Establishing Trust in Motion: A Saudi Regulatory Sandbox for Data Provenance in AI-Enabled Healthcare

Executive Summary

In an age when health systems increasingly rely on data to drive diagnostics, research, and policy making, the trustworthiness of that data is more than a technical issue—it's a national priority. For Saudi Arabia, whose Vision 2030 charts a course toward a fully digitized, AI enabled healthcare ecosystem, the question of how to ensure verified, privacy-compliant, and policy-aligned data exchange is both urgent and complex.

This case study presents a pioneering collaboration between Alfaisal University, SDM Diagnostics, and LabTrace, which piloted the Data & Trust Alliance (D&TA) data provenance standards in a real-world, regulatory sandbox.

Together, the partners designed and tested a blockchain-powered prototype that enabled machine-verifiable provenance of healthcare datasets while maintaining full compliance with Saudi data localization laws and the Personal Data Protection Law (PDPL).

The result was not simply a successful technical test—but a transformational demonstration of how Saudi Arabia can lead in building scalable, ethical, and globally interoperable data infrastructure for healthcare. It is the first localized use of international provenance standards to certify metadata across institutional boundaries, setting a precedent for trusted cross-border data exchange in alignment with national policy.

Problem Definition:
Why Trust in Healthcare Data Must Be Proven, Not Presumed

In traditional health systems, data exchange between trusted institutions often depends on personal relationships and signed agreements. However, the introduction of AI into healthcare workflows has shifted the stakes entirely. AI models, particularly those used for diagnostic or predictive tasks, are only as good as the data they are trained on. Without clear metadata—information on where the data originated, how it was collected, and what legal basis governs its use—such models can introduce systemic bias, legal vulnerabilities, and ethical blind spots.

Moreover, Saudi Arabia's regulatory landscape, shaped by PDPL and guided by SDAIA, places strict limitations on how sensitive data can move and be used. These requirements—designed to safeguard privacy, uphold sovereignty, and control risk—have unintentionally created a bottleneck for innovation. Cross-institutional data flows are slow, opaque, and dependent on manual documentation that cannot scale to the demands of AI or global research collaboration.

Thus, the central challenge emerged: How can healthcare data be made verifiably trustworthy in a way that meets national compliance while enabling the real-time, large-scale usage AI systems require? And how can this be done not only between unknown parties, but even between known, trusted institutions operating under a shared national mandate?

Context and Background:
Saudi Arabia at the Crossroads of Sovereignty and Scalability

Saudi Arabia's Vision 2030 reform agenda positions digital health not only as a modernization goal but as a pillar of national capability. The Kingdom has made significant strides in developing national platforms such as NPHIES, fostering AI talent through SDAIA, and investing in future-ready infrastructure. Yet, the global shift toward AI-enabled care has revealed a gap: while the infrastructure to collect and analyze health data exists, the governance models to trust that data at scale are still evolving.

In parallel, global initiatives like the Data & Trust Alliance have developed frameworks to standardize data provenance for ethical AI development. However, these standards are often created with Western legal frameworks in mind (e.g., GDPR), and may not align naturally with the data localization and consent regimes specific to Saudi Arabia.

In this context, Alfaisal University recognized the need to bridge these two realities. As an academic institution with a strategic mandate in health innovation, Alfaisal was uniquely positioned to act as a neutral testing ground—a regulatory sandbox—for adapting and operationalizing international data governance standards within Saudi Arabia's legal, cultural, and technological context.

Designing the Sandbox:
A Controlled Testbed for Trust

The sandbox initiative was conceptualized as a real-world experiment, not a simulation. The participating institutions—Alfaisal University and SDM Diagnostics—already had an established working relationship. This made the experiment even more relevant, as it tested whether verifiable trust could be introduced even in environments where implicit trust existed, thereby proving scalability beyond known entities.

SDM, as the data provider, submitted live healthcare datasets used in AI-based diagnostics for chronic and ocular diseases. These datasets were accompanied by structured metadata following the D&TA's provenance standards, which include attributes such as data origin, collection method, consent framework, and legal use conditions.

To transform these metadata fields into immutable, machine-verifiable certificates, the project employed LabTrace, a blockchain-based verification layer built on the Algorand platform. Rather than storing sensitive data on the blockchain, the system generated cryptographic hash identifiers (CIDs) for each dataset's metadata and notarized them on-chain. Actual patient data remained off-chain and fully localized, stored through IPFS within Saudi jurisdiction. The process enabled real-time, privacy-preserving, independently verifiable trust in data origin and usage conditions.

Methodology and Implementation:
Technology Anchored in Policy

Unlike many blockchain experiments in healthcare, the technology in this sandbox was not layered onto existing systems as an abstraction. It was integrated directly into the data flow between two institutions, operating within live regulatory constraints.
The methodology included:

- Translating D&TA provenance fields into programmable metadata schema.
- Attaching provenance metadata to datasets at the point of exchange.
- Hashing and notarizing metadata on Algorand, chosen for its energy efficiency and suitability for permissioned environments.
- Managing access via smart contracts and Role-Based Access Control (RBAC).
- Storing hashed records on IPFS while keeping raw data within Saudi servers.

This design ensured full compliance with PDPL, including principles such as data minimization, purpose limitation, and consent verification. It also aligned with SDAIA's localization rules, ensuring that all personal data remained physically within the Kingdom—even while its trust markers became auditable from anywhere.
The result was an end-to-end, policy-aligned system that proved data provenance could be both technically rigorous and operationally simple.

Outcomes and Key Insights:
Trust, Proven at Scale

The sandbox delivered measurable improvements in data verification and cross-institutional trust:

01. Speed: Verification of dataset provenance and consent reduced from days to minutes.
02. Transparency: Immutable audit trails eliminated ambiguity in cross-team collaboration.
03. Compliance Confidence: Institutions could demonstrate adherence to PDPL and SDAIA rules without manual reporting.
04. Scalability: The model proved feasible even in known relationships, supporting future adoption across unknown or global partners.

From SDM's perspective, the process did not disrupt existing workflows—it enhanced them. For Alfaisal, the prototype provided not only data for research but also evidence of trust that could be reported to regulators or passed on to future collaborators. For both institutions, the project served as an operational and strategic win.
Equally important were the non-technical insights: education and change management were essential. Familiarizing stakeholders with the concept of verifiable provenance—not as a security measure but as an enabler of innovation—was key to adoption.

Strategic Implications:
Saudi Arabia as a Model for Ethical, Sovereign Data Governance

This sandbox, while rooted in Saudi regulatory realities, presents lessons with global relevance.
First, it demonstrates that data localization and global interoperability are not mutually exclusive. By separating sensitive data from its trust metadata, and by embedding compliance directly into the data layer, nations can assert sovereignty while participating in global health ecosystems.
Second, the case proves that blockchain in healthcare is viable—not as a buzzword, but as infrastructure for regulatory assurance, especially when paired with standards like those from D&TA.
Third, the initiative offers a blueprint for other sectors. The same model could be adapted for use in pharmaceuticals, insurance, education, or energy—anywhere data flows need to be verified, governed, and scaled responsibly.
And finally, it positions Saudi Arabia not just as an adopter of global health tech trends, but as a shaper of global standards. The Kingdom is building models that other countries with strong privacy and localization requirements may soon follow.

Conclusion:
From Prototype to Policy Infrastructure

This case study illustrates a turning point in Saudi Arabia's digital health transformation. Alfaisal University's leadership in initiating a regulatory sandbox for verifiable data trust demonstrates how academia can operate as a policy lab, testing not only technologies but the future of governance.
By implementing international standards within a Saudi framework, and by proving that trust can be engineered, verified, and scaled, this initiative lays the groundwork for a national health data ecosystem where data is not only secure, but certifiably ethical and AI-ready.
Saudi Arabia is not merely adopting global health innovation—it is now authoring its own chapter.

APPENDIX:
TECHNOLOGY IN A NUTSHELL

What is a Blockchain?
A blockchain is a distributed public ledger that is immutable and held by a very large network of users (nodes). Nodes in the network continuously verify the truthfulness of the ledger and grow the same ledger through a consensus protocol that adds novel records (blocks) that are securely linked together via cryptographic hashes. While blockchain is most commonly associated with bitcoin, the digital currency, this kind of record does not have to utilise or be related to currencies or payment. Instead, the blockchain functions as a secure record of history.

Labtrace: Blockchain Technology
Labtrace has recently developed and demonstrated the use of a block-chain based platform to ensure data and data-processing integrity in a research clinical trial. Once a file is uploaded on the platform, a file identifier (content-hash) is created that is uniquely linked to the file content. The hash is then uploaded on the blockchain together with any relevant information. The user receives a certificate containing the hash, the ancillary information and the link to the block in the ledger. While the certificate remains public, the user has the option of either publishing the file or keeping the file private within its own firewalls (the whole process is GDPR compliant); in the latter case, proof of true certification can be provided at any time.

Labtrace: The Process
The flexibility of the Labtrace platform allows the recording of a unique data identifier for any type of file at a certain time (timestamp) and share proof of its veracity (certification). The platform also allows the creation of a chain of evidence with secondary data, that is those data that are obtained from raw data (for example images) through some software. The product(s) of such processing, that we call secondary files, can then be linked via the blockchain to the primary data and to the software used. These chains of evidence can be easily inspected by a reviewer