SSR Journal of Artificial Intelligence (SSRJAI)

OPEN ACCESS

Homepage: https://ssrpublisher.com/ssrjai/

Volume 3, Issue 8, 2026 ISSN: 3049-0413

An Adaptive Quantum-Resilient Self-Sovereign Identity (AQR-SSI) Framework Using Blockchain, Zero-Knowledge Proofs, and Explainable AI for Secure Electronic Health Records Sharing

Peter Wonah Odey, Abdulazeez Ogunlade Babatunde and Myedan Nahum Eli

Centre for Cyberspace Studies, Nasarawa State University, Keffi, Nasarawa State, Nigeria

Received: *05.07.2026* | ***Accepted:*** *25.07.2026* | ***Published:*** *16.08.2026*

***Corresponding author:** Abdulazeez Ogunlade Babatunde

DOI: *10.5281/zenodo.21967154*

Abstract **Original Research**

The digitization of Electronic Health Records (EHRs) has significantly improved care coordination, but has concurrently aggregated highly sensitive identity and clinical access data into centralized repositories. These repositories serve as attractive targets for unauthorized breach, insider misuse, and, with the rapid development of cryptographically relevant quantum computers, future retrospective decryption (commonly known as "harvest-now, decrypt-later" attacks). This paper presents AQR-SSI, an Adaptive Quantum-Resilient Self-Sovereign Identity framework designed to address these fundamental vulnerabilities. The architecture seamlessly integrates: (i) a permissioned blockchain ledger for decentralized identifiers (DIDs) and verifiable credentials (VCs); (ii) zero-knowledge proofs (ZKPs) for unlinkable, selective disclosure of patient attributes during Electronic Health Record (EHR) sharing; (iii) a hybrid post-quantum cryptographic (PQC) suite combining NIST-standardized lattice- and hash-based primitives (ML-DSA, Falcon, SPHINCS+, ML-KEM) with classical algorithms to facilitate a secure migration period; and (iv) an explainable AI (XAI) monitoring layer that continuously audits access requests and generates human-readable explanations for anomaly flags via SHAP and LIME-based attributions. We detailed the structural reference architecture. We also described the actor and trust model, cryptographic protocol design, smart-contract-based consent and revocation mechanisms, and conducted an in-depth, threat-model-driven security analysis. Additionally, we provide an illustrative, literature-grounded performance analysis estimating the computational and storage overhead of the hybrid signature scheme relative to classical elliptic-curve baselines, and discuss strict compliance alignment under HIPAA and GDPR. The AQR-SSI framework is structurally adaptive: its built-in cryptographic agility ensures algorithm suites can be easily swapped as NIST standards evolve, while the XAI layer remains fully auditable as the underlying anomaly detection models undergo periodic retraining. We conclude by identifying key open problems to prioritize in empirical follow-up work, specifically focusing on on-chain ZKP verification cost at scale, PQC signature bloat, and the real-world usability of patient-held digital wallets.

Keywords: Self-sovereign identity, Blockchain, Zero-knowledge proofs, Explainable AI, Electronic health records.

1. INTRODUCTION

In a modern landscape where, digital interactions are technology driven and increasingly predominant, the secure management and absolute confidentiality of consumers' digital identities have emerged as critical pillars of societal infrastructure [1]. This is particularly crucial in the healthcare sector, where individuals are routinely required to disclose highly sensitive personal and clinical information in diverse digital environments [2]. The contemporary reliance on centralized Identity Providers (IdPs) and monolithic healthcare databases presents a severe systemic risk: a compromise of the central identity provider exposes not only basic login credentials but directly links a patient's real-world identity to their entire history of clinical interactions [1, 2].

To overcome these challenges, Self-Sovereign Identity (SSI) has emerged as a paradigm-shifting model. SSI places cryptographic custody of decentralized identifiers (DIDs) and verifiable credentials (VCs) directly in the hands of the patient [3]. This allows patients to present clinical or identity assertions to relying parties (such as hospitals or insurers) without relinquishing ownership or control of their underlying records. Although recent reviews of blockchain enabled SSI in healthcare show promising concepts, most existing work remains stuck at the prototype stage, focusing on narrow identity verification or basic credential exchange [4]. Critical medical assessment gaps persist in areas of cryptographic longevity, system interoperability, governance, and explanatory oversight.

To the best of current scientific knowledge, no existing digital identity framework simultaneously satisfies the four core requirements necessary for a clinically deployable, future proof medical access system:

- Patient-controlled, highly portable identity and credential management across multiple health systems.
- Selective and cryptographically unlinkable disclosure of clinical information at the point of data sharing to prevent profiling.
- Thorough protection against both classical and quantum adversaries, mitigating the immediate threat of retrospective decryption ('harvest-now, decrypt-later' attacks) on sensitive clinical records.
- Transparent, fully explainable auditing of automated data access, allowing compliance officers to investigate suspicious events without relying on blind trust in black-box machine learning models.

While individual studies address these issues in isolation such as blockchain based SSI for access control, zero-knowledge proofs for selective disclosure, post-quantum signatures for ledger security, or explainable AI for system auditing these four domains have developed independently. This fragmented landscape leaves an urgent, unaddressed research gap in designing an integrated, unified architectural framework capable of executing all four requirements simultaneously [4].

The primary objective of this research is to design and propose the AQR-SSI framework using blockchain, zero-knowledge proofs, and explainable AI for secure electronic health records sharing.

The objectives include:

- Layered Reference Architecture (AQR-SSI): Establishing a unified trust model that integrates decentralized identifiers, verifiable credentials, zero-knowledge proofs, hybrid classical/post-quantum cryptography, and explainable AI monitoring.
- Preserving Privacy Authentication Protocol: Formulating an anonymous, unlinkable patient authentication and consent-scoped credential presentation protocol that integrates trusted-setup-free zero-knowledge proofs with a hybrid classical/post-quantum signature scheme to support seamless cryptographic migration.
- Lifecycle Management based on Smart Contract: Designing standardized smart contracts for consent management, credential revocation registries, and on-chain cryptographic commitment storage to minimize ledger footprints.

- Access Auditing Pipeline based on XAI: Constructing an explainable anomaly-detection monitoring pipeline that translates complex SHAP and LIME attributions into natural, clinician-interpretable rationales.
- Security and Performance analysis: Executing an all-inclusive threat model driven security analysis, discussing HIPAA/GDPR regulatory compliance, and producing grounded performance estimate while comparing proposed hybrid signature suites against classical Elliptic Curve Digital Signature Algorithm (ECDSA) baselines based on literature.

This work focuses on the architectural design, cryptographic protocol formulation, and analytical validation of the AQR-SSI framework. The framework is designed for permissioned healthcare consortia (e.g., regional hospital networks, insurers, and regulators). Endpoint device security (e.g., patient mobile wallet physical theft, malware on user devices) is treated as a complementary hardware-level concern (such as secure enclaves and biometrics) and is explicitly out of scope for this protocol-level study. Furthermore, the performance characterization is literature grounded and analytical, designed to guide implementers rather than report empirical results from a multi-node clinical deployment.

2: LITERATURE REVIEW

A 2026 scoping review of blockchain-enabled self-sovereign identity in healthcare demonstrated that current applications focus heavily on identity verification, credential management, and privacy preserving data exchange across EHR, mobile health, and access-control domains [5,6,7]. These systems are commonly built on decentralized identifiers (DIDs), verifiable credentials (VCs), smart contracts, and basic privacy-enhancing technologies like selective disclosure [5]. However, the review highlights that the field remains highly early-stage, with limited real-world validation and a major structural gap between conceptual designs and clinical-grade deployment [5]. A taxonomy of 390 papers on distributed-ledger identity decomposed the design space into trust-anchor implementation and identity architecture, revealing high fragmentation of design choices. Parallel efforts have proposed non-fungible patient tokenization for health information exchange and blockchain-based medical record management systems, and have begun exploring SSI-secured federated learning across hospitals using DIDs and VCs to authenticate participating research institutions [6, 7].

Zero-Knowledge Proofs (ZKPs) allow one party (the prover) to convince another (the verifier) of the absolute truth of a mathematical statement without revealing any underlying information [5]. A recent survey on ZKP usage in identity and blockchain applications showed that zk-SNARKs, zk-STARKs, and bulletproofs are the dominant constructions [4, 5, 6]. Critically, SNARK-based schemes usually require a complex trusted setup (ceremony), whereas STARK-based schemes are fully transparent (require no trusted setup) and provide additional post-quantum-plausible security based on secure hash functions rather than pairing-based assumptions [4, 5]. Directly relevant to EHR sharing is a 2025 study that combined STARK proofs with anonymous credentials to provide unlinkable patient authentication on both public and private ledgers, detailing highly optimized on- and off-chain algorithms [5, 6]. Consent-management surveys also highlight selective disclosure (the ability to prove a specific attribute, such as age > 18, without revealing the actual birth date) as the fundamental building block for privacy preservation in modern clinical systems [4].

The National Institute of Standards and Technology (NIST) has standardized the ML-DSA, SLH-DSA [8], and is finalizing FN-DSA as post-quantum digital signature algorithms, alongside ML-KEM for key encapsulation [8]. Under NIST Level 2 benchmarks, ML-DSA is significantly faster than Falcon or SPHINCS+ signing (approx. 0.65 ms vs 3.28 ms and 131.9 ms, respectively), but carries substantial signature and public key sizes. For comparison, classical ECDSA/EdDSA signatures are a compact 64–73 bytes, while Dilithium2 (ML-DSA) requires 2,420 bytes, Falcon-512 requires 666 bytes, and SPHINCS+ requires up to 41

kilobytes [8]. Because a naive, direct replacement of classical signatures with PQC primitives in blockchains leads to intolerable storage, bandwidth, and consensus-liveness overhead, the literature converges on architectural mitigations [4, 8]. These include signature aggregation, Merkle-tree compression, and hybrid dual-signature transactions (carrying both classical and PQC signatures) during a migration window [9, 10]. Healthcare-specific proposals include lattice threshold signcryption schemes based on the Small Integer Solution (SIS) and Learning-With-Errors (LWE) assumptions, and lattice-based group signature schemes with backward unlinkability for medical blockchains storing encrypted records in IPFS while anchoring access credentials on-chain [6, 7].

Opaque, black-box anomaly detection models are notoriously difficult for clinicians, security auditors, and healthcare regulators to trust or legally challenge [11]. To resolve this, recent literature combines unsupervised or supervised anomaly detectors with post-hoc explanation methods. In a landmark study [1, 6, 7, 12] of AI-based anomaly detection for health-record access auditing, SHAP (SHapley Additive exPlanations) attributions surfaced provider mismatch as the strongest anomaly indicator, while off-hours or repeated access patterns also carried high SHAP importance, corroborating qualitative security findings with quantitative evidence [12]. Broader scoping reviews on XAI in EHR research found that SHAP is the most common attribution method (used in the vast majority of reviewed studies) [1, 6, 7, 11], usually applied on top of gradient-boosted trees or random forests for structured EHR logs, whereas gradient- and propagation-based methods dominate for unstructured imaging data. Evaluation metrics for explanation faithfulness, stability, and comprehensibility in auditing contexts are provided by evaluation frameworks for white-box XAI methods like Integrated Gradients, Layer-wise Relevance Propagation, and Deep-SHAP [13].

The literature review reveals a clear and limited number of studies that integrate transparent, quantum-resistant Zero-Knowledge Proofs (ZKPs), post-quantum cryptography, Self-Sovereign Identity (SSI), and Explainable Artificial Intelligence (XAI) within a single unified framework for secure EHR management [11, 12]. While existing studies [1, 11, 12,] have addressed these technologies individually or in partial combinations, there is no evidence of an integrated approach that simultaneously:

- Employs transparent, quantum-resistant ZKPs for selective disclosure of clinical attributes.
- Combines NIST-standardized post-quantum digital signatures with DID/VC-based trust anchors for SSI.
- Provides explainable access-control decisions that are both cryptographically verifiable and interpretable by human auditors.

3: METHODOLOGY

Principles of Design - Theoretical and conceptual framework: proposed AQR-SSI architecture

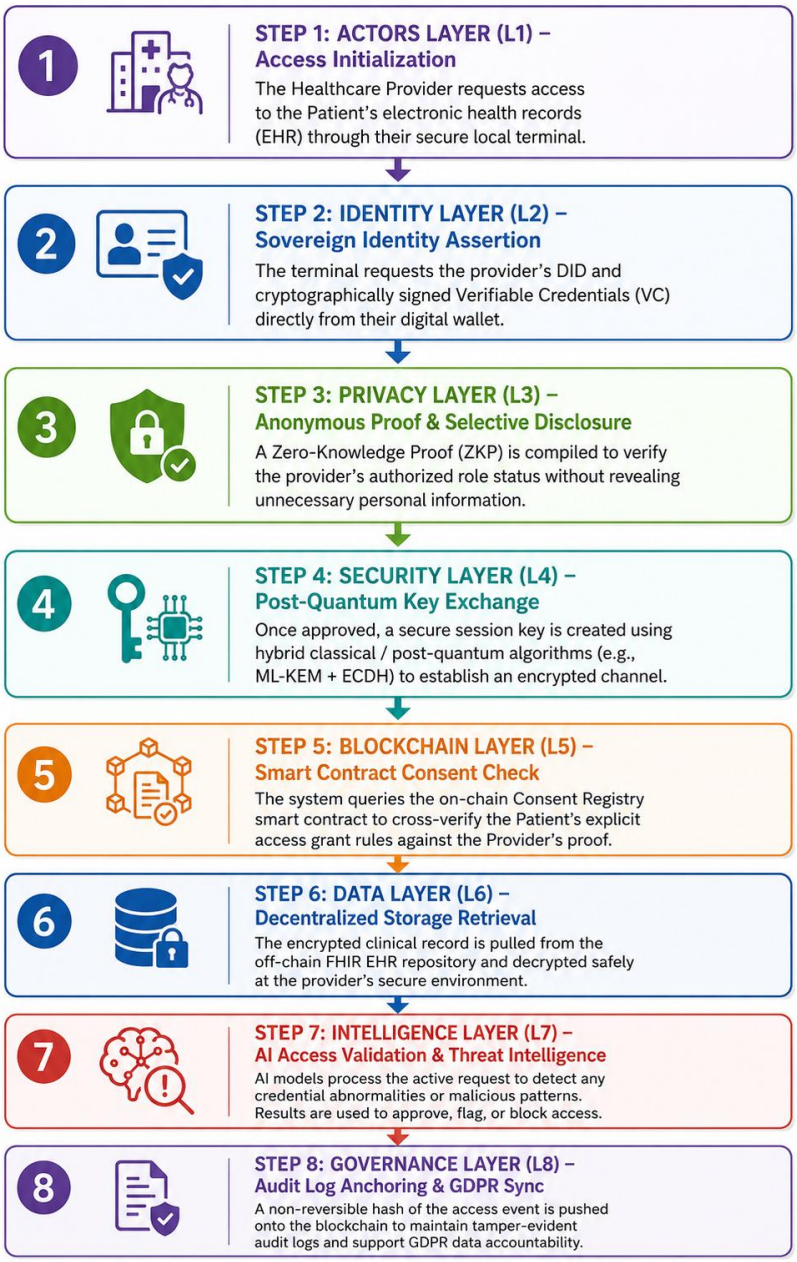
The AQR-SSI framework is designed upon five foundational design principles to ensure clinical viability, longevity, and trust [5]:

1. Patient Custody: Cryptographic private keys and credentials are saved strictly within a patient-controlled wallet (mobile or hardware-backed). No third party can unilaterally access or disclose a patient's credential.
2. Data Minimization by Construction: Verifiers learn only the minimum necessary mathematical predicate (e.g., proving 'I am over 18' without revealing the birth date), ensuring that actual clinical or personal attribute values are never unnecessarily shared.
3. Cryptographic Agility: Signature and key encapsulation (KEM) algorithms are registered per-DID and per-credential, allowing the seamless coexistence, rotation, and upgrading of classical, post-quantum, and hybrid suites without requiring a disruptive ledger fork.
4. Off-chain Payload, On-chain Proof: Clinical EHR payloads are never stored on the ledger. The blockchain ledger anchors only DIDs, credential-schema hashes, revocation states, consent-policy commitments, and ZKP verification keys.

5. Explainability as a First-Class Requirement: Any automated access-control decision that deviates from an established consent policy must generate a human-readable rationale, not just an opaque allow/deny bit.

Healthcare Architecture Interaction Flow - End-to-end operational lifecycle & sequence model

Figure 1: An 8-layer Sovereign Identity, Cryptographic, And Auditability Framework



When a Healthcare Provider requests access to a Patient's EHR, the flow initiates at the Actors Layer. The provider presents their Verifiable Credential from the Identity Layer. The Privacy Layer performs selective disclosure and verifies the Patient's active consent stored in the Blockchain Layer's consent registry. Upon positive validation, the provider receives a decryption key protected by Post-Quantum Key Management (Security Layer), allowing them to retrieve and decrypt the off-chain FHIR record from the Data Layer. Simultaneously, the

Intelligence Layer continuously monitors this transaction for anomalous access vectors, while the Governance Layer outputs a signed regulatory event log [3, 4, 6, 7, 9].

Table 1: AQR-SSI layered reference architecture structured into six cooperating and integrated layers

Layer	Function	Key Components
Identity (SSI) Layer	Creation and lifecycle management of patient, provider, and device identities	DIDs, DID documents, verifiable credentials, patient wallet
Blockchain Ledger Layer	Tamper-evident anchoring of identity, revocation, and consent policy states	Permissioned DLT (Hyper ledger Fabric) or L2 rollup (Starknet) with public anchoring
ZKP Layer	Selective disclosure and anonymous patient authentication	zk-STARK circuits, commitment schemes, nullifiers for replay/unlinkability control
PQC / Hybrid Crypto Layer	Signature and key-exchange primitives resistant to quantum adversaries	ML-DSA, FN-DSA (Falcon), SLH-DSA (SPHINCS+), ML-KEM, dual-signature transition
Consent & Revocation Layer	Smart-contract enforcement of patient-granted access policies	Consent registry contract, revocation registry contract, policy-hash commitments
XAI Monitoring Layer	Continuous auditing of access events with human-readable justifications	Feature pipeline, anomaly detection model, SHAP/LIME explainer, alert dashboard

AQR-SSI Actors and Trust Model: The AQR-SSI framework defines four distinct classes of actors operating within a healthcare consortium:

- The Patient: The patient holds secure digital wallet containing their **DIDs (Decentralized Identifiers)** and **VCs (Verifiable Credentials),** this represents the sole authority capable of authorizing data disclosure or updating consents as in SSI. By functioning as the sole gatekeeper, the patient ensures total privacy and compliance with data regulations, and safety against unauthorized breaches [6, 7].
- The Issuers: They are the trusted institutions (hospitals, laboratories, pharmacies, insurers) that verify real world identities and also sign verifiable credentials attesting to clinical or demographic facts. This tamper proof information verification, checkmates fraud, and enables real-time sharing of health or demographic facts [6, 7].
- The Verifiers (Relying Parties): The Verifiers handles clinicians, referred specialists, or clinical research organizations. They validate cryptographic proofs (such as in zero-knowledge proofs) submitted by clinicians, researchers, and specialists. This ensures access is granted only when specific mathematical predicates (e.g., specific demographic thresholds or diagnosis criteria) are met, protecting patient privacy [6, 7].

- The Ledger Network: This is a consortium of healthcare sector nodes (regional health authorities, accredited networks) that execute BFT (Byzantine Fault Tolerance) consensus, anchor revocation states, and enforce smart contracts. Because no single entity holds ultimate control, this setup eliminates data silos, enforces rigid privacy (like HIPAA/GDPR), and prevents malicious tampering [6, 7].

The trust model assumes that issuers are honest but can be compromise (their signing keys can be cleanly revoked and rotated), verifiers are honest but inquisitive (they follow the protocol but may attempt to link patient profiles across sessions unless restricted by ZKP unlinkability), and the ledger network is BFT up to the standard consensus threshold.

Identity Life Cycle: The identity life cycle initiates when a patient generates a DID and registers the corresponding DID Document containing public keys and service endpoints on the ledger [8]. Following an out of band identity verification, an issuer signs a verifiable credential binding attributes to the patient's DID using a post-quantum or hybrid signature key [8, 9]. The raw credential resides solely in the patient's wallet, while a revocation registry entry (e.g., a Merkle-tree leaf) is anchored on-chain. This ensures that the patient can construct ZK proofs of non-revocation without revealing which specific credential is being verified [8].

EHR Sharing Data Flow with Selective Disclosure: When a verifier requests access, the patient's wallet gets a proof request that says what predicate is needed (e.g., "active coverage under a recognized insurer" or "authorized to see cardiology records for encounter X") [9, 10]. The wallet generates a zk-STARK proof such that it [9, 10]; (a) has a validly signed credential satisfying the predicate, (b) the credential is not on the revocation registry, and (c) a nullifier is fresh, thus preventing replay while avoiding cross session linkability. The verifier then checks the proof against the on-chain verification key, and if successful is given a decryption ability (via a proxy re-encryption or attribute-based encryption key) to the specific off-chain EHR payload stored in an encrypted object store or IPFS-style content addressing system, in keeping with designs in which desensitized, encrypted records are stored off-chain and only access credentials anchored on-ledger [10]. The XAI monitoring layer logs every access event, both successful events and denied events.

Quantum Resilient Cryptographic Suite: To achieve balance in core engineering tradeoffs, such as signature size, computation cost, and long-term security margin, AQR-SSI adopts a hybrid per role signature strategy instead of a single algorithm, following the migration guidance in the post-quantum blockchain literature [11, 12, 13].

Table 2: Role assignment and hybrid crypto suite

Primitive	Type	Assigned Role in AQR-SSI	Representative Published Figures	Source / Context
ML-DSA (Dilithium)	Lattice based signature	High-frequency signing: session proofs, transaction endorsement	~0.65 ms sign (Level 2); ~2,420 B signature [8, 13, 14]	NIST FIPS 204
FN-DSA (Falcon)	NTRU-lattice signature	DID document signing where compact signatures matter (mobile wallets)	~3.28 ms sign; ~666 B signature [8, 13]	NIST FIPS 206
SLH-DSA (SPHINCS+)	Hash-based signature	Long-lived issuer root keys and revocation-registry roots (maximal conservatism)	~131.9 ms sign; up to ~41 KB signature [8, 12, 15]	NIST FIPS 205

Primitive	Type	Assigned Role in AQR-SSI	Representative Published Figures	Source / Context
ML-KEM (Kyber)	Lattice based KEM	Establishing quantum safe channels for offchain payload key exchange	768 B–1 KB ciphertext/key range [8, 11]	NIST FIPS 203
ECDSA/EdDSA (classical)	Elliptic-curve signature	Carried alongside PQC signature during migration window (dual signature transactions)	~64–73 B signature [8, 13]	Classical Baseline

Table 2 summarizes the roles assigned to each NIST-standardized primitive in the framework, along with representative published performance figures used.

During the transition period, dual-signature transactions, where a ledger entry is protected by both a classical and a post-quantum signature at the same time, offer backward compatibility with existing verifier tooling, while guaranteeing that any archived cipher text or credential remains protected against future quantum cryptanalysis of the classical component alone, an approach consistent with proposed hybrid migration patterns for blockchains generally [11, 14, 17].

Explainable AI Access-Monitoring Module: The access-event log (requester role, timestamp, requested predicate, proof verification result, patient consent-policy version, and contextual metadata such as facility and shift schedule) is fed into the XAI layer, which scores each event via an anomaly-detection model. For each flagged event, we report a binary or scalar risk score and a SHAP-based local attribution that indicates the features that contributed to the score (e.g., off-hours access and provider-patient mismatch), in agreement with the feature-importance patterns seen in health record access-auditing studies, where provider mismatch, off-hours timing, and access frequency were consistently the dominant SHAP contributors [18, 19]. We present this attribution to the compliance officer who is investigating the alert as a short natural language explanation, not as a raw feature vector, as recommended for XAI in clinical and security settings [18, 20].

4: DISCUSSION

Cryptographic protocol and smart contract design - Threat Model and Adversarial Modeling

To formalize the security boundaries of AQR-SSI, we define four distinct adversarial classes:

- External Network Attacker (T1): Who is capable of observation, interception, and injection ledger and API traffic, but has inadequate control over any consortium consensus nodes.
- Malicious or Coerced Verifier (T2): This Verifier aims to link patient proofs across different clinical sessions or over-collect data beyond the strictly authorized predicate.
- Compromised Issuer (T3): A compromised issuer whose signing keys have been exposed, necessitating immediate, verifiable revocation without invalidating unaffected credentials.
- Impending Considerable Attack (T4): A future attacker who possesses a cryptographically relevant quantum computer and seeks to retrospectively decrypt archived ledger and TLS traffic.

We explicitly do not consider endpoint compromise of the patient's own device (wallet key theft via malware) within the scope, treating this as a separate, complementary hardware-security concern (e.g., secure enclaves, biometric wallet unlock) rather than a protocol-level defence.

Cryptographic Protocols: Our selective disclosure protocols are built as transparent, trusted-setup-free zk-STARK circuits over a

Merkle-tree-based credential commitment scheme. To prevent cross-session profiling by verifiers (T2), the protocol incorporates a session-scoped nullifier generated using a pseudorandom function keyed on the patient's credential secret and the verifier's session challenge. The verifier's smart contract maintains a nullifier registry to reject replays, ensuring that two proofs generated by the same patient for different verifiers (or different sessions) remain cryptographically unlinkable.

To support an orderly migration to quantum-safe standards, the protocol utilizes hybrid dual-signature transactions. During the transition, transactions carry a classical signature such as Elliptic Curve Digital Signature Algorithm (ECDSA) and a post-quantum signature such as Module-Lattice-Based Digital Signature Algorithm (ML-DSA), ensuring backward compatibility with legacy verifier systems while guaranteeing long-term security against quantum cryptanalysis [21].

Signatures are issued based on the role assignment of Table 2. An algorithm identifier is stored in the credential schema per credential, so that the proof verification circuit of a verifier chooses the corresponding verification routine. In case of an issuer compromise (T3), the key rotation procedure is the standard DID key rotation semantics: a new key entry is added to the DID-document with an activation ledger-height and the validity-interval of the compromised key is marked closed. Credentials signed before the compromise are still valid if their signing key’s validity-interval contained their issuance time.

Smart Contract Construction: The ledger hosts two critical stateful smart contracts:

1. Consent Registry Contract: Maintains a secure hash commitment of each patient's DID to their current active access-policy document (e.g., specifying which providers can access which records). Updates require a valid patient signature and trigger events consumed by the XAI monitoring layer.
2. Revocation Registry Contract: Maintains an append-only cryptographic accumulator (or sparse Merkle tree) containing revoked credential identifiers. ZK proof circuits require a non-membership witness against this accumulator to verify validity.

Neither contract ever stores clinical data, both stores only commitments and status flags, restricting on-chain storage to fixed-size commitments, which keeps the cryptographic overhead of post-quantum signatures firmly confined.

Explainable AI Module Flow & Technology Stack: The XAI pipeline processes access-event logs containing features such as requester relationship distance, time-of-day/day-of-week deviation, request frequency in a rolling window, cross-facility flags, and predicate sensitivity [22, 23] These structured features are fed into an anomaly detection model (e.g., a gradient-boosted tree or isolation forest). When an access event scores above a defined risk threshold, the module computes local SHAP or LIME attributions to identify the dominant anomaly contributors. These mathematical weights are translated into a natural-language summary (e.g., "Flagged primarily due to access outside the requester's typical facility and shift window") displayed on the compliance dashboard. A feedback loop records the compliance officer's disposition to periodically retrain the model while ensuring explanation stability.

1. Feature engineering: This builds structured features from the access-event log distance of requester-patient relationship, requester's time-of-day/day-of-week deviation from historical pattern, request frequency in rolling window, cross-facility flag, and predicate sensitivity tier.
2. Model training: The Trains a gradient-boosted tree ensemble (or, in the unsupervised setting, an isolation forest / auto encoder) on historical access logs, as is typical for the model families most often used with SHAP in EHR-focused XAI research [17].
3. Explanation generation: For each event scoring above the alert threshold, calculate local SHAP values and choose the top contributing features.
4. Justification rendering: translate top SHAP features to a short natural-language summary (e.g., “Flagged primarily due to access outside the requester's typical facility

and shift window") for the compliance dashboard.

5. Feedback loop: Compliance-officer dispositions (anomaly confirmed vs. false positive) are logged and used to retrain the model periodically, tracking explanation stability across retraining cycles so that justifications don't drift unpredictably.

Implementation Technology Stack: The reference implementation would consist of a W3C DID Core compliant DID method, a permissioned ledger (e.g., Hyper-ledger Fabric or an Ethereum compatible L2 such as Starknet, which has been used in prior work for STARK based patient authentication [4]) for the consent and revocation contracts, an off-chain encrypted object store or IPFS-compatible network for EHR payloads, the Open Quantum Safe project's liboqs library (or equivalent) for ML-DSA/FNDSA/SLH-DSA/ML-KEM primitives, and a Python-based XAI pipeline using SHAP and LIME libraries integrated with the access-event log via a message queue. This stack is proposed as an implementation target; results from a deployed instance of it are not reported in this paper.

Security and Privacy Analysis: This section provides a demanding safety and privacy assessment of the proposed AQR-SSI framework. We formalize our specific security objectives, deliver a thorough threat-by-threat mitigation evaluation directly mapped to our defined adversarial profiles, and broadly analyze systemic regulatory alignment with international healthcare privacy standards, specifically focusing on HIPAA and GDPR mandates.

Security Objectives and Formal Definitions: To guarantee medical assessment and ensure hardened protection of confidential patient credentials under diverse network operational environments, the AQR-SSI framework is cryptographically and scientifically designed to satisfy six fundamental security and privacy objectives:

1. Confidentiality: Electronic Health Record (EHR) payloads are end-to-end encrypted and never exposed in plaintext on the immutable ledger. Decryption keys (managed off-chain via secure proxy re-encryption or attribute-based encryption mechanisms) are accessible solely to verifiers who present a cryptographically verified, active proof of patient authorization.
2. Integrity and Authenticity: All Decentralized Identifier (DID) documents, verifiable credentials (VCs), state registries, and patient consent policies are digitally signed using a hybrid classical/post-quantum signature scheme. This design guarantees total non-repudiation and cryptographic unforgeability against both current classical network adversaries and future quantum cryptanalysis during the transitional migration window.
3. Unlinkability: External relying parties, network observers, or colluding verifiers are accurately prevented from linking separate credential presentations, transaction traces, or clinical access sessions to the same patient. Unlinkability is cryptographically enforced via session-specific nullifiers computed within the zk-STARK proof generation loop.
4. Forward and Long-Term (Quantum) Security: Optimal assurance identity trust anchors, including issuer root public keys, certificate-authority certificates, and revocation accumulators, are secured using maximally conservative hash-based (SLH-DSA) and lattice-based (ML-DSA) post-quantum primitives. This proactive approach removes the systemic 'harvest-now, decrypt-later' (T4) exploit window by ensuring long-term ledger resistance.
5. Surveillance-Free Accountability: The Explainable AI (XAI) monitoring pipeline audits, logs, and flag access patterns based strictly on non-sensitive structural metadata (e.g., temporal offsets, spatial deviations, or relationship distances). This ensures real-time operational compliance auditing and automated threat hunting without decrypting any sensitive clinical payloads or violating patient anonymity.
6. Revocability: If an issuer private key is compromised, or if a user device hosting a digital wallet is physically lost or damaged, the affected key records can be independently revoked, rotated, or blacklisted within a single ledger block-confirmation interval. This lifecycle phase is executed via gas-efficient smart-contract state updates on the Revocation

Registry, avoiding the wholesale invalidation of unaffected credentials.

Threat-by-Threat Evaluation: In order to evaluate the adversarial resilience of the AQR-SSI framework, we map our layered defense architecture against the four major threat profiles (T1 through T4) formalized in Chapter 3. This ensures that every vector is systematically addressed by a corresponding cryptographic or analytical mitigation.

Table 3: Formal threat-by-threat mitigation matrix under the AQR-SSI model.

Threat / Adversary Profile	Attack Vector & Mechanism	Primary Mitigation Strategy	Residual Risk Handling
Network Interception & Tampering (T1) - External Attacker	Man-in-the-middle (MitM) attacks, transaction modification, or active eavesdropping on TLS or public ledger endpoint traffic.	Deployment of hybrid TLS channels utilizing ML-KEM for quantum-safe key exchange. Every ledger state transaction is secured using dual signatures (classical ECDSA paired with post-quantum ML-DSA).	Eavesdroppers might analyze structural metadata (timing patterns or payload volumes). Mitigated via traffic-padding and random routing (out of protocol scope).
Cross-Session Profiling (T2) - Malicious/Curious Verifier	Correlating and linking unique patient transactions or credential presentations across different clinical facilities to profile user identity.	Verification of credential attributes is conducted off-chain via zk-STARKs. Presentation proofs incorporate a fresh, session-specific nullifier (Nullifier = H(Secret || Challenge)), ensuring complete mathematical unlinkability.	If a verifier possesses exclusive access to an extremely rare clinical or diagnostic predicate, the patient's identity could be inferred by exclusion. Mitigated by generalizing predicate categories.
Issuer Key Compromise (T3) - Compromised Issuer Authority	An adversary compromises the private signing key of an authorized medical facility, attempting to forge credentials.	The Revocation Registry smart contract maintains an active, append-only accumulator. Proof circuits verify non-revocation against this state. Key rotation limits key validity to targeted ledger block-heights.	Credentials forged and presented during the short exploit window (prior to detection and revocation) must be flagged and addressed via manual back-end

Threat / Adversary Profile	Attack Vector & Mechanism	Primary Mitigation Strategy	Residual Risk Handling
			clinical audit procedures.
Harvest-Now, Decrypt-Later (T4) - Impending Considerable Attack	Recording current encrypted network traffic and ledger transactions with the intent of decrypting them once cryptographically relevant quantum computers (CRQCs) become accessible.	PQC signature schemes (ML-DSA, Falcon) and lattice-based KEMs (ML-KEM) are applied to all newly issued credentials and session negotiations from system inception.	Legacy health records encrypted using legacy pre-quantum standards (e.g., standard RSA or ECC) before framework migration remain exposed. These must undergo immediate bulk re-encryption.

Regulatory Compliance Analysis: Deploying decentralized identity architectures in clinical and research healthcare networks necessitates strict compliance alignment with dominant regional legal mandates such as in the United states and the European Union. AQR-SSI reconciles blockchain's immutability with privacy laws by design.

HIPAA Security & Privacy Rule Alignment (United States)

The Health Insurance Portability and Accountability Act (HIPAA) [24] impose strict security requirements regarding the handling of Protected Health Information (PHI). AQR-SSI satisfies these standards through several architectural controls:

1. The 'Minimum-Necessary' Standard: By validating claims using zk-STARK proofs, medical verifiers only learn specific, abstract clinical predicates (e.g., verifying active health insurance eligibility or proving a patient's age threshold) without gaining exposure to unnecessary personal demographic or clinical diagnostic values.
2. Administrative and Security Safeguards (§ 164.312): All active access permissions are mapped to the patient's on-chain Consent Registry contract. This promises that healthcare providers cannot retrieve clinical payloads without cryptographically validating the patient's current active consent policy. Concurrently, the XAI layer translates complex machine-learning anomaly detections into clear, clinician-readable explanations, providing a detailed, non-repudiable audit trail.
3. Systemic De-risking of Centralized Repositories: Because actual clinical EHR records are hosted exclusively in encrypted, off-chain storage systems (such as secure IPFS networks or private object stores) while keeping only public keys, consent-policy hashes, and credential revocation status on-chain, centralized repository risk is entirely streamlined.

GDPR Compliance Mapping (European Union)

The General Data Protection Regulation (GDPR) [25] establishes strong privacy rights for individuals inside the European Union. The audit trail generated by XAI satisfies the accounting of disclosures expectations with

justifications attached rather than opaque log lines. Under GDPR, DID- and VC-based patient custody is designed to support data subject rights such as access and, via non-issuance or revocation of credentials in the future, a practical approximation of erasure for future disclosures – though the immutability of ledger-anchored commitments means that erasure of on-chain state itself must be carefully scoped (e.g., anchoring only hashes/commitments, never personal data, so that off-chain deletion is sufficient for GDPR purposes).

This paper describes these properties at the design level; a full compliance determination would require legal review specific to the jurisdiction, which is outside the scope of a technical framework paper. AQR-SSI maps directly onto these regulatory mandates through clear technological boundaries:

1. The Right to Erasure ('Right to be Forgotten' - Article 17): Due to the permanent, tamper-resistant nature of blockchain, storing personal health information directly on-chain violates GDPR's mandate to erase data upon request. AQR-SSI circumvents this issue by storing strictly zero personal data or direct identifiers on the ledger. When a patient requests the deletion of a specific health record, the off-chain encrypted medical payload is permanently erased from the storage layer. Consequently, the isolated on-chain cryptographic commitment hashes become completely un-linkable, precisely satisfying the legal standard for irreversible pseudonymization and erasure.
2. Data Minimization and Purpose Limitation (Article 5): Because the patient maintains direct cryptographic custody of their credentials inside a hardware-secured or mobile wallet, they dictate exactly when, with whom, and under what constraints their attributes are shared. This hands-on control ensures that no bulk data scraping or unauthorized lateral processing can occur without active, signed user consent.
3. Explainable Auditing and Human-in-the-Loop Governance (Article 22): Under Article 22 of the GDPR, data subjects have the right not to be subjected to fully automated decision-making that significantly affects them. The AQR-SSI framework enforces a human-in-the-loop governance structure: the XAI anomaly-detection module does not unilaterally lock or black-list provider credentials. Instead, it serves as an interpretive advisory system, supplying compliance officers with clear, SHAP-derived explanations that illustrate *why* a particular clinical access request deviated from standard baseline parameters, enabling rapid and informed human-led investigations.

Cryptographic Suite Role Assignment: To achieve optimal balance between signature size, computational latency, and long-term security, AQR-SSI rejects a single-algorithm design in favor of a hybrid, role-specific signature strategy, as detailed in Table 2, Section 2.6.

Cryptographic and Storage Overhead Estimates: An illustrative, literature-grounded analysis indicates that incorporating a dual ML-DSA-plus-ECDSA signature on a single AQR-SSI credential presentation adds approximately 2.4 kilobytes of signature material over a classical ECDSA baseline (2,420 bytes for ML-DSA vs. 64–73 bytes for ECDSA). In terms of computation, it adds roughly 0.65 ms of signing latency on the credential holder's mobile device. SLH-DSA signing latency (~131.9 ms) and large signature size (~41 KB) are restricted strictly to long-lived issuer roots and revocation registries, ensuring these expensive parameters are heavily amortized over long validity intervals rather than incurred during individual, real-time patient interactions.

Scalability Challenges and Explainability Fitness: The dominant scalability risk is on-chain ZK-proof verification cost at consortium scale, as verification is computationally heavier than signing for STARK-family proofs. To mitigate this, AQR-SSI advocates verification on a Layer-2 rollup (such as Starknet) with periodic state checkpointing to a permissioned L1. In terms of explainability, the framework proposes measuring faithfulness, stability, and comprehensibility metrics against white-box standards (e.g., Deep-SHAP) to ensure explanations remain stable and highly

interpretable for compliance reviewers.

Proposed Future Experimental Validation Setup: To transition the AQR-SSI framework from a theoretical design to a clinical-grade reality, we outline a detailed future experimental setup. This plan details the exact empirical validation steps required for a production-ready pilot:

1. Reference Stack Deployment: Deploy the proposed implementation stack on a permissioned network of at least four consortium nodes, utilizing Hyper-ledger Fabric or Starknet for contract execution.
2. Workload Replay: Replay a de-identified, high-fidelity access-log workload derived from public EHR datasets to evaluate system throughput and latency under clinical operational load.
3. Edge Device Measurement: Measure end-to-end proof-generation latency directly on mid-range and low-end patient smartphone hardware to determine mobile wallet usability.
4. L2 Rollup Throughput: Benchmark proof verification costs, gas usage, and maximum transaction throughput (proofs verified per second) on Layer-2 runtimes.
5. User Usability Study: Conduct a double-blind usability study of the XAI natural-language dashboard with active clinicians and compliance officers to measure explanation clarity and decision-making accuracy.

Contribution to Knowledge

1. Brings together four separate design literatures (SSI, ZKP, PQC, XAI) into one aligned trust and data-flow model rather than one-off add-ons.
2. Cryptographic agility is baked into the credential schema from day one so future migration is non-disruptive.
3. Storing the payload off-chain reduces the on-chain footprint and thus the PQC signature overhead, regardless of the amount of clinical data.
4. Explanations associated with each access-anomaly flag help satisfy clinical governance and regulatory audit requirements beyond a simple allow/deny decision.

6. CONCLUSIONS AND RECOMMENDATION

Limitations

1. The paper presents architecture and protocol design with a literature grounded overhead analysis, but does not report results from a working implementation, and real-world proof-generation latency on low-end patient devices, ledger throughput under adversarial load, and XAI explanation quality on real access logs all require empirical study.

2. Wallet key management on patient-held devices remains a practical usability and recovery challenge that this framework does not solve; social or institutional key-recovery mechanisms are needed and were explicitly out of scope. •

3. Even with amortization, the large signature size of SLH-DSA may still be prohibitive for constrained IoT-class medical devices acting as issuers or verifiers, motivating further work on signature aggregation for healthcare identity workloads. Regulatory analysis here is at the design-intent level only and is not a substitute for jurisdiction-specific legal review.

Conclusions

The AQR-SSI framework successfully bridges a critical integration gap in digital health systems. By unifying blockchain-based self-sovereign identity, transparent zero-knowledge proofs, hybrid post-quantum cryptography, and explainable artificial intelligence under a single clinical trust model, the framework provides a highly protected, future-resistant approach for electronic health records sharing. The analytical performance estimates confirm that while post-quantum signatures introduce non-trivial storage requirements, they remain computationally feasible for contemporary clinical ecosystems, provided Layer-2 rollups and cryptographic agility are leveraged.

Recommendations

Based on the findings of this study, the following recommendations are proposed:

1. Prototype Development and Validation: Future research should implement the proposed AQR-SSI framework as a functional prototype within a real or simulated healthcare environment. Such implementation would enable empirical evaluation of its security, scalability, usability, and interoperability under realistic operational conditions.
2. Clinical Pilot Studies: This study recommends that future research should look into the feasibility of pilot implementations of the proposed framework within healthcare institutions to evaluate how to use, interpret and improve operability.
3. Comprehensive Performance Evaluation: Subsequent studies should conduct experimental performance evaluations using benchmark datasets and representative healthcare workloads to measure authentication latency, credential verification time, blockchain transaction throughput, computational overhead, and resource utilization.
4. Integration with Existing Healthcare Standards: Future implementations should investigate interoperability with established healthcare information standards, including HL7 FHIR, DICOM, and other nationally adopted health information exchange frameworks, to facilitate seamless integration into existing healthcare infrastructures.
5. Enhanced Explainable AI Models: Additional research should evaluate alternative Explainable Artificial Intelligence techniques beyond SHAP and LIME to determine their effectiveness in improving the transparency, interpretability, and trustworthiness of anomaly detection for healthcare access monitoring.
6. Regulatory and Policy Evaluation: Future studies should examine the regulatory implications of adopting decentralized identity systems employing post-quantum cryptography and assess their alignment with applicable healthcare data protection regulations, including HIPAA and GDPR.
7. Broader Security Assessment: Future investigations should evaluate the resilience of the proposed framework against emerging cyber threats, including sophisticated insider attacks, adversarial machine learning techniques, and evolving quantum computing capabilities, through comprehensive penetration testing and formal security verification.

References

1. Pokharel, A. & Kathayat S. (2026). Blockchain-Enabled Self-Sovereign Identity Applications in Health Care: Scoping Review. J Med Internet Res Vol. 28:89574. DOI: 10.2196/89574
2. Gupta, S. (2025). Zero-Knowledge Proofs for Privacy-Preserving Systems: A Survey across Blockchain, Identity, And Beyond. Engineering and Technology Journal, Vol. 10(7), 5755-5761. DOI:10.47191/etj/v10i07.23
3. Tertulino, R., Almeida, R. & Alencar, L. (2026). Trustworthy blockchain-based federated learning for electronic health records: Protecting participant identity using decentralised identifiers and verifiable credentials. Cryptography and Security. https://doi.org/10.48550/arXiv.2602.02629
4. Myeong , G. E. & Ram, K. S. (2025). Blockchain Based Zero Knowledge Proof Protocol For Privacy Preserving Healthcare Data Sharing. Journal of Technology Informatics and Engineering (JTIE), Vol. 4(1), 171 – 189. DOI: 10.51903/jtie.v4i1.296
5. Dieye, M., Valiorgue, P., Gelas, J., Diallo, E., Ghodous, P., Biennier, F. & Peyrol, E. (2023). A Self-Sovereign Identity based on Zero-Knowledge Proof and blockchain. IEEE Access, Vol. 4(11), 49445-49455. DOI:10.1109/ACCESS.2023.3268768
6. Ezz, M., Alaerjan, A. S. & Mostafa, A. M. (2025). Ethical AI in Healthcare: Integrating Zero-Knowledge Proofs and Smart Contracts for Transparent Data Governance. Bioengineering, 12(11), 1236-1266. https://doi.org/10.3390/bioengineering12111236
7. Adeoye, S. (2025). Blockchain-Enabled, Post-Quantum Cryptographic Framework for Securing Electronic Health Records: A Next-Generation Approach to Healthcare Data

Protection. Cognizance Journal of Multidisciplinary Studies, Vol. 5(4), 77-104. DOI: 10.47760/cognizance.2025.v05i04.006

8. National Institute of Standards and Technology. (2024). Module-lattice-based digital signature standard (Federal Information Processing Standards Publication [FIPS] 204). U.S. Department of Commerce. https://doi.org/10.6028/NIST.FIPS.204

9. López Martínez, A., Gómez, J. A., & Martínez, S. (2025). Empower healthcare through a self-sovereign identity infrastructure for secure electronic health data access. Cryptography and Security. https://doi.org/10.48550/arXiv.2501.12229

10. Goswami, P. & Sirsikar, S. (2024). Self-Sovereign Identity to Secure Digital Identity using Blockchain Technology. 5th International Conference on Image Processing and Capsule Networks (ICIPCN), Dhulikhel, Nepal, 826-831. DOI:10.1109/ICIPCN63822.2024.00143.

11. Madine, M., Salah, K., Jayaraman, R. & Yaqoob, I. (2025). Zero-knowledge proofs for anonymous authentication of patients on public and private blockchains. Array, 28. 100590. DOI:10.1016/j.array.2025.100590.

12. Hoang, C. T. A. (2026). Adoption and Effectiveness of AI-Based Anomaly Detection for Cross Provider Health Data Exchange. Research paper on AI-based anomaly detection in healthcare audit logs using simulation and scoping review. Computers and Society. https://doi.org/10.48550/arXiv.2604.09630

13. Garcia, R. D., Ramachandran, G., Dunnett, K., Jurdak, R., Ranieri, C., Krishnamachari, B. & Ueyama, J. (2024). Survey of Blockchain-Based Privacy Applications: An Analysis of Consent Management and Self-Sovereign Identity Approaches. Cryptography and Security. https://doi.org/10.48550/arXiv.2411.16404

14. Gummadi, A. N, Arreche, O. & Abdallah, M. (2025). A systematic evaluation of white-box explainable AI methods for anomaly detection in IoT systems. Internet of Things, Vol. 30, 101505. https://doi.org/10.1016/j.iot.2025.101505

15. Sourav, R. I. & Ali, R. (2025). Post-quantum secure health records: a blockchain-based lattice threshold signcryption scheme. Cluster Computing, 28, 450 https://doi.org/10.1007/s10586-025-05117-2

16. Hasib, S., Rasool, A., Gyanchandani, M. Haripriya, R. & Kumar, L. (2026). A structured review of lattice based attribute based encryption methods for post quantum security. Discover Computing, 29, 175. https://doi.org/10.1007/s10791-026-09965-3

17. Liu, Z., Chen, X. & Xie, Y. (2025). A lattice-based group signature with backward unlinkability for medical blockchain systems. Journal of Information Security and Applications, Vol. 94, 104226. https://doi.org/10.1016/j.jisa.2025.104226

18. Emergent Mind (2025). Post-Quantum Blockchain Cryptography. Emergent Mind Topic Review, accessed July 2026. https://www.emergentmind.com/topics/post-quantum-blockchain-cryptography

19. Chhetri, G., Somvanshi, S., Hebli, P., Brotee, S. & Das, S. (2026). Post-Quantum Cryptography and Quantum-Safe Security: A Comprehensive Survey V2. Cryptography and Security. arXiv:2510.10436. https://doi.org/10.48550/arXiv.2510.10436

20. Schemitt, A.G., Fan da Silva, H., Lunardi, R. C., Kreutz, D., Mansilha, R. B. & Zorzo, A. F. (2025). Assessing the Impact of Post-Quantum Digital Signature Algorithms on Blockchains. IEEE 24th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom 2025), Cryptography and Security. https://doi.org/10.1109/Trustcom66490.2025.00276

21. Yang, Z., Alfauri, H., Farkiani, B., Jain, R., Di Pietro, R. & Erbad, A. (2024). A Survey and Comparison of Post-quantum and Quantum Blockchains. Cryptography and Security. https://doi.org/10.48550/arXiv.2409.01358

22. Agbo, E. O., Azim, S. M., & Dehzangi, I. (2026). Explainable AI Applications in

Healthcare: A Systematic Review. Algorithms, 19(6), 488. https://doi.org/10.3390/a19060488

23. Caterson, J., Lewin, A. & Williamson, E. (2024) .The application of explainable artificial intelligence (XAI) in electronic health record research: A scoping review. Digit Health. Vol. 10: 1–12. DOI:10.1177/20552076241272657.

24. HIPAA. (2026). HIPAA Security Rule Update: New Requirements Every Healthcare Organization Must Prepare For. Health Insurance Portability and Accountability Act of 1996, 45 C.F.R. Pts. 160 & 164. https://medcurity.com/hipaa-security-rule-2026-update/

25. General Data Protection Regulation [GDPR]. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119/1 (2016). europa.eu