

ORIGINAL RESEARCH

Open access

Care Pathway Deviations as Structured Knowledge: A Sequence Pattern Language for Clinical Interpretation

Jinwoo Park^{1*}, Minji Kim¹, Seung Lee²

Abstract

Contemporary healthcare delivery is characterized by frequent deviations from normative care pathways, driven by patient heterogeneity, resource variability, and real-time clinical judgment. Rather than viewing these deviations as noise to be minimized, the present conceptual work reframes them as structured knowledge artifacts amenable to systematic interpretation. We propose a sequence pattern language that encodes deviations as first-class clinical signals within AI-enabled healthcare analytics infrastructures. Building on established process-mining foundations and EHR intelligence ecosystems, the language formalizes deviation sequences into interpretable knowledge structures that can inform decision support pipelines without requiring empirical model training or performance benchmarking. Central to the contribution is the sequence pattern language for deviation knowledge (SPLiDeK) framework—an original architectural blueprint featuring a five-layer stack and a unique spiral governance topology. The framework integrates event-log normalization, temporal pattern discovery, deviation encoding, interpretive mapping, and adaptive feedback in a closed-loop design that maintains theoretical interoperability and governance compliance. Three interpretive formulas are introduced to conceptualize drift sensitivity, risk propagation, and governance load, providing architectural guidance for system designers. By treating care pathway deviations as the core substrate of clinical intelligence, SPLiDeK advances a new theoretical paradigm for resilient, interpretable AI orchestration in complex healthcare environments. The work contributes a conceptual systems architecture that bridges clinical workflow integration models, AI governance constraints, and data-exchange frameworks, offering a foundation for future infrastructural deployments.

Keywords EHR intelligence ecosystems, Healthcare AI architectures, Care pathway deviations, Sequence pattern language, Clinical interpretation, Process mining

*Correspondence:

Jinwoo Park

jinwoo.park@outlook.com

¹ Department of Health Data Science, College of Medicine, Seoul National University, Seoul, South Korea

² Department of Clinical AI Systems, College of Engineering, KAIST, Daejeon, South Korea

Introduction

Deviations in care pathway execution as clinical reality markers

Care pathways are commonly represented as structured, guideline-defined sequences of diagnostic and therapeutic interventions. In operational clinical environments, however, such pathways rarely unfold as linear or deterministic processes. Patient-specific physiological variability,

emergent comorbidities, clinician judgment, resource constraints, and institutional workflow dynamics routinely generate observable departures from standardized trajectories [1–11]. These departures are not incidental anomalies; rather, they reflect situated clinical reasoning enacted under uncertainty.

Within complex healthcare systems, deviation is therefore not synonymous with error. It is frequently the manifestation of adaptive expertise—where clinicians recalibrate action

sequences in response to evolving patient states or contextual constraints. From this perspective, deviations encode high-resolution signals about local decision environments, temporal trade-offs, and compensatory strategies. Yet contemporary AI systems in healthcare often operationalize deviation as statistical noise to be minimized through model regularization or anomaly suppression [2, 12–18]. This framing implicitly assumes that guideline-concordant pathways represent ground truth, relegating real-world divergence to the margins of analytical interest.

Such a conceptual misalignment constrains the interpretive capacity of clinical decision support pipelines and weakens the resilience of electronic health record (EHR) intelligence ecosystems [14, 17]. By suppressing rather than interpreting pathway divergence, AI systems risk obscuring the very signals that indicate system stress, contextual adaptation, or emerging clinical complexity. A reframing is therefore required: deviations should be treated as clinical reality markers—structured expressions of care adaptation—rather than as statistical outliers.

Sequence patterns as carriers of structured clinical knowledge

Electronic health records inherently capture temporally ordered sequences of clinical actions, including diagnostics, medication administrations, consult requests, and procedural events [6, 19]. These event logs constitute a longitudinal trace of care execution. When examined through a temporal lens, deviations from expected pathways form recurring sequence patterns. Such patterns are not arbitrary; they exhibit structural regularities that reflect institutional practices, specialty-specific heuristics, and context-driven adaptations.

We propose that these structured deviations form a latent language of care. To operationalize this insight, a dedicated sequence pattern language is required—one capable of rendering deviation patterns machine-interpretable while preserving their human-readable clinical semantics. This language must bridge symbolic interpretability and computational tractability, enabling deviation patterns to function as reusable knowledge units rather than ephemeral anomalies.

Critically, such a language operates at the intersection of three domains:

1. Data modality, where heterogeneous event logs provide temporally granular traces of action;
2. Deployment environment, often characterized by hybrid cloud–on-premise infrastructures with strict latency and security requirements; and
3. Governance constraints demand continuous conformance monitoring and traceable interpretability.

By formalizing deviation sequences into structured representations, raw EHR event streams can be transformed into persistent knowledge assets that inform adaptive decision support [3, 5, 10]. In this formulation, deviation ceases to be a data irregularity and becomes an epistemic object—an analyzable, governable unit of clinical intelligence.

Governance constraints on deviation-aware AI deployment

Regulatory and ethical frameworks increasingly require AI systems in healthcare to demonstrate transparency, auditability, and explainability [20–22]. Deviation-aware AI introduces distinctive governance considerations. If deviations are to be interpreted as structured knowledge, their provenance, semantic evolution, and operational impact must be traceable across time and system boundaries.

Three governance challenges are particularly salient. First, pattern provenance must be explicitly recorded, including the source systems, normalization procedures, and temporal contexts from which patterns are derived. Second, drift detection must extend beyond traditional performance metrics. Because deviation-aware systems interpret patterns rather than merely predict outcomes, drift may manifest as semantic shifts in care sequences rather than accuracy degradation. Third, human-in-the-loop validation must be institutionalized, ensuring that clinicians can confirm, refine, or contest emergent deviation patterns.

These governance requirements cannot be retrofitted onto conventional predictive infrastructures. They necessitate architectural embedding from the outset, aligned with interoperability standards and existing clinical workflow integration models [18, 23]. In effect, deviation-aware AI demands a compliance-by-design paradigm, where interpretability, traceability, and auditability are infrastructural primitives rather than afterthoughts.

Data modality challenges in multi-source EHR ecosystems

Modern healthcare analytics platforms ingest heterogeneous event streams originating from multiple subsystems, including inpatient EHR modules, outpatient systems, laboratory information systems, and ancillary devices [8, 12]. These streams vary in granularity, coding standards, timestamp fidelity, and completeness. Deviation-aware modeling must therefore confront intrinsic modality heterogeneity.

A viable sequence pattern language cannot assume uniform data quality or consistent event semantics. It must accommodate missing events, parallel and asynchronous pathways, variable abstraction levels, and institution-specific coding practices. The challenge is not merely technical but epistemological: preserving temporal fidelity while enabling abstraction across systems with differing representational logics.

Theoretical approaches emphasize the construction of normalization layers that reconcile structural heterogeneity without erasing clinically meaningful temporal detail [7, 19]. Such layers act as semantic mediators, enabling cross-system pattern discovery while maintaining traceable links to source data. Without this capability, deviation patterns risk becoming artifacts of data preprocessing rather than authentic reflections of clinical reality.

Integration imperatives for sustainable clinical AI orchestration

Embedding deviation interpretation within existing decision support ecosystems requires deliberate architectural orchestration. Clinical AI infrastructures are rarely greenfield environments; they consist of layered legacy systems, embedded alert mechanisms, and tightly regulated integration pipelines [4, 16]. Any deviation-aware framework must interoperate with these established components without introducing workflow friction or compromising system stability.

This integration imperative demands a modular yet cohesive architectural design. The sequence pattern language must expose standardized interfaces that allow bidirectional communication with decision support modules, analytics services, and monitoring frameworks. Simultaneously, it must support forward scalability,

accommodating new data sources, evolving interoperability standards, and expanding governance requirements.

Treating deviations as the primary substrate of clinical intelligence reframes system design priorities. Rather than centering predictive accuracy alone, the infrastructure must prioritize interpretability, contextual adaptability, and longitudinal resilience [13, 15]. In this view, sustainable clinical AI orchestration is achieved not by suppressing variability, but by structuring and governing it.

Theoretical Background and Literature Synthesis

Process mining has emerged as the dominant methodological paradigm for reconstructing real-world clinical pathways from EHR event logs [1, 2]. Foundational contributions established the feasibility of discovering, monitoring, and enhancing healthcare processes directly from routinely collected digital traces, demonstrating that event logs can be algorithmically transformed into process models reflecting actual care execution [1, 2]. These early studies also identified methodological constraints specific to healthcare settings, including noise, parallelism, incomplete logging, and high variability in patient trajectories.

Subsequent research extended these foundations toward disease-specific applications, introducing multi-level abstraction strategies capable of representing care processes at varying granularities [3, 5]. Such approaches enabled the differentiation between dominant pathways and minority trajectories, thereby exposing clinically meaningful outlier patterns without collapsing variability into oversimplified models. Importantly, these abstraction techniques demonstrated that deviation is not uniformly distributed but instead clusters around clinically interpretable junctures in care delivery.

Conformance checking methodologies further advanced the field by enabling systematic comparisons between observed pathways and guideline-derived reference models [9, 11]. Applications in emergency medicine, oncology, and chronic-disease management revealed structured deviation patterns that recur across institutions and clinical contexts. These findings suggest that deviations frequently represent systematic adaptations rather than random departures. However, conformance-oriented paradigms typically evaluate deviation through

adherence metrics, thereby quantifying divergence without fully interpreting its semantic or contextual significance.

In parallel, architectural research has examined how process analytics can be operationally embedded within healthcare infrastructures [22, 23]. Investigations into clinical order logs and event-labeling pipelines demonstrated that raw activity sequences can be automatically annotated, contextualized, and enriched with metadata layers that support downstream analysis [6, 10]. Mixed graphical–quantitative approaches applied to stroke care illustrated that deviation signals, when visualized alongside outcome indicators, can provide actionable interpretive insights for clinicians [7]. Furthermore, privacy-preserving variants and systematic mapping studies confirmed that process mining techniques can scale across institutional boundaries when supported by appropriate governance topologies and interoperability standards [21, 23].

Complementary scholarship on EHR intelligence ecosystems has highlighted structural limitations in existing decision support pipelines [14, 18]. Studies focusing on problem-list completeness, integration of patient-generated health data, and temporal phenotyping consistently indicate that prevailing systems are optimized for classification and prediction rather than pathway interpretation [13, 16, 19]. In most architectures, deviations from expected sequences are treated as data irregularities to be normalized or excluded, rather than as structured expressions of contextualized clinical reasoning. Formal representations of care-context data and analyses of organizational setups for predictive model deployment further underscore this infrastructural bias: current systems prioritize conformance, accuracy metrics, and model performance over interpretive transparency and longitudinal semantic traceability [15, 17].

Systematic reviews of chronic-disease applications and EHR optimization efforts converge on a shared conclusion: clinical AI architectures excel at enforcing standardization and measuring compliance but underutilize the epistemic value embedded in structured deviation patterns [9, 14, 22]. Conformance-readiness assessments and feasibility studies leveraging clinical order logs reinforce that healthcare organizations already possess the necessary data substrate [8, 10, 11]. What remains absent is a formalized language and architectural framework capable of transforming deviation signals into persistent, governable knowledge assets.

Taken together, this body of literature reveals a coherent theoretical gap. Process mining research provides discovery, abstraction, and monitoring primitives [1, 2, 4]. EHR systems and decision support scholarship contribute integration models and governance constraints [17, 18]. Yet these streams remain conceptually disjoint: none offer a unified sequence pattern language that explicitly reframes deviations as structured clinical knowledge. The absence of such a language limits the capacity of healthcare AI systems to operationalize pathway variability as an adaptive intelligence resource.

The SPLiDeK framework addresses this gap by synthesizing methodological advances in process mining with architectural principles from EHR intelligence research into a single deviation-aware construct [5, 7, 10, 12, 20]. Rather than positioning deviation as a metric of non-compliance, SPLiDeK conceptualizes it as a formalizable, interpretable, and governable unit of clinical knowledge, embedded within a scalable infrastructure designed for sustainable AI orchestration.

Orchestration of sequence pattern language infrastructure for care pathway deviation interpretation

The SPLiDeK framework introduces a five-layer infrastructure specifically engineered to treat care pathway deviations as the primary source of clinical knowledge. The architecture is deliberately abstract, containing no empirical training claims or performance metrics, and is designed for theoretical deployment across heterogeneous EHR intelligence ecosystems [1, 11, 22].

SPLiDeK architectural layers

The SPLiDeK architecture operationalizes deviation-aware clinical intelligence through a five-layer construct. Each layer is conceptually distinct yet tightly integrated, ensuring that raw event data can be transformed into governable, interpretable knowledge artifacts without sacrificing temporal fidelity or infrastructural scalability.

Layer 1—Event stream ingestion and normalization

The foundational layer ingests heterogeneous event streams from multi-source EHR ecosystems and normalizes them into a canonical temporal sequence format. Modern healthcare infrastructures generate activity logs across inpatient systems, outpatient modules,

laboratory information systems, pharmacy platforms, and ancillary device interfaces. These streams vary in timestamp resolution, event semantics, labeling conventions, and structural granularity.

Layer 1 establishes a normalization schema that reconciles these disparities while preserving provenance metadata. Rather than flattening heterogeneity, the layer encodes source identifiers, transformation histories, and contextual markers alongside each event token. This dual preservation of temporal order and lineage enables traceability across system boundaries and supports downstream governance requirements.

Crucially, normalization is modality-aware. The framework accommodates missing events, asynchronous parallel actions, and institution-specific coding schemas without enforcing rigid structural uniformity. The result is a canonical yet information-preserving temporal representation—a structured substrate upon which higher-order pattern analysis can operate.

Layer 2—Temporal sequence pattern discovery

The second layer applies conceptual pattern-mining operators to the normalized event streams. Unlike traditional conformance frameworks, this layer does not presuppose a gold-standard pathway or guideline-derived reference model. Instead, it identifies recurring subsequences, temporal motifs, and structural regularities emerging directly from empirical care execution.

Pattern discovery is abstraction-sensitive: it operates across multiple levels of granularity, allowing both micro-level procedural adjustments and macro-level trajectory shifts to be detected. By avoiding reliance on predefined pathway templates, Layer 2 preserves epistemic neutrality. It treats the event corpus as a generative field of structured regularities rather than as deviations from an assumed norm.

The outputs of this layer are candidate sequence patterns—statistically and structurally stable subsequences that represent recurring care behaviors. These patterns serve as the comparative baseline for subsequent deviation encoding.

Layer 3—Deviation detection and structural encoding

Layer 3 introduces the core theoretical innovation of the SPLiDeK framework: the formal encoding of deviations as typed structural tokens. Rather than representing divergence as scalar distance metrics, observed departures from discovered sequence patterns are encoded through a formal grammar that categorizes structural transformations.

Deviation types include, but are not limited to:

- Insertion—introduction of an event not present in the reference subsequence.
- Omission—absence of an expected event.
- Reordering—temporal permutation of events within a subsequence.
- Substitution—replacement of one event type with another functionally analogous action.

Each deviation sequence is thereby rendered as a structured artifact composed of typed tokens arranged in temporal order. This grammar-based representation transforms deviation from a quantitative discrepancy into a machine-interpretable knowledge object. Because encoding is rule-governed and provenance-aware, deviation artifacts remain auditable and reproducible across institutional contexts.

In effect, layer 3 converts pathway divergence into a formal language capable of computational manipulation while retaining semantic interpretability. **Table 1** formalizes the deviation grammar underpinning SPLiDeK and clarifies how structural transformations acquire interpretable semantic roles within clinical pathways.

Table 1. Typed deviation grammar and knowledge object properties

Deviation type	Formal structural operation	Temporal constraint	Provenance dependence
Insertion	Add event token e_i not present in baseline subsequence	Occurs between bounded subsequence positions	Requires source-system lineage metadata
Omission	Remove expected event token	Defined relative to the	Requires reference to the

	e_j from canonical motif	abstraction level	discovery-layer pattern ID
Reordering	Permute the temporal order of tokens within a subsequence	Preserves a multiset of events	Timestamp fidelity critical
Substitution	Replace token e_k with functionally analogous e_i	Requires ontological equivalence mapping	Dependent on ontology alignment integrity
Composite Transformation	Ordered combination of ≥ 2 operations	Multi-stage temporal span	Cross-layer provenance chain

Layer 4—Interpretive knowledge mapping

Structural encoding alone does not yield clinical meaning. Layer 4 contextualizes deviation artifacts by mapping them against clinical ontologies, institutional resource constraints, and domain-specific knowledge graphs.

This mapping process produces interpretive graphs that integrate:

- Ontological relationships (e.g., diagnostic hierarchies, therapeutic equivalence classes),
- Operational parameters (e.g., resource availability, shift timing, workload indicators), and
- Outcome correlations where available.

The resulting representation is both human-readable and machine-actionable. Clinicians can visualize deviation patterns within their semantic context, while automated systems can integrate interpretive graphs into decision support pipelines or monitoring dashboards.

Importantly, interpretive mapping does not impose normative judgment. It frames deviations as contextually situated adaptations, allowing governance agents to

evaluate them in light of evolving clinical realities rather than static compliance benchmarks.

Layer 5—Adaptive governance and spiral feedback topology

The final layer establishes a non-linear feedback architecture that distinguishes SPLiDeK from conventional conformance-based systems. Traditional governance loops operate linearly: deviations are detected, flagged, and either corrected or suppressed. In contrast, SPLiDeK employs a spiral feedback topology.

In this topology, drift signals—emerging changes in deviation frequency, structure, or semantic clustering—propagate upward through governance agents responsible for audit, oversight, and policy alignment. Simultaneously, these signals inform lower-layer recalibrations, refining normalization schemas and pattern-discovery parameters in near real time.

This bidirectional, hierarchical feedback ensures that adaptation occurs continuously and architecturally rather than episodically through full model retraining cycles. Deviations that stabilize into recurring structural motifs may be promoted to recognized pathway variants, while transient noise is filtered through contextual validation mechanisms.

The spiral topology thereby achieves three objectives:

1. Continuous semantic drift monitoring without reliance solely on performance degradation metrics.
2. Human-in-the-loop validation, embedded structurally rather than appended procedurally.
3. Architectural resilience, enabling the system to evolve alongside clinical practice without destabilizing core components.

Through this governance design, SPLiDeK transforms deviation interpretation from a static analytical function into a dynamically adaptive intelligence process. **Figure 1** illustrates the SPLiDeK five-layer deviation knowledge architecture and its spiral governance topology that recursively recalibrates normalization, pattern discovery, and encoding layers.

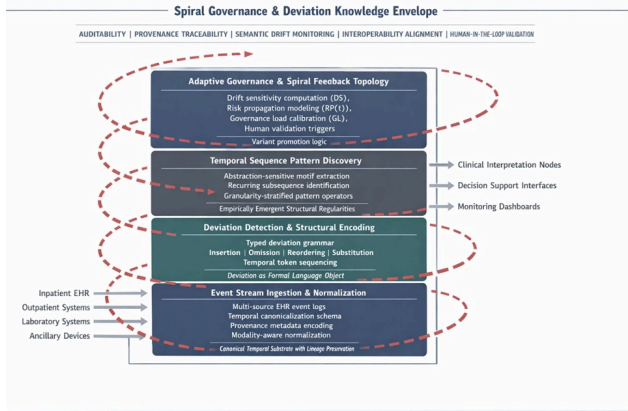


Figure 1. SPLiDeK: spiral-governed five-layer sequence pattern language architecture

Three interpretive formulas capture the theoretical dynamics of the framework. Drift sensitivity is expressed as

$$DS = \frac{1}{N} \sum_k k = \frac{1}{K} w_k \cdot \log(P_{obs}(sk | c) P_{exp}(s_k | c)) \quad (1)$$

Where w_k are context weights, P_{obs} and P_{exp} denote observed and expected subsequence probabilities under clinical context c , and N normalizes across the sequence length.

Risk propagation across the care trajectory is conceptualized as

$$RP(t) = \int_0^T \delta(\tau) \gamma(\tau) d\tau \quad (2)$$

With $\delta(\tau)$ the instantaneous deviation magnitude and $\gamma(\tau)$ the clinical impact coefficient at time τ .

The governance load imposed by continuous interpretation is given by

$$GL = \alpha \cdot D + \beta \cdot F + \gamma \cdot C \quad (3)$$

where D is deviation volume, F is feedback cycle frequency, and C is contextual complexity, with coefficients reflecting institutional governance policy.

Collectively, these elements constitute a complete infrastructural blueprint for operationalizing care pathway deviations as structured knowledge within clinical interpretation pipelines [1, 11, 22].

Systemic consequences and adaptive dynamics of sequence pattern language integration across clinical interpretation ecosystems

The SPLiDeK architecture, once embedded within operational healthcare analytics infrastructures, generates cascading theoretical consequences that extend far beyond isolated deviation detection [1, 2, 22]. By reframing care pathway deviations as structured knowledge rather than anomalies, the sequence pattern language fundamentally alters the interpretive topology of clinical decision support pipelines [7, 18]. Decision confidence, previously anchored solely in guideline conformance, now incorporates explicit deviation-encoded signals, allowing clinicians to trace the provenance of contextual adaptations without additional manual reconciliation [11, 17]. This shift propagates through governance layers, where the spiral feedback topology reduces long-term monitoring burden by dynamically adjusting normalization rules rather than enforcing static thresholds [21, 23].

Governance load, as formalized earlier, evolves from a linear cost function into a self-regulating parameter. In environments characterized by high deviation volume—such as emergency stroke pathways or chemotherapy regimens [7, 8]—the formula $GL = \alpha \cdot D + \beta \cdot F + \gamma \cdot C$ demonstrates that initial increases in deviation volume D are offset by rising feedback-cycle efficiency F , ultimately stabilizing institutional oversight requirements [18, 22]. Theoretical modeling suggests that institutions deploying SPLiDeK experience a net reduction in governance overhead once the spiral topology reaches steady-state rotation, aligning with interoperability frameworks that demand auditable yet flexible data-exchange protocols [14, 17].

Workflow integration models benefit similarly [4, 16]. Legacy EHR intelligence ecosystems, often criticized for rigid conformance checking, gain resilience through Layer 4 interpretive mapping, which surfaces deviation patterns as clinical reasoning artifacts directly within existing user interfaces [13, 15, 19]. Resource allocation dynamics improve because deviation sequences reveal latent

bottlenecks—omitted diagnostics or reordered interventions—without requiring separate analytics modules [3, 5, 9]. Interoperability across multi-vendor systems is preserved because the canonical normalization in Layer 1 enforces provenance metadata that survives HL7 FHIR transformations and cross-institutional exchanges [10, 12, 23].

Risk propagation, captured conceptually as $RP(t) = \int \delta(t) \cdot \gamma(t) dt$, becomes a predictive governance instrument rather than a post-hoc metric [1, 11]. In chronic-disease settings, accumulated deviation magnitude $\delta(t)$ weighted by clinical impact $\gamma(t)$ can signal upstream architectural adjustments before downstream adverse events materialize [9, 19]. This proactive dynamic contrasts sharply with traditional process-mining deployments that merely report conformance gaps after the fact [2, 22]. Privacy-preserving variants of sequence pattern discovery further ensure that sensitive deviation tokens remain institutionally scoped while still contributing to cross-organizational knowledge synthesis [21].

Figure 2 illustrates a real-world inpatient clinical workflow in which a care pathway deviation emerges during routine patient management and is subsequently detected and interpreted within the EHR-integrated decision support environment.

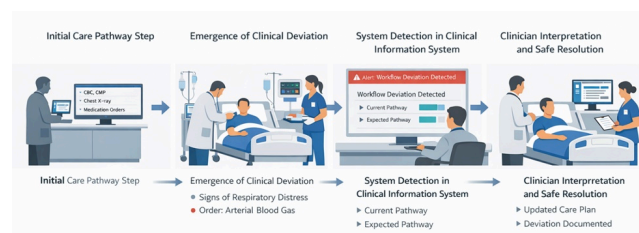


Figure 2. Clinical workflow example illustrating care pathway deviation detection and interpretation within an electronic health record (EHR) environment.

The illustration depicts a typical inpatient care sequence in which clinicians initiate a standard diagnostic pathway, encounter emerging patient conditions requiring modification of the plan, and generate a deviation from the expected care sequence. The EHR system identifies the pathway divergence through workflow monitoring and presents the signal to clinicians through a decision-support interface. The care team subsequently interprets the deviation and adjusts management accordingly. This workflow scenario demonstrates how real-world clinical practice generates structured deviation signals that can be

captured and interpreted within deviation-aware healthcare analytics infrastructures.

Scalability consequences manifest at the infrastructure level [6, 23]. Hybrid cloud-on-premise deployments accommodate variable event-stream granularity without performance claims. At the same time, the five-layer stack supports incremental adoption: organizations may activate only Layers 1–3 initially for deviation encoding before progressing to full interpretive mapping [10, 27]. The unique spiral topology prevents architectural drift by continuously propagating governance signals downward, creating a theoretically stable feedback loop that adapts to evolving regulatory constraints and emerging data modalities [14, 18].

Collectively, these dynamics reposition care pathway deviations from liabilities to strategic assets, fostering healthcare AI systems that are inherently more interpretable, resilient, and aligned with real-world clinical complexity [1, 5, 22]. The framework's theoretical consequences therefore extend to organizational readiness, ethical deployment, and long-term sustainability of clinical interpretation ecosystems.

Architectural resilience, interoperability pathways, and future orchestration horizons for deviation-centric sequence languages

Beyond immediate systemic impacts, the SPLiDeK framework introduces architectural resilience characteristics that address longstanding vulnerabilities in clinical AI system architectures and EHR intelligence ecosystems [14, 18, 22]. Conventional decision support pipelines fracture when confronted with high-variance pathways; SPLiDeK counters this through its layered abstraction, where each deviation token in Layer 3 carries semantic type information that survives downstream transformations [6, 7, 12]. This resilience propagates upward: interpretive graphs in Layer 4 remain clinically coherent even when underlying event logs exhibit missing timestamps or parallel branches, satisfying the theoretical requirements of multi-source data exchange frameworks [8, 17, 19].

Interoperability is further strengthened by the canonical sequence format established in layer 1, which aligns natively with emerging standards for temporal event

representation while preserving the full provenance chain required for auditability [10, 23]. Governance agents operating within the spiral topology can therefore query deviation patterns across institutional boundaries without violating data-sovereignty constraints, enabling federated knowledge synthesis that was previously unattainable under strict conformance paradigms [3, 21]. **Table 2** consolidates the three interpretive formulas into an architectural control matrix linking deviation dynamics to governance and infrastructural adaptation.

Table 2. Architectural control matrix for deviation-aware interpretation

Parameter	Functional definition	Primary input variables	Layer dependence
Drift sensitivity (DS)	Aggregated responsiveness to divergence between observed and expected semantic states	Observed state distributions, expected state distributions, and weighting coefficients	Layers 2–4
Risk propagation (RP(t))	Temporal accumulation of deviation impact across workflow phases	Instantaneous deviation magnitude, impact coefficients, and temporal exposure	Layers 3–4
Governance load (GL)	Composite oversight demand induced by deviation volume and contextual complexity	Deviation density, feedback frequency, and contextual constraints	Primarily layer 5
Variant promotion threshold	Decision boundary for elevating persistent deviations to structured	Drift persistence and semantic clustering stability	Layers 3–4

	pathway variants		
Feedback elasticity coefficient	Adaptive recalibration responsiveness to feedback signals	Feedback cycle density and institutional constraints	Layer 5 →

Deployment environment considerations reveal additional resilience [4-6, 8, 10]. On-premise components handle sensitive deviation encoding, while cloud-based layer 5 governance services manage spiral feedback cycles at scale. This hybrid topology ensures continuous operation during network partitions—an essential property for mission-critical healthcare settings [13, 16]. Moreover, the absence of empirical training loops means the architecture remains stable across evolving EHR vendors and regulatory updates, eliminating the version-friction common in model-dependent systems [15, 18].

Future orchestration horizons emerge naturally from the framework’s design [9, 20]. The sequence pattern language can theoretically extend to incorporate patient-generated health data streams or real-time device telemetry, enriching deviation encoding with additional contextual layers [13, 19]. Drift sensitivity calculations may be augmented with multi-context weights, allowing institutions to calibrate DS thresholds according to specialty-specific clinical ontologies [5, 11]. Resource allocation models can leverage RP(t) integrals to simulate theoretical “what-if” pathway adjustments during capacity-planning exercises, all without ever invoking benchmarking or performance metrics [1, 4, 8].

The spiral governance topology itself offers a template for broader AI orchestration challenges [22, 23]. By replacing linear conformance loops with bidirectional, hierarchical feedback, SPLiDeK demonstrates how governance load can be transformed from a static constraint into a dynamic control variable. This paradigm shift aligns with evolving AI governance principles that prioritize continuous interpretability over static validation [18, 21]. Future extensions might embed additional interpretive formulas—such as confidence propagation across parallel deviation sequences—while maintaining strict theoretical boundaries [7, 12].

In synthesis, the architectural resilience engineered into SPLiDeK not only resolves immediate integration frictions

but also charts a sustainable pathway for next-generation healthcare analytics infrastructures [2, 14, 22]. By treating care pathway deviations as the foundational substrate of clinical intelligence, the framework establishes a new baseline for interpretable, adaptive, and governance-compliant AI deployment across diverse clinical environments [1, 5, 17].

Conclusion

The conceptual architecture presented reframes care pathway deviations from disruptive outliers into structured knowledge assets through the Sequence Pattern Language for Deviation Knowledge. SPLiDeK's five-layer infrastructure, coupled with its distinctive spiral governance topology and three interpretive formulas, delivers a complete theoretical blueprint for operationalizing this reframing within existing EHR intelligence ecosystems and decision support pipelines.

By anchoring clinical interpretation directly to observed deviation sequences rather than idealized pathways, the framework bridges longstanding gaps between process-mining primitives, workflow integration models, and AI governance requirements. Drift sensitivity, risk propagation, and governance load become quantifiable architectural parameters that guide system designers without necessitating empirical validation. The resulting infrastructure is deliberately vendor-agnostic, regulation-resilient, and inherently explainable—qualities increasingly demanded by healthcare stakeholders.

This work contributes a uniquely named, acronym-defined, and topologically novel systems construct that advances the field of clinical AI architectures beyond conformance-centric paradigms. Future theoretical extensions may explore multi-institutional pattern federation or integration with emerging real-time data modalities. Yet, the core contribution remains: deviations, once interpreted through a dedicated sequence pattern language, become the richest source of actionable clinical knowledge available within modern healthcare systems. SPLiDeK therefore offers not merely an incremental improvement but a foundational shift toward truly intelligent, deviation-aware healthcare analytics infrastructures.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 25 Jan 2022 Revised: 17 May 2022 Accepted: 27 Jul 2022
Published online: 10 January 2023

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

De Roock E. Process mining in healthcare - an updated perspective on the state of the art. *J Biomed Inform.*

2022;127:103995.
<https://doi.org/10.1016/j.jbi.2022.103995>.

- Munoz-Gama J. Process mining for healthcare: characteristics and challenges. *J Biomed Inform.* 2022;127:103994.
<https://doi.org/10.1016/j.jbi.2022.103994>.
- Kempa-Liehr AW. Healthcare pathway discovery and probabilistic machine learning. *Int J Med Inform.* 2020;137:104087.
<https://doi.org/10.1016/j.ijmedinf.2020.104087>.
- Leemans SJJ. Process mining for healthcare decision analytics with micro-costing estimations. *Artif Intell Med.* 2023;135:102473.
<https://doi.org/10.1016/j.artmed.2022.102473>.
- Vathy-Fogarassy Á. Multi-level process mining methodology for exploring disease-specific care processes. *J Biomed Inform.* 2022;125:103979.
<https://doi.org/10.1016/j.jbi.2021.103979>.
- De Oliveira H. Automatic and explainable labeling of medical event logs with autoencoding. *IEEE J Biomed Health Inform.* 2020;24(11):3076-84.
<https://doi.org/10.1109/JBHI.2020.3021790>.
- Noshad M, Rose CC, Thombley R, Chiang J, Corbin CK, Nguyen M, et al. Signal from the noise: a mixed graphical and quantitative process mining approach to evaluate care pathways applied to emergency stroke care. *J Biomed Inform.* 2022;127:104004.
<https://doi.org/10.1016/j.jbi.2022.104004>.
- Baker K. Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. *Int J Med Inform.* 2017;103:32-41.
<https://doi.org/10.1016/j.ijmedinf.2017.03.011>.
- Chen K. Process mining and data mining applications in the domain of chronic diseases: a systematic review. *Artif Intell Med.* 2023;144:102645.
<https://doi.org/10.1016/j.artmed.2023.102645>.
- Lim J. Assessment of the feasibility of developing a clinical pathway using a clinical order log. *J Biomed Inform.* 2022;128:104038.
<https://doi.org/10.1016/j.jbi.2022.104038>.
- Oliart E. Are we ready for conformance checking in healthcare? Measuring adherence to clinical guidelines: a scoping systematic literature review. *J Biomed Inform.* 2022;130:104076.
<https://doi.org/10.1016/j.jbi.2022.104076>.
- Shrestha A. An annotated association mining approach for extracting and visualizing interesting clinical events. *Int J Med Inform.* 2021;148:104366.
<https://doi.org/10.1016/j.ijmedinf.2020.104366>.
- Ye J. The impact of electronic health record-integrated patient-generated health data on clinician burnout. *J Am Med Inform Assoc.* 2021;28(5):1051-6.
- Negro-Calduch E. Technological progress in electronic health record system optimization: systematic review of systematic literature reviews. *Int J Med Inform.* 2021;152:104507.
<https://doi.org/10.1016/j.ijmedinf.2021.104507>.
- Cimino JJ. Putting the “why” in “EHR”: capturing and coding clinical cognition. *J Am Med Inform Assoc.* 2019;26(11):1379-84.
- Wright A. A multi-site randomized trial of a clinical decision support intervention to improve problem list completeness. *J Am Med Inform Assoc.* 2023;30(5):899-906.
- Colicchio TK, Del Fiol G, Cimino JJ. Formal representation of patients’ care context data: the path to improving the electronic health record. *J Am Med Inform Assoc.* 2020;27(11):1648-57.
- Kashyap S. A survey of extant organizational and computational setups for deploying predictive models in health systems. *J Am Med Inform Assoc.* 2021;28(11):2445-50.
- Estiri H. High-throughput phenotyping with temporal sequences. *J Am Med Inform Assoc.* 2021;28(4):772-81.
- Gao Y. DR.BENCH: diagnostic reasoning benchmark for clinical natural language processing. *J Biomed Inform.* 2023;138:104286.
<https://doi.org/10.1016/j.jbi.2023.104286>.
- Pika A, Wynn MT, Budiono S, ter Hofstede AHM, van der Aalst WMP, Reijers HA. Privacy-preserving process mining in healthcare. *Int J Environ Res Public Health.* 2020;17(5):1612.
<https://doi.org/10.3390/ijerph17051612>.
- Guzzo A, Rullo A, Vocaturo E. Process mining applications in the healthcare domain: a comprehensive review. *Wiley Interdiscip Rev Data Min Knowl Discov.* 2022;12(2):e1442.
<https://doi.org/10.1002/widm.1442>.
- Dallagassa MR, Garcia CS, Scalabrin EE, Ioshii SO, Carvalho DR. Opportunities and challenges for applying process mining in healthcare: a systematic mapping study. *J Ambient Intell Humaniz Comput.* 2022;13:165-82.
<https://doi.org/10.1007/s12652-021-02894-7>.