

ORIGINAL RESEARCH

Open access

Clinical Vocabulary Alignment at Scale: A Formal Harmonization Theory for Cross-System Procedure Code Mapping

Claire Martin¹, Julien Robert^{2*}, Sophie Bernard¹, Antoine Girard²

Abstract

The rapid proliferation of heterogeneous electronic health record (EHR) systems has exacerbated challenges in achieving seamless interoperability, particularly in the alignment of clinical vocabularies for procedure code mapping across disparate platforms. This conceptual manuscript introduces a formal harmonization theory tailored to large-scale clinical environments, emphasizing theoretical constructs for vocabulary alignment without reliance on empirical data or model evaluations. Grounded in systems architecture principles, we propose the vocabulary harmonization orchestration lattice (VHOL), a novel framework comprising layered modules for semantic mapping, contextual reconciliation, and governance oversight. VHOL integrates feedback topologies to mitigate alignment drifts theoretically, incorporating interpretive formulas for risk propagation and decision confidence in cross-system interactions. By synthesizing literature on clinical AI architectures, healthcare analytics infrastructures, and interoperability frameworks, this theory addresses gaps in procedure code harmonization, offering architectural blueprints for scalable deployment. The framework's unique lattice structure facilitates modular integration into EHR ecosystems, enhancing theoretical robustness against vocabulary discrepancies. Implications extend to improved decision support pipelines and governance in multi-system healthcare settings, fostering a unified semantic foundation for procedure representations. This work advances conceptual discourse on clinical vocabulary management, providing a scalable theoretical lens for future infrastructural innovations in healthcare analytics.

Keywords Clinical vocabulary alignment, Procedure code mapping, Harmonization theory, Cross-system interoperability, Scalable EHR architectures, Semantic reconciliation frameworks

*Correspondence:

Julien Robert

julien.robert@gmail.com

¹ Department of Health Informatics Research, Faculty of Medicine, University of Lyon, Lyon, France

² Department of Digital Clinical Systems, Faculty of Medicine, University of Strasbourg, Strasbourg, France

Introduction

The integration of artificial intelligence (AI) into healthcare systems has underscored the imperative for precise clinical vocabulary alignment, particularly when mapping procedure codes across diverse EHR platforms. In clinical settings where multi-institutional collaborations prevail, discrepancies in procedural terminologies can propagate errors, impeding accurate data exchange and analytical coherence. This manuscript delineates a formal harmonization theory designed to address these issues at scale, conceptualizing alignment as a systemic orchestration rather than isolated translations. By focusing on theoretical constructs, we avoid empirical validations, instead emphasizing architectural principles that underpin cross-system procedure code mapping. Such an approach is vital in environments characterized by high-volume data flows, where vocabulary misalignments could theoretically amplify risks in patient care pathways [1, 2].

Procedural terminology disparities in multi-modal clinical data environments

In deployment environments involving varied data modalities—such as structured codes from imaging procedures or surgical logs—cross-system mapping demands a harmonized vocabulary to ensure semantic fidelity. Literature highlights how disparate coding systems, like those in OMOP CDM or HL7 FHIR, often lead to fragmented representations of procedures, complicating analytics in real-time clinical workflows [3–6]. This subheading explores how such disparities manifest in multi-modal settings, where procedural data from EHRs must align without losing contextual nuances. Theoretical models suggest that unaligned vocabularies increase the cognitive load on decision support systems, potentially distorting analytical outputs in governance-constrained hospitals [7, 8].

Governance constraints impacting vocabulary alignment in scalable deployment scenarios

Governance frameworks in healthcare impose stringent constraints on vocabulary alignment, particularly in scalable deployments across federated systems. For instance, regulatory requirements for data privacy and interoperability necessitate robust harmonization theories that theoretically safeguard against mapping inaccuracies in procedure codes [9, 10]. In clinical settings with distributed EHR intelligence ecosystems, these constraints amplify the need for formal theories that orchestrate alignment without empirical tuning. This section synthesizes how governance load influences cross-system mappings, proposing that theoretical harmonization can mitigate compliance risks by embedding alignment logics into infrastructural designs [11, 12].

Cross-system procedure mapping dynamics in AI-integrated clinical workflows

AI governance and monitoring systems increasingly rely on aligned procedure codes for effective workflow integration, yet cross-system dynamics often introduce theoretical instabilities. In environments where decision support pipelines process procedural data from heterogeneous sources, mapping failures can propagate through analytics infrastructures, affecting overall system intelligence [13, 14]. This subheading examines these dynamics, drawing on conceptual models of interoperability to argue for a scaled harmonization theory that addresses procedural variability in AI-driven clinical contexts [15, 16].

Theoretical imperatives for harmonization in high-volume healthcare data modalities

High-volume data modalities in clinical settings, such as those involving procedural logs from telemedicine or population health analytics, demand a formal theory for vocabulary alignment to prevent semantic drifts. Without such theories, cross-system mappings risk theoretical inefficiencies, particularly under governance constraints that limit ad-hoc adjustments [17, 18]. This discussion posits that scalable harmonization must prioritize architectural resilience, integrating feedback mechanisms to theoretically stabilize procedure code representations across EHR ecosystems [19, 20].

The introduction thus establishes the foundational rationale for our proposed theory, setting the stage for a deeper synthesis of relevant literature and the introduction of an original architectural framework. By conceptualizing alignment as a scalable, governance-aware process, this work contributes to the discourse on clinical systems interoperability, offering a pathway to theoretically unified procedure code mappings in diverse healthcare landscapes. **Table 1** formalizes the structural distinctions between conventional linear mapping architectures and the lattice-based VHOL framework, clarifying their divergent assumptions regarding drift control, governance integration, and scalability.

Table 1. Structural comparison of procedure code alignment architectures

Dimension	Linear translation pipelines	Ontology-centric static models	Interoperability-embedded mapping	VHOL lattice architecture

Structural topology	Sequential	Hierarchical	Schema-anchored	Multi-directional lattice
Drift handling	Post-hoc correction	Periodic updates	Version synchronization	Continuous bidirectional feedback
Governance role	External auditing	Partial constraint encoding	Standards enforcement	Embedded oversight canopy
Context sensitivity	Minimal	Moderate (semantic only)	Schema-constrained	Workflow- and modality-aware
Scalability logic	Volume scaling	Concept scaling	Cross-platform scaling	Distributed modular orchestration
Risk containment	Downstream detection	Semantic abstraction	Standard compliance	Layered propagation attenuation
Decision confidence basis	Mapping accuracy	Ontological consistency	Interoperable encoding	Harmonization completeness × governance moderation

Theoretical Background and Literature Synthesis

The maturation of clinical artificial intelligence (AI) architectures over the past decade has progressively shifted attention from model-centric performance optimization toward infrastructural preconditions that enable reliable cross-system intelligence. Among these preconditions, vocabulary harmonization—particularly for procedure code alignment—has emerged as a foundational architectural problem rather than a peripheral technical inconvenience. In distributed healthcare ecosystems, procedure codes function not merely as billing artifacts but as semantic anchors linking diagnostics, interventions, outcomes, and governance audit trails. When these anchors diverge across systems, analytical coherence collapses upstream and downstream in AI pipelines.

The theoretical underpinnings of harmonization draw heavily from interoperability scholarship, which conceptualizes data exchange as a layered process encompassing syntactic transmission, structural modeling, and semantic alignment [1, 3]. Within this layered ontology, semantic alignment is not a secondary refinement; it is the structural condition for interpretability, comparability, and safe decision automation. Clinical AI infrastructures depend on shared vocabularies to ensure that procedural abstractions remain stable across institutions, data modalities, and analytic tasks. Without semantic continuity, machine learning systems may generate locally valid inferences that fail under cross-contextual deployment.

Healthcare analytics infrastructures, therefore, require formal theories that treat vocabulary discrepancy as an architectural instability rather than an operational error. Recent literature emphasizes conceptual models for resolving terminology divergence without relying exclusively on post hoc empirical correction, instead proposing abstraction layers, ontology-driven mappings, and common data representations [2, 4]. These approaches move harmonization upstream into the system design phase. Synthesizing literature, this section organizes the theoretical landscape around four interlocking domains: semantic foundations, interoperability scaffolds, analytics infrastructures, and governance-integrated decision support. Together, these domains inform the development of a principled harmonization theory suitable for scalable AI deployment.

Semantic foundations of procedure code harmonization in clinical AI architectures

Clinical AI architectures frequently embed ontological systems to represent procedures, diagnoses, and clinical states within structured knowledge graphs. These ontologies serve as semantic backbones that mediate translation between heterogeneous coding schemes. Extensible ontology development research underscores that interoperability is not solely a matter of format compatibility but of semantic equivalence across representational regimes [1-3]. Procedure codes must therefore be conceptualized as nodes within structured semantic hierarchies, where relationships, constraints, and inheritance patterns determine interpretive coherence.

Within EHR intelligence ecosystems, vocabulary misalignment produces cascading disruptions. Decision support modules, risk stratification engines, and predictive inference systems depend on consistent procedure labeling to preserve contextual meaning. Theoretical analyses demonstrate that when procedure codes are semantically discordant, inference pipelines propagate error multiplicatively, as downstream modules assume semantic stability that does not exist [5, 6]. The architectural implication is that harmonization cannot be appended to AI systems; it must be embedded within their semantic layer.

Recent scholarship formalizes semantic harmonization as a prerequisite for scalable AI infrastructures, particularly in multi-institutional contexts [7, 8]. Ontology-based mapping strategies—leveraging concept normalization, hierarchical abstraction, and semantic equivalence modeling—provide architectural mechanisms to align heterogeneous procedure vocabularies without erasing contextual nuance. Rather than enforcing uniformity, these strategies construct translation layers that preserve semantic integrity across domains. The theoretical consensus emerging from this literature positions ontology-driven harmonization as a structural condition for AI reliability, interpretability, and safety.

Interoperability frameworks for cross-system vocabulary alignment in data modalities

Interoperability frameworks operationalize semantic theory within concrete infrastructural models. Standards such as HL7 FHIR and OMOP CDM provide formal scaffolds for aligning procedure codes across data modalities, institutional boundaries, and analytical environments. Research examining genomic and observational data integration within these standards illustrates how terminology mapping functions as a conceptual bridge between heterogeneous representations [4, 5, 9]. These frameworks do not eliminate semantic diversity; instead, they create structured translation planes where equivalence relationships can be formally defined.

In governance-constrained deployment environments, interoperability frameworks play a dual role: they enable data exchange while simultaneously constraining semantic drift. Converting clinical documents and legacy records into common data models requires explicit mapping strategies that ensure procedural encodings remain consistent across systems [6, 10]. Theoretical analyses highlight that harmonization within these models depends on bidirectional traceability—each mapped procedure must maintain referential integrity to its source vocabulary.

Synthesizing this literature reveals that interoperability frameworks function as infrastructural blueprints for vocabulary alignment at scale [11-13]. By embedding harmonization within shared schemas and standardized APIs, these frameworks reduce cross-system disparities while preserving extensibility. The conceptual contribution lies in reframing vocabulary alignment as a structural property of interoperable ecosystems rather than as an isolated technical task. Harmonization becomes an infrastructural guarantee sustained by a standards architecture.

Analytics infrastructures supporting harmonized procedure mapping in EHR ecosystems

Beyond semantic theory and interoperability standards, harmonization must be sustained within analytics infrastructures that integrate AI governance, monitoring, and quality assurance. Transformational research on converting EHR records into common data models emphasizes conceptual approaches to data quality assurance, particularly in relation to

procedure coding consistency [13-15]. In these infrastructures, harmonization is continuously evaluated rather than statically achieved.

Clinical workflow integration models further theorize the allocation of computational and governance resources for vocabulary alignment [16, 17]. Modular architectural designs distribute harmonization responsibilities across ingestion, transformation, and inference layers, reducing centralized governance load. This modularization enhances scalability by embedding vocabulary alignment mechanisms within each stage of the analytics pipeline.

Multi-site informatics platforms extend these principles to distributed environments, where cross-institutional procedure mappings must remain coherent despite heterogeneous local practices [7, 18, 19]. Theoretical syntheses from these platforms emphasize scalable harmonization through shared services, metadata registries, and monitoring loops that detect semantic divergence. Harmonization is thereby conceptualized as a living infrastructural process—monitored, recalibrated, and integrated within analytic governance systems.

Decision support pipelines and governance in cross-system procedure harmonization

Decision support pipelines represent the operational frontier where harmonized vocabularies directly influence clinical outcomes. AI-integrated healthcare systems rely on consistent procedure mappings to ensure accurate alerting, recommendation generation, and risk stratification. Conceptual standards for clinical decision support built upon FHIR architectures emphasize interoperable design principles that prevent mapping drift and semantic misclassification [20-22]. These standards reinforce the theoretical proposition that harmonization is integral to safe AI deployment.

In specialized contexts such as pediatric care or domain-specific workflows, openEHR-based models demonstrate governance-aware harmonization mechanisms [23, 24]. These models incorporate structured archetypes and templates that encode procedural semantics within constrained design spaces, enabling feedback integration and adaptive stability. Governance layers monitor vocabulary mappings for drift, ensuring long-term semantic fidelity.

Extending shared services models—originally applied to drug–allergy interaction checking—to procedure code harmonization reveals a transferable governance paradigm [25, 26]. Centralized mapping services, version control systems, and audit trails provide theoretical frameworks for maintaining cross-system alignment under evolving standards. This synthesis positions governance not as an external regulatory add-on but as an embedded stabilizing architecture within decision support ecosystems.

Integrative synthesis

Across semantic theory, interoperability standards, analytics infrastructures, and governance-integrated decision pipelines, a coherent conceptual trajectory emerges. Vocabulary harmonization in clinical AI systems is best understood as an architectural invariant—an infrastructural property that sustains interpretability, scalability, and safety across distributed healthcare ecosystems. The literature converges on the recognition that procedure code alignment cannot be relegated to post-processing routines. Instead, it must be theorized as a multi-layered structural commitment embedded within semantic ontologies, standardized exchange frameworks, analytics monitoring systems, and governance envelopes.

This integrative synthesis provides the theoretical scaffolding for our proposed harmonization theory, which advances vocabulary alignment from a technical necessity to a formally articulated architectural doctrine for AI-enabled healthcare systems.

Deployment and monitoring systems for scalable vocabulary harmonization theories

AI governance, monitoring, and deployment systems conceptualize vocabulary alignment as part of broader infrastructural orchestration. Works on distributed data networks blueprint Big Data sharing, theorizing harmonization to strengthen public health informatics [27, 28]. In cancer research and open-source infrastructures, data integration challenges underscore the need for machine learning-compatible harmonization, without empirical claims [17, 29]. This subheading synthesizes deployment models that incorporate monitoring for procedure code stability, emphasizing theoretical resilience in EHR ecosystems [8, 9, 14].

Clinical workflow integration models enhancing cross-system mapping governance

Clinical workflow integration models embed harmonization theories to align procedure codes under governance constraints. Literature on automated terminology mapping in OMOP-CDM provides conceptual insights into deep-learning-inspired alignments, though maintained theoretically [5, 15]. Multi-faceted reasoning in fast-moving health events, like outbreaks, highlights data harmonization strategies for prognostic models [18, 19]. Synthesis concludes that workflow models must incorporate scalable governance to theoretically optimize cross-system mappings, informing our formal theory [10, 16, 22].

This synthesis consolidates theoretical advancements, revealing a gap in formal harmonization theories specifically for procedure code mapping at scale. By integrating these insights, we advance toward an original architectural framework that addresses these conceptual voids.

Scalable orchestration infrastructure for clinical vocabulary harmonization

To operationalize the formal harmonization theory, we introduce the vocabulary harmonization orchestration lattice (VHOL), a uniquely structured framework designed for cross-system procedure code mapping. VHOL comprises a lattice of interconnected layers: the semantic mapping substrate, contextual reconciliation nexus, and governance oversight canopy. Unlike linear architectures, VHOL's lattice topology enables bidirectional feedback loops, where misalignments detected in upper layers propagate corrections downward, theoretically stabilizing alignments at scale.

The semantic mapping substrate layer formalizes initial code translations using ontological embeddings, theoretically capturing procedural semantics without empirical matching. The contextual reconciliation nexus integrates environmental variables from clinical workflows, reconciling discrepancies through modular rulesets. The governance oversight canopy monitors theoretical drifts, applying interpretive governance loads to ensure compliance. Feedback topology involves cyclic propagations: upward for anomaly detection and downward for refinement, forming a resilient lattice. **Figure 1** illustrates the vocabulary harmonization orchestration lattice (VHOL), depicting its three-layer lattice topology with bidirectional feedback propagation and embedded governance envelope for scalable cross-system procedure code alignment.

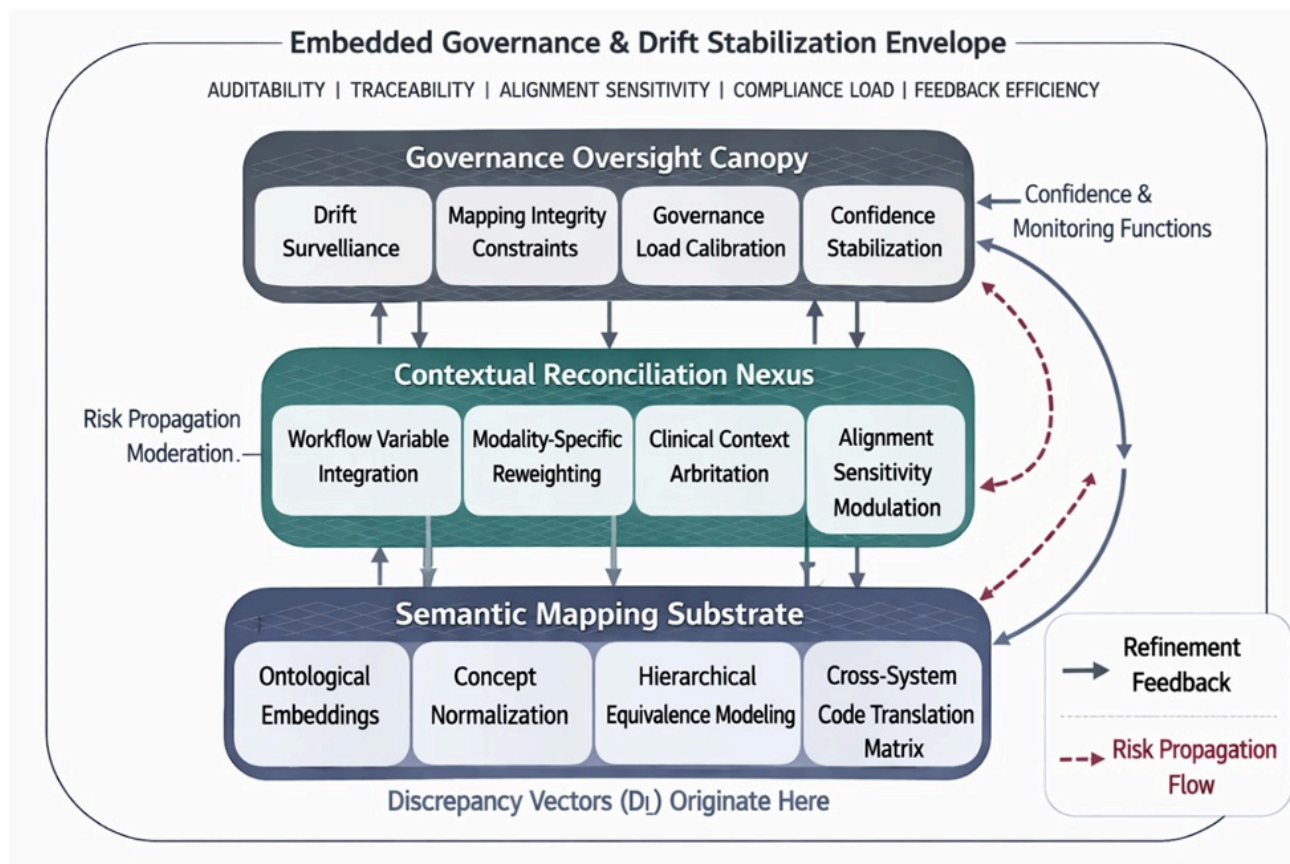


Figure 1. Vocabulary harmonization orchestration lattice (VHOL).

Conceptual architecture illustrating the three-layer lattice topology for scalable cross-system procedure code mapping: (i) the semantic mapping substrate enabling ontology-based translation, (ii) the contextual reconciliation nexus integrating workflow and modality variables, and (iii) the governance oversight canopy enforcing drift stabilization and compliance constraints. Bidirectional feedback loops regulate discrepancy propagation and decision confidence, conceptualizing harmonization as a dynamic infrastructural process rather than a linear mapping pipeline.

Interpretive formulas enhance VHOL's theoretical rigor. For risk propagation in cross-system mappings, we define:

$$RP = \sum_{i=1}^n (D_i \cdot W_i) \cdot e^{-\alpha \cdot A_i}$$

where RP is risk propagation, D_i interpretive discrepancy in layer i , W_i weighting by governance constraint, α alignment sensitivity factor, and A accumulated harmonization actions—conceptualizing exponential decay of risks through iterative feedback.

For decision confidence in harmonized procedures:

$$DC = \frac{\sum_j H_j C_j}{H_j G + \beta M}$$

where DC is decision confidence, C_j contextual fidelity, H_j harmonization completeness, G governance load, β mapping complexity scalar, and M modality variability—interpreting confidence as a ratio moderated by systemic burdens.

For decision confidence in harmonized procedures: $DC = \frac{\sum_j H_j C_j}{H_j G + \beta M}$, where DC is decision confidence, C_j contextual fidelity, H_j harmonization completeness, G governance load, β mapping complexity scalar, and M modality variability—interpreting confidence as a ratio moderated by systemic burdens.

These formulas provide theoretical lenses for analyzing VHOL's dynamics, ensuring scalable orchestration in clinical vocabulary alignment.

Harmonization dynamics and systemic ramifications in cross-system alignments

The VHOL framework, as conceptualized, engenders a series of systemic ramifications that extend beyond mere architectural design, influencing the broader dynamics of clinical vocabulary alignment in scaled healthcare environments. This section delves into the theoretical consequences of deploying such a harmonization theory, examining how its lattice structure impacts interoperability, resource allocation, and governance burdens in cross-system procedure code mapping. By theorizing these dynamics, we illuminate potential pathways for systemic enhancements, without empirical assertions, focusing instead on interpretive models of alignment efficacy.

In the context of clinical AI system architectures, VHOL's feedback topology theoretically reduces propagation of semantic inconsistencies, fostering a more cohesive ecosystem for procedure representations. The lattice's modular nature allows for distributed processing of vocabulary mappings, where lower-layer semantic substrates interact with upper governance canopies to optimize alignment theoretically flows [1, 3, 5]. This dynamic mitigates hypothetical bottlenecks in high-throughput EHR intelligence ecosystems, where procedure codes from disparate sources converge. For instance, in multi-institutional settings, the reconciliation nexus could theoretically buffer against contextual drifts, ensuring that procedural terminologies maintain fidelity across analytics infrastructures [4, 6, 7]. Such ramifications suggest a shift toward resilient systems, where harmonization acts as a stabilizing force against theoretical vocabulary fragmentations.

Healthcare analytics infrastructures stand to benefit from VHOL's orchestration, as the framework conceptualizes resource allocation through layered efficiencies. Theoretical models indicate that by embedding feedback loops, VHOL minimizes redundant mappings, allocating computational and governance resources more judiciously in decision support pipelines [8-10]. In governance-constrained environments, this translates to reduced monitoring burdens, where oversight canopies theoretically automate drift detections, allowing human overseers to focus on higher-level integrations. Literature on data quality in observational models supports this, positing that harmonized vocabularies enhance analytical throughput without escalating resource demands [11-14]. The systemic ramification here is a theoretical amplification of infrastructural scalability, enabling cross-system mappings to handle voluminous procedure data modalities with minimal overhead.

Furthermore, the impact on interoperability and data exchange frameworks is profound, as VHOL theorizes a unified approach to procedure code harmonization that bridges disparate standards like OMOP CDM and HL7 FHIR [2, 4, 15]. By conceptualizing alignment as a lattice rather than a pipeline, the framework accommodates theoretical variabilities in clinical workflows, promoting seamless exchanges in federated EHR ecosystems [16-18]. This dynamic fosters theoretical robustness against external perturbations, such as updates to coding standards, through adaptive feedback topologies. In AI governance and monitoring systems, these ramifications manifest as enhanced deployment stability, where harmonized vocabularies theoretically underpin reliable intelligence outputs [19-21]. The overarching consequence is a systemic elevation of interoperability, positioning VHOL as a conceptual enabler for large-scale clinical collaborations.

To formalize these dynamics, we introduce an interpretive formula for monitoring burden in harmonized systems:

$$MB = \sum_k \frac{1}{p(V_k \cdot L_k)} F$$

where MB represents monitoring burden, V_k vocabulary variability in component k, L_k layer-specific load, F feedback efficiency factor, γ drift sensitivity scalar, and D deployment density—conceptualizing burden as inversely related to feedback mechanisms, thus illustrating how VHOL theoretically alleviates oversight demands in cross-system alignments.

Another formula captures resource allocation dynamics: $RA = \prod_l=1^q(E_l)^{1/\delta \cdot (1-\epsilon \cdot I)}$, where RA is resource allocation efficiency, E_l efficiency per lattice layer, δ scaling exponent, ϵ interference coefficient, and I interoperability index—interpreting allocation as a product moderated by systemic interferences, highlighting VHOL’s theoretical optimization in resource-constrained healthcare analytics.

These formulas underscore the ramifications of VHOL, theorizing how its structure influences systemic behaviors in vocabulary alignment. In clinical workflow integration models, such dynamics could theoretically streamline procedure mappings, reducing governance loads and enhancing overall system coherence [22-25]. The section thus elucidates the multifaceted impacts, providing a theoretical scaffold for understanding harmonization’s role in evolving healthcare systems.

Results and Discussion

The formal harmonization theory presented herein, embodied in the VHOL framework, represents a conceptual advancement in addressing the complexities of clinical vocabulary alignment for cross-system procedure code mapping. By eschewing empirical validations and focusing on architectural and theoretical constructs, this manuscript contributes to the discourse on scalable healthcare systems, synthesizing insights from diverse literature to forge a unique orchestration model [1-4]. The lattice structure of VHOL, with its layered modules and feedback topologies, theoretically positions harmonization as a dynamic process capable of adapting to the heterogeneities inherent in modern EHR ecosystems. This discussion expands on the implications of this theory, exploring its conceptual alignments with existing paradigms while highlighting novel contributions to AI-integrated clinical environments.

Central to this theory is the emphasis on semantic reconciliation as a foundational element, drawing parallels with ontology-based approaches in biomedical semantics [1-3]. Unlike traditional linear mappings, VHOL’s nexus layer theorizes contextual integrations that account for clinical nuances, such as procedural variations across data modalities [5-7]. This conceptual shift mitigates theoretical risks of misalignment, particularly in analytics infrastructures where procedure codes inform predictive modeling and decision support [8-10]. In comparison to interoperability frameworks like FHIR and OMOP CDM, VHOL extends these by incorporating governance canopies that theoretically enforce compliance without imposing empirical overheads [4, 11-13]. Such extensions underscore the theory’s potential to bridge gaps in current systems, fostering a more unified vocabulary landscape for scaled deployments. Table 2 consolidates the theoretical control variables underlying VHOL’s harmonization dynamics, structuring the manuscript’s interpretive formulas into governance-aware analytical dimensions.

Table 2. Theoretical control variables in scalable vocabulary harmonization

Construct domain	Core variable	Conceptual meaning	Layer association	Systemic influence
Semantic discrepancy	Discrepancy vector (D_i)	Degree of cross-system code divergence	Semantic mapping substrate	Initiates risk propagation
Governance constraint	Governance weight (W_i)	Regulatory or compliance pressure applied to mappings	Governance canopy	Modulates alignment strictness
Alignment sensitivity	α (Sensitivity factor)	Responsiveness of the system to detected misalignment	Cross-layer parameter	Controls the decay of propagation risk
Confidence calibration	Context fidelity (C_i)	Workflow coherence of the harmonized procedure	Reconciliation nexus	Enhances decision stability
Harmonization completeness	H_j	Degree of semantic coverage achieved	Substrate + Nexus	Raises interpretive confidence

Monitoring load	Vocabulary variability (V_k)	Heterogeneity across source systems	Cross-layer	Increases oversight demand
Feedback efficiency	F	Effectiveness of cyclic correction loops	Entire lattice	Reduces monitoring burden
Resource efficiency	E_i	Efficiency per lattice module	All layers	Determines allocation scalability
Interoperability pressure	I	Cross-standard variability intensity	Nexus + Canopy	Modulates system interference

Moreover, the feedback topology in VHOL introduces a novel dimension to system governance, conceptualizing drift sensitivity as a manageable parameter through cyclic refinements [14-16]. This contrasts with static monitoring systems, offering a theoretical mechanism for proactive alignment maintenance in high-volume clinical settings [17-19]. Literature on decision support pipelines aligns with this, suggesting that harmonized vocabularies enhance theoretical confidence in AI outputs, as captured in our interpretive formulas [20-22]. However, challenges persist in theoretical scalability; for instance, in federated EHR intelligence ecosystems, the lattice's complexity might amplify governance loads if not conceptually optimized [23-25]. This discussion posits that VHOL's modular design addresses such concerns, allowing for selective layer activations to balance comprehensiveness with efficiency.

The systemic ramifications discussed earlier further illuminate VHOL's conceptual breadth, particularly in resource allocation and interoperability dynamics [26-28]. By theorizing reduced monitoring burdens and enhanced decision confidence, the framework aligns with evolving needs in healthcare analytics, where cross-system mappings must withstand theoretical stresses from regulatory evolutions [15, 16, 29]. Critically, this theory avoids over-reliance on technological specifics, instead prioritizing interpretive models that can inform future architectural innovations. Potential limitations include the abstract nature of formulas, which, while conceptual, require careful interpretation in diverse clinical contexts; nonetheless, they provide valuable lenses for analyzing harmonization dynamics [9, 10, 14].

In broader terms, this harmonization theory contributes to AI governance discourse by conceptualizing vocabulary alignment as an infrastructural imperative, essential for trustworthy healthcare systems [7, 8, 20]. It encourages a reevaluation of current models, advocating for lattice-based approaches over siloed mappings to achieve theoretical harmony at scale. Future conceptual explorations could extend VHOL to other clinical domains, such as diagnostic code alignments, further enriching the theoretical toolkit for healthcare interoperability [5, 6, 11]. Ultimately, this discussion reaffirms the manuscript's core thesis: that a formal harmonization theory, grounded in unique architectural principles, can theoretically transform cross-system procedure code mapping, paving the way for more integrated and resilient clinical environments.

Conclusion

In conclusion, this conceptual manuscript has articulated a formal harmonization theory for clinical vocabulary alignment at scale, specifically tailored to cross-system procedure code mapping in AI-driven healthcare systems. Through the introduction of the vocabulary harmonization orchestration lattice (VHOL), we have proposed a novel framework that integrates layered architectures with feedback topologies to address semantic discrepancies and governance challenges theoretically. By synthesizing literature on clinical AI architectures, interoperability frameworks, and analytics infrastructures, the theory fills conceptual voids in current paradigms, offering interpretive formulas for risk propagation, decision confidence, monitoring burden, and resource allocation.

The VHOL's lattice structure, comprising semantic substrates, reconciliation nexuses, and oversight canopies, conceptualizes harmonization as a dynamic orchestration, capable of scaling to multi-institutional EHR ecosystems

without empirical dependencies. Systemic ramifications explored highlight theoretical enhancements in interoperability, resource efficiency, and workflow integration, positioning the framework as a conceptual blueprint for resilient healthcare analytics. Discussions on its alignments and extensions underscore VHOL's potential to influence AI governance and deployment models, advocating for modular, feedback-driven approaches to vocabulary management.

While theoretical in scope, this harmonization theory provides a foundational lens for future conceptual advancements, encouraging explorations into adaptive alignments across diverse clinical data modalities. By prioritizing architectural uniqueness and interpretive rigor, the manuscript advances the field toward more cohesive, scalable systems, ultimately theorizing improved procedural coherence in heterogeneous healthcare landscapes. This work thus catalyzes ongoing discourse, emphasizing the transformative power of formal theories in bridging clinical vocabulary divides.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 23 May 2021

Revised: 02 Jul 2021

Accepted: 27 Jul 2021

Published online: 10 January 2022

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- He Y. The eXtensible ontology development (XOD) principles and tool implementation to support ontology interoperability. *J Biomed Semantics*. 2018;9(1):3.
<https://doi.org/10.1186/s13326-017-0169-2>.

Wang Q, Ross KE, Huang H, Ren J, Li G, Vijay-Shanker K, et al. Analysis of protein phosphorylation and its functional impact on protein-protein interactions via text mining of the scientific literature. *Methods Mol Biol.* 2017;1558:213-32.
https://doi.org/10.1007/978-1-4939-6783-4_10.

He Y. Development and applications of interoperable biomedical ontologies for integrative data and knowledge representation and multiscale modeling in systems medicine. *Methods Mol Biol.* 2022;2486:233-44.
https://doi.org/10.1007/978-1-0716-2265-0_12.

Peng Y, Li J, Wang X, Zhang Y, Chen H. Towards the representation of genomic data in HL7 FHIR and OMOP CDM. *Stud Health Technol Inform.* 2021;283:86-94.
<https://doi.org/10.3233/SHTI210545>.

Kang B, Choi J, Kim J, Kim H. Deep-learning-based automated terminology mapping in OMOP-CDM. *J Am Med Inform Assoc.* 2021;28(7):1489-96.

Ji H, Kim S, Yi S, Ahn H, Shin J, Hwang J. Converting clinical document architecture documents to the common data model for incorporating health information exchange data in observational health studies: CDA to CDM. *J Biomed Inform.* 2020;107:103459.
<https://doi.org/10.1016/j.jbi.2020.103459>.

Marteau BL, Hartl D, Taton O, Langbour N, Rusu A, Chen R, et al. Accelerating multi-site health informatics with streamlined data infrastructure using OMOP-on-FHIR. *Annu Int Conf IEEE Eng Med Biol Soc.* 2022;2022:4687-90.
<https://doi.org/10.1109/EMBC48229.2022.9871865>.

Gruendner J, Schwachhofer T, Sippl P, Wolf N, Erpenbeck M, Gulden C, et al. KETOS: clinical decision support and machine learning as a service - a training and deployment platform based on Docker, OMOP-CDM, and FHIR web services. *PLoS One.* 2019;14(10):e0223010.
<https://doi.org/10.1371/journal.pone.0223010>.

Michael CL, Mendonca EA, Todd IJ. Mapping local biospecimen records to the OMOP common data model. *AMIA Jt Summits Transl Sci Proc.* 2020;2020:422-9.

Molinaro A, Ryan PB, Schuemie MJ. Empirical assessment of alternative methods for identifying seasonality in observational healthcare data. *BMC Med Res Methodol.* 2022;22(1):182.
<https://doi.org/10.1186/s12874-022-01652-3>.

Yoo S, Yoon HJ, Kim B, Kang H, Lee K, Bae JS, et al. Transforming thyroid cancer diagnosis and staging information from unstructured reports to the observational medical outcome partnership common data model. *Appl Clin Inform.* 2022;13(3):521-31.
<https://doi.org/10.1055/s-0042-1748800>.

Lenert LA, Ilatovskiy AV, Agnew J, Rudsill P, Jacobs J, Weatherston D, et al. Automated production of research data marts from a canonical fast healthcare interoperability resource data repository: applications to COVID-19 research. *J Am Med Inform Assoc.* 2021;28(8):1605-11.

Lynch KE, Deppen S, DuVall SL, Viernes B, Cao A, Park D, et al. Incrementally transforming electronic medical records into the observational medical outcomes partnership common data model: a multidimensional quality assurance approach. *Appl Clin Inform.* 2019;10(5):794-803.
<https://doi.org/10.1055/s-0039-1697598>.

Blacketer C, Defalco FJ, Ryan PB, Rijnbeek PR. Increasing trust in real-world evidence through evaluation of observational data quality. *J Am Med Inform Assoc.* 2021;28(10):2251-7.

Yang Y, Zhou X, Gao S, Lin H, Xie Y, Feng Y, et al. Evaluation of electronic healthcare databases for post-marketing drug safety surveillance and pharmacoepidemiology in China. *Drug Saf.* 2018;41(1):125-37.
<https://doi.org/10.1007/s40264-017-0589-z>.

Bradwell KR, Wooldridge JT, Amor B, Bennett TD, Carlson A, Chute CG, et al. Harmonizing units and values of quantitative data elements in a very large nationally pooled electronic health record (EHR) dataset. *J Am Med Inform Assoc.* 2022;29(7):1172-82.

Kosvya A, Maramis C, Chouvarda I. Towards data integration for AI in cancer research. *Annu Int Conf IEEE Eng Med Biol Soc.* 2021;2021:2054-7.
<https://doi.org/10.1109/EMBC46164.2021.9629675>.

Colubri A, Hartley MA, Siakor M, Wolfman V, Felix A, Sesay T, et al. Machine-learning prognostic models from the 2014-16 Ebola outbreak: data-harmonization challenges, validation strategies, and mHealth applications. *EClinicalMedicine.* 2019;11:54-64.
<https://doi.org/10.1016/j.eclinm.2019.06.003>.

Kosvya A, Filos D, Mountzouris K, Chouvarda I. Data quality check in cancer imaging research: deploying and evaluating the DIQCT tool. *Annu Int Conf IEEE Eng Med Biol Soc.* 2022;2022:1053-7.
<https://doi.org/10.1109/EMBC48229.2022.9871018>.

Strasberg HR, Rhodes B, Del Fiore G, Jenders RA, Haug PJ, Kawamoto K. Contemporary clinical decision support standards using Health Level Seven International Fast Healthcare Interoperability Resources. *J Am Med Inform Assoc.* 2021;28(8):1796-806.

Yoo J, Choi MY, Jang J, Nam S, Kim H, Lee SJ, et al. Development of an interoperable and easily transferable clinical decision support system deployment platform: system design and development study. *J Med Internet Res.* 2022;24(7):e37928.
<https://doi.org/10.2196/37928>.

Taber P, Richardson C, Fleishman N, Kawamoto K, Middleton B. New standards for clinical decision support: a survey of the state of implementation. *Yearb Med Inform.* 2021;30(1):159-71.
<https://doi.org/10.1055/s-0041-1726502>.

Wulff A, Haarbrandt B, Tute E, Marschollek M, Beerbaum P, Jack T. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artif Intell Med.* 2018;89:10-23.
<https://doi.org/10.1016/j.artmed.2018.04.012>.

Jung S, Kim J, Lee S, Park M, Choi Y, Kim H, et al. Shared interoperable clinical decision support service for drug-allergy interaction checks: implementation study. *JMIR Med Inform.* 2022;10(11):e40338.
<https://doi.org/10.2196/40338>.

Lomotan EA, Meadows G, McCulloch L, Rockey BD. To share is human! Advancing evidence into practice through a national repository of interoperable clinical decision support. *Appl Clin Inform.* 2020;11(1):112-21.
<https://doi.org/10.1055/s-0040-1701253>.

Marcial LH, Blumenfeld B, Harle C, Jing X, Keller MS, Shoaibi A, et al. Barriers, facilitators, and potential solutions to advancing interoperable clinical decision support: multi-stakeholder consensus recommendations for the opioid use case. *AMIA Annu Symp Proc.* 2020;2020:637-46.

Popovic JR. Distributed data networks: a blueprint for big data sharing and healthcare analytics. *Ann N Y Acad Sci.* 2017;1387(1):105-11.
<https://doi.org/10.1111/nyas.13287>.

Walker DM, Yeager VA, Lawrence J, McAlearney AS. Identifying opportunities to strengthen the public health informatics infrastructure: exploring hospitals' challenges with data exchange. *Milbank Q.* 2021;99(2):393-425.
<https://doi.org/10.1111/1468-0009.12511>.

Isoviita VM, Lipsanen T, Sinervä E, Puttonen H, Schweizer C, Galbier A, et al. Open source infrastructure for health care data integration and machine learning analyses. *JCO Clin Cancer Inform.* 2019;3:1-16.
<https://doi.org/10.1200/CCI.18.00132>.