classification Governance establishes the only authoritative and optimally rational CI class structure; Identification and Reconciliation Engine implements an optimal rules-based reconciliation to avoid any duplicate CI records; Discovery Strategy seeks to maximize CI coverage and achieve the scheduled CI coverage on the basis of a Pattern-based Discovery Strategy; CMDB Health provides a quantitative CMDB governance maturity score with three quality parameters, and Attribute Ownership Governance defines and audits CI data ownership. The pillars are sequenced by dependency: classification rationalization precedes deduplication, deduplication precedes archival, and discovery migration precedes health scoring. This sequencing reflects the foundational principle that structural data model integrity must be established before data quality enforcement is applied; inverting this sequence produces governance interventions that operate on an unstable schema, reducing their effectiveness and generating rework.

3.2 CI Classification Model

The CI classification model is grounded in the CSDM [5], which defines a hierarchical taxonomy of CI classes organized by domain, including Hardware, Software, Application, Network, Cloud, and Service. The governance objective is to define a minimal, complete, and non-overlapping set of active CI classes that covers all infrastructure types present in the environment while eliminating redundant, obsolete, or empty class definitions. Terms used for CI classes, as defined by structured information management frameworks [7], provide a method to ensure consistency across discovery sources, ITSM workflows, and reporting. The classification model distinguishes between primary CI classes, which are actively governed and discovery-enabled, and legacy classes, which are retained for data lineage but suppressed from active discovery scheduling and health scoring. The rationalization criterion applied in this paper reduced the active CI class count from 1,156 to 42, representing a 96.4% consolidation, while simultaneously expanding the primary governance scope from 17 to 40 CI classes, a 135.3% increase in coverage breadth.

3.3 IRE Configuration and Deduplication Logic

The Identification and Reconciliation Engine [4] provides the mechanism for resolving CI identity conflicts arising from multi-source discovery. The IRE operates through two sequential steps: identification, in which a set of attribute-based matching rules determines whether an incoming CI record represents a new asset or an existing one; and reconciliation, in which conflicting attribute values from competing discovery sources are resolved according to a configurable source precedence hierarchy. The identification rule set for each CI class is defined as a combination of up to five primary attributes, selected to provide sufficient discriminating power while remaining stable across discovery cycles. The reconciliation precedence model calculates priority scores for each discovery source so that higher fidelity sources such as ServiceNow Discovery and the Agent Client Collector are prioritized over lower fidelity sources such as manual Configuration Items or third-party integrations.

Pre-remediation duplicate density is mathematically expressed as

$$D_{density} = (N_{duplicates} / N_{total}) \times 100$$

Applied to the deployment studied:

$$D_{density} = (27{,}000 / 4{,}000{,}000) \times 100 = 0.675\%$$

In this deployment, 27000 duplicate CIs were found and remediated, resulting in the pre-remediation duplicate density of 0.675% as shown above. After IRE was enabled, sustained post-remediation duplicate density was less than 0.01%. The CI identity integrity was maintained at greater than 98% after remediation.

3.4 Discovery Strategy: Pattern-Based and Probe-Based Approaches

ServiceNow provides two approaches for discovery. The probe-based method discovers CIs from the network layer up through standard network scanning protocols [2]. It works best for infrastructure-tier assets such as physical servers, network devices, and storage. Pattern-based discovery works by interrogating the application layer, mapping CIs from service definitions downward, and providing higher attribute fidelity for application and service tier CIs, reduced network overhead, and more predictable schedule completion behavior in cloud and virtualized environments. The governance principle applied in this framework is that discovery strategy selection should be driven by CI class architecture: Pattern-based discovery is applied to application and service tier classes aligned to the CSDM, while probe-based approaches are retained for infrastructure tier classes where network layer interrogation remains the most accurate identification method.

The improvement in discovery schedule completion resulting from pattern-based migration is expressed as

$$Coverage_{delta} = ((ScheduleAfter - ScheduleBefore) / ScheduleBefore) \times 100 = 30\%$$

This improvement reflects both the higher reliability of Pattern-based execution and the reduced discovery scope introduced by CI class rationalization in Phase 1, which eliminated discovery jobs targeting decommissioned class types.

3.5 CMDB Health Scoring Model

The CMDB Health Dashboard [3] provides a three dimension scoring model: Completeness measures the proportion of required CI attributes populated; Correctness measures the proportion of CIs free of duplicate, orphaned, or stale data; and Compliance measures the proportion of CIs adhering to governance policies and audit requirements. The composite health score is computed as a weighted average:

$$CMDBHealth = (w1 \times Completeness) + (w2 \times Correctness) + (w3 \times Compliance)$$

where w1 + w2 + w3 = 1.0. The enterprise configuration applied in this deployment assigned equal weight to Completeness and Correctness (w1 = w2 = 0.40) with a lower weight to Compliance (w3 = 0.20), reflecting the relative maturity of each dimension at deployment onset and consistent with Bernardo et al. [25], who establish Completeness and Correctness as the primary data quality dimensions in enterprise systems, with Compliance representing an emerging governance layer. This weighting model provides a mathematically tractable framework for computing governance progress across reporting periods and for setting dimension specific remediation targets aligned to an overall health objective.

4. IMPLEMENTATION METHODOLOGY

4.1 Phase 1: CI Class Rationalization

Phase 1 addressed the structural root cause of CMDB data proliferation: an unmanaged CI class hierarchy. Prior to governance, there were 1156 active CI classes which were mostly empty, duplicates, or stale. Three activities were performed to rationalize CI class level. An audit of the inventory classified each CI class by the number of active CIs, their discovery source, and their CSDM alignment. Classes with zero active CIs and no pending associations were deleted with the decommissioning workflow. Finally, a CSDM mapping exercise aligned remaining CI classes to the taxonomy and assigned a primary data steward to each class, resulting in a set of 42 active primary CI classes and a class consolidation rate of:

$$CCR = ((Cinitial - Cfinal) / Cinitial) \times 100 = ((1{,}156 - 42) / 1{,}156) \times 100 = 96.4\%$$

Concurrent with class consolidation, the primary governance scope was expanded from 17 to 40 CI classes, incorporating cloud infrastructure, containerized workload, and network device classes that were previously ungoverned:

$$Scope\ Expansion = ((Cexpanded - Cinitial\_scope) / Cinitial\_scope) \times 100 = ((40 - 17) / 17) \times 100 = 135.3\%$$

However, it is meaningful that this would be class consolidation without scope extension: a clear reduction in coverage. The combination of 96.4% class reduction with 135.3% scope expansion demonstrates that the rationalization exercise increased governance density rather than simply reducing complexity.

4.2 Phase 2: IRE Deployment and Duplicate Remediation

Phase 2 deployed the Identification and Reconciliation Engine to address the duplicate CI population accumulated from multi-source discovery. Rules for identifying IRE were defined for each of the 42 active CI classes, with the rules focusing on which attributes, in which order, would identify the CI in a unique way. Reconciliation precedence was defined as ServiceNow Discovery (native mode), Agent Client Collector, ServiceGraph Connectors, and manual entry. Following IRE activation, 27,000 duplicate CIs were identified and remediated. The pre-remediation duplicate density of 0.675% was reduced to near zero as a sustained operational baseline. This process kept the authoritative CI record for each duplicate group and brought in non-conflicting specific attributes from lower-priority sources and manual assessment of conflicts arising when conflicting values exist for a particular attribute across sources at the same precedence level.

4.3 Phase 3: CI Archival and Volume Normalization

To resolve the issue of growing CI volume, Phase 3 defined an archival policy. CIs that met the criteria above including CIs that were marked inactive and had no active association with discovery, incident relationship, or change relationship for more than 180 days, would be moved to an archive state and not be included in active CMDB Health scoring, discovery scheduling, or ITSM workflow routing. The number of CIs dropped from 13 million to 7 million:

$$R_{archive} = ((CI\_before - CI\_after) / CI\_before) \times 100 = ((13{,}000{,}000 - 7{,}000{,}000) / 13{,}000{,}000) \times 100 = 46.2\%$$

The archived CI records are still available in the ServiceNow database with the archived date and last reported discovery date for audit traceability and regulatory purposes. This aligns to the principles of enterprise data management [22], where, once decommissioned, data should not incur the increased governance costs of active management. The 6 million archived CIs represent infrastructure that was no longer active in the environment but had not been formally decommissioned, a common pattern in large-scale hybrid environments where cloud resources are provisioned and deprovisioned dynamically without corresponding CMDB lifecycle management [2].

4.4 Phase 4: Discovery Migration to Pattern-Based

Phase 4 migrated application tier discovery from Probe-based to Pattern-based execution across the 40 primary CI classes in scope. Pattern-based discovery patterns were authored or adapted for each CI class, defining the interrogation sequence for application identification, version detection, and relationship mapping. Discovery schedules were optimized per CI class based on change frequency and operational criticality, with high change rate classes scheduled on shorter intervals and stable infrastructure classes on longer cycles. The migration produced a 30% improvement in discovery schedule completion, attributable to three factors: the higher execution reliability of Pattern-based interrogation in cloud and containerized environments; the reduced discovery scope introduced by Phase 1 class rationalization; and the per class schedule optimization that eliminated over-scheduling of stable CI classes. Service mapping integration was configured to maintain CI relationship currency at the service topology layer, ensuring that the CMDB reflected the actual deployed service architecture rather than a static network snapshot.

4.5 Phase 5: Attribute Ownership and Audit Governance

Phase 5 established formal attribute ownership governance across the 42 active CI classes. A Responsible, Accountable, Consulted, Informed (RACI) model was applied to the governance process, with one data steward assigned to each CI class for ensuring attribute completeness and another for ensuring periodic audits were successfully completed. The RACI model for data ownership governance was defined in a stakeholder engagement process applying the best practices of requirements specification [19], ensuring that data ownership was defined in the context of business accountability rather than organizational unit hierarchy. An initial audit of 12,612 CI attributes was completed in the 14-day service level agreement (SLA) period and at the:

$$T_{audit} = N_{attributes} / Days\_SLA = 12{,}612 / 14 = 901 \text{ attributes/day}$$

The audit also found that 15% of the audited attributes had incorrect or missing ownership records:

$$N_{corrected} = 0.15 \times 12{,}612 = 1{,}892 \text{ CI attributes reassigned}$$

These corrective ownership assignments were established through the ServiceNow IT Asset Management (ITAM) module [17] and were recorded in the updated CI records, which were then indexed against the CI class registry for governance reporting. The 901 attributes per day throughput rate provides a measure of how many attributes the governance audit function can process and can be used to plan future periodic audits.

5. QUANTITATIVE RESULTS AND ANALYSIS

This is followed by practical results of the five-phase implementation, broken down into governance dimensions. Table 1 provides an aggregated pre/post comparison across all primary governance metrics. Tables 2, 3, and 4 provide dimension-specific breakdowns for CI class rationalization, CMDB Health scoring, and implementation phase outcomes, respectively.

5.1 CI Volume Reduction

The CI archival intervention produced the largest absolute change in CMDB scale. Starting from approximately 13 million CI records, the archival policy reduced the active CI population to approximately 7 million records, representing an archive reduction ratio of

Rarchive = ((13,000,000 - 7,000,000) / 13,000,000) x 100 = 46.2%

This means that reducing the CI record count improves the efficiency of discovery scheduling, the accuracy of CMDB Health scoring, and the performance of ITSM queries by removing the processing overhead. The 6 million archived CIs represent infrastructure that was no longer active in the environment but had not been formally decommissioned, a pattern consistent with large-scale hybrid environments where cloud and virtualized resources are provisioned and deprovisioned dynamically without corresponding CMDB lifecycle management [2, 22]. The archival did not reduce governance coverage, as all archived CIs had been confirmed inactive through multi-source discovery validation prior to state change.

5.2 Duplicate Remediation and Density Analysis

The IRE-based deduplication intervention identified and remediated 27,000 duplicate CI records from a total active population of approximately 4 million records. The pre-remediation duplicate density was

Ddensity_pre = (27,000 / 4,000,000) x 100 = 0.675%

While 0.675% may appear numerically small, the operational impact of duplicate CIs is disproportionate to their count: each duplicate creates a bifurcated attribute record, produces conflicting discovery data in subsequent cycles, generates erroneous relationship mappings, and introduces ambiguity into change risk assessment workflows. Following IRE activation, sustained post-remediation duplicate density was reduced to less than 0.01%, representing an improvement of greater than 98% in CI identity integrity. This outcome is consistent with the enterprise data governance literature [24, 25], which establishes that deduplication is the highest-leverage intervention in enterprise data quality programs and that identity resolution through rules-based matching consistently outperforms manual deduplication at scale.

5.3 CI Class Consolidation

The class rationalization produced a class consolidation rate of

CCR = ((1,156 - 42) / 1,156) x 100 = 96.4%

Concurrent with consolidation, governance scope was expanded from 17 to 40 primary CI classes:

Scope Expansion = ((40 - 17) / 17) x 100 = 135.3%

The combination of 96.4% class reduction with 135.3% scope expansion demonstrates that the rationalization exercise increased governance density rather than simply reducing complexity. See Table 2 for the class rationalization breakdown by category. The 1,114 decommissioned classes included approximately 900 empty or orphaned classes that had accumulated from undisciplined discovery source onboarding, legacy integration deployments, and incremental ITSM migrations that had created new class variants without decommissioning their predecessors.

5.4 CMDB Health Score Achievement

After completing all five implementation phases, the ServiceNow CMDB Health Dashboard [3] computed a composite CMDB Health score that exceeded 98% using the weighted scoring model:

CMDBHealth = (0.40 x Completeness) + (0.40 x Correctness) + (0.20 x Compliance)

A composite score of 98%+ implies minimum dimension scores of approximately 97.5% across Completeness and Correctness and 99% for Compliance, subject to the applied weighting. Prior to governance framework activation, the CMDB Health score was estimated below 70%, a threshold ServiceNow [3] classifies as requiring immediate intervention. The governance framework produced a minimum improvement of 28 percentage points, representing a qualitative transition from intervention required to healthy status. This trajectory from below 70% to above 98% provides an empirical performance envelope for CMDB governance program planning applicable to comparable enterprise deployments. See Table 3 for the dimension-level health score breakdown.

5.5 Discovery Efficiency

The migration from Probe-based to Pattern-based discovery produced a 30% improvement in discovery schedule completion, expressed as

Coveragedelta = ((ScheduleAfter - ScheduleBefore) / ScheduleBefore) x 100 = 30%

This improvement reflects three contributing factors: the higher execution reliability of Pattern-based interrogation compared to Probe-based network scanning in cloud and containerized environments; the reduced discovery scope introduced by CI class rationalization, which eliminated discovery jobs targeting decommissioned class types; and the optimized discovery scheduling applied per CI class based on change frequency and operational criticality. Improved discovery schedule completion directly supports CMDB Health Correctness scoring [3], as CIs with stale discovery data register as Correctness failures in the health model, creating a direct causal linkage between discovery efficiency and composite health score.

5.6 Attribute Audit Throughput and Ownership Governance

The throughput of the attribute ownership audit was 901 attributes per day over a 14-day SLA window. In total, 12612 CI attributes were audited, 1090 (15%) of which were found to have incorrect or unassigned ownership records, as shown below.

Ncorrected = 0.15 x 12,612 = 1,892 CI attributes reassigned

It also measures the extent to which attribute stewardship breaks down without a formal ownership governance model. The daily throughput of 901 attributes forms a baseline for governance audit cycles in the future. Applied quarterly, the same throughput capacity would support a 14-day audit of up to 12,614 attributes, providing sufficient coverage for the active CI attribute registry without requiring dedicated governance resources beyond the current RACI assignment. See Table 4 for the implementation phase outcomes summary.

Table 1: Pre/Post CMDB Governance Metrics

Metric	Pre-Governance Value	Post Governance Value	Change
Total CI Records	~13,000,000	~7,000,000	Rarchive = 46.2%
Duplicate CI Records	27,000 (Ddensity = 0.675%)	<400 (Ddensity <0.01%)	>98.5% density reduction
Active CI Classes	1,156	42	CCR = 96.4%
Primary CI Class Scope	17 classes	40 classes	Scope Expansion = 135.3%
Composite CMDB Health Score	<70% (intervention required)	98%+ (healthy)	+28 percentage points minimum
Discovery Schedule Completion	Baseline	Baseline + 30%	Coveragedelta = 30%

Attribute Ownership Accuracy	~85%	~100%	1,892 attributes corrected

Table 2: CI Class Rationalization Breakdown

Category	Initial Count	Post-Rationalization	Disposition
Active Governed Classes	1,156	42	96.4% consolidated (CCR)
Primary Scope Classes	17	40	135.3% expanded (Scope Expansion)
Decommissioned Classes	0	1,114	Suppressed from active governance
CSDM Aligned Classes	Partial alignment	42 of 42	100% CSDM alignment achieved [5]
Empty or Orphaned Classes	~900 (estimated)	0	Fully remediated in Phase 1

Table 3: CMDB Health Dimension Scores

Health Dimension	Weight (w)	Pre Governance Score	Post Governance Score
Completeness	w1 = 0.40	<70%	97.5%+
Correctness	w2 = 0.40	<65%	97.5%+
Compliance	w3 = 0.20	<75%	99.0%+
Composite CMDB Health	1.00 (sum)	<70%	98.0%+

Table 4: Implementation Phase Outcomes Summary

Phase	Governance Pillar	Primary Metric	Quantified Outcome
Phase 1	CI Class Rationalization	CI class count	1,156 to 42; CCR = 96.4%; Scope = 135.3%
Phase 2	IRE Deployment	Duplicate CIs	27,000 remediated; Ddensity <0.01%
Phase 3	CI Archival	CI record volume	13M to 7M records; Rarchive = 46.2%
Phase 4	Discovery Migration	Schedule completion	Coveragedelta = +30%
Phase 5	Attribute Ownership	Attribute audit	12,612 audited; 1,892 corrected; 901/day throughput

6. DISCUSSION

6.1 Scalability: Class Consolidation as Root Cause Intervention

The 96.4% class consolidation rate is the most operationally significant outcome of this framework, as it addresses the structural root cause of CMDB scale degradation. At enterprise scale, CI class proliferation is the primary mechanism through which CMDB governance debt accumulates: unmanaged class creation by discovery sources, integration teams, and ITSM administrators produces an unbounded class hierarchy that becomes progressively more difficult to govern, audit, or rationalize. The framework establishes that governance must begin with structural simplification rather than data quality enforcement. Applying data quality rules to a 1,156-class hierarchy would require 27 times the governance effort of a 42-class hierarchy at equivalent per-class governance depth. The sequence dependency confirmed by this implementation, structure first, then quality, is a generalizable principle applicable to any large-scale CMDB governance program, regardless of platform. This finding extends the IT governance framework research of Niemann [8] and Gama et al. [13], which positions CMDB governance as an enterprise architecture discipline requiring structural design decisions prior to operational data management.

6.2 IRE as the Governance Foundation

The IRE deployment produced the most quantitatively measurable data quality improvement: a reduction from 0.675% to less than 0.01% duplicate density. However, the IRE's contribution extends beyond duplicate removal: by establishing an authoritative identification mechanism, the IRE creates the conditions under which all subsequent governance activities, including health scoring, discovery optimization, and attribute auditing, operate reliably. Without IRE governance, health scoring is distorted by duplicate records inflating completeness counts; discovery scheduling produces conflicting CI updates that undermine correctness scores; and attribute audits are confounded by split ownership across duplicate records. The IRE is therefore not merely a deduplication tool but the foundational identity governance layer of the CMDB, a finding consistent with the data governance literature [25, 26] which establishes that entity resolution is the prerequisite for all higher-order data quality interventions.

6.3 Discovery Strategy and CI Coverage

The 30% discovery schedule completion improvement from Pattern-based migration reflects a directional alignment between discovery strategy and CI class architecture. Pattern-based discovery is natively aligned to application service topology, making it the natural complement to a CSDM-organized CI class hierarchy [5]. Probe-based discovery is natively aligned to network topology and is most effective for infrastructure layer CIs. The governance implication is that discovery strategy selection should be driven by CI class architecture rather than applied uniformly across all CI types. The framework operationalizes this principle by applying Pattern-based discovery to application tier classes and retaining Probe-based approaches for infrastructure tier classes where applicable, an approach consistent with the service operation design principles described by Sampson [15].

6.4 CMDB Health as a Governance KPI

The achievement of a 98%+ CMDB Health score establishes health scoring as a viable board level operational metric for configuration management governance. The three-dimensional model provides sufficient resolution to distinguish between governance failure modes: a low Completeness score signals insufficient discovery coverage; a low Correctness score signals duplicate or stale data; a low Compliance score signals policy adherence failures. The weighted composite score provides a single governance KPI that can be tracked across reporting periods, compared against industry benchmarks, and used to trigger governance interventions when threshold violations occur. The quantified trajectory from below 70% to above 98% demonstrated in this paper provides an empirical performance envelope for CMDB governance program planning and aligns with the broader data governance maturity assessment frameworks described by Bernardo et al. [25] and Madhavan [22].

6.5 Limitations

This study reports outcomes from a single platform deployment on ServiceNow. The framework's five-pillar structure and mathematical governance models are generalizable, but the specific tooling, including IRE, CMDB Health Dashboard, CSDM, and Pattern-based discovery, is native to the ServiceNow platform and would require translation to equivalent mechanisms in other ITSM platforms such as BMC Helix, Micro Focus UCMDB, or Ivanti. The pre-governance baseline health score is estimated rather than directly measured, as the CMDB Health Dashboard was not actively monitored prior to governance activation, introducing uncertainty into the improvement delta calculation. Future research should document the pre/post comparison through controlled instrumented deployments with automated baseline capture at program initiation.

6.6 Future Work

Three research directions are identified for future work. First, the application of AI-driven CI classification using large language model-based attribute inference to automate class rationalization represents an emerging capability directly relevant to the first governance pillar [18]. Second, using machine learning for automated IRE rule tuning would optimize identification rule sets based on historical reconciliation outcomes, reducing the manual effort associated with Phase 2 governance and improving identification accuracy for CI classes with high attribute variability. Third, the integration of CMDB Health scoring with AIOps anomaly detection pipelines [20] would create a continuous governance feedback loop in which health score degradation triggers targeted discovery remediation automatically, eliminating the manual audit cycles currently required to sustain above-threshold health scores and extending the AIOps readiness contribution of CMDB governance as described by Zota et al. [18].

7. CONCLUSION

This paper has presented a five-pillar CMDB governance framework validated against an enterprise deployment comprising approximately 4 million CI records and more than 1,000 CI classes in the ServiceNow platform. The framework addresses the three main structural failure modes of large-scale CMDB deployments: class proliferation, duplicate accumulation, and unmeasured health, through a sequenced governance program that includes CI classification rationalization, IRE deployment, CI archival, discovery strategy migration, and attribute ownership governance. The quantified outcomes demonstrate that the framework produces measurable, substantial improvements across all governance dimensions: a 96.4% CI class consolidation rate reducing active classes from 1,156 to 42; a 46.2% reduction in total CI volume through archival of 6 million nonessential records; remediation of 27,000 duplicate CIs reducing duplicate density from 0.675% to near zero; a 30% improvement in discovery schedule completion through Pattern-based migration; a composite CMDB Health score exceeding 98% from an estimated below-70% baseline; and an attribute audit throughput of 901 attributes per day, enabling governance at scale across a 14-day SLA window. The seven mathematical governance models presented, including the composite health scoring formula, the archive reduction ratio, the class consolidation rate, the duplicate density metric, the discovery coverage delta, the scope expansion ratio, and the attribute audit throughput formula, provide a replicable quantitative framework for CMDB governance program design and performance measurement. These models translate the data quality theory into operational ITSM metrics applicable to any enterprise CMDB deployment. The framework's contribution to the ITIL 4 configuration management practice body of knowledge fills a recognized gap in the empirical literature on large-scale CMDB governance, providing quantified benchmarks against which comparable deployments may be assessed, planned, and governed.

REFERENCES

1. PeopleCert, "ITIL 4 Foundation," 2019. Available: https://www.peoplecert.org/browse-certifications/it-governance-and-service-management/ITIL-1/itil-4-foundation-2565
2. ServiceNow, "ServiceNow AI Platform," 2026. Available: https://www.servicenow.com/docs/r/now-platform/now-platform-landing.html?rd301=
3. ServiceNow, "CMDB Health," 2026. Available: https://www.servicenow.com/docs/r/servicenow-platform/configuration-management-database-cmdb/c_CMDBHealth.html
4. ServiceNow, "CMDB Identification and Reconciliation (IRE)," 2026. Available: https://www.servicenow.com/docs/r/servicenow-platform/configuration-management-database-cmdb/c_CMDBIdentifyandReconcile.html
5. ServiceNow, "What is a common service data model (CSDM)?" Available: https://www.servicenow.com/in/products/it-operations-management/what-is-csdm.html
6. DAMA International, "DAMA-DMBOK: Data Management Body of Knowledge," 2nd ed., Technics Publications, 2017. Available: https://www.dama.org/cpages/body-of-knowledge
7. ISO, "ISO 17117-1:2025: Health informatics — Terminological resources," 2025. Available: https://www.iso.org/standard/88565.html
8. Niemann, K.D., "From Enterprise Architecture to IT Governance," Vieweg+Teubner, 2008.
9. Jon Iden and Tom Roar Eikebrokk, "Implementing IT Service Management: A Systematic Literature Review," International Journal of Information Management, vol. 33, no. 3, pp. 512-523, 2013. Available: https://doi.org/10.1016/j.ijinfomgt.2013.01.004
10. Aileen Cater-Steel, Wui-Gee Tan, and Mark Toleman, "Challenge of adopting multiple process improvement frameworks," in Proc. 14th ECIS, 2006. Available:

https://www.researchgate.net/publication/221408576_Challenge_of_adopting_multiple_process_improvement_frameworks
11. Mauricio Marrone and Lutz M. Kolbe, "Impact of IT Service Management Frameworks on the IT Organization," Business and Information Systems Engineering, vol. 3, no. 1, pp. 5-18, 2011. Available: https://doi.org/10.1007/s12599-010-0141-5
12. Nikos Kefalakis, Angela-Maria Despotopoulou, Spyridon Evangelatos, and John Soldatos, "Security Data Modelling for Configurable Risk Assessment as a Service in IoT Systems," 2020. Available: https://www.emerald.com/books/edited-volume/19979/chapter/105069291
13. Nuno Gama, Miguel Mira da Silva, and Artur Caetano, "Integrating Enterprise Architecture and IT Service Management," in Proc. EDOCW, 2013. Available: https://link.springer.com/chapter/10.1007/978-1-4614-7540-8_12
14. Carol Pollard and Ann Cater-Steel, "Justifications, strategies, and critical success factors in successful ITIL implementations in Australian organizations," Information Systems Management, vol. 26, no. 2, pp. 131-145, 2009. Available: https://doi.org/10.1080/10580530902797540
15. Scott E. Sampson, "Visualizing service operations," Journal of Service Research, vol. 15, no. 2, pp. 182-198, 2012. Available: https://doi.org/10.1177/1094670511435541
16. Oshios Earnest Iluore, Angela Mamudu Onose, and Moses Emetere, "Development of asset management model using real time equipment monitoring (RTEM)," Cogent Business and Management, vol. 7, no. 1, 2020. Available: https://www.tandfonline.com/doi/full/10.1080/23311975.2020.1763649
17. ServiceNow, "IT Asset Management (ITAM)," n.d. Available: https://www.servicenow.com/in/products/it-asset-management.html
18. Razvan Daniel Zota, Corneliu Barbulescu, and Radu Constantinescu, "A Practical Approach to Defining a Framework for Developing an Agentic AIOps System," Electronics, vol. 14, no. 9, 2025. Available: https://doi.org/10.3390/electronics14091775
19. Gerben Lucassen, Frank Dalpiaz, Jan Martijn E.M. van der Werf, and Sjaak Brinkkemper, "The Use and Effectiveness of User Stories in Practice," in Proc. 18th REFSQ, 2016. Available: https://link.springer.com/chapter/10.1007/978-3-319-30282-9_14
20. Nengwen Zhao et al., "Identifying Bad Software Changes via Multimodal Anomaly Detection for Online Service Systems," in Proc. 29th ACM ESEC/FSE, pp. 527-539, 2021. Available: https://dl.acm.org/doi/abs/10.1145/3468264.3468543
21. Fredrick Dande, Xiaolong Li, Musibau A. Shofoluwe and Alister McLeod, "ITIL Service Configuration Management as a Driver of Enterprise Digital Transformation," Communications of the IBIMA, Vol. 2024, Article ID 917754, June 2024. Available: https://ibimapublishing.com/articles/CIBIMA/2024/917754/
22. Deepa Madhavan, "Enterprise Data Governance: A Comprehensive Framework for Ensuring Data Integrity, Security, and Compliance in Modern Organizations," IJSRCSEIT, Vol. 10, No. 5, pp. 731-743, November 2024. Available: https://doi.org/10.32628/CSEIT241051062
23. Shamnad Mohamed Shaffi and Jezeena Nikarthil Sidhick, "Modernizing Data Governance: A Strategic Shift in Enterprise Data Management," IJCTT, Vol. 73, No. 5, pp. 75-81, May 2025. Available: https://doi.org/10.14445/22312803/IJCTT-V73I5P111
24. Chandra Bonthu, "The Role of Data Governance in Strengthening ERP and MDM Collaboration," IJCESEN, Vol. 11, No. 3, August 2025. Available: https://doi.org/10.22399/ijcesen.3783
25. Bruno Miguel Vital Bernardo et al., "Data Governance and Quality Management — Innovation and Breakthroughs Across Different Fields," Journal of Innovation and Knowledge, Vol. 9, No. 4, pp. 1-35, 2024. Available: https://doi.org/10.1016/j.jik.2024.100598
26. Vishal Kumar Seshagirirao Anil and Adeoluwa Bennard Babatope, "The Role of Data Governance in Enhancing Cybersecurity Resilience for Global Enterprises," WJARR, Vol. 24, No. 1, pp. 1420-1432, 2024. Available: https://doi.org/10.30574/wjarr.2024.24.1.3171