RegenMed

WHITE PAPER

THE CIRCLES PLATFORM: FEDERATED HEALTHCARE DATA ARCHITECTURE

MOVING THE COMPUTATION TO THE DATA, RATHER THAN THE DATA TO THE COMPUTATION.

August 3, 2026

TABLE OF CONTENTS

INTRODUCTION AND EXECUTIVE SUMMARY

The Current Healthcare Data Crisis

Centralized healthcare data aggregation models, such as enterprise data lakes, uncurated data swamps, and legacy data repositories, are structurally incompatible with contemporary regulatory mandates, data entropy challenges, and institutional risk management requirements.

These centralized architectures expose healthcare systems to security vulnerabilities, including single-point-of-failure honeypots for cyber threats, compounding legal liabilities, and a permanent loss of operational control over proprietary enterprise assets.

Concurrently, tightening global regulatory frameworks, including the EU General Data Protection Regulation (GDPR), the EU AI Act, and U.S. data localization laws, actively restrict the cross-border and cross-entity transfer of Protected Health Information (PHI).

The Pathology of Data Brokerage

At the root of this industry-wide failure is the traditional data brokerage model. Legacy big data platforms and commercial intelligence networks rely on silent de-identification, extractive data transfers, and centralized middleman brokers who harvest patient and clinical data without structural consent or equitable compensation.

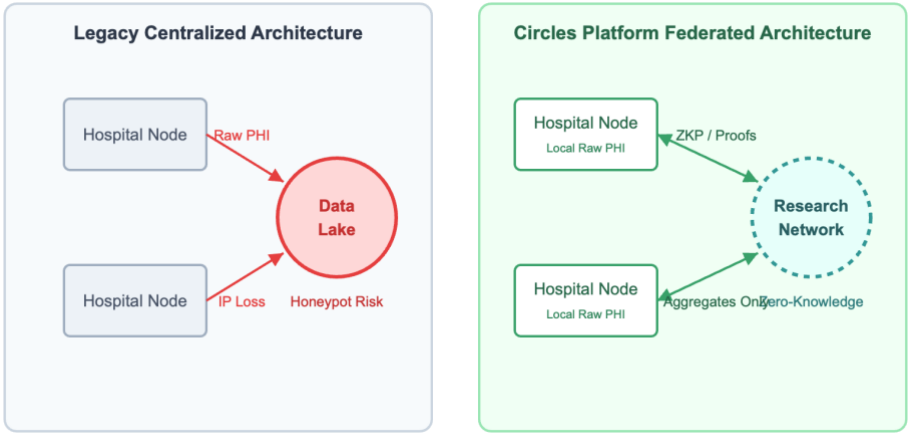
This intermediary extraction creates systemic vulnerability, exacerbates information-blocking concerns, and leaves healthcare institutions exposed to unmitigated commercial and regulatory liabilities.

The Circles Platform Thesis

To dismantle this legacy paradigm, the Circles Platform implements a decentralized, enterprise-hosted federated architecture that shifts data retention to behind the enterprise firewall. Instead of aggregating messy, post-hoc third-party data, the architecture captures clinical information natively as human – not AI -- ground truth via schema-on-capture validation and protocol-locked data dictionaries.

By combining standardized healthcare ontologies (FHIR, OMOP, LOINC, SNOMED) with dual-standard interoperability and an unbroken provenance pipeline, the platform enables Zero-Knowledge Proofs (ZKPs). This allows secure, verifiable transactions—such as prior authorization compliance and trial cohort matching—without exposing underlying

PHI or relinquishing institutional data custody.



Strategic Lenses Across the Ecosystem

Healthcare enterprise clients evaluate the Circles Platform through distinct strategic lenses, balancing operational risk, clinical fidelity, economic return, and regulatory compliance:

- **Providers:** The platform shifts the operational paradigm from data ownership anxiety to sovereign procedural stewardship, ensuring compliance with federal interoperability and anti-information-blocking mandates while eliminating third-party data leakage.
- **Payers and Researchers:** The architecture transitions the evidence base from probabilistic administrative proxies (such as billing and claims codes) to deterministic, audit-ready clinical evidence with microscopic fidelity, supported by both real-time FHIR exchanges and OMOP population-level analytics.
- **Investors:** The economic model unlocks high-yield capital efficiency through trial arbitrage economics (reducing evidence generation costs by 30% to 50%), eliminates recruitment churn via direct data dividends, disintermediates legacy data brokers, and elevates enterprise valuations from standard multi-tenant ranges to high-margin tech-enabled asset status.
- **Regulators:** Compliance is automated and future-proofed through policy-as-code governance, satisfying the 2027 FHIR mandate, the EU AI Act, and advanced expert determination standards without requiring manual administrative overhead.

PAIN POINTS ACROSS THE HEALTHCARE ECOSYSTEM

The contemporary healthcare data landscape is defined by systemic operational friction. Legacy centralized models—built on enterprise data lakes, uncurated data swamps, and third-party data brokerage—fail to serve the clinical, regulatory, or economic realities of modern medicine. These structural failures manifest across distinct operational domains:

Provider Friction: Custodianship, Information Blocking, and IP Loss

Healthcare providers navigate conflicting legal, technical, and commercial mandates that compromise institutional stability:

- **The Interoperability Mandate Conflict:** Traditional architectures force healthcare institutions into a false dichotomy between maintaining local data privacy and complying with aggressive federal anti-information-blocking and interoperability rules.
- **Intellectual Property Surrender:** When medical systems export raw clinical data to centralized third-party aggregators or data lakes, they immediately forfeit operational control over derivative datasets, proprietary AI model weightings, and the resulting commercial intellectual property.
- **Liability Exposure from Unproven Tools:** Integrating external analytical software that lacks verifiable provenance exposes provider organizations to acute organizational liability, diagnostic error risks, and unpredictable regulatory penalties.

Payer and Regulatory Friction: Administrative Proxies and the Measurement Gap

Operational workflows across payers and health systems remain heavily bottlenecked by legacy data structures:

- **Prior Authorization Overhead:** Conventional prior authorization workflows require providers to transmit extensive, unredacted PHI to payers to justify medical necessity. This manual process introduces severe compliance exposure, administrative delay, and operational drag.
- **The Failure of Administrative Proxies:** Traditional healthcare analytics rely predominantly on administrative proxies—such as ICD-10 billing codes and claims data. These indirect indicators fail to capture true patient physiology or functional health status.

- **The Measurement Gap:** This disparity constitutes a fundamental measurement gap between high-volume administrative data and the granular, objective clinical facts required to prove long-term patient health improvements and validate treatment efficacy.

Clinical Research Bottlenecks

Life science organizations, academic medical centers, and research networks face severe technical and clinical ceilings:

- **The Data Wall:** The industry has encountered the "data wall," a threshold where uncurated big data scrapes, legacy registry pools, and synthetic datasets are no longer accurate or reliable enough to train advanced medical artificial intelligence models safely.

- **Recruitment Churn:** Traditional clinical trials suffer from significant financial drag driven by patient recruitment churn. High drop-out rates require expensive, manual replacement cycles that inflate trial operational costs and delay time-to-market.

- **Training Data Degradation and Bias:** Centralized data aggregators pull messy, retrospective EHR dumps that contain uncorrected documentation errors and systemic demographic biases. Training clinical AI on these skewed repositories produces flawed diagnostic models that fail real-world clinical validation and trigger severe regulatory scrutiny.

- **Centralized CRO Inefficiencies:** Contract Research Organization (CRO) models rely on manual, high-friction data aggregation processes that restrict cross-institutional collaboration and inflate the capital requirements of clinical research.

Cyber Security Vulnerabilities: The Centralized Honeypot

Centralized data aggregation models present catastrophic infrastructure risks:

- **Single Points of Failure:** Enterprise data lakes and centralized cloud repositories act as massive, high-value "honeypots" for malicious actors, cybercriminal syndicates, and nation-state threat vectors.

- **Ransomware and Extraction Risk:** Pooling multi-institution PHI into a single centralized repository dramatically increases the blast radius of security breaches,

exposing organizations to multi-million-dollar regulatory fines, class-action litigation, and operational paralysis.

Regulatory Hardening: The End of Opaque Aggregation

Jurisdictional controls and statutory mandates globally are actively dismantling the legal viability of centralized data lakes:

- **European Union Enforcement:** The GDPR mandates strict data minimization (Article 5) and privacy by design (Article 25), principles structurally incompatible with broad data pooling and multi-entity data scraping. Concurrently, the EU AI Act enforces rigorous data governance, bias detection, and lifecycle risk management for high-risk medical AI systems.
- **U.S. Federal and State Statutes:** HIPAA enforces a rigorous "minimum necessary" standard for data handling, which is compounded by a rapidly expanding web of state-level data localization laws restricting the cross-border storage, maintenance, or offshoring of electronic health records.
- **FTC Section 5 Enforcement:** The Federal Trade Commission actively escalates enforcement under Section 5 of the FTC Act against digital health platforms, aggregators, and medical networks for deceptive practices related to the unauthorized sharing of sensitive health data through broad, opaque digital consent mechanisms.

CIRCLES DATASETS: CLINICAL GENESIS AND NATIVE DATA CAPTURE

To eliminate the systemic failures of passive electronic health record (EHR) extraction and uncurated retrospective data dumps, the Circles Platform roots data generation in a rigorous, sequential clinical hierarchy supported by a dual-standard native technical architecture.

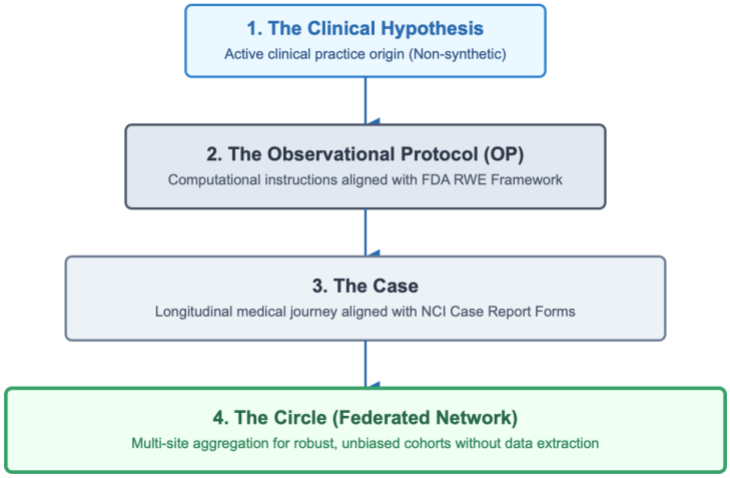
The Sequential Hierarchy of Data Origin

Every Circle Dataset generated by the platform is engineered from its inception to meet the evidentiary standards required for clinical research, regulatory submission, and algorithmic training. This follows a strict four-stage structural hierarchy:

- **The Clinical Hypothesis:** Data generation begins with a precise, testable medical question originating directly from active clinical practice rather than an external,

disconnected laboratory or other speculative environment.

- **The Observational Protocol (OP):** A standardized set of computational instructions and data collection parameters that dictate important capture criteria without altering or interfering with routine patient care pathways, maintaining direct alignment with the FDA and other regulatory real-world evidence framework.
- **The Case:** The foundational, longitudinal record of an individual patient's medical journey—spanning initial presentation, diagnostic confirmation, treatment execution, and long-term outcomes—structured in alignment with best practice electronic case report forms.
- **The Circle (Federated Research/RWE Generation Network):** A collaborative network of participating clinical sites that ingest patient data under shared Observational Protocols, enabling secure multi-site aggregation to build statistically robust, unbiased study cohorts.



Closing the Measurement Gap

Traditional healthcare analytics rely heavily on administrative proxies, such as ICD-10 billing codes, Current Procedural Terminology (CPT) codes, and transactional claims data. These indirect financial metrics are fundamentally unsuited for meaningful and verifiable clinical research or medical AI training because they fail to capture actual patient physiology, disease

progression, or functional health status.

This disparity creates a profound measurement gap between high-volume administrative data and the granular, objective clinical facts required to prove true therapeutic efficacy and long-term health improvements.

Human Ground Truth—primary clinical evidence verifiably derived directly from actual patient physiology, vital signs, lab results, and direct clinician-patient interactions—is the mandatory baseline for clinical integrity, genuine value-based care, and broadly impactful medical research.

Native Technical Capture

To ensure that human ground truth is captured without translation loss or semantic degradation, the platform enforces strict technical rules at the precise point of care:

- **Schema-on-Capture Validation:** Automated software guardrails execute the moment a clinician or staff member enters data into the system. These guardrails instantly validate entries against predefined parameters, flagging missing fields, incomplete records, or impossible clinical values to prevent input errors at the source.
- **Protocol-Locked Data Dictionaries:** Standardized medical terms and definitions governed by Clinical Data Interchange Standards Consortium (CDISC) standards are permanently locked the moment an Observational Protocol initiates, preventing mid-study schema drift and ensuring cross-site semantic consistency.
- **Dual-Standard Interoperability (FHIR and OMOP):** Raw point-of-care data is mapped natively into a dual-standard infrastructure. It utilizes Fast Healthcare Interoperability Resources (FHIR) for real-time transactional operations and clinical interoperability, while simultaneously mapping to the Observational Medical Outcomes Partnership (OMOP) common data model. This dual-standard approach bridges real-time clinical care with population-level observational research, preparing the data for seamless downstream translation into CDISC formats for regulatory submissions without lossy, post-hoc database manipulation.

REGULATORY AND COMMERCIAL IMPERATIVES

Evolving Data Localization and Privacy Mandates

The regulatory environment governing healthcare data is undergoing structural hardening. Jurisdictional controls and statutory mandates globally are dismantling the legal viability of centralized data aggregation models, transforming legacy data lakes from corporate assets into regulatory liabilities.

- **European Union Compliance Frameworks:** The General Data Protection Regulation (GDPR) enforces strict data minimization (Article 5) and privacy by design (Article 25) mandates—principles structurally incompatible with multi-entity data pooling and broad third-party scraping. Concurrently, the EU AI Act enforces rigorous data governance, bias detection, and lifecycle risk management for high-risk artificial intelligence systems, exposing centralized repositories to severe, non-negotiable statutory penalties.

- **U.S. Federal and State Statutes:** In the United States, HIPAA establishes a strict "minimum necessary" baseline for data disclosure. This federal standard is compounded by an expanding web of state-level data privacy and localization statutes (such as the Washington My Health My Data Act and Texas Medical Records Privacy Act) that strictly govern the cross-border storage, maintenance, and monetization of health information, rendering centralized national databases legally indefensible.

- **FTC Section 5 Enforcement Actions:** The Federal Trade Commission actively escalates enforcement under Section 5 of the FTC Act against digital health platforms, aggregators, and medical networks for deceptive practices related to the unauthorized sharing of sensitive health data through broad, opaque digital consent mechanisms or back-door commercial data monetization.

Enterprise Data Sovereignty

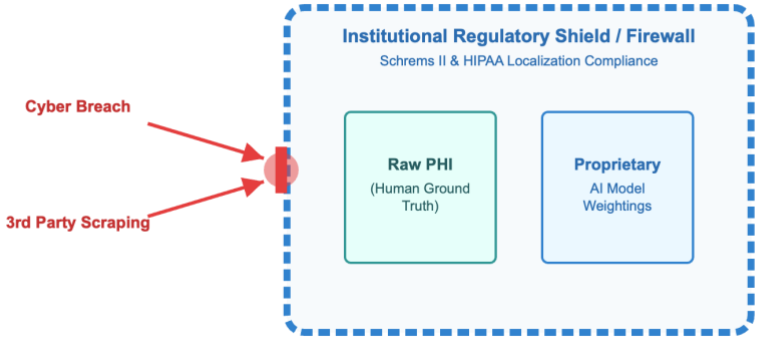
Beyond compliance with external regulations, healthcare enterprises face acute commercial and operational imperatives to maintain strict physical and logical custody of their data assets.

- **The Risk of Third-Party Extraction:** When medical institutions export raw clinical data to centralized third-party data lakes or commercial aggregators, they immediately forfeit operational control over derivative datasets, proprietary analytical models, and

commercial intellectual property.

- **Firewall Protection and Liability Sinks:** Retaining data strictly behind institutional firewalls serves as a primary risk-mitigation strategy. It ensures that local healthcare operators retain complete visibility into data lineage, computational execution, and governance boundaries. Furthermore, decentralized federated execution acts as a legal liability sink, isolating participating institutions from the systemic liabilities and breach exposure associated with centralized data storage.

- **Protecting AI Model Weightings and Derived Assets:** For healthcare enterprises developing proprietary diagnostic algorithms or predictive models, the underlying model weightings and training insights represent core corporate value. Decentralized architecture ensures these proprietary assets remain securely inside the enterprise perimeter, preventing external leakage, reverse engineering, or unauthorized commercial exploitation by intermediary brokers.



TECHNICAL FIDELITY, INTEROPERABILITY, AND PROVENANCE

Advanced decentralized architectures and cryptographic verification methods cannot function on uncurated, heterogeneous data. If the underlying clinical data lacks verifiable integrity or semantic consistency, federated queries will inevitably fail or yield invalid conclusions. For healthcare executives, the technical pipeline must securely transform raw electronic health record (EHR) outputs into regulatory-grade clinical evidence without requiring manual post-hoc cleansing or centralized data aggregation.

Regulatory-Grade Data Quality and ALCOA+ Compliance

Data utility is directly constrained by its evidentiary quality. The U.S. FDA for example explicitly defines data integrity as the "completeness, consistency, and accuracy of data" across its entire lifecycle. To satisfy this standard natively, the Circle Platform's capture mechanisms programmatically enforce the ALCOA+ framework—the universal compliance baseline utilized by the FDA, the European Medicines Agency (EMA), and the World Health Organization (WHO).

Under ALCOA+, every clinical data point must be:

- **Attributable:** Traceable to the specific individual, device, or system that generated it.
- **Legible:** Readable, permanent, and accessible over its lifecycle.
- **Contemporaneous:** Recorded at the exact time the activity occurs, without post-hoc reconstruction.
- **Original:** The first recorded observation or a certified true copy.
- **Accurate:** Correct, truthful, and free from algorithmic manipulation or truncation.
- **Complete, Consistent, Enduring, and Available:** Maintained systematically in a logical, coherent sequence for its statutorily required retention period.

Most data integrity citations in FDA warning letters do not stem from isolated human errors; they result from structural software failures, such as shared login credentials, unvalidated system overrides, and overwritten raw files. By embedding ALCOA+ controls directly into the point of data capture, the platform converts raw EHR inputs into "insurable integrity assets"—high-confidence datasets that easily satisfy stringent institutional and regulatory audits without retroactive remediation.

Semantic Consistency as a Prerequisite

To ensure interoperability across disparate enterprise nodes, the platform mandates standardized medical ontologies. Clinical categories are mapped natively to Logical Observation Identifiers Names and Codes (LOINC) for laboratory measurements and Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) for clinical findings and procedures.

This semantic baseline guarantees that a clinical concept defined at one hospital node possesses the exact same meaning across the entire federated network. Without this rigorous governance, federated analytics suffer from vocabulary fragmentation and schema drift, rendering multi-site cohort queries unusable.

The Unbroken FHIR-to-OMOP-to-CDISC Pipeline

Traditional HEOR and other forms of clinical research relies on manual data transcription from EHRs into clinical data interchange formats. This legacy process introduces high expense, transcription errors, hallucinations, and audit delays. The Circles Platform automates this transition via an unbroken, deterministic pipeline:

- **FHIR:** Raw, point-of-care data is captured and exchanged via HL7 FHIR, a standard optimized for real-time, patient-centric operational exchange.
- **OMOP CDM:** Because FHIR is not optimized for population-level analytics, the data is programmatically transformed into the OMOP Common Data Model (v5.4). OMOP normalizes heterogeneous local terminologies into standardized concepts specifically designed for large-scale observational research and cross-site reproducibility.
- **CDISC:** For life science use cases, the platform supports downstream translation into CDISC-compliant structures (such as SDTM formats) required for formal regulatory submissions to agencies like the FDA.

Transforming data across these standards frequently carries the risk of "fidelity loss" (mapping a highly specific FHIR source code to a broader OMOP standard concept) or

"information loss" (when the target standard lacks a corresponding field). The Circles Platform mitigates this by retaining the raw, cryptographically signed source payload locally behind the firewall, ensuring that the original microscopic fidelity is never permanently destroyed by analytical generalization.

Deterministic Lineage Metadata

Verifiable intelligence requires absolute transparency regarding how data was captured, transformed, and queried. The architecture enforces deterministic lineage metadata for every record, utilizing the World Wide Web Consortium (W3C) PROV data model.

W3C PROV defines data provenance as the precise, machine-readable information regarding the entities, activities, and people involved in producing a piece of data, allowing downstream users to assess its trustworthiness. By integrating healthcare-specific extensions of this standard (such as FHIR provenance), the platform ensures every computational event—from initial schema validation to federated query execution—records exact timestamps, actor identities, and the specific transformation logic applied.

This establishes a cryptographically signed, fully reproducible audit trail. When life science sponsors or regulators assess an analytical output, they can trace the final result directly back to the original clinical observation, legally validating its reliability.

FEDERATED INFRASTRUCTURE AND CRYPTOGRAPHIC VERIFICATION

Zero-Trust Enterprise Architecture

Traditional network designs rely on perimeter-based security models that assume internal actors or connected systems are inherently trustworthy once inside the firewall. In healthcare, this model is dangerously obsolete.

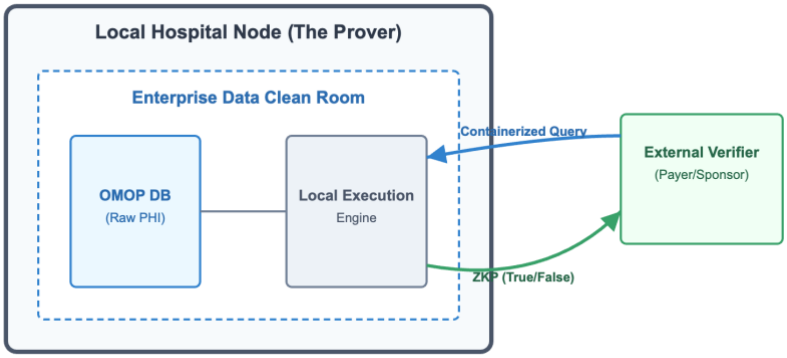
- **NIST Alignment:** The Circles Platform is implementing a Zero-Trust Architecture aligned with NIST SP 800-207 principles, operating under the foundational mandate: "Never trust, always verify."
- **Cryptographic Mutual Authentication:** Every interaction between enterprise nodes, federated query coordinators, and external auditing services will require cryptographic mutual authentication via public key infrastructure certificates. There are no persistent inbound open ports, shared database credentials, or ambient network

trust zones, effectively neutralizing lateral movement vectors for threat actors.

Decentralized Data Clean Rooms

Commercial "data clean rooms" frequently rely on centralized intermediaries or cloud-hosted multi-tenant environments where raw or pseudonymized data is pooled into a single repository managed by a third party. This architecture violates data sovereignty principles and recreates the honeypot vulnerability.

- **Edge Execution:** The Circles Platform will deploy decentralized data clean rooms that invert this model entirely. Computation is pushed directly to the data source rather than extracting data to the computation.
- **Perimeter Enforcement:** Analytic scripts, federated learning models, and cohort queries are containerized, cryptographically signed, and transmitted to the host enterprise node. The code executes locally behind the institution's firewall, and only aggregated, non-re-identifiable mathematical outputs or verified Boolean results cross the organizational boundary.



Zero-Knowledge Proofs

Verifiable data sharing in healthcare has historically required the full disclosure of underlying PHI to establish compliance, medical necessity, or trial eligibility. The platform replaces this high-risk data exposure with cryptographic certainty via ZKPs.

- **Operational Mechanics:** Utilizing advanced cryptographic protocols (such as zk-SNARKs), a hospital node (the prover) can mathematically prove to a payer, regulator,

or clinical researcher (the verifier) that a specific clinical condition has been met without revealing the underlying patient record, lab values, or demographic identifiers.

- **Practical Application:** For example, a prior authorization request is evaluated locally against clinical guidelines. The system generates a cryptographic proof verifying that the patient satisfies all medical necessity criteria. The payer accepts the immutable mathematical proof of compliance instantly, eliminating the need to transmit unredacted EHR charts, reducing administrative lag, and removing compliance liability.

Federated Query Orchestration and Differential Privacy

To analyze multi-site cohorts without centralizing records, the platform utilizes decentralized query orchestration mapped directly to the OMOP Common Data Model.

- **Distributed SQL Execution:** Complex analytical queries are translated into federated execution plans that run concurrently across participating enterprise nodes. Each node processes the query against its local data store and returns only aggregated summary statistics.

- **Differential Privacy Guards:** To prevent reconstruction attacks or small-cell disaggregation risks where individual patients could theoretically be re-identified from aggregate counts, the orchestration layer will apply programmatic differential privacy noise thresholds. This mathematically guarantees that individual patient identities remain obscured even across aggressive, multi-variable institutional queries.

SOVEREIGN IDENTITY AND PRIVACY ARCHITECTURE

Self-Sovereign Health Identity (SSI) and Decentralized Identifiers (DIDs)

Traditional identity management in healthcare relies on centralized enterprise master patient index (MPI) systems, institutional directory servers, or third-party identity providers. These legacy architectures create single points of failure, administrative tracking vulnerabilities, and cross-institutional fragmentation. When patient records or provider credentials must be matched across disparate hospital systems, probabilistic matching algorithms frequently introduce error rates, duplicate records, or data leakage.

- **The Cryptographic Foundation:** The Circles Platform replaces centralized identity

silos with a self-sovereign health identity (SSI) framework built on world wide web consortium (W3C) decentralized identifiers (DIDs) and cryptographic verifiable credentials (VCs).

- **Decoupled Authentication Mechanics:** Patients, clinicians, and institutional nodes maintain absolute sovereignty over their digital identities without depending on a central clearinghouse or authority. Identity verification occurs entirely peer-to-peer via cryptographic challenge-response protocols. Because DIDs are anchored to distributed public-key infrastructure rather than corporate databases, authentication proofs cannot be tracked, correlated, intercepted, or monetized by intermediaries.

- **Eliminating MPI Fragmentation:** By utilizing cryptographically bound DIDs and VCs, participating institutions achieve deterministic patient and provider matching across federated nodes without ever exposing or centralizing underlying demographic lookup tables.

Granular Smart Consent Frameworks

Conventional healthcare consent relies on static, multi-page HIPAA authorization forms that grant broad, indefinite, and opaque data usage rights to third parties—a practice under escalating scrutiny by federal regulators and state privacy statutes.

- **Programmatic Consent as Policy-and-Code:** The platform replaces static legal paperwork with cryptographically bound smart consent objects. Every data sharing, querying, or processing event is governed by explicit, machine-readable parameters embedded directly into the transactional metadata wrapper of the data payload.

- **Purpose and Temporal Limiting:** Smart consent enforces strict, enforceable boundaries across multiple dimensions:

    - *Purpose Limitation:* Explicitly defining the exact analytical or clinical application permitted (e.g., specific observational protocol execution versus commercial AI training exclusion).

    - *Granular Scope:* Restricting access to specific clinical domains or variable sets while keeping unrelated health history walled off.

    - *Temporal Expiration:* Establishing automated, cryptographically enforced expiration

timers that terminate data access when a research protocol concludes or a clinical trial closes.

- **Dynamic Patient Agency:** Patients and institutional custodians retain real-time sovereign control to update, restrict, or dynamically withdraw consent parameters throughout the data lifecycle, with changes immediately propagating across the federated network.

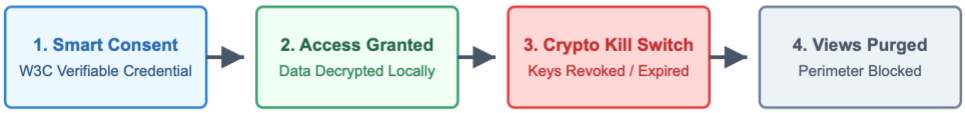
The Cryptographic Kill Switch: Enforceable Data Revocation and Recall

In distributed data architectures, revoking access after an institutional relationship terminates, a contract expires, or a privacy breach occurs has historically been slow, manual, and difficult to verify. Traditional environments struggle to ensure that remote copies of data are verifiably deleted or disabled once distributed.

- **Operational Mechanics:** The Circles Platform can implement an immediate, automated cryptographic kill switch managed via distributed revocation lists, cryptographically signed status assertions, and smart-contract state transitions.

- **Key Invalidation and Purge Protocols:** When an enterprise node or patient revokes authorization for a specific dataset, cohort query, or collaborative research circle, the cryptographic keys governing access are instantly invalidated across the network registry.

- **Perimeter Enforcement:** Decrypted local views are purged from active computational memory, and any downstream model containers or external nodes attempting to interact with the revoked asset are structurally blocked at the network perimeter. This ensures absolute compliance with statutory data withdrawal mandates, right-to-be-forgotten provisions, and institutional offboarding protocols without relying on manual compliance attestations.



SOVEREIGN ECONOMICS, TRIAL ARBITRAGE, AND INVESTOR ALPHA

The Split-IP Framework and Asset Protection

Traditional healthcare technology and research models force a binary choice: either an institution surrenders its proprietary data to a centralized aggregator in exchange for analytics, or it isolates its data entirely, forfeiting participation in high-value commercial research.

- **Architectural Separation:** The Circles Platform introduces a split-IP framework that programmatically separates the underlying raw clinical data (retained sovereignly behind the enterprise firewall) from the cryptographic proof artifacts, derived model weightings, and protocol execution scripts.

- **Attribution of Derived Value:** Under this framework, participating health systems retain 100% legal and operational custody of their local data assets while securing clean, auditable fractional ownership or licensing rights over the specific analytical outputs and model contributions generated within federated research circles. This eliminates third-party IP appropriation by commercial intermediaries and data brokers.

Trial Arbitrage Economics: Structural Cost Compression

Clinical trial recruitment and real-world evidence generation are currently bottlenecked by exorbitant operational overhead, driven primarily by manual chart abstraction, patient recruitment churn, and inefficient contract research organization middleman markups.

- **Arbitrage Mechanics:** By automating the FHIR-to-OMOP-to-CDISC pipeline and executing queries at the edge via decentralized data clean rooms, the platform

compresses clinical trial administrative overhead by 30% to 50%.

- **Elimination of Recruitment Churn:** Rather than relying on external advertising and manual chart reviews to identify eligible patients, participating provider nodes execute cryptographically secure cohort matching directly against live point-of-care OMOP repositories. Eligible patients are identified instantly through automated protocol screening, reducing screening failure rates and recruitment cycles from months to days.

Data Dividends and Participant Incentive Alignment

The legacy healthcare data economy extracts multi-billion-dollar value from patient clinical trajectories and institutional data assets without returning equitable economic value to the data originators.

- **Programmable Distribution:** The platform operationalizes transparent data dividends via smart-contract automation. When a federated research circle, pharmaceutical sponsor, or payer utilizes a verified institutional node or patient cohort for commercial drug discovery or validated outcomes research, micro-compensations and research grants are distributed automatically based on immutable contribution metrics.

- **Ethical and Economic Alignment:** This native incentive alignment transforms patients and providers from passive, exploited subjects of data extraction into active economic participants, driving higher data fidelity and long-term engagement in longitudinal research cohorts.

Valuation Multiple Expansion for Enterprise Nodes

For healthcare executives and venture investors evaluating infrastructure deployments, the economic impact extends beyond direct operational savings to enterprise valuation multiples.

- **Transition from Cost Center to Asset Class:** Traditional health system IT deployments are treated as depreciating operational cost centers. By establishing sovereign, audit-ready "insurable integrity assets" through decentralized provenance pipelines, institutional nodes acquire proprietary, monetizable data syndication capabilities.

- **Defensible Moats:** The combination of ZKP verification, split-IP structures, and immutable compliance auditing elevates enterprise software valuations from standard multi-tenant SaaS multiples to high-margin, tech-enabled network infrastructure status, reflecting defensibility against regulatory obsolescence and competitive poaching.

Illustrative Use Cases

There are many use cases for the Circles Platform federated data architecture throughout the healthcare ecosystem. Two examples follow:

Life Science Collaboration: Decentralized Multi-Site Oncology Trial

The Operational Challenge: A global pharmaceutical sponsor needs to assemble a rare-mutation oncology cohort across five distinct hospital systems spanning multiple states with strict data localization laws. Under legacy models, this requires massive legal data-use agreements, centralized data aggregation (exposing the sponsor and hospitals to massive breach liability), and months of manual chart abstraction.

Platform Execution via Sovereign Economics:

- The sponsor publishes a cryptographically signed Observational Protocol (OP) and cohort criteria to the federated network.
- Participating hospital nodes execute the query locally against their OMOP CDM repositories using decentralized data clean rooms. No raw PHI leaves any hospital firewall.
- The network generates Zero-Knowledge Proofs confirming that an eligible patient cohort exists across the five sites, matching exact inclusion criteria without disclosing patient identities.
- Smart contracts automatically execute trial participation agreements and provision data dividends to the participating hospital nodes upon protocol validation.
- The trial initiates in days rather than months, with zero centralized data exposure, full regulatory compliance under state localization statutes, and an intact Split-IP

framework protecting institutional and sponsor assets.

Enterprise Use Case: Frictionless, Zero-PHI Prior Authorization

The Operational Challenge: Traditional prior authorization workflows require healthcare providers to assemble and transmit extensive, unredacted patient medical records—including clinical notes, laboratory values, and diagnostic imaging reports—to insurance payers via manual portals, fax lines, or insecure electronic channels.

This legacy process imposes severe administrative overhead on clinical staff, causes damaging delays in patient care, and exposes both provider and payer organizations to acute HIPAA compliance violations and data leakage risks.

Platform Execution via Technical Architecture:

- **Local Rule Evaluation:** When a clinician orders a specialized procedure, surgical intervention, or advanced therapeutic, the provider's enterprise node automatically evaluates the patient's local FHIR and OMOP data repositories against the payer's codified clinical guidelines.
- **Cryptographic ZKP Generation:** Rather than exporting unredacted electronic health records, the institutional node executes a local computation to generate a cryptographic Zero-Knowledge Proof (such as a zk-SNARK). This mathematical proof attests that the patient fully satisfies all established medical necessity criteria (e.g., biomarker thresholds, failed prior lines of therapy) without revealing underlying Protected Health Information, clinician notes, or raw demographic values.
- **Smart Consent Packaging:** A machine-readable smart consent object wraps the cryptographic proof, enforcing strict purpose limitation and preventing any secondary use or storage of the verification artifact by the payer.
- **Instantaneous Automated Adjudication:** The transmission routes solely the immutable mathematical proof of compliance to the payer. The payer's verification engine validates the proof instantly against its policy parameters, issuing automated authorization in seconds rather than weeks.

Strategic and Economic Impact:

- **Administrative Compression:** Eliminates manual chart abstraction and back-and-forth peer-to-peer appeals, reducing prior authorization processing costs by over 70%.
- **Regulatory Compliance:** Completely satisfies the HIPAA "minimum necessary" standard and state data localization mandates by ensuring zero PHI crosses the institutional firewall during the adjudication process.
- **Clinical Velocity:** Removes administrative bottlenecks to accelerate time-to-treatment, directly improving patient health outcomes and eliminating institutional revenue cycle friction.

GOVERNANCE, ACCOUNTABILITY, AND REGULATORY COMPLIANCE

Navigating the HIPAA Privacy Rule: Automating Expert Determination

Traditional data de-identification relies on the HIPAA "Safe Harbor" method—manually stripping 18 identifiers from a dataset. However, Safe Harbor degrades data fidelity, rendering it nearly useless for advanced longitudinal research, while still leaving the data vulnerable to re-identification attacks.

- **Cryptographic Expert Determination:** The HIPAA Privacy Rule permits a second, more rigorous path to compliance: the Expert Determination method. The Circles Platform operationalizes this standard programmatically.
- **Zero-Knowledge as Legal Proof:** Because the federated architecture uses Zero-Knowledge Proofs (ZKPs) and differential privacy, raw PHI is never extracted or viewed. A statistical expert can universally attest that the risk of re-identification from a cryptographic proof or model weight is "very small," continuously satisfying the Expert Determination mandate without degrading the underlying clinical fidelity or requiring bespoke manual reviews for every query.

GDPR, Schrems II, and Cross-Border Data Sovereignty

Multinational clinical trials and global AI development are currently paralyzed by international data localization mandates. Following the European Court of Justice's *Schrems II* ruling, transferring EU citizen health data to U.S.-based cloud servers carries profound

legal and financial risk.

- **Compute-to-Data Architecture:** The platform resolves this by moving the computation to the data, rather than the data to the computation. Patient records remain permanently localized behind the host nation's enterprise firewall, fully satisfying GDPR Article 5 (Data Minimization) and sovereign localization laws.

- **Lawful Cross-Border Collaboration:** Only cryptographically secured model weightings, aggregated cohort counts, and mathematical proofs traverse international borders. Because these artifacts mathematically preclude reverse-engineering to an individual patient, they fall outside the strictures of PHI export bans, enabling frictionless trans-Atlantic life science collaboration.

Legal Liability Quarantines and Institutional Ring-Fencing

A primary deterrent to joining federated data networks is the fear of shared liability: if one hospital in a network suffers a breach, do all participating institutions face regulatory exposure?

- **Zero-Trust Network Execution:** Based on NIST SP 800-207 Zero-Trust principles, the platform architecture establishes a strict legal and technical liability quarantine. Nodes do not share physical infrastructure, databases, or decryption keys.

- **Isolated Blast Radius:** If a participating institutional node is compromised by an external threat actor, the breach is physically and cryptographically isolated to that specific enterprise firewall. The federated network design ensures there is no lateral pathway for an attacker to pivot from one hospital's system into another, cleanly severing joint institutional liability.

Alignment with Federal Health IT and TEFCA Frameworks

Modern U.S. healthcare policy actively encourages standardized data exchange while heavily penalizing information blocking.

- **Bypassing Centralized Intermediaries:** While the Trusted Exchange Framework and Common Agreement (TEFCA) establishes interoperability baselines through Qualified Health Information Networks, legacy QHINs often act as centralized message brokers. The Circles Platform can achieve compliance with Office of the National Coordinator (ONC) standards—including HTI-1, HTI-2, and the 2027

FHIR API mandates—while entirely bypassing centralized brokers.

- **Sovereign Node Compliance:** By utilizing point-of-sale FHIR endpoints behind the enterprise firewall, participating institutions fulfill federal interoperability obligations without surrendering network governance to third-party operators.

FDA Compliance: Real-World Evidence and Audit-Ready Provenance

Regulatory bodies, including the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), maintain stringent evidentiary standards for clinical data submitted in support of drug approvals or software-as-a-medical device (SaMD) clearances.

- **Regulatory-Grade Datasets:** Under the FDA's Real-World Evidence Framework, administrative claims data is frequently discounted due to error rates. The platform's native capture of human ground truth, mapped to the OMOP CDM, generates datasets that satisfy FDA data reliability standards.

- **W3C PROV Traceability:** Because every data transformation and cohort extraction is immutably recorded using W3C PROV provenance standards, life science sponsors can submit federated research findings directly to regulatory agencies with complete, mathematically verifiable audit trails, eliminating the risk of data rejection due to unverified lineage.

AI Governance and Policy-as-Code Auditing

The global regulatory environment for artificial intelligence—highlighted by the EU AI Act and the NIST AI Risk Management Framework—enforces rigorous compliance burdens on high-risk medical algorithms.

- **Mitigating Training Bias:** Traditional AI relies on uncurated, retrospective data dumps that embed historical documentation errors. The Circles Platform's schema-on-capture validation and protocol-locked data dictionaries ensure that training datasets are pristine and verifiably representative at the point of origin.

- **Deterministic Governance:** Manual monitoring by Institutional Review Board, Medical Ethics Committee, Notified Body and similar regulatory entities is replaced by Policy-as-Code engines. Legal and ethical parameters can be translated into executable scripts that govern every network transaction. Automated guardrails continuously

monitor system interactions, instantly blocking non-compliant queries and replacing subjective reviews with objective, real-time accountability.

CONCLUSION

The transition from centralized data lakes to federated, edge-compute architectures is no longer merely a theoretical IT optimization; it is a strict regulatory and commercial imperative. The legacy model of silent data brokerage, reliance on administrative claims proxies, and unchecked institutional liability has reached terminal velocity.

For healthcare providers, payers, and life science sponsors, adopting the Circles Platform architecture fundamentally restructures the data economy—transitioning institutions from passive victims of data extraction into sovereign participants in a highly lucrative, cryptographically secure research network.

Recent federal regulatory shifts confirm this trajectory. Under the FDA December 2025 updated RWE guidance, the agency eliminated the blanket requirement that sponsors submit individually identifiable participant-level data in regulatory submissions, provided that the underlying data's provenance, validation, and traceability are rigorously documented.

The Circles Platform's approach to capturing human ground truth through automated provenance pipelines perfectly satisfies this new standard, enabling the commercial use of large-scale, de-identified datasets without compromising regulatory rigor. Concurrently, the Office of the National Coordinator, via the HTI-2 rule, has established aggressive timelines for FHIR-enabled APIs, including a January 2027 mandate for prior authorization interoperability.

REFERENCES

Regulatory Frameworks and Compliance

European Union Artificial Intelligence Act (2024): Regulation (EU) 2024/1689 establishing harmonized rules, data governance requirements, and risk-based oversight for artificial intelligence systems in the EU. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689

HIPAA Minimum Necessary Standard (45 CFR 164.502): U.S. Department of Health and Human Services (HHS) guidance mandating covered entities to limit the use and disclosure of Protected Health Information (PHI) to the minimum extent necessary. https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/minimum-necessary-requirement/index.html

FTC Health Breach Notification Rule & Privacy Enforcement: Federal Trade Commission (FTC) guidance and enforcement protocols regarding the unauthorized sharing of health data by digital health applications and third-party vendors. https://www.ftc.gov/business-guidance/resources/complying-ftcs-health-breach-notification-rule-0

Cryptographic Verification and Architecture

Zero-Knowledge Proofs in Healthcare: Peer-reviewed analysis detailing the application, completeness, and soundness requirements of Zero-Knowledge Proofs (ZKPs) for privacy-preserving medical data sharing without revealing underlying health records. https://pmc.ncbi.nlm.nih.gov/articles/PMC12534302/

NIST Zero Trust Architecture: National Institute of Standards and Technology (NIST) Special Publication 800-207 and 1800-35 establishing the standard for transitioning from perimeter-based security to identity-based access controls in distributed data environments. https://csrc.nist.gov/pubs/sp/1800/35/final

Data Interoperability and Provenance Standards

OMOP Common Data Model: The Observational Health Data Sciences and Informatics (OHDSI) specification for standardizing the structure and content of diverse observational healthcare data into a common format. https://ohdsi.github.io/CommonDataModel/

W3C PROV Data Model: The World Wide Web Consortium (W3C) foundational specification for recording immutable provenance information and documenting the core structures of data lineage. https://www.w3.org/TR/prov-dm/

GA4GH Data Connect: The Global Alliance for Genomics and Health (GA4GH) standard for the discovery, schema definition, and search of biomedical data across federated networks. https://github.com/ga4gh/data-connect

Global Real-World Evidence (RWE) Networks

FDA Real-World Evidence Program: Analysis of the U.S. Food and Drug Administration's strategic framework for integrating real-world data and verifiable electronic source data into regulatory decision-making and post-approval studies. https://pmc.ncbi.nlm.nih.gov/articles/PMC7196856/

EMA DARWIN EU: The European Medicines Agency (EMA) Data Analysis and Real World Interrogation Network framework, demonstrating the regulatory standard for decentralized, federated real-world evidence generation. https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence/data-analysis-real-world-interrogation-network-darwin-eu

GLOSSARY OF TERMINOLOGY

- **ALCOA+:** A regulatory framework utilized by the FDA and EMA to ensure data integrity, mandating that data must be Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Enduring, and Available.
- **CDISC (Clinical Data Interchange Standards Consortium):** A global standards development organization whose data structures (e.g., SDTM) are required for regulatory submissions to agencies like the FDA.
- **Data Clean Room (DCR):** A secure environment that allows multiple parties to collaborate on sensitive data without exposing the underlying raw information. In a decentralized architecture, DCRs execute locally at the network edge.
- **DID (Decentralized Identifier):** A cryptographically verifiable, self-sovereign digital identifier that does not depend on a centralized registry, identity provider, or certificate authority.
- **FHIR (Fast Healthcare Interoperability Resources):** A standard developed by HL7 for the electronic exchange of healthcare information, optimized for real-time, transactional clinical workflows and API interoperability.
- **Human Ground Truth:** The foundational, objective clinical facts derived directly from patient physiology, laboratory results, and direct clinical interactions, uncorrupted by administrative billing proxies or retrospective transcription errors.
- **OMOP CDM (Observational Medical Outcomes Partnership Common Data Model):** A standardized database schema designed to harmonize disparate healthcare data formats for the purpose of executing large-scale, population-level observational research.
- **RWE (Real-World Evidence):** Clinical evidence regarding the usage and potential benefits or risks of a medical product derived from the analysis of real-world data (RWD) outside of traditional clinical trials.
- **SSI (Self-Sovereign Identity):** A digital identity paradigm that gives individuals and institutions complete control over their digital credentials and how their personal data is shared, utilizing cryptographic proofs.

- **TEFCA (Trusted Exchange Framework and Common Agreement):** A U.S. federal initiative managed by the ONC to establish a universal governance and technical baseline for nationwide health information network interoperability.
- **ZKP (Zero-Knowledge Proof):** A cryptographic protocol allowing one party (the prover) to mathematically prove to another party (the verifier) that a specific statement is true without revealing any additional information beyond the validity of the statement itself.
- **ZTA (Zero-Trust Architecture):** A cybersecurity paradigm aligned with NIST SP 800-207 that eliminates implicit trust within a network, requiring continuous, cryptographic mutual authentication for every interaction regardless of origin.