

WHITE PAPER
DECENTRALIZED CLINICAL EVIDENCE ARCHITECTURE

July 16, 2026

TABLE OF CONTENTS

EXECUTIVE SUMMARY.....	3
RETHINKING REAL-WORLD EVIDENCE PARADIGMS	3
THE EXECUTIVE IMPERATIVE	4
THE CORE CHALLENGE: LEGACY DATA BROKERAGE AND THE AI PROVENANCE CRISIS	4
THE CIRCLES PLATFORM: DECENTRALIZED SOVEREIGNTY AND IN-SITU COMPUTE	5
THE COMMERCIAL AND REGULATORY VALUE PROPOSITION	6
STRUCTURAL DEFICIENCIES OF CENTRALIZED DATA BROKERAGE	7
FINANCIAL AND STRUCTURAL LIABILITIES OF AGGREGATED DATASETS	7
REGULATORY COMPLIANCE LIABILITIES AND ALGORITHMIC OPACITY	8
RE-IDENTIFICATION RISKS AND DOWNSTREAM CORPORATE LIABILITY	9
REGULATORY NON-COMPLIANCE OF STATIC CONSENT FRAMEWORKS	9
HEALTHCARE EVIDENCE IN THE WORLD OF AI	10
ENTERPRISE PROLIFERATION OF ALGORITHMIC SYSTEMS	10
MODEL FRAGMENTATION AND TRAINING REALITIES	11
RECURSIVE ALGORITHMIC ERROR PROPAGATION	11
THE GOVERNANCE, LIABILITY, AND REGULATORY MANDATE	12
STRATEGIC ALIGNMENT WITH THE CIRCLES PLATFORM	13
THE POWER OF DECENTRALIZED EVIDENCE ARCHITECTURE.....	13
ARCHITECTURAL PARADIGM: ACTIVE EVIDENCE GENERATION VS. PASSIVE REPOSITORIES	13
PROTOCOL-DRIVEN DATA DESIGN.....	14
THE LONGITUDINAL, INTEGRATED AND VERIFIABLE CASE RECORD	15
POINT-OF-CARE CAPTURE AND WORKFLOW INTEGRATION	16
LOCAL CUSTODY AND LEDGER VALIDATION	16
COMPLIANT VALUE EXCHANGE FOR HIGH-QUALITY EVIDENCE	17
THE SHARED INTELLECTUAL PROPERTY MODEL.....	17
QUANTIFICATION OF DATA CONTRIBUTIONS.....	18
ENTERPRISE QUERY INGESTION MECHANICS	18
AUTOMATED CLEARING AND TRANSACTION SETTLEMENT	19

NETWORK GOVERNANCE AND LEGAL COMPLIANCE INFRASTRUCTURE	19
EXAMPLE: ORPHAN ORTHOPEDIC THERAPEUTIC VALIDATION	20
CLINICAL AND FINANCIAL RISK CONTEXT	20
POINT-OF-CARE PROTOCOL EXECUTION	20
CRYPTOGRAPHIC QUERY VALIDATION	21
CONTRACT SETTLEMENT AND YIELD DISTRIBUTION	22
APPENDIX: SECURE DISTRIBUTED COMPUTE INFRASTRUCTURE.....	23
IN-SITU COMPUTATIONAL QUERIES.....	23
CRYPTOGRAPHIC VERIFICATION VIA ZERO-KNOWLEDGE PROOFS.....	24
DISTRIBUTED ANALYTIC FRAMEWORKS	24

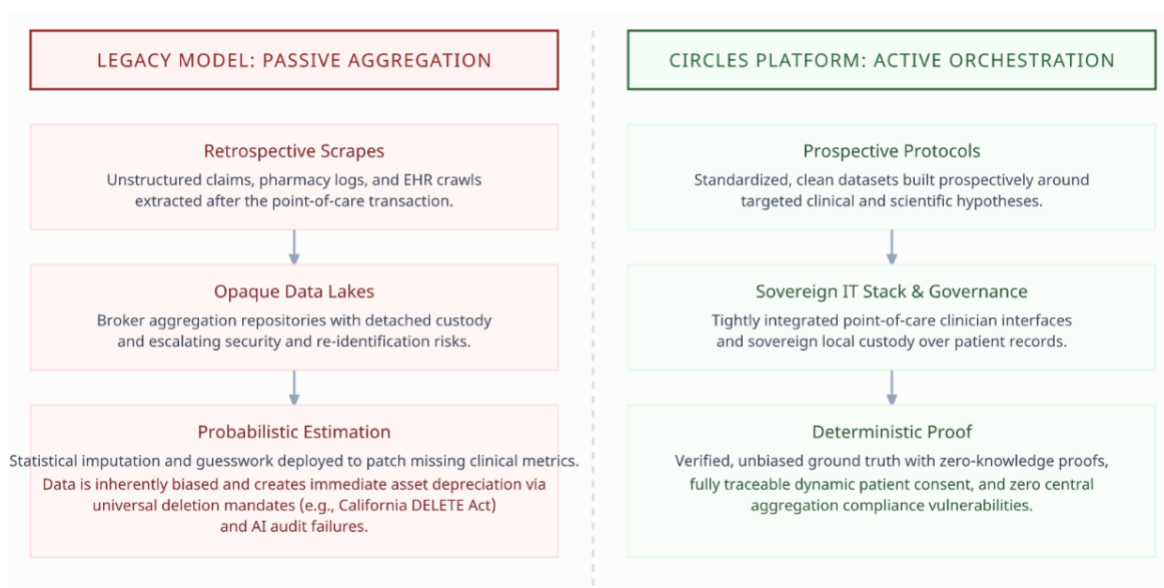
EXECUTIVE SUMMARY

Rethinking Real-World Evidence Paradigms

The real-world evidence (RWE) marketplace operates under a fundamental structural error: it conflates the volume of retrospective billing data with clinical evidence. To build a secure, compliant, and high-value healthcare data economy, organizations must distinguish between current administrative data capture and the deterministic metrics required for clinical, reimbursement, and regulatory validation:

Administrative Data Convention (Retrospective Aggregation): Today, evidence is treated as a passive byproduct of administrative billing workflows, consisting of unstructured electronic medical record (EMR) scrapes, fragmented pharmacy logs, and insurance claims collected post-hoc. Because this data was generated to secure a financial transaction rather than validate a scientific hypothesis, it is inherently incomplete. Relying on this model forces researchers and payers to accept probabilistic estimation—using statistical imputation to patch missing variables and broken patient timelines.

Deterministic Evidence Standard (Prospective Design): Genuine clinical evidence must be prospective, structured, and deterministic. True evidence requires a validated, unbroken link between a patient's baseline state, a targeted clinical intervention, and a longitudinal outcome measured through objective clinical metrics and standardized Patient-Reported Outcome Measures (PROMs). It must prove clinical utility with mathematical and observational certainty, rather than statistical approximation.



The Executive Imperative

For health system executives, payers, and regulators, this distinction dictates specific financial and operational outcomes:

For Providers: Treating clinical data as a low-fidelity commodity to be scraped by third parties depreciates the hospital's native intellectual property. True evidence must remain an institutional asset, prospectively curated to prove the value of care delivered.

For Payers: Settling outcomes-based risk structures or authorizing high-cost specialty therapeutics requires granular, verifiable clinical metrics (such as joint range of motion measured in degrees). Probabilistic data dumps fail to provide the deterministic certainty required to execute and settle these contracts.

For IT and Regulatory Officers: As computational cross-referencing renders traditional "de-identified" datasets highly vulnerable to re-identification, the only compliant path forward is an architecture that verifiably links prospective protocols to dynamic patient consent.

The Core Challenge: Legacy Data Brokerage and the AI Provenance Crisis

The legacy healthcare data ecosystem relies on an obsolete operational model: retrospective data brokerage. Commercial brokers pull uncurated data into centralized repositories to license them for commercial research and artificial intelligence development. This approach has created two severe structural liabilities:

Structural Liabilities of Centralized Aggregation: Centralized data scraping relies on a probabilistic patchwork that lacks the accuracy required to train safe AI models or automate prior authorizations. Furthermore, computational mosaic linkage allows anonymized patient records to be re-identified with success rates exceeding 80%, exposing downstream licensees to statutory penalties and class-action litigation. Concurrently, legacy data capture relies on single-point-in-time signatures, violating modern regulatory frameworks that grant patients a continuous legal right to modify or revoke data access.

Recursive Algorithmic Error Propagation: As AI becomes pervasive across clinical decision-making, drug development, and prior authorizations, models are increasingly trained on centralized data pools characterized by uncertain provenance. This establishes a process of recursive error propagation: models are trained on the noisy inputs of other algorithms, multiplying underlying statistical assumptions, baseline inaccuracies, and generative hallucinations. In clinical workflows and patient-facing applications, these compounding error loops introduce severe operational, financial, and clinical risks.

The Circles Platform: Decentralized Sovereignty and In-Situ Compute

The Circles Platform is a decentralized evidence architecture built on the concept of a Sovereign Case Node. The platform consists of a highly interoperable IT stack tightly integrated with active point-of-care clinician workflows via standard SMART on FHIR sidecar applications, patient-facing interfaces, and an independent multi-party governance framework designed to eliminate explicit and implicit data bias.

The architecture operationalizes data integrity and security through three core components:

Protocol-Driven Data Design: Datasets are built prospectively around independent clinical hypotheses. Every variable is pre-structured before observation begins, eliminating data fragmentation at the point of capture within the native EMR

workflow.

Decoupled Storage and Verification: Raw medical records and protected health information (PHI) remain permanently localized within the health system's secure infrastructure. The shared blockchain network acts exclusively as a tamper-evident digital notary, recording only cryptographic hashes of record states, clinician credentials, and active patient consent updates.

In-Situ Computational Queries: The platform reverses the traditional data-sharing workflow by bringing the clinical query or algorithm directly to the local node. Computations execute entirely behind the institutional firewall, and only aggregate statistical results and cryptographic proofs are transmitted back to the purchaser.



The Commercial and Regulatory Value Proposition

This architecture establishes the Shared Intellectual Property Model, allowing healthcare providers to commercialize the clinical utility of their records without surrendering ownership of the underlying assets.

Strategic Capability	Operational Mechanism	Institutional Benefit
Asset Retention	Licensing of aggregated, derivative evidence products rather than raw databases.	Primary clinical records and proprietary diagnostic expertise remain on the provider's balance sheet.
Automated Financial Settlement	Smart contract rules verify active consent and execute financial distribution based on LDI metadata weights calculated locally by the node.	Revenue is distributed directly to patients, clinicians, and node operators via standard fiat or stablecoins.

Cryptographic Process Audits	Local nodes utilize Zero-Knowledge Proofs (ZKPs) to mathematically verify cohort inclusion and protocol compliance.	Commercial buyers receive procedural assurance without the administrative overhead or privacy exposure of manual file audits.
------------------------------	---	---

By linking point-of-care clinical validation with automated financial settlement, this architecture transforms routine medical tracking into a regulatory-grade, legally compliant, and unbiased real-world evidence product. This framework provides health systems, payers, and pharmaceutical sponsors with the deterministic truth required to support outcomes-based reimbursement contracts, protect operational margins, and advance clinical research without compromising institutional data sovereignty.

STRUCTURAL DEFICIENCIES OF CENTRALIZED DATA BROKERAGE

Financial and Structural Liabilities of Aggregated Datasets

The legacy commercial data brokerage model relies on purchasing raw, uncurated data feeds from disparate institutional silos, aggregating them into centralized data lakes, and licensing access to insurers, pharmaceutical firms, and artificial intelligence developers. This model introduces significant administrative capital requirements and structural data deficiencies that diminish the commercial and clinical value of the asset:

Escalating Curation Expenses: Data brokers pay high premiums to acquire uncurated electronic medical records (EMRs), pharmacy logs, and insurer databases. Normalizing these disparate systems requires substantial administrative capital due to a systemic lack of shared clinical ontologies and standardized data dictionaries. This extensive manual curation overhead results in inflated licensing fees passed directly to the buyer for datasets that remain structurally incomplete.

Statistical Imputation and Technical Approximation: Broker algorithms utilize automated parsing to estimate missing clinical variables and reconstruct fragmented patient timelines when a patient moves outside a specific health system's firewall. This reliance on a probabilistic patchwork fails to deliver the deterministic clinical evidence required for regulatory or reimbursement validation.

Absence of Baseline-to-Outcome Lineage: Aggregated broker data pools cannot provide longitudinal patient outcomes that are verifiably correlated back to specific baseline metrics and exact clinical interventions for individual cases. Consequently, data consumers reach an operational utility limit where uncured scrapes and synthetic data records lack the accuracy required to train safe artificial intelligence models, support precision clinical decision systems, or automate prior authorizations.

Regulatory Compliance Liabilities and Algorithmic Opacity

The convergence of global artificial intelligence governance and aggressive state-level privacy mandates introduces severe operational cost structures that legacy data brokers are unequipped to meet:

Algorithmic Transparency and Data Lineage Mandates: Emerging AI and machine learning regulatory frameworks require data providers to deliver absolute transparency regarding data lineage, training inputs, and preprocessing validation. Because legacy brokers rely on proprietary, black-box normalization algorithms and probabilistic imputation to patch missing data fields, they cannot deliver the verifiable mathematical audit trails required by regulators. This technical opacity prevents institutional buyers from validating the provenance of the dataset, rendering the asset non-compliant for regulated AI application training.

The Universal Deletion Vector (The California DELETE Act): Modern personal data privacy trends—exemplified by statutory frameworks such as the California DELETE Act (SB 362)—fundamentally disrupt the economics of centralized data aggregation. These mandates enforce a centralized, one-click mechanism allowing consumers to command all registered data brokers to continuously delete their personal, diagnostic, and behavioral records.

Downstream Operational Invalidation: For a centralized broker, complying with continuous, automated deletion requests requires massive engineering overhead to locate, isolate, and purge individual data points across massive, historical repositories. Furthermore, because these repositories lack traceable baseline-to-outcome data architecture, a single deletion command can compromise the statistical integrity of an entire synthetic or aggregated cohort, invalidating the data assets previously licensed to downstream enterprise buyers.

Balance Sheet Asset Depreciation: The financial risk of relying on legacy brokers is no longer limited to data quality. Enterprise buyers face immediate capital depreciation when a licensed dataset becomes legally unusable due to a retrospective deletion request or an audit failure under AI transparency laws. Institutional data strategies must shift away from aggregated pools that cannot survive continuous statutory erasure loops.

Re-Identification Risks and Downstream Corporate Liability

Licensing health datasets under the assumption that traditional de-identification completely eliminates regulatory risk creates immediate legal and financial exposure for the purchasing organization. Modern computational capacity has neutralized legacy anonymization protections through three specific vectors:

Mosaic Linkage Vulnerabilities: De-identified clinical files are highly vulnerable to reverse-engineering through cross-referencing against external data sources. Automated linkage algorithms match medical records with public registries, consumer purchasing histories, and public genetic records to re-identify specific individuals, achieving real-world success rates exceeding 80%.

Transfer of Legal Liability: Once a dataset is reverse-engineered, legal liability shifts directly to the commercial consumer that licensed and utilized the data asset. Regulatory bodies and courts treat data misuse and downstream re-identification resulting from a broken chain of custody as an actionable corporate violation rather than an external vendor issue.

Regulatory Sanctions and Operational Disqualification: The Federal Trade Commission (FTC) and the HHS Office for Civil Rights (OCR) treat the commercial sale and use of re-identifiable data as a deceptive and non-compliant practice. This exposure subjects corporate buyers and health networks directly to class-action litigation and statutory civil penalties, while the loss of compliance value renders the asset legally unusable for commercial research, AI model training, or formal regulatory submissions.

Regulatory Non-Compliance of Static Consent Frameworks

Modern regulatory frameworks grant patients continuing legal rights to modify, restrict, or

revoke access to their medical records over time. Static consent mechanisms—such as point-in-time digital signatures or paper forms—cannot operationalize these continuous rights, creating systemic compliance exposure for both data holders and downstream purchasers:

Lack of Dynamic Revocation Mechanisms: Point-in-time digital signatures authorize data reuse only at a single instance. They lack the technical capacity to dynamically update, segment, or withdraw access permissions as a patient's treatment path, clinical trial status, or data preferences evolve.

Chain of Custody Disruption: If a patient exercises their statutory right to withdraw information, centralized data repositories lack the infrastructure to trace, isolate, or propagate that revocation across third-party licensees. This leaves the purchaser holding an un-auditable, non-compliant data asset.

Asset Disqualification Under Privacy Frameworks: Utilizing datasets with broken or static consent trails guarantees non-compliance under modern state and federal privacy frameworks. This structural exposure renders the dataset legally unusable for payer reimbursement audits, AI model training, or formal regulatory submissions.

HEALTHCARE EVIDENCE IN THE WORLD OF AI

Enterprise Proliferation of Algorithmic Systems

Artificial Intelligence (AI) has transitioned to an infrastructure-level layer operating across multiple domains of the healthcare ecosystem:

Product and Therapeutic Development: Predictive models are deployed to accelerate molecular target discovery, simulate clinical trials, optimize cohort selection, and analyze drug-to-drug interactions.

Clinical Decision Support (CDS): Systems analyze structured records and diagnostic imaging to flag anomalies, generate differential diagnoses, and deliver real-time insights to clinicians at the point of care.

Utilization Management and Reimbursement: Payer networks rely on automated

algorithms to conduct clinical reviews, assess medical necessity, and execute prior authorization workflows.

Clinical and Patient Education: Large language models (LLMs) synthesize medical literature for clinicians and deliver natural-language health education materials directly to patients.

Model Fragmentation and Training Realities

The healthcare AI landscape is characterized by deployment fragmentation rather than centralized standardization, operating across a highly diverse matrix:

Open vs. Closed-Source Configurations: Institutions operate a mix of proprietary commercial APIs and highly customized self-hosted configurations to balance processing capacity against strict data security requirements.

Model Scaling and Architectures: Infrastructure scales from multi-billion parameter foundation models down to localized, small language models (SLMs) trained to execute specialized tasks with low latency.

Institutional Customization and Weightings: Pharmaceutical companies, hospital systems, and payers independently build internal models, assigning proprietary neural weightings and training systems on historical, institutional data pools to drive localized operational efficiency.

Recursive Algorithmic Error Propagation

A critical vulnerability undermines this algorithmic layer: the data fueling modern healthcare AI models is frequently derived from retrospective scrapes, administrative billing feeds, and uncured electronic health records (EHR), resulting in uncertain data provenance. This lack of deterministic ground truth establishes a cycle of recursive error propagation:

Downstream Model Degradation: Models are increasingly trained on synthetic or derivative datasets generated by prior algorithms, initiating a process of recursive error amplification and model degradation.

Compounded Statistical Assumptions: Secondary predictive networks are layered

on top of primary estimation models, multiplying the underlying statistical assumptions and blind spots.

Systemic Input Bias: Noisy, unstructured historical inputs introduce systemic bias and baseline inaccuracies that degrade predictive accuracy over longitudinal operational cycles.

Generative Clinical Variance: Generative systems can fabricate clinically plausible but completely erroneous medical histories, diagnostic metrics, and false treatment paths.

Operational Risk Profile: While algorithmic errors in consumer-facing applications present minor inconveniences, clinical and administrative healthcare workflows cannot tolerate these systemic compounding errors due to direct patient safety and financial liabilities.

The Governance, Liability, and Regulatory Mandate

Operating AI systems in high-stakes clinical and administrative environments demands absolute verification. To mitigate operational, financial, and clinical risks, institutions must observe strict operational standards:

Deterministic Ground Truth: AI systems must be trained and updated using prospective, structured, and verifiable clinical data with an unbroken chain of custody.

Clinical Oversight Controls: System architecture must enforce human-in-the-loop constraints, embedding licensed clinicians at key analytical intervals to validate algorithmic recommendations before they influence clinical or reimbursement decisions.

Computational and Process Transparency: Computational paths cannot function through opaque architectures. Every calculation, data lineage path, and model weight adjustment must be transparent, traceable, and auditable.

Regulatory Enforcement Exposure: Federal regulatory bodies—including the Food and Drug Administration (FDA) and the HHS Office for Civil Rights (OCR)—strictly require high levels of operational provenance. Failure to establish these

transparency and verification standards introduces significant financial, legal, and reputational liabilities for healthcare executives and institutions.

Strategic Alignment with the Circles Platform

The Circles Platform architecture satisfies these operational and regulatory requirements, providing the utility needed to secure and validate healthcare AI applications:

Securing AI Data Provenance: By enforcing prospective, protocol-driven data capture through the inCytes™ and Benchmarc™ interfaces, the platform ensures that the data used to train and validate AI models is verified at the point of capture, eliminating recursive error loops at the source.

Integrating Verified Human Oversight: The platform organizes clinical records into structured, longitudinal Sovereign Case Files that are curated and validated by verified clinical contributors, guaranteeing that computational inputs are rooted in active clinical judgment.

Enforcing Mathematical and Computational Transparency: By utilizing the shared ledger as a tamper-evident notary and running in-situ queries backed by Zero-Knowledge Proofs (ZKPs), the platform mathematically proves that the defined clinical algorithms ran correctly against a committed set of records. This delivers full, auditable process transparency to regulators and purchasers without exposing protected health information (PHI).

To better understand the evolving federal standards around algorithmic lifecycle management and data provenance, this [FDA AI Medical Device Guidance Review](#) provides a detailed breakdown of the latest regulatory requirements for AI-enabled devices and continuous real-world performance monitoring.

THE POWER OF DECENTRALIZED EVIDENCE ARCHITECTURE

Architectural Paradigm: Active Evidence Generation vs. Passive Repositories

The Circles Platform replaces passive retrospective data aggregation with an active decentralized evidence architecture. The architecture rejects standard commercial data

conventions across three specific domains:

Heterogeneous Storage Exclusions: The platform does not ingest, centralize, or pool uncured databases derived from disconnected third-party scrapes or unstructured historical text.

Therapeutic and Cohort Flexibility: The software stack operates independently of specific clinical disciplines, enabling cross-functional configuration across any therapeutic area, clinical site, or tracking timeline.

Deterministic Computation: The architecture replaces black-box predictive modeling and statistical imputation with verifiable, deterministic clinical ground truth.

By capturing clinical variables at the point of origin, the resulting datasets eliminate the explicit selection biases, geographical gaps, and socioeconomic skewing inherent in legacy data broker feeds. This integrated stack operationalizes data capture to deliver verified clinical, regulatory, and commercial utility:

For Clinicians and Health Systems: Integrates directly into native workflows to transform routine clinical observations into structured institutional assets without increasing documentation overhead.

For Patients: Establishes data sovereignty and dynamic consent tracking, aligning longitudinal participation with transparent economic recognition.

For Payers and Pharmaceutical Sponsors: Delivers verifiable clinical outcomes that eliminate manual auditing overhead and provide the deterministic proof required to settle outcomes-based reimbursement contracts.

Protocol-Driven Data Design

Circle Datasets are constructed from the ground up around pre-structured, independent clinical and scientific hypotheses, reversing the traditional workflow of post-hoc data scraping. This structural methodology enforces precise data quality through specific technical constraints:

Pre-Structured Collection Rules: The data schema is fully defined before patient

observation begins, eliminating data fragmentation and missing clinical variables by architecture.

Targeted Observational Protocols: Every collected variable is engineered to support defined scientific or contractual endpoints, removing the capital expenditure required for post-hoc data cleaning and normalization.

Traceable Data Hierarchy: Structural information flow follows a strict linear path: an individual data event maps directly to a Case, which compiles into a virtual cohort dataset, ultimately forming a verified evidence product.

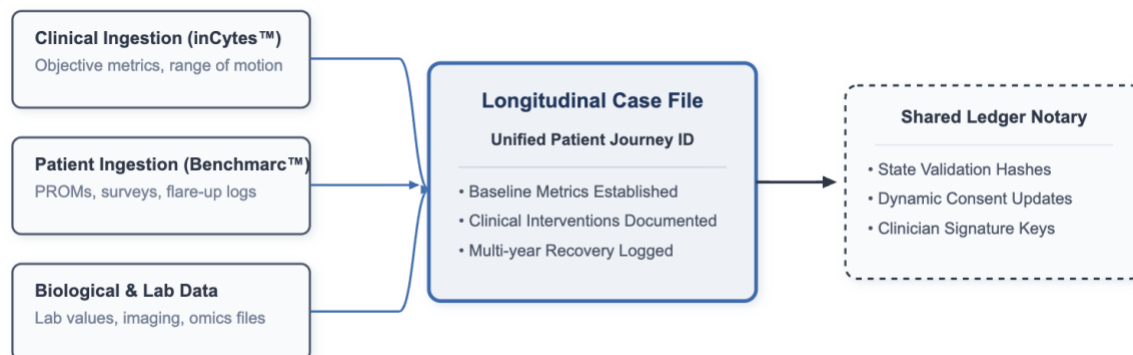
The Longitudinal, Integrated and Verifiable Case Record

The core organizing unit of the architecture is the Case, which aggregates all longitudinal data associated with an individual patient's therapeutic journey into a single structured file. This record architecture organizes clinical data through three specific mechanisms:

Multi-Modal Data Ingestion: Each Case groups distinct clinical and phenotypic data streams over time, combining clinician-curated diagnoses, objective laboratory and imaging results, biological omics data, and standardized Patient-Reported Outcome Measures (PROMs).

Defined Evaluation Timelines: The record binds one patient, a specific clinical condition, and an explicit tracking path to measure functional recovery against the patient's original baseline metrics.

Permanent Data Contextualization: Every entry remains permanently anchored to the specific Case record, the validated clinical contributor, and the corresponding observational protocol, preventing the loss of clinical context during secondary analysis.



Point-of-Care Capture and Workflow Integration

To secure data provenance and enforce strict quality controls, data entry is restricted to a closed-loop software framework operating at the point of care. Rather than utilizing retrospective extraction, collection occurs during routine medical visits via synchronized interfaces designed to minimize administrative friction:

The Clinician Interface (inCytes™): Operating natively within existing electronic medical records via a standard SMART on FHIR sidecar or bi-directional APIs, this portal allows healthcare providers to log objective clinical metrics, validate record completeness, and verify protocol compliance without dual-documentation workflow duplication.

The Patient Interface (Benchmarc™): Patients utilize this secure portal to complete standardized functional outcome surveys and update dynamic consent settings in real time.

Provenance Verification: Capturing clinical observations directly inside these integrated portals prevents the data dilution, transcription errors, and loss of clinical context that occur when extracting records from external EMR databases post-hoc.

Local Custody and Ledger Validation

The architecture protects institutional clinical assets by separating physical data storage from the shared network validation layer. Raw medical records and protected health information (PHI) remain permanently within the hospital's local, encrypted databases behind

institutional firewalls. This division of storage and verification utilizes three control mechanics:

The Ledger as a Cryptographic Notary: The shared blockchain operates strictly as a tamper-evident notary. It records only cryptographic hashes of record states, timestamped validations, clinician credentials, and active patient consent updates, keeping all clinical content off-chain.

Local Compute Sovereignty: Participating networks retain total control over the analytical tools, risk models, and processing rules applied to their records, preventing the surrender of proprietary clinical algorithms or institutional diagnostic expertise.

Auditable Correctability: Clinical files remain adjustably correctable at the local node to allow for accurate diagnostic updates or laboratory amendments. The ledger maintains an immutable provenance history by permanently logging exactly who modified the file, when it occurred, and what specific changes were executed.

COMPLIANT VALUE EXCHANGE FOR HIGH-QUALITY EVIDENCE

The Shared Intellectual Property Model

The Shared Intellectual Property framework defines the precise contractual boundaries governing data utilization, preventing the permanent extraction or depreciation of institutional clinical assets. Rather than requiring health systems to surrender raw data to centralized repositories, this model separates underlying record custody from the derivative evidence products licensed to commercial buyers:

Retention of Primary Assets: Participating health systems and medical practices retain absolute physical control and balance-sheet ownership over their primary medical records, proprietary risk models, and institutional diagnostic expertise.

Licensing of Derivative Products: Commercial purchasers do not buy or access raw patient-level databases. Instead, they contract exclusively for structured, aggregated evidence products, such as verified clinical answers, cohort descriptions, and validated model calibrations.

Non-Dilutive Asset Monetization: The framework allows clinicians and health systems to realize financial yields for validating clinical data without transferring underlying data ownership or corporate intellectual property to external brokers.

Quantification of Data Contributions

The platform separates the assessment of clinical data quality from its commercial market value through two distinct calculations:

The Contribution Score: Calculated locally by the sovereign node at data ingestion using the Longitudinality, Depth, and Integrity (LDI) index, this metric commits cryptographic metadata scores to the ledger notary to award non-transferable entitlement units based strictly on protocol compliance.

The Revenue Allocation Formula: Calculated at the transaction level, this formula determines which specific Cases and clinical contributors were material to a purchaser's commercial query and coordinates distribution accordingly.

Dynamic Market Pricing: By separating baseline utility scoring from transactional settlement, the architecture prevents a quality score from being misrepresented as a fixed market price, recognizing that the economic value of clinical variables fluctuates depending on the specific study parameters.

Enterprise Query Ingestion Mechanics

To initiate a network-wide analysis without compromising node sovereignty, commercial buyers utilize an inbound query ingestion workflow:

Protocol Specification: Pharmaceutical sponsors or payers utilize a standardized protocol builder interface to define explicit cohort parameters, including inclusion/exclusion criteria, targeted clinical variables, and longitudinal tracking timelines.

Query Compilation: The defined protocol is compiled into a sandboxed, low-latency analytical script (such as a standardized SQL query or a federated machine learning script) that contains zero extraction commands for raw patient identifiers.

Decentralized Distribution: The secure coordinator distributes this compiled query package simultaneously to all participating network nodes for local validation and execution behind their respective institutional firewalls.

Automated Clearing and Transaction Settlement

The contracting platform automates commercial transactions by linking financial clearing directly to protocol compliance. Rather than replacing legal frameworks, automated software rules function strictly as operational execution tools to verify permissions and distribute payments:

Real-Time Consent Auditing: The software runs automated compliance checks against the ledger to confirm an active, dynamic consent record is in force for every individual case before any computational query executes.

Monetary Clearing and Distribution: Commercial buyers pay for verified evidence products using standard fiat currency or regulated stablecoins. The settlement engine automatically routes these proceeds directly to verified patients, clinical validators, and local node operators according to their ledger-recorded entitlements.

Contractual Enforceability: These automated execution rules operate strictly within the boundaries of formal clinical evidence licenses, data-use agreements (DUAs), liability indemnities, and statutory healthcare regulations.

Network Governance and Legal Compliance Infrastructure

The network architecture is overseen by an independent, multi-party governance committee that operates as the fundamental control plane across the platform, defining automated parameters, human-in-the-loop operational intervals, and local institutional controls. The legal and operational framework manages three specific responsibilities:

Reciprocal Legal Architecture: Node participation is governed by a unified legal framework consisting of reciprocal Business Associate Agreements (BAAs) and standardized Data Use Agreements (DUAs) that ensure cross-node compliance with HIPAA and state-level privacy statutes.

Participant Credentialing and Protocol Validation: The committee manages the

onboarding and credential verification of all clinical participants and reviews and approves all prospective Observational Protocols before data capture can begin.

Operational and Formulaic Adjustments: The governing body actively monitors network parameters, including systematic updates to the Longitudinality, Depth, and Integrity (LDI) contribution scoring formulas and model-version controls.

Financial Oversight and Dispute Resolution: The control plane coordinates the auditing of platform reserve accounts and executes formal dispute resolution procedures to settle structural or transactional conflicts across the network.

EXAMPLE: ORPHAN ORTHOPEDIC THERAPEUTIC VALIDATION

Clinical and Financial Risk Context

The management of progressive orthopedic conditions—such as Fibrodysplasia Ossificans Progressiva (FOP)—requiring high-cost orphan therapeutics introduces specific financial risk for health systems and payers. Validating treatment efficacy under these conditions requires protocol-specified clinical metrics that administrative billing systems cannot capture:

Therapeutic Durability Metrics: Payers require objective, multi-year proof of therapeutic durability, such as tracking the exact range of joint motion measured in degrees, to justify the reimbursement of high-cost specialty pharmaceuticals.

Institutional Risk Alignment: Health systems participating in outcomes-based risk structures require precise, prospective tracking to protect operational margins against financial penalties associated with treatment failure.

Retrospective Deficiencies: Standard hospital billing records and retrospective claims datasets are blind to long-term functional improvement metrics, leaving the ecosystem without the deterministic evidence required to settle insurance contracts.

Point-of-Care Protocol Execution

Operationalizing the platform for an FOP cohort begins by deploying localized protocols across participating clinical networks, enforcing standardized data capture at the point of

care during routine medical visits:

Clinician Metric Capture: Treating physicians utilize the integrated inCytes™ EMR sidecar to log specific anatomical variables, focusing on heterotopic ossification sites and joint range of motion measured in degrees. The interface applies schema-on-capture validation rules to block incomplete entries or anomalous data at the millisecond of entry.

Longitudinal Patient Ingestion: Through the Benchmarc™ interface, patients log localized flare-ups, functional quality-of-life adjustments, and standardized PROMs directly into the system. This protocol-driven participation secures continuous real-world data points omitted by standard retrospective workflows.

Localized Custody Maintained: Raw clinical observations, survey inputs, and protected health information (PHI) remain localized inside the hospital network's secure data vaults behind institutional firewalls, with zero raw data extracted to external data repositories.

Cryptographic Query Validation

When a payer or pharmaceutical sponsor deploys an inbound query to validate the FOP cohort, the decentralized network delivers procedural verification without exposing underlying patient identifiers:

Zero-Knowledge Inclusion Proofs: The local hospital node uses Zero-Knowledge Proofs (ZKPs) to mathematically prove to external partners that all cohort parameters are met—verifying that the cases strictly matched FOP inclusion criteria, possessed active consent, and met completeness standards—without exposing patient identities or raw clinical variables.

Procedural Assurance: The commercial purchaser receives the aggregate clinical analysis alongside a verified cryptographic proof package. The system mathematically guarantees that the approved protocol was followed precisely, providing complete auditability without manual file inspection.

Source and Computational Integrity: The network pairs the ZKP with signed digital credentials and ledger hashes to verify source authenticity and data integrity,

ensuring complete computational truth.

Contract Settlement and Yield Distribution

The final phase of the FOP reference scenario translates verified clinical milestones into automated financial clearing, executing outcomes-based contracts without manual auditing overhead:

Outcomes-Based Reimbursement Triggers: When the decentralized compute network delivers mathematical proof that the patient cohort has achieved pre-specified durability thresholds—such as the preservation of joint range of motion—the system validates the contractual milestone. This validation allows payers and health systems to settle risk-sharing agreements or execute therapeutic performance guarantees automatically, protecting provider operating margins.

Automated Financial Distribution: Once the commercial query settles, the platform's clearing engine processes the transaction using standard fiat currency or regulated stablecoins. The system parses the ledger-recorded entitlements to route non-dilutive data yields directly to participating patients, validating clinicians, and health system node operators based on their calculated LDI data contributions.

Operational Constraints: The underlying smart contracts function strictly as execution tools to automate active consent validation, document transaction states, and calculate revenue allocations. They operate entirely within the boundaries of formal clinical evidence licenses, data-use agreements, and statutory healthcare regulations, ensuring financial settlement occurs without exposing protected health information.



APPENDIX: SECURE DISTRIBUTED COMPUTE INFRASTRUCTURE

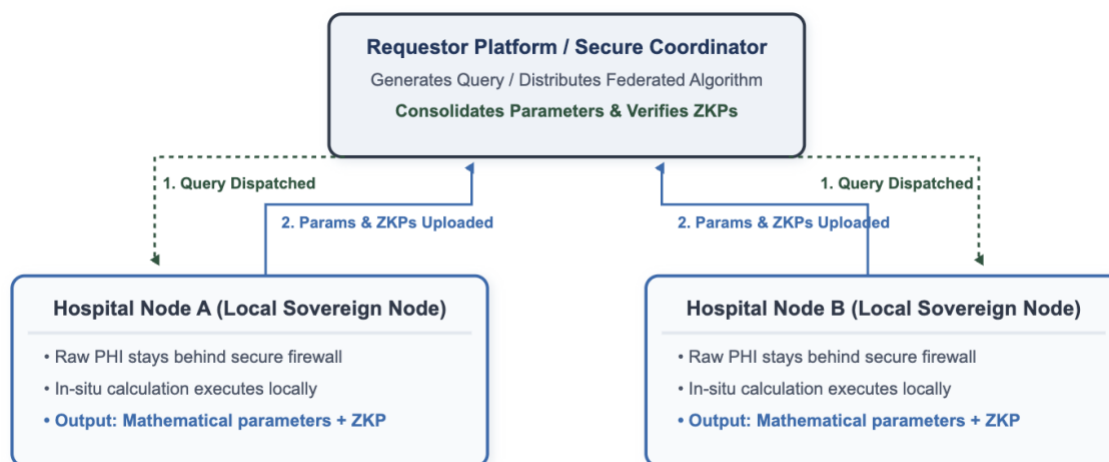
In-Situ Computational Queries

Traditional real-world evidence extraction requires pooling raw datasets into a centralized repository. This architecture reverses that workflow by transmitting the approved clinical query or algorithm directly to the locally controlled data nodes. This decentralized query framework operates under three operational constraints:

Localized Processing: Raw patient-level data and protected health information (PHI) remain permanently within the physical infrastructure and firewall of the originating institution.

Transaction-Level Authorization: Participating nodes evaluate inbound queries on an individual, transaction-by-transaction basis, retaining absolute authority over cohort inclusion, variable access, and local execution rules.

Ephemeral Data Aggregation: Once local calculations execute within the secure environment, only aggregate statistical results and cryptographic proofs of execution are transmitted back to the secure coordinator. The multi-center dataset remains entirely virtual, existing only for the duration of the approved transaction rather than as a permanent database under external control.



Cryptographic Verification via Zero-Knowledge Proofs

To establish auditability across federated nodes without manual file verification, the platform utilizes Zero-Knowledge Proofs (ZKPs) to verify procedural compliance. This cryptographic verification architecture operates through three primary mechanisms:

Procedural Compliance Verification: Local nodes utilize non-interactive ZKPs to mathematically prove to external entities that specific pre-contracted conditions were met—such as verifying that a patient cohort matched inclusion criteria, possessed active consent, and met completeness standards—without revealing patient-level identifiers or private clinical variables.

Computational Integrity Confirmation: The local node generates a cryptographic proof establishing that the approved query, algorithm, or model was executed accurately against the committed dataset and was not altered during or after execution.

Distinction Between Computational and Clinical Validity: The ZKP verifies computational and procedural truth, mathematically guaranteeing that a defined calculation ran correctly against a committed set of records according to protocol. Because a ZKP cannot independently verify the accuracy of original clinician data entry or patient survey inputs, the network combines ZKPs with systematic source authenticity verification via signed digital credentials and data integrity checks via immutable ledger hashes.

Distributed Analytic Frameworks

To extract network-wide insights or train predictive models without pooling patient records, the platform utilizes a federated analytics architecture, allowing multiple sovereign nodes to run complex mathematical computations locally and aggregate the resulting mathematical parameters safely. This distributed computation is governed by three primary mechanisms:

Federated Algorithm Execution: Rather than extracting patient records to a central processor, the platform transmits standardized computational algorithms—such as regression models, survival analyses, or machine learning training scripts—to each local node for execution within its secure firewall.

Secure Parameter Averaging: Nodes do not transmit patient-level clinical variables.

Instead, they export only mathematical gradients, model weights, or aggregated statistical summaries. A central coordinator aggregates these parameters (such as via federated averaging algorithms) to update the global model without gaining visibility into the underlying case records.

Information Leakage Prevention: Individual nodes never establish direct data-sharing connections with one another. Communication occurs strictly through secure aggregate protocols that prevent the reverse-engineering of local patient datasets, guaranteeing that institutional intellectual property and individual patient privacy remain intact at every node.
