



State-Scale Digital Archiving with IPFS and Hyperledger: A Sovereign, Secure Infrastructure for Public Records.

White Paper - MindStack Research Division (Nov. 2025)

Prepared for government leaders, regulators, and technical architects responsible for modernising national digital archiving systems.

Executive Summary

Governments must keep critical digital records - laws, court decisions, tax data, civil status, health information, and more - safe and usable over very long periods. Centralised archiving platforms, often based on proprietary products, create single points of failure, are hard to audit independently, and can weaken digital sovereignty. At the same time, regulations on data protection, electronic identification, and trusted services demand strong guarantees of integrity, confidentiality, traceability, and long-term access.

This paper proposes a state-scale digital archiving architecture built on three open and complementary technologies: IPFS for distributed, content-addressed storage of encrypted documents; Hyperledger Fabric for a permissioned ledger that manages metadata, access policies, and audit trails; and Hyperledger Indy for self-sovereign identity and verifiable credentials for citizens, public servants, and institutions. Together, they provide cryptographic proof of integrity and origin, fine-grained and verifiable access control, and resilient storage that reduces dependence on any single vendor or infrastructure.

The document explains the context and requirements of public digital archiving, introduces the three technologies in accessible terms, and describes a reference architecture with key data models and end-to-end workflows. It examines security, privacy, and regulatory compliance, proposes a governance model adapted to a governmental consortium, and outlines priority use cases and a pragmatic roadmap for progressive deployment at the scale of a State.

1. Introduction

States are now producing almost all their important information in digital form: laws and court decisions, tax records, civil status documents, social protection and health data, education records, and more. These are not simple files; they are part of the legal memory of the State and the rights of citizens. They must stay readable, authentic, and trustworthy for very long periods, while technologies, vendors, and cyber-threats change rapidly. Relying only on centralised, proprietary archiving platforms creates single points of failure, vendor lock-in, and makes it harder to prove independently that records have not been altered.



At the same time, regulations on data protection, electronic identification, electronic signatures, and public archives require strong guarantees: integrity, confidentiality, controlled access, long-term preservation, and clear audit trails. Public administrations also need better interoperability between ministries and agencies, and more control over where and how sensitive data is stored. Digital archiving is therefore no longer just a technical question; it is a matter of sovereignty, trust, and long-term resilience for the State.

This paper explores how a combination of three open technologies - IPFS, Hyperledger Fabric, and Hyperledger Indy - can support a new generation of digital archiving systems at state scale. The key idea is to separate and secure three core functions: storage of encrypted documents (IPFS), registration and governance of metadata and policies (Fabric), and management of identities and authorisations using self-sovereign identity (Indy).

Our objective is to propose a reference architecture that is technically realistic, legally aware, and understandable for non-specialists. The paper briefly sets the context and requirements for state digital archives, introduces the three technologies in accessible terms, then describes how they can be combined into a coherent system with clear data flows. It also highlights security and privacy implications, governance options for a governmental consortium, key use cases, and a pragmatic roadmap for progressive deployment.

2. Technology overview: IPFS & Hyperledger (Fabric + Indy)

2.1 IPFS, Hyperledger Fabric and Hyperledger Indy

IPFS (InterPlanetary File System) is a distributed file system where content is addressed by its cryptographic hash instead of by its location on a specific server. When a document is stored, it is split into chunks, hashed, and given a unique content identifier (CID). Any later retrieval uses this CID, and the system automatically verifies that the content has not changed. For a State archive, this means that integrity checking is built in: if the bits change, the CID changes. IPFS can run as a private network, with nodes controlled by public institutions and policies for replication and retention, while the archives themselves are encrypted before storage.

Hyperledger Fabric is a permissioned blockchain platform designed for consortia of known organisations. Instead of a public, anonymous network, Fabric is run by identified members (for example, ministries and agencies), each operating nodes that participate in consensus and host smart contracts ("chaincodes"). Fabric provides a shared, tamper-evident ledger where all participants can see and verify the same history of transactions, subject to access rules. It supports private data, channels, and endorsement policies to control who can see what and who must approve which operations. In an archiving system, Fabric is used to register archives (linking to IPFS CIDs), manage metadata and access policies, and keep a verifiable audit trail of all actions.

Hyperledger Indy is a distributed ledger and toolkit specialised in decentralised, self-sovereign identity (SSI). It introduces decentralised identifiers (DIDs) and verifiable credentials (VCs) that allow entities (citizens, civil servants, organisations, services) to prove who they are and what roles or rights they have, without exposing more personal data than necessary. Proofs can be built using advanced cryptography such as zero-knowledge proofs, so that an agent can demonstrate they are, for example, a "tax officer with level X clearance" without revealing their full identity details. In our context, Indy provides the identity layer for the archiving system: it underpins authentication, authorisation, and delegation of rights in a way that is auditable and privacy-preserving.



2.2 Complementarity of IPFS + Fabric + Indy for State archiving

Combined, these three technologies map naturally to the core functions of a state-scale digital archiving system. IPFS offers resilient, scalable storage for encrypted documents, with strong guarantees of integrity through content addressing. Hyperledger Fabric acts as the trusted coordination and governance layer: it records which document exists, under which identifier, with which metadata, policies, and lifecycle events, and it ensures that all participating institutions share the same, tamper-evident view. Hyperledger Indy provides the identity and credential layer that expresses who is allowed to perform which actions on which archives, and allows these rights to be checked cryptographically at each step.

This separation of responsibilities makes the overall system more robust and flexible. Storage can evolve (new hardware, new data centres) without breaking integrity, because the CIDs and ledger records remain stable. Governance rules and access policies can change over time through chaincode upgrades and new credentials, while the history remains auditable. Identity management can integrate with existing national eID schemes and future standards, without hard-coding user accounts into each application. For a State, this tri-layer approach supports sovereignty (control over infrastructure and rules), trust (cryptographic proofs and shared ledger), and adaptability (ability to extend to new use cases and regulations).

3. Reference architecture for a State-scale archiving system

3.1 Architectural principles and threat model

The proposed architecture is built on a few simple principles. First, it clearly separates three core functions: storing encrypted documents (IPFS), governing metadata and policies (Hyperledger Fabric), and managing identities and rights (Hyperledger Indy). This separation makes the system easier to reason about and easier to evolve: storage can change without breaking identity, and identity can change without rewriting the storage layer. Second, the design follows a “zero trust” approach: no single component, network zone, or administrator is trusted by default. Every access to an archive must be authenticated, authorised, and logged, and all critical operations must be verifiable by independent parties. Third, the architecture is distributed but permissioned. Multiple public institutions operate the nodes, share responsibility, and validate each other’s actions, while access to the network remains restricted to known and controlled actors.

From a threat point of view, the system assumes that it may face both external attackers (criminal groups, hacktivists, even hostile states) and internal threats (malicious or careless administrators, compromised user accounts, misconfigured systems). Possible attacks include attempts to alter or delete archives, inject fake documents, read confidential data, forge access rights, or deny access to legitimate users. The architecture therefore relies on strong cryptography, redundancy, and transparent logs to make such attacks difficult to execute and easy to detect. We also assume that some infrastructure components can fail or be compromised over time. For this reason, no single node, database, or data centre should be able to silently rewrite history or gain uncontrolled access to archives. Instead, the combination of IPFS, Fabric, and Indy is used to distribute trust, limit damage when something goes wrong, and provide clear evidence of what happened.

3.2 African Identity Ecosystems: The Present Fragmentation



At a high level, the architecture can be seen as three main layers sitting on top of secure government infrastructure: a storage layer that holds encrypted documents, a ledger and metadata layer that records what exists and under which rules, and an identity and access layer that decides who can do what. Around these layers, a set of application and integration services expose the system to ministries, agencies, and end-users (citizens, public servants, auditors) through APIs and user interfaces.

The storage layer, based on IPFS, is responsible only for storing and serving encrypted binary content. When an administration wants to archive a document, the document is first prepared (normalised format, optional compression) and encrypted. The encrypted file is then added to a private IPFS network operated by public institutions. IPFS splits the file into chunks, computes a content identifier (CID) based on its hash, and replicates it according to configured policies (for example, at least N copies in different data centres). The storage layer does not “know” what the document means, who owns it, or who can read it. Its job is simply to ensure that, when asked for a given CID, it can return the corresponding encrypted content and that any modification would be immediately visible because the hash would change.

The ledger and metadata layer, built with Hyperledger Fabric, keeps the “official memory” of the system. For each archived item, it stores a structured record that includes at least: a unique archive identifier, one or more IPFS CIDs, basic descriptive metadata (type of document, issuing authority, dates), the applicable access and retention policies, and a history of operations (creation, updates to metadata, accesses, legal events such as legal hold or destruction orders). These records are written and updated through smart contracts (chaincodes) that implement the business rules agreed by the participating institutions. Because Fabric is a permissioned distributed ledger, every participating organisation hosts nodes that validate and record transactions. No single actor can secretly modify or delete an archive record without detection, and all participants can verify the consistency of the ledger.

The identity and access layer, provided by Hyperledger Indy and self-sovereign identity tools, manages the digital identities and credentials used to interact with the archiving system. Citizens, public servants, services, and institutions are represented by decentralised identifiers (DIDs). Trusted authorities (for example, the civil service HR authority, a bar association, a health regulator) issue verifiable credentials that state roles and rights such as “judge of court X”, “tax officer with level Y clearance”, or “citizen owning identifier Z”. When an actor wants to perform an action (archive a document, consult a file, approve a policy change), their client proves to the system, using these credentials and cryptographic proofs, that they have the required attributes. The archiving applications do not need to manage long lists of roles and permissions internally; they rely on verifiable credentials checked against policies encoded in smart contracts.

On top of these three core layers, application and integration services provide a bridge with existing information systems. Business applications in justice, taxation, health, education, or civil status connect to the archiving platform through secured APIs or middleware. When they send a document for archiving, the platform orchestrates the steps: user or system is authenticated via Indy, the document is encrypted and stored in IPFS, a new record is created on Fabric with the CID and metadata, and a confirmation is returned. For consultation, the reverse happens: the requesting user proves their rights with credentials, Fabric is queried to check policies and log the access, the encrypted content is fetched from IPFS, decrypted, and delivered to the application. This orchestration can be exposed through generic services (for example, “archiveDocument”, “getDocument”, “searchArchive”) that different ministries can reuse.

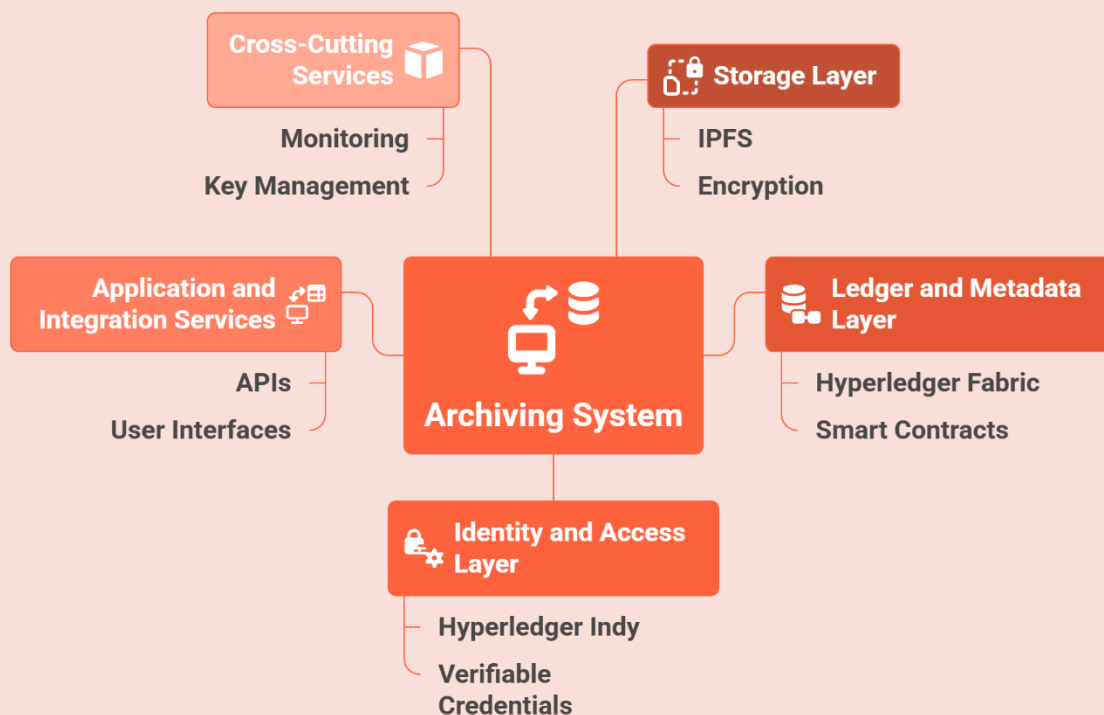
Finally, the logical view includes cross-cutting services such as monitoring, logging, key management, and administrative consoles. Monitoring services collect health and performance information from IPFS, Fabric, and Indy nodes to detect incidents and capacity issues. A central or federated key management service (potentially backed by hardware security modules) manages encryption keys for documents and credentials, enforcing policies such as key rotation and separation of duties. Administrative consoles allow authorised operators to configure organisations, onboarding new nodes, update chaincodes, define credential schemas, or adjust



replication and retention policies. All these actions themselves go through the same identity and ledger mechanisms, so that configuration changes are traceable and auditable.

In summary, the logical architecture clearly assigns responsibilities: IPFS handles “what bits are stored and where”, Fabric maintains “what archives exist and under which rules”, Indy governs “who you are and what you are allowed to do”, and application services tie everything together into usable workflows for administrations and citizens. This separation makes the system more understandable, more adaptable to change, and more resistant to both technical failures and intentional attacks.

Logical Architecture of Archiving System



3.3 Data model

At the centre of the system is a single main object: the Archive. Conceptually, an Archive is not just a file; it is a bundle that links together (1) the encrypted document stored in IPFS, (2) the metadata and rules stored in Hyperledger Fabric, and (3) the identities and rights expressed through Indy credentials. Every Archive receives a stable, system-level identifier (for example `ArchivelD`) that is used by business applications, while the IPFS content identifier (CID) and other technical details are kept inside the model. This separation allows the State to change storage technology or migrate data without breaking the external reference to the Archive.

For each Archive, the data model includes one or more content references. The most common case is a single IPFS CID for the encrypted document. In some situations, several CIDs may be needed, for example different versions of the same document, language variants, or associated attachments. These CIDs are treated as opaque technical values: the Fabric record simply states “this Archive is linked to these CIDs”, and any future integrity check consists in re-hashing the content and comparing it with the stored CIDs. Information about encryption (algorithm, key identifier, optional format information) is also part of the Archive record, but the keys themselves are managed in a separate key management system, not in the ledger.



The metadata section of the Archive captures what the document is and in which context it exists. Typical attributes include: type of document (judgment, tax declaration, birth certificate, diploma, administrative decision, etc.), issuing authority, dates (creation, registration, expiry, retention end), classification level (public, restricted, secret), and links to related archives (for example, all documents belonging to the same court case or the same administrative file). This metadata is structured and stored on Fabric so that it can be searched, filtered, and used to drive policies. It is important to keep this model relatively generic and extensible, with optional fields and possibly domain-specific extensions, so that different sectors (justice, health, taxation) can reuse the same core structure.

Associated with metadata, the model includes a set of policies and lifecycle rules. Policies describe who is allowed to perform which actions (read, update metadata, place legal hold, request destruction) under which conditions. They refer not to named users but to roles and attributes that can be carried in Indy verifiable credentials (for example, "any judge of court X", "tax officer with clearance level ≥ 3 ", "the citizen to whom this record belongs"). Lifecycle rules define how long the archive must be kept, whether it can be destroyed, and under which legal or administrative events its status may change (for example, when a case is reopened or when a minimum retention period expires). These rules are encoded in Fabric smart contracts, but their parameters are visible in the Archive record so that they can be inspected and audited.

Finally, each Archive carries a history of key events. Instead of storing full logs inside the Archive object, the ledger records a sequence of transactions linked to its ArchiveID: creation, updates to metadata, changes to policies, access events (possibly summarised or aggregated), and lifecycle events such as legal hold or destruction orders. Each transaction references the DID or credential of the actor, the action performed, and a timestamp. This creates a tamper-evident audit trail that can be reconstructed for any Archive when needed. In this way, the data model ties together content (IPFS), context and rules (Fabric), and actors (Indy) in a coherent, verifiable representation of what an "Archive" means at the scale of a State.

Here are two formal JSON examples of the Archive object: one for a court judgement, one for a birth certificate.

Court judgement: https://docs.origamis.pro/workspace/c1b666af-6c21-46da-b670-22212c5904a3/UX_2bL3UnAjYtj-S2CDSM

Birth certificate: <https://docs.origamis.pro/workspace/c1b666af-6c21-46da-b670-22212c5904a3/rGhqwWlvgee5QnOOa35pF>

3.4 End-to-end business flows

3.4.1 Archiving a document

In the proposed architecture, archiving a document is not a single technical action but an orchestrated sequence of steps that involve identity verification, encryption, distributed storage, and registration on the ledger. The process typically starts in a line-of-business application, such as a court case management system, a tax declaration portal, or a civil status application. From the user's point of view, they are simply "sending a file to the archive". Under the surface, the archiving platform coordinates IPFS for content storage, Hyperledger Fabric for metadata and policies, and Hyperledger Indy for identity and authorisation checks.

First, the user (a judge, tax officer, civil status officer, or a backend process acting on behalf of a citizen) authenticates in the business application using credentials that are ultimately backed by Indy verifiable credentials. The application or a dedicated "wallet" component obtains cryptographic proofs that the user has the required roles and attributes (for example, "civil status officer at office X"). The archiving service validates these proofs against the Indy network before



accepting any operation. This ensures that only authorised actors can create archives in a given domain and that their identity is strongly bound to the operation.

Once identity and authorisation are established, the business application sends the document and its initial metadata to the archiving service through a secure API. The document is normalised if needed (for example, converted to a long-term format like PDF/A) and then encrypted. The encryption keys are generated or retrieved from a government key management service (KMS), which enforces policies such as per-domain keys and key rotation. The archiving service never stores raw keys in application code or logs; it uses the KMS to encrypt and decrypt content when required.

The encrypted document is then added to the private IPFS cluster. The IPFS nodes chunk the file, compute the content identifier (CID), and replicate the data according to policies (for example, at least three copies in different data centres). The resulting CID is returned to the archiving service. At this point, the document's bits are durably stored, but from a governance perspective nothing exists yet: no archive will be visible or usable until it is registered on the Fabric ledger.

Next, the archiving service creates a new **Archive** object, including the business-level **archiveld**, the IPFS CID and encryption metadata, the descriptive metadata provided by the business application, and the initial policies and lifecycle rules. It then submits a transaction to the relevant Hyperledger Fabric channel, invoking a chaincode such as **CreateArchive**. This transaction includes the **Archive** data and is endorsed by the required set of organisations according to Fabric's endorsement policies. After endorsement, the transaction is ordered, committed to the ledger, and becomes part of the shared, tamper-evident history.

Finally, once the transaction is confirmed, the archiving service returns a success response to the business application with the **archiveld** and any relevant references. From that moment, the Archive is an official part of the State's digital memory: its existence, content reference, and initial rules are recorded on Fabric; its encrypted content is stored and replicated on IPFS; and the identity of the actor who performed the archiving is captured through their DID and credentials. Any later consultation, policy change, or lifecycle event will refer back to this initial record and build on it through additional Fabric transactions.

Below is a simplified sequence diagram of this flow:

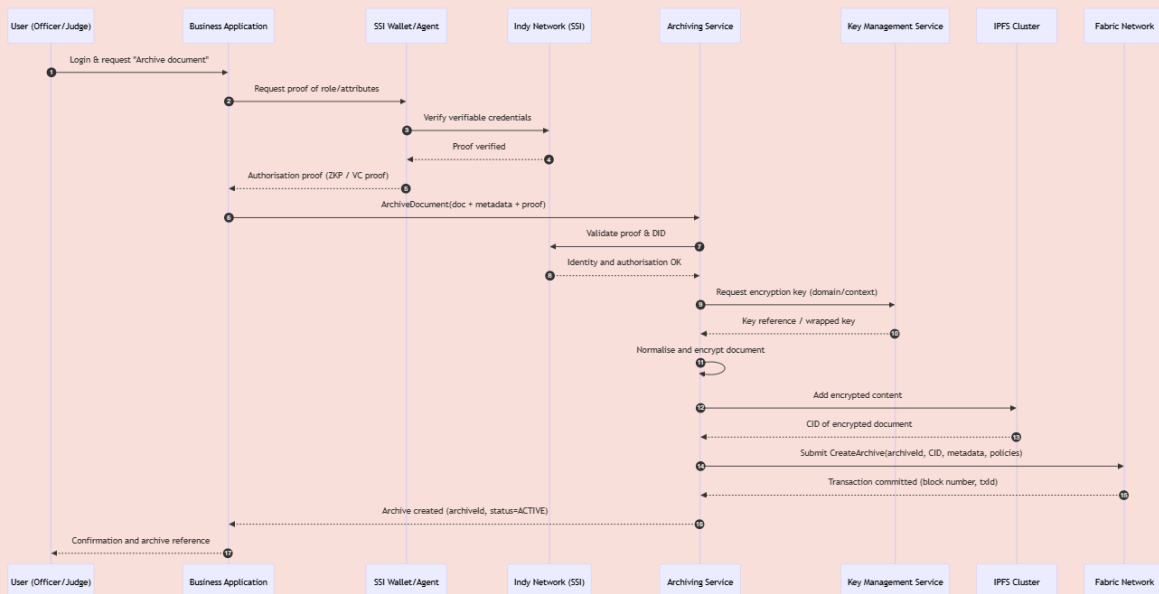


Figure 1: Document archiving flow diagram



3.4.2 Consulting a document and verifying its integrity

The consultation of an archived document follows a similar pattern to archiving: it is a controlled, auditable sequence where identity, authorisation, and integrity checks are all enforced. The process usually starts from a business application – for example, a court case system, a tax auditor’s tool, or a portal where a citizen requests a copy of a birth certificate. From the user’s point of view, they are simply “opening” or “downloading” a document. In reality, the request triggers interactions with Indy for identity, Fabric for policies and audit, and IPFS for content retrieval.

First, the user authenticates in the business application. As in the archiving flow, the application (with the help of an SSI wallet or agent) obtains verifiable credentials proving the user’s role and attributes: for example, “judge of court X”, “tax officer with level 3 clearance”, or “citizen with national identifier Y”. The application calls the archiving service with a GetDocument or GetArchive request, providing the target `archiveId` and a cryptographic proof of the user’s identity and rights. The archiving service verifies this proof against the Indy network before going further. If the proof is invalid, expired, or does not contain the required attributes, the request is rejected.

Once identity and authorisation are checked, the archiving service queries Hyperledger Fabric to retrieve the Archive record associated with the `archiveId`. The chaincode enforces access policies: it evaluates whether the attributes presented by the user’s credentials match the rules attached to the Archive (for example, “any judge of court X”, “the data subject of this record”, “civil status officer of office Z”). It can also check lifecycle constraints (for example, the archive is not destroyed and not under a special restriction). If the policy conditions are not met, the chaincode rejects the access and records a failed access attempt. If they are satisfied, Fabric returns the relevant data: the Archive metadata, the IPFS CID(s), the encryption information, and a confirmation that access is authorised. A new ledger entry is created to log the access event, referencing the user’s DID, the action performed, and a timestamp.

With the CID in hand, the archiving service contacts the IPFS cluster to retrieve the encrypted content. IPFS locates and assembles the file chunks, then returns the encrypted document. As part of this step, the archiving service can recompute the hash of the content and compare it with the CID and/or with a hash stored in the Archive record on Fabric. If there is any mismatch, the content is rejected and an integrity incident is logged. If the integrity check passes, the archiving service requests the decryption key (or key material) from the key management service (KMS). KMS enforces additional rules (for example, only decrypting content for authorised processes in secure environments). The document is then decrypted and optionally transformed (for example, watermarked or rendered as a view-only format) before being returned to the business application.

Throughout this flow, the separation of responsibilities remains clear: Indy proves who is requesting access and under which rights; Fabric decides whether access is allowed and records the decision; IPFS supplies what content is stored and allows integrity checking through the CID; KMS controls the use of decryption keys; and the business application presents the result to the user. The combination of these steps ensures that consultations are both user-friendly and strongly controlled, with cryptographic evidence for any later audit or dispute.

A simplified sequence diagram of the consultation and integrity verification flow is shown below:

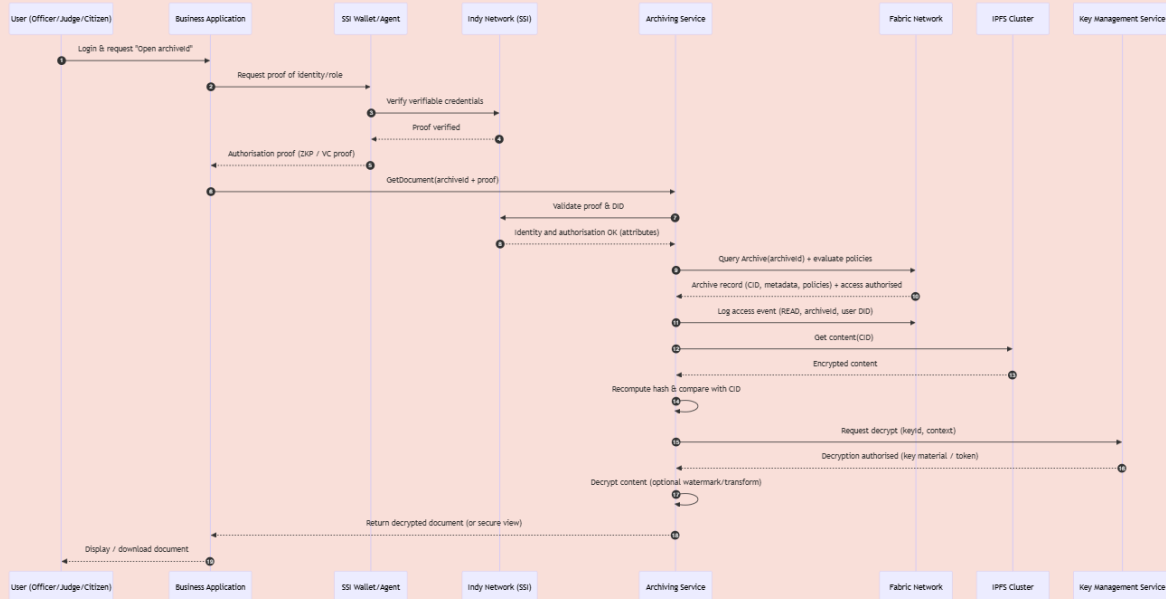


Figure 2: Document access and verification flow diagram

3.4.3 Lifecycle management: retention, legal hold and destruction

Beyond simple storage and consultation, a state-scale archiving system must manage the entire lifecycle of each archive: how long it must be kept, under which conditions it can be accessed, when it can or must be destroyed, and how exceptional events (such as legal disputes) temporarily “freeze” it. In the proposed architecture, lifecycle management is driven by rules encoded in Hyperledger Fabric smart contracts, while IPFS, Indy and the key management service (KMS) enforce the technical consequences of these rules.

When an archive is created, the business application or domain authority provides retention parameters: legal basis, minimum retention period, and, where applicable, an initial retention end date. These values become part of the Archive record on Fabric and are interpreted by chaincode responsible for lifecycle. Periodically, or when explicitly triggered, this chaincode evaluates archives whose retention thresholds are approaching or have passed. For example, it can flag an archive as “eligible for destruction”, “to be reclassified” (e.g. from confidential to public), or “to be transferred” to a different storage class. All state changes are written to the ledger, ensuring that every lifecycle step is traceable.

In parallel, external events can impose legal holds or other forms of freeze. A court order, an internal investigation, or a regulatory inquiry may require that certain archives cannot be destroyed or altered, even if their retention period has expired. In this architecture, authorised actors (for example, legal departments or regulators) hold specific credentials that allow them to issue “legal hold” transactions on Fabric, targeting individual archives or entire classes of archives. The chaincode updates the Archive’s lifecycle status (e.g. legalHold = true) and prevents any destruction or irreversible change while the flag is active. Removing a legal hold requires similar credentials and is also logged as a distinct event.

Destruction itself is handled carefully, especially given the tension between long-term integrity and principles such as the right to erasure. Rather than deleting all traces of an archive, the system follows a “controlled disappearance” approach. At the content level, the primary mechanism is cryptographic erasure: the keys needed to decrypt the document are destroyed or made inaccessible by the KMS once the destruction conditions are met and no legal hold is active. Without those keys, the encrypted content stored on IPFS becomes practically unreadable, even though the chunks may still exist for some time. Depending on regulatory

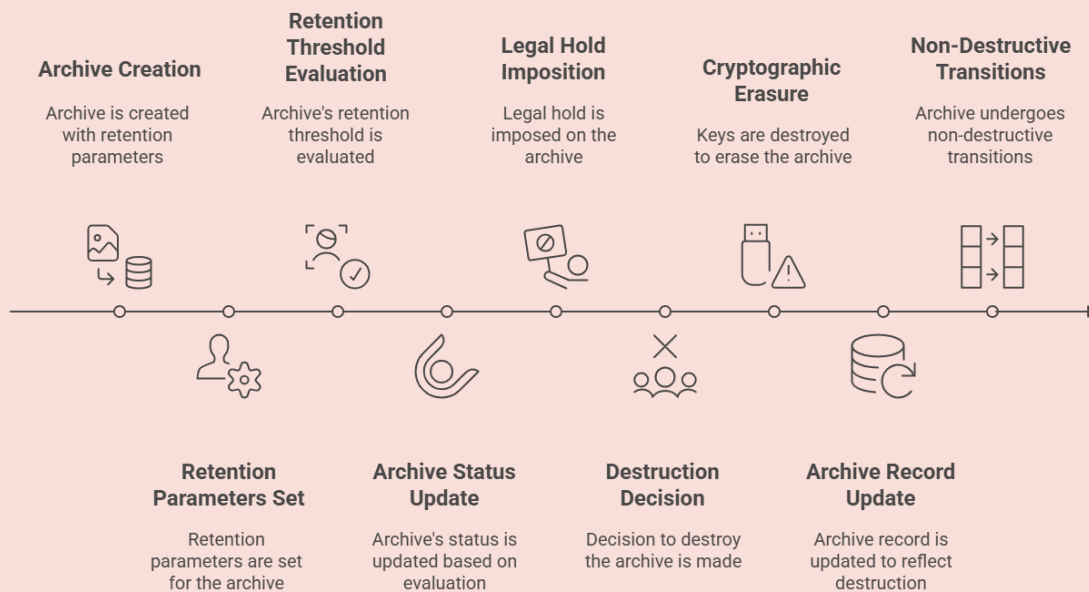


requirements, the IPFS nodes may also unpin the content so that it is eventually garbage-collected and removed from storage. On Fabric, the Archive record is not silently deleted; instead, its status is changed to something like **DESTROYED**, with a reference to the legal basis, the requestor, and the timestamp of the destruction decision. This preserves an auditable trace that the archive once existed and was destroyed according to defined rules, without keeping the underlying personal or sensitive data accessible.

Lifecycle management also covers non-destructive transitions, such as reclassification (making an archive public after a delay), migration to different storage profiles (e.g. from “hot” to “cold” archives), or linking an archive to new legal contexts (for example, when an old case is reopened). In each case, the same pattern applies: a credentialed actor or automated process requests a change; the Fabric chaincode checks policies and conditions; if accepted, it updates the Archive’s lifecycle fields and writes an event to the ledger; IPFS, KMS, and business applications adjust their behaviour accordingly (for instance, changing replication level or access rules).

By centralising lifecycle rules and decisions in Fabric, while executing their technical effects through IPFS and the KMS under the control of Indy-based identities, the architecture ensures that retention, holds, and destruction are not ad hoc operations but part of a coherent, auditable, and legally aligned process covering the full life of each archive.

Archive Lifecycle Management Process



3.5 Physical view and network topology

3.5.1 General physical architecture and network segmentation

At the physical level, the proposed archiving system is deployed across several sovereign data centres and connected through a secure governmental backbone network. The goal is to distribute responsibilities and resilience-no single building, rack, or cluster should be able to compromise the integrity or availability of the State’s archives. Typically, at least two national data centres (primary and secondary) operate in active–active or active–standby mode, with



additional regional sites depending on scale and risk appetite. All critical components-IPFS, Hyperledger Fabric, Hyperledger Indy, key management, and monitoring-are replicated across these sites.

The network itself is organised into security zones, separated by firewalls and gateways, to enforce the "zero trust" principle in the physical infrastructure. Facing the outside world, an external zone hosts citizen and partner portals, API gateways, and reverse proxies that expose well-controlled services to the internet or to inter-administration networks. Behind this, a DMZ (demilitarised zone) receives incoming traffic and terminates TLS, performs basic filtering, rate limiting, and threat detection. No direct connection from the internet to the ledger, storage, or database layers is allowed; all flows must pass through the DMZ and application gateways.

Deeper inside the infrastructure, an application and services zone hosts the archiving backend components: the orchestration services that handle "archiveDocument" and "getDocument" APIs, format normalisation services, audit and reporting components, and integration middleware to connect with line-of-business systems (justice, taxation, civil status, health, etc.). These services communicate with the ledger, identity, and storage layers using mutual TLS (mTLS) and strict firewall rules: only the necessary ports and protocols are opened, and every machine is authenticated by certificates issued from a government PKI.

The ledger zone is where the Hyperledger Fabric and Hyperledger Indy nodes live. For Fabric, this includes orderer nodes (forming the ordering service), peer nodes for each participating organisation, and certificate authority (CA) services for issuing identities to peers and clients. These nodes are spread across multiple data centres and, when appropriate, across different ministries or agencies, so that consensus and validation are not controlled by a single operator. Similarly, the Indy network is composed of validator nodes (stewards) operated by trusted public institutions (for example, interior, justice, finance, national IT agency), plus one or more agencies or wallet services that support SSI interactions for applications and users. The ledger zone is highly protected: only authorised peers, orderers, and agents can connect, and administrative access is tightly controlled and logged.

The storage zone hosts the private IPFS cluster and associated storage hardware. IPFS nodes run on servers that have access to large, redundant storage pools (SAN/NAS or object storage) and are distributed across data centres for resilience. Each node participates in the private IPFS swarm, identified by a shared swarm key and restricted to government-controlled infrastructure. Gateways between the application zone and the IPFS cluster are again protected by mTLS and firewall policies. In some scenarios, read-only IPFS gateways can be deployed closer to certain agencies or internal networks to optimise bandwidth and latency, while still enforcing strict controls on who can request which CIDs.

Cross-cutting services are grouped into a management and security zone. Here we find the key management service (KMS) and hardware security modules (HSMs) for encryption keys (or another system using MPC and ZKP), centralized multilevel logging and SIEM (Security Information and Event Management) platforms, monitoring and observability stacks, configuration management, and backup/restore tooling. Access to this zone is limited to a small number of administrators under strong procedural and technical controls (multi-factor authentication, jump hosts, bastion servers, just-in-time access). Even administrators do not bypass the ledger rules: actions such as key destruction, node onboarding, or policy changes must still go through Fabric/Indy processes and are recorded on the ledger.

Connectivity between ministries, departments, municipalities, and central data centres uses a government backbone or VPN overlay with QoS and segmentation. From a logical point of view, each institution participates in the same archiving system; physically, their systems connect via secure tunnels or dedicated links to the central services, ledger, and storage zones. Where local infrastructure is present (for example, regional data centres or municipal server rooms), it may host edge components such as local IPFS nodes, Fabric peers, or caching proxies, but always under central policies for certificates, routing, and security.



Overall, the physical view can be summarised as a layered, segmented architecture:

- Edge / External: citizen and partner portals, API gateways.
- DMZ: reverse proxies, WAF, first-line security filters.
- Application zone: archiving services, business integration, normalisation and presentation layers.
- Ledger zone: Fabric orderers and peers, Indy validators and agents.
- Storage zone: private IPFS cluster and storage systems.
- Management & security: KMS/HSM/MPC+ZKP, monitoring, logging, backup, configuration.

Each zone is isolated but interconnected through controlled, authenticated, and logged paths. This structure creates a physical foundation that matches the logical separation of responsibilities: storage, ledger/governance, and identity, all supported by a secure, sovereign infrastructure operated by multiple public actors.

Security Zone Comparison

Characteristic	Purpose	Components	Security	Connectivity
External Zone	Hosts portals, APIs	Portals, API gateways	Well-controlled services	Internet, inter-administration
DMZ	Receives traffic, filters	Reverse proxies, WAF	TLS termination, filtering	DMZ to application gateways
Application & Services Zone	Hosts backend components	Orchestration, normalisation	Mutual TLS, firewall rules	mTLS to ledger, storage
Ledger Zone	Hosts ledger nodes	Fabric, Indy nodes	Authorised peers, logging	Authorised peers, orderers
Storage Zone	Hosts IPFS cluster	IPFS nodes, storage	mTLS, firewall policies	mTLS to application zone
Management & Security Zone	Provides key management	KMS, HSM, logging	Limited admin access	Limited admin access

Figure 3: Security Zones and layers

3.5.2 Concrete deployment example: 7 departments and 26 municipalities

To make the physical view more tangible, we consider a country organised into 7 departments and 26 municipalities. The objective is to distribute the nodes so that:

- no single data centre or institution can compromise the system,
- latency stays acceptable for local administrations,
- cryptographic control is shared using MPC (Multi-Party Computation) and ZKP (Zero-Knowledge Proofs) instead of a single hardware security device.

We assume:

- 2 national data centres (DC-A and DC-B),
- 3 regional data centres (RDC-North, RDC-Center, RDC-South),



- 7 departmental capitals (one per department),
- 26 municipal IT sites (often lighter infrastructure).

A. Central level: national and regional data centres

At the national level (DC-A and DC-B), we host the core of the archiving infrastructure:

- Fabric ordering service:
 - Raft cluster distributed between DC-A and DC-B (e.g. 3–5 orderer nodes).
- Fabric peers for transversal “State” organisations:
 - National IT agency, Ministry of Justice, Ministry of Finance, Ministry of Interior.
- Indy validator nodes (“stewards”):
 - 4–6 validator nodes, split across DC-A / DC-B and one regional data centre.
- IPFS “backbone” nodes:
 - High-capacity IPFS nodes that guarantee long-term replication across the country.
- Archiving core services:
 - API gateways, orchestration services, normalisation services, audit dashboards.
- MPC key management cluster:
 - Instead of a central HSM, the root cryptographic material is split using MPC between several national entities (e.g. national IT agency, justice, finance, interior).
 - Each entity holds a share; key usage (e.g. decrypt, rotate, destroy) requires a threshold of shares and is accompanied by ZK proofs that the protocol was correctly executed without revealing the underlying key.

At the regional level (RDC-North, RDC-Center, RDC-South), we reinforce resilience and reduce latency:

- Additional Fabric peers for regional branches of key ministries.
- Additional Indy agents / edge services to support local identity operations.
- Regional IPFS nodes configured to keep copies of archives relevant to their departments (pinning policies per department).
- Caching gateways for read operations from departments and municipalities.

This central/regional backbone ensures that even if a whole department loses connectivity or suffers an outage, archives remain accessible from other regions, and the consensus of Fabric/Indy is not broken.

B. Department level: distributed trust across 7 departments

Each of the 7 departments plays an active role in the consortium:

- One Fabric peer per department
 - Operated by the departmental administration (or prefecture).
 - Participates in one or more channels (e.g. justice, civil status, taxation) depending on competences.
 - Keeps a local copy of the ledger segments relevant to its domain.
- One IPFS node per department
 - Hosts frequently used archives for that department, reducing bandwidth and latency.
 - Applies pinning policies coordinated with regional/national IPFS nodes (e.g. at least one departmental copy plus backbone copies).
- One Indy edge agent per department



- Used by departmental applications to interact with the Indy network (issue / verify credentials for local officers, courts, etc.).
- Can be co-located with the Fabric peer or in a separate departmental service zone.
- Local MPC key share
 - Departments do *not* hold root keys, but they may hold MPC shares for domain-specific keys (e.g. "department-justice-key-cluster").
 - When decrypting or rotating keys for archives under their responsibility, their share participates in an MPC protocol together with national/regional shares, producing a ZK proof that the operation followed policy.

Network-wise, each department connects to the governmental backbone via a VPN or dedicated link, with strict segmentation:

- Business systems (justice, civil status, taxation) are in departmental application zones.
- Departmental Fabric peers, IPFS nodes, and Indy agents are in protected infrastructure zones, only reachable via authenticated channels from central/regional services and local applications.

This gives departments real technical weight in the system (validation, storage, identity), instead of being mere clients of a national black box.

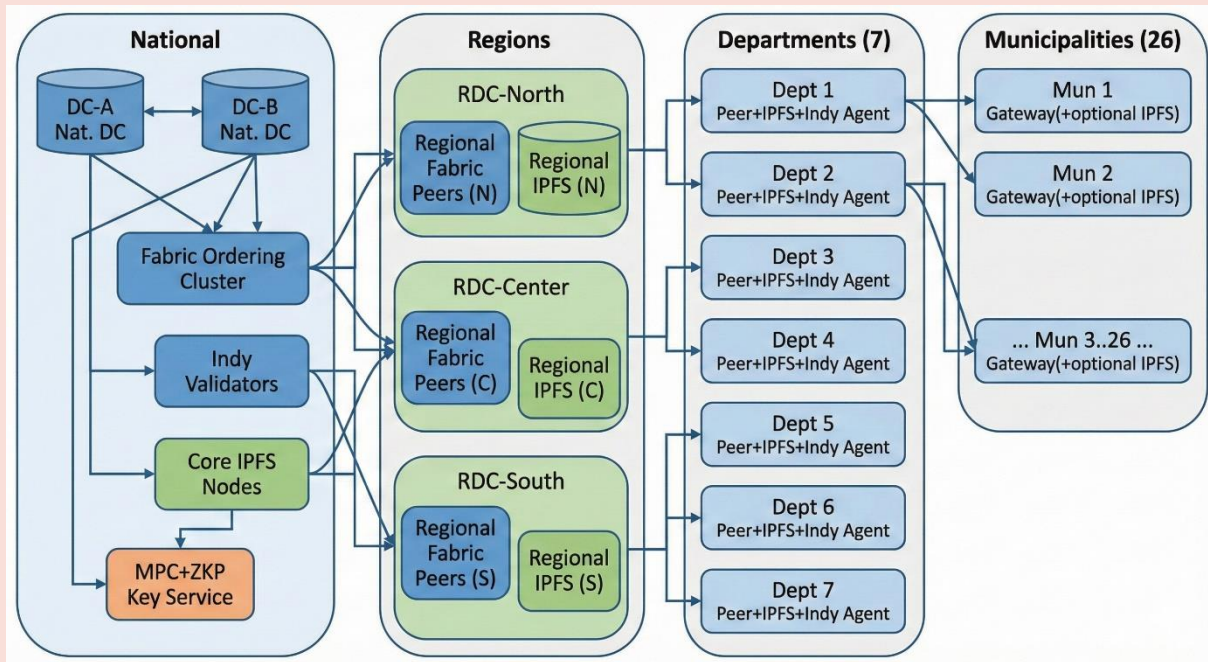
C. Municipal level: 26 municipalities as edge participants

The 26 municipalities have lighter infrastructure and usually do not need to run full ledger or storage nodes. Instead they act as edge participants:

- Municipal gateway / proxy
 - A secure gateway in each municipality connects local business applications (civil status, local taxes, permits) to the central archiving services.
 - Uses mTLS to talk to departmental/regional API gateways.
- Optional lightweight IPFS nodes
 - Some larger municipalities may run a small IPFS node for caching recently used archives (birth certificates, local permits), improving responsiveness.
 - These nodes are *not* authoritative for long-term replication; they simply mirror what is pinned at departmental or regional level.
- Local credential issuance (Indy)
 - Municipal HR/administration may issue verifiable credentials to municipal officers (e.g. "civil status officer, municipality X"), using a delegated authority model from the Ministry of Interior.
 - The issuing agent can be hosted centrally, but municipal offices manage which employees are active, again with MPC+ZKP-secured operations on key material (e.g. revocation keys).

From the municipality's point of view, the experience is simple: their applications call standardized APIs to archive and retrieve documents; identity and access rules are enforced centrally, but reflect local roles and structures; and performance is improved by departmental and regional caching.

Putting it together (high-level topology)



In this layout:

- National and regional sites provide backbone consensus, storage, and MPC+ZKP-based key services.
- Each of the 7 departments runs at least one Fabric peer, one IPFS node, and an Indy edge agent, plus holds a share in departmental key MPC protocols.
- The 26 municipalities connect as edge clients and optional caching nodes, leveraging the security and governance defined by the central consortium.

This gives you a concrete, scalable pattern that you can adapt to a real country, with clear roles for each level of administration and a modern cryptographic foundation based on MPC and zero-knowledge proofs rather than a single central HSM.

3.6 Comparative Synthesis

Performance and scalability in a state-scale archiving system are less about “raw speed” and more about predictable response times and stable behaviour under growth. The architecture must absorb millions of archives, thousands of concurrent users, and decades of accumulation without forcing disruptive redesigns every few years. In practice, this means designing for horizontal scaling from day one, separating hot paths from cold ones, and avoiding single global bottlenecks in IPFS, Fabric, or Indy.

From the user’s point of view, two operations dominate: *archiving a document* and *consulting a document*. Archiving is typically latency-tolerant (hundreds of milliseconds to a few seconds are acceptable), but must handle large volumes and peak periods (end of fiscal year, examination periods, etc.). Consultation is more latency-sensitive: judges, tax officers, or civil status agents expect documents to open in roughly the same time as from a local DMS, even when the underlying storage is distributed. The design therefore pushes heavy work (encryption, replication, consensus) into background or parallel steps when possible, while keeping the synchronous path for the user as short as possible.

On the IPFS side, performance is primarily driven by network topology, pinning strategy, and chunking parameters. Reads become fast when content is available on local or regional IPFS nodes close to the requesting administration; this is why the topology in 3.5.2 places nodes at



national, regional, and departmental levels. Commonly used archives (e.g. recent court cases, frequently requested civil status records) can be pinned on departmental nodes so that consultations do not always traverse the backbone. For writes, encryption and chunking are done once, then IPFS takes care of replication asynchronously: the user gets a confirmation as soon as at least one reliable node has pinned the content, while additional replicas are created in the background. Chunk size and blockstore configuration can be tuned so that typical document sizes (a few MBs) do not result in an excessive number of small blocks.

For Hyperledger Fabric, throughput and latency depend on endorsement policies, block size, chaincode complexity, and the number of peers. The paper's architecture uses multiple channels (for example, justice, civil status, taxation) and possibly private data collections to avoid one global channel becoming a bottleneck. Read-heavy operations, such as consulting archive metadata or checking access policies, can be served from local peers in each department, reducing cross-country traffic. Write operations (CreateArchive, lifecycle updates, access logs) are batched into blocks, so choosing appropriate block sizes and timeouts is important: too small and the network is noisy; too large and latency grows. Chaincodes are designed to stay relatively simple-no heavy computation, limited joins-while more complex analytics (global reporting, large audits) are offloaded to off-chain indexing systems that read from Fabric events.

Hyperledger Indy has very different performance characteristics: identity operations (writing DIDs, schemas, credential definitions) and on-ledger updates are relatively infrequent compared to document archiving; most interactions are off-ledger proof exchanges between wallets and verifiers. To keep latency acceptable, the architecture uses edge agents at departmental level and caches Indy data (schemas, credential definitions, revocation registries) locally, so that verifying a credential or a ZK proof does not require remote lookups for every request. Revocation and status checks can be batched, and common proofs (e.g. "civil status officer of office X") can be reused across multiple operations during a session, within defined security limits.

Scalability is achieved by horizontal replication and partitioning rather than vertical scaling of a few big machines. IPFS nodes can be added at new regional or departmental sites, pinning policies adjusted, and traffic naturally rebalanced as more replicas exist. Fabric peers can be added for new organisations or for load-sharing, with channels partitioned by domain or geography to contain state size. Indy validators and agents can likewise be extended as more institutions join. Indexing and search are explicitly moved to separate, scalable services (for example, search engines that subscribe to Fabric events and maintain searchable views of metadata), so that reporting and discovery do not overload the core ledger.

Finally, performance and scalability are continuously observed and tuned: monitoring collects metrics such as transaction latency on Fabric, IPFS fetch times, proof verification time for Indy, and end-to-end response times for archiving and consultation flows. These metrics feed capacity planning (adding peers or IPFS nodes before saturation), policy adjustments (changing block sizes, pinning strategies), and, when needed, refactoring of business processes that generate unnecessary load (for example, replacing repeated full downloads by incremental views). In this way, the architecture is not a static blueprint but a living system that can grow with the State's digital workload while keeping user experience and legal guarantees under control.

4. Security, confidentiality and regulatory alignment

4.1 Technical security measures



The main technical controls that protect confidentiality, integrity, and availability in the proposed architecture are based on :

A. Encryption and cryptographic design

- End-to-end content encryption
 - Every archived document is encrypted before it enters IPFS.
 - Use a standard envelope approach:
 - per-document symmetric key (e.g. AES-256-GCM or ChaCha20-Poly1305)
 - encrypted with a higher-level *domain key* (e.g. "justice", "civil status", "health").
 - Integrity is provided both by authenticated encryption (GCM tag) and by the IPFS CID (hash of the content).
- Hashing and content addressing
 - For CIDs and integrity checks, use modern hash functions (e.g. SHA-256 or BLAKE2) with multihash format so algorithms can evolve.
 - The archive record stores both the CID and, if needed, a separate hash field to allow parallel verification or algorithm migration.
- Key scopes and segregation
 - Different domains (justice, taxation, health, civil status) get separated key hierarchies.
 - Keys for test, pre-production, and production environments are strictly segregated.
 - Keys for content, signatures, and infrastructure (e.g. TLS) are never reused across purposes.

B. MPC + ZKP-based key management

Rather than relying on a single HSM, the State operates a distributed key management service based on Multi-Party Computation (MPC) and Zero-Knowledge Proofs (ZKP):

- Sharded key custody
 - Root and high-value keys (e.g. domain master keys, revocation keys) are split into shares held by independent institutions: national IT agency, ministries, possibly central bank or data protection authority.
 - No single institution can decrypt or rotate keys alone; operations require a threshold of shares (e.g. 3-of-5).
- Policy-driven MPC protocols
 - Operations such as "derive a new content-encryption key", "rotate a domain key", or "destroy a key" are expressed as high-level policies (who must approve, under which conditions).
 - MPC participants execute protocols that implement these policies without ever reconstructing the full key on one machine.
- Zero-knowledge proofs of correct key use
 - Every MPC operation produces a ZK proof that the protocol followed the expected algorithm and inputs (e.g. correct threshold, correct key identifier) without revealing the underlying secret.
 - The hash of this proof, or a reference to it, is stored on Fabric as part of the lifecycle events for keys and archives. This gives auditors evidence that cryptographic operations complied with policy.
- Auditable key destruction



- For archive destruction, KMS executes an MPC protocol that renders specific content keys unrecoverable.
- A ZK proof and a Fabric transaction record the event, providing strong evidence that decryption is no longer possible while preserving an audit trail.

C. Node and service authentication

- Mutual TLS everywhere
 - All communication between services (application ↔ archiving service ↔ Fabric peers ↔ IPFS nodes ↔ Indy agents) uses TLS 1.2+ or preferably TLS 1.3 with mutual authentication.
 - Certificates are issued by a state PKI or Fabric CAs; each node and service has a unique identity.
- Consistent identity for machines and services
 - Fabric's MSP identities and Indy DIDs are used not only for human users but also for service accounts (batch jobs, integration services).
 - Policies in chaincode can require that certain actions be initiated only by specific service identities (e.g. "only the civil status archiving service of Ministry X may create birth certificate archives").
- Short-lived credentials and rotation
 - Client certificates, access tokens, and session keys are short-lived (hours or days) and renewed automatically.
 - Revocation of compromised identities is propagated via Fabric and Indy (e.g. revocation registries for credentials, CRLs/OCSP for certificates).

D. Network and infrastructure protection

- Strict segmentation and micro-perimeters
 - As described in 3.5, the infrastructure is divided into zones (DMZ, application, ledger, storage, management).
 - Within zones, critical components are further segmented with firewalls and software-defined networking (for example, separate subnets for Fabric peers vs. orderers vs. IPFS).
- Least-privilege connectivity
 - Firewall rules are defined per service: "service A can talk to service B on port X" is explicit; nothing is open "just in case".
 - Default-deny policies at all layers (network, Kubernetes/VM security groups, service meshes).
- DDoS and abuse controls
 - At the edge/API gateway: rate limiting, abuse detection, and WAF (Web Application Firewall) rules prevent bulk scraping or brute-force attempts to enumerate archives.
 - Internal protection: safeguards against misbehaving internal services (e.g. circuit breakers, quotas per client).
- Hardening of nodes
 - IPFS, Fabric peers, and Indy validators are deployed on hardened OS images (minimal packages, secure kernels, disabled unused services).
 - Configuration is managed by code (Infrastructure as Code) so that deviations can be detected and rolled back.

E. Application, chaincode and wallet security



- Secure chaincode/smart contract lifecycle
 - Chaincode is treated as critical infrastructure code: peer review, static analysis, and security testing before deployment.
 - Each chaincode is code-signed, and Fabric endorsement policies ensure only approved versions can be instantiated or upgraded.
- Input validation and policy enforcement
 - Chaincode validates all inputs (metadata, policy updates, lifecycle commands) against schemas and role/attribute rules.
 - No business logic related to integrity or access control is left only in the off-chain application layer; the ledger always reconfirms.
- Secure SSI wallets and agents
 - Citizen and staff wallets use secure storage for private keys (device secure enclaves where available, protected keystores otherwise).
 - For server-side agents (e.g. departmental Indy agents), keys are managed by the MPC-KMS where feasible, reducing the risk of large key theft.
 - Phishing-resistant flows (e.g. QR-code based proof requests in physical offices, signed proof requests from official domains) reduce social engineering.
- File normalisation sandboxing
 - Document conversion (to PDF/A, for example) is a frequent attack vector. These services run in sandboxed environments (containers or VMs) with no direct access to IPFS, Fabric, or KMS.
 - Only the resulting normalised file and a checksum leave the sandbox; any malware embedded in originals cannot escape to the core system.

F. Logging, monitoring and incident response

- Multi-layer logging
 - Application logs: record API calls, validation errors, and user-facing failures.
 - Ledger events (Fabric): every archive creation, update, and access decision is recorded, including actor identifiers and reasons for rejection.
 - MPC/ZKP logs: record which institutions participated in key operations, plus references to ZK proofs stored or hashed on Fabric.
- Security monitoring
 - Central SIEM aggregates logs from nodes, gateways, OS, chaincode, and MPC services.
 - Alerts on anomalies: unusual access patterns (e.g. massive reads by one officer), repeated failed proof verifications, abnormal IPFS traffic, unusual number of key operations.
- Forensic-ready design
 - Time synchronisation (NTP with authenticated sources) across all components to make timelines coherent.
 - Logs are write-once where possible (append-only stores, WORM object storage), and hashes of sensitive logs can be anchored periodically on Fabric to detect tampering.
- Incident response playbooks
 - Predefined procedures for credential theft, node compromise, key share compromise, data exfiltration suspicion, or malicious chaincode deployment.
 - Automated or semi-automated actions: revocation of credentials, isolation of nodes, temporary tightening of policies (e.g. disabling certain high-risk operations during investigation).

G. Resilience and backup



- Multi-site replication
 - IPFS pinning policies ensure that at least N geographically separated nodes hold each archive.
 - Fabric peers and orderers are distributed across data centres; a single site outage does not stop consensus.
 - Indy validators follow similar placement rules to preserve ledger availability.
- Backups with cryptographic consistency
 - Off-line or off-site backups of ledger state and configuration are taken regularly.
 - Backups include enough information to verify, via hashes or checkpoints, that restored states match the on-chain history.
- Graceful degradation
 - Critical read operations should remain possible even under partial outage (e.g. via regional IPFS and peers), with clear indicators when some services (analytics, batch lifecycle evaluations) are temporarily unavailable.

Together, these technical measures ensure that security is enforced at every layer: cryptography, identity, network, software, and operations. They provide a concrete foundation on which the following subsections (4.2, 4.3, etc.) can show how confidentiality and regulatory obligations (privacy, archiving law, eIDAS, etc.) are met in practice.

4.2 Functional Components and Technical Standards

In this architecture, access control is entirely driven by identity attributes and verifiable credentials, not by hard-coded user lists or roles buried in each application. Hyperledger Indy and self-sovereign identity (SSI) provide the foundation: every actor-citizen, public servant, institution, or technical service-has a decentralised identifier (DID) and can be equipped with verifiable credentials (VCs) that state their roles, rights, and relations to specific archives or domains.

At a high level, three categories of actors need to be distinguished:

- Citizens and legal entities: holders of credentials proving their civil identity, national identifier, and relationship to specific records (e.g. "parent of child X").
- Public servants and professionals: judges, clerks, tax officers, doctors, civil status officers, archivists, auditors, each with credentials describing their function, organisation, and level of clearance.
- Systems and services: business applications, batch jobs, integration services that act on behalf of organisations and must be recognised and constrained just like human users.

Each of these actors has one or more DIDs and receives VCs from trusted issuers: ministries, professional orders, regulators, or delegated authorities (e.g. departmental HR). The schemas of these credentials (what fields exist, what they mean) are registered on the Indy ledger, so that verifiers across the State interpret them consistently.

A. From roles to policies: how rights are expressed

Instead of saying "user X can read archive Y", the system expresses rights as policies that refer to attributes inside credentials. For example:

- "Any holder of a credential of type `vc:judge` with `courtCode = COA-X` may read judgements issued by Court of Appeal X."
- "The person whose `nationalIdentifier` matches the one in the birth certificate metadata may request a copy of that certificate."



- "Any holder of a vc:civilStatusOfficer with officeCode = MUNI-12 may create or update civil status archives for municipality 12."

These policies are attached to the Archive object (in its policies.access section) and implemented inside Fabric chaincode. When a request arrives, the chaincode does not care about the user's name; it asks: *"Does this proof show that the actor satisfies the attributes required by the policy?"*

Because credentials are cryptographically signed by their issuers, and because issuer keys and schemas are anchored on Indy, the verifier (archiving service + chaincode) can trust them without maintaining central role tables for the whole State.

B. Proof-based access flow

Concretely, an access request always involves a proof exchange:

1. The business application sends a proof request (structured JSON) to the user's wallet or to a server-side agent. The request specifies which credential types and which attributes or predicates are required.
2. The wallet selects appropriate credentials, builds a zero-knowledge proof that it satisfies the request (for example, "I have a vc:judge with courtCode = COA-X, issued by Ministry of Justice, not revoked"), and sends back the proof.
3. The archiving service verifies this proof using Indy: it checks issuer keys, schemas, revocation status, timestamps, and the mathematical correctness of the proof.
4. If valid, the service extracts only the attributes the policy needs (e.g. courtCode, clearanceLevel, nationalIdentifier), never the full content of all credentials.
5. Fabric chaincode then evaluates policies against these attributes and the Archive's metadata and lifecycle state to decide whether to authorise the requested action.

This design has several advantages:

- It supports fine-grained, attribute-based access control (ABAC) without centralising all user data.
 - It naturally enforces data minimisation: the verifier sees only what is necessary (and possibly even just proof of a predicate like "clearanceLevel \geq 3").
 - It works across organisational boundaries: a credential issued by one ministry can be used to access archives controlled by another, as long as the policy recognises that schema and issuer.
-

C. Delegation, separation of duties and emergency access

Real public administration workflows require more than simple "can read/can write" rules. Three important patterns must be supported:

- Delegation: A senior officer may temporarily delegate some rights to a colleague or a replacement. This can be modelled as an additional credential:
 - A delegator requests a new credential for the delegate (e.g. "acting clerk for Court X until date Y"), issued by a trusted authority;
 - Policies in Fabric require that either a "permanent" or an "acting" credential be presented;
 - The delegated credential naturally expires, with no need to remove the delegator's rights.
- Separation of duties: Certain actions (e.g. key destruction for archive deletion, applying legal hold, modifying lifecycle rules) should require two or more independent approvals. The policy can state that:



- Two different credentials of type `vc:legalSupervisor` from two distinct organisations must approve the transaction;
- The MPC+ZKP key management layer only executes the requested operation if the Fabric transaction showing both approvals has been committed.
- Emergency or break-glass access: Some legal frameworks require that, in very specific situations (e.g. disaster, urgent medical access), normal rules can be bypassed under strict conditions. In this architecture:
 - Emergency credentials are rare and highly controlled (e.g. "crisis officer with emergency access rights").
 - Policies can encode special branches like "permit access if user holds `vc:emergencyAccess` AND event X is active";
 - Any emergency access triggers mandatory logging and notification and may require post-hoc justification recorded in Fabric.

D. Governance of schemas, issuers and revocation

Strong access control depends on who is allowed to issue which credentials and how those credentials are managed over time. This is a governance question as much as a technical one:

- Credential schemas are defined and approved at national level (e.g. by a cross-ministerial body). Once published on Indy, they serve as standards across all sectors (justice, taxation, civil status, health).
- Issuers (ministries, departments, professional bodies) have DIDs with roles that allow them to register credential definitions for given schemas. For example, only the Ministry of Justice may issue `vc:judge` credentials.
- Revocation is critical: when a judge retires, an officer moves, or a system account is compromised, associated credentials must be revoked.
 - Revocation registries on Indy allow verifiers to check if a credential is still valid without exposing all revoked IDs.
 - Fabric can keep a mirror view of revocation events relevant to archiving policies, making it easier to audit and to correlate with access logs.

Policies on Fabric always refer to schemas and attribute names, not to specific credential definitions. This ensures that issuers can rotate keys, update revocation registries, or even change technical details of credential issuance without breaking high-level access rules.

E. Binding identities to archives and operations

Finally, SSI-based identity is not used only at the boundary; it is deeply embedded in the audit trail:

- Every significant operation on an Archive—creation, metadata update, policy change, lifecycle event, access—carries references to:
 - the DID of the actor (human or system),
 - the type of credentials used,
 - optional hashes or identifiers of the proof (without exposing personal data).
- These elements are written into Fabric transactions, making it possible to reconstruct who did what, based on which rights, at which time.

Because proofs are generated and verified with well-specified cryptographic protocols, and because issuer keys and schemas are anchored on Indy, this audit trail is verifiable by third parties (auditors, regulators, courts) without requiring blind trust in the operators of the system.

In this way, SSI and Indy do not just "log users in": they become the core mechanism that drives, constrains, and explains access control across the entire state-scale archiving platform.



4.3 Legal and regulatory compliance

From a legal point of view, a state-scale archiving system must fit into a dense and heterogeneous set of laws, not just data protection rules. Public archives are at the intersection of:

- data protection and privacy laws,
- public archives and records management laws,
- procedural and evidence law,
- trust services and electronic identification frameworks,
- sector-specific regulations (health, taxation, justice, education, etc.),
- and sometimes national security, cybersecurity and critical infrastructure obligations.

The proposed architecture is designed so that these requirements can be mapped to technical controls in a systematic way, and so that the system can be adapted to the specifics of each country's legal framework.

4.3.1 Data protection and privacy

Most modern legal systems now include data protection principles similar to those in the GDPR, even if the exact text differs:

- lawful basis for processing,
- purpose limitation,
- data minimisation,
- storage limitation,
- integrity and confidentiality,
- rights of the data subject (access, rectification, erasure, restriction, portability, objection, etc.).

The architecture responds to these principles as follows:

- Lawful basis and purpose limitation:
 - Each Archive includes metadata about the legal basis and the purpose of the processing (e.g. tax assessment, civil status registration, judicial decision).
 - Policies in Fabric can prevent the use of archives for incompatible purposes (for example, blocking bulk reuse for unrelated profiling).
- Data minimisation:
 - At access time, SSI and ZK proofs mean that applications see only the attributes strictly necessary to decide whether a user may access an archive; they do not need full identity dossiers.
 - For citizens' access, views can be limited to the parts of the document that concern them, with masking/pseudonymisation for other parties, enforced by business logic and policies.
- Storage limitation and erasure:
 - Retention and destruction rules are modelled explicitly in the Archive's lifecycle.
 - Cryptographic erasure via MPC-managed keys supports practical implementation of "erasure" or "restriction", while maintaining evidence that something existed and was destroyed according to law.
- Integrity and confidentiality:
 - Confidentiality is enforced by encryption + access control, as described in 4.1 and 4.2.
 - Integrity is supported by content-addressing (IPFS CIDs), ledger immutability (Fabric), and signed credentials (Indy), providing strong technical guarantees that records have not been silently altered.



The important point is that data protection is treated as a structural constraint on design and operation, not as a later compliance layer.

4.3.2 Public archives and records laws

Most States also have specific public archives laws or records management regulations that define:

- which documents are public archives,
- who is responsible for them,
- how long they must be kept,
- when and how they may be made public,
- and under which conditions they can be destroyed or transferred to historical archives.

The architecture reflects these requirements directly in the data model and lifecycle:

- Responsibility and ownership:
 - The Archive metadata includes fields for issuing authority, owner organisation, and sometimes a chain of custody (e.g. initial ministry → national archives).
 - Fabric channels can correspond to archival responsibilities (e.g. "current archives" vs "intermediate archives" vs "historical archives"), and transitions between channels can reflect legal transfer processes.
- Retention schedules:
 - Retention categories and durations are explicitly encoded (e.g. "permanent", "30 years after event X", "until case closure + 10 years").
 - Automated lifecycle evaluations by chaincode ensure that archives are flagged for transfer or destruction when the legal retention period is reached, not arbitrarily.
- Opening and public access rules:
 - Some archives must become publicly accessible after a delay (e.g. 25, 50, 75 years).
 - Policies can automatically relax access restrictions when the "opening date" is reached, while still logging who accesses the archives and under what conditions.

By embedding these concepts into Fabric and the Archive model, the system is aligned with archival law by design and can produce clear evidence that archives have been handled according to legal rules.

4.3.3 Evidence law, procedural rules and trust services

Digital archives frequently appear as evidence in court proceedings or in administrative disputes. Their legal value depends on the ability to demonstrate:

- that the document comes from the claimed authority,
- that it has not been altered since a given date,
- that the chain of custody (who accessed or manipulated it) is clear,
- and that signatures or seals meet the relevant trust service standards (e.g. qualified electronic signatures/timestamps, national equivalents of eIDAS-like rules).

The architecture supports these requirements in several ways:

- Proof of origin and authorship:
 - Issuing authorities (courts, ministries, municipalities) have their own DIDs and signing keys.



- Archives can include digital signatures or seals on the original document, and the Fabric record can store references to those signatures and to qualified timestamp tokens.
- Proof of integrity over time:
 - The IPFS CID and additional hashes stored on Fabric allow recomputation and verification at any point.
 - The ledger gives a chronological, tamper-evident record of when an archive was created, which CIDs were associated, and when (if ever) changes were made to metadata or policies.
- Chain of custody:
 - Every significant access or modification is logged with the DID and role of the actor, plus time and context.
 - This helps demonstrate to a court how a document was handled from creation to presentation as evidence.
- Compliance with trust service rules:
 - Where the law requires the use of qualified signatures, seals or timestamps, the architecture can integrate with national trust service providers.
 - References to such qualified evidence (e.g. token IDs, certificate chains) are stored in the Archive metadata so that they can be verified later, even if cryptographic algorithms evolve.

This design allows the State to argue that archived electronic documents fulfill the same, or higher, evidentiary standards as their paper equivalents.

4.3.4 Sector-specific regulations and professional secrecy

Many legal obligations are sector-specific:

- medical secrecy and health data regulations,
- tax secrecy, banking secrecy,
- professional secrecy for lawyers and notaries,
- specific justice/penal codes governing access to criminal records,
- education, social protection, and child protection laws, etc.

These rules often define who may see what, under which circumstances, with strong sanctions for violations. The architecture uses per-domain policies and credentials to map these constraints:

- Domain-specific credentials:
 - Health professionals, tax officers, social workers, judges, police officers, lawyers, etc. receive credentials that encode both profession and scope (hospital X, tax office Y, court Z, bar association W).
 - Policies for archives in each domain reference these credentials and attributes explicitly.
- Fine-grained policies per archive type:
 - A "tax audit report" may be accessible to a wider group than a "suspicion report on money laundering".
 - A "psychiatric report" may have stricter conditions than a general health record.
 - These distinctions are reflected in policy templates that can be applied automatically based on Archive metadata (documentType, sensitivityLevel, etc.).
- Multi-party access conditions:
 - Some laws require that access be granted only if several roles are involved (e.g. a judge + a medical expert, or an investigator + a prosecutor).
 - Fabric policies can enforce such joint conditions before granting access or unsealing a document.



The combination of SSI credentials and policy-driven Fabric chaincodes allows legal rules that are currently applied in organisational procedures to be translated into verifiable, machine-enforced access rules.

4.3.5 Cybersecurity, critical infrastructure and sovereignty

Finally, many States classify their digital infrastructures-especially those holding justice, security, or health records-as critical information infrastructures. Laws may require:

- specific security measures (encryption, logging, incident reporting, penetration testing),
- certification against standards (ISO 27001, national cybersecurity frameworks),
- localisation or sovereignty requirements (data stored in-country, under public control),
- and strict conditions for outsourcing or use of cloud services.

The proposed architecture supports these requirements by design:

- Sovereign control:
 - Core components (IPFS nodes, Fabric peers, Indy validators, MPC key services) are deployed in State-controlled data centres or in trusted sovereign cloud environments.
 - Cryptographic keys are split across national institutions via MPC, reducing the risk that one provider or foreign authority can unilaterally access archives.
- Alignment with security regulations:
 - Technical security measures described in 4.1 (encryption, segmentation, logging, monitoring, incident response) can be directly mapped to national cybersecurity controls.
 - The ledger-based audit trails and ZK proofs for key operations make it easier to demonstrate compliance to regulators and auditors.
- Controlled use of external providers:
 - If external or cloud services are used (for example, for non-critical analytics), they can be kept outside the core ledger/storage zones and only receive pseudonymised or aggregated data.
 - The architecture keeps the most sensitive elements (keys, raw archives, core ledgers) under direct sovereign control.

Overall, point 4.3 shows that the IPFS + Hyperledger (Fabric + Indy) architecture is not only about technical robustness; it is a compliance-by-design framework. By explicitly modelling legal obligations-across data protection, archival law, evidence law, sector regulations and cybersecurity-inside the data structures, policies, and cryptographic mechanisms, the State can demonstrate that its digital archiving system is both legally aligned and technically enforceable, and can be adapted to the specific laws and institutions of any given country.

4.4 Probative value, global trust frameworks and evidentiary requirements

For a State, a digital archiving system is only useful if its records can stand in court or before oversight bodies. The question is simple but strict: *"Can this document be accepted as reliable evidence of what it claims to be?"* Different jurisdictions use different terms and frameworks (eIDAS in the EU, other trust and evidence laws elsewhere), but they share a set of common expectations:

- identifiable and legitimate source (who issued the record),
- integrity over time (no undetected alteration),
- clear and auditable chain of custody,



- and technically robust signatures, seals, timestamps and logs.

The proposed IPFS + Hyperledger Fabric + Indy architecture is designed to meet these expectations in a way that can align with any national or regional trust framework, not just one specific regulation.

4.4.1 Proving origin and authorship

Courts and regulators want to know who really created or approved a record and whether that authority was legitimate at the time.

- Institutional identities
 - Ministries, courts, agencies and municipalities have DIDs and associated signing keys.
 - These keys are recorded and governed on Indy (and/or a national root of trust), so anyone can verify that "this key belonged to Court X on date Y".
- Document-level signatures and seals
 - Original documents (judgements, certificates, tax notices) can be digitally signed or sealed by the issuing authority.
 - The Archive metadata in Fabric stores:
 - the type of signature/seal used (simple, advanced, qualified, or equivalent under local law),
 - certificate identifiers and chains,
 - and references to trusted timestamp tokens when applicable.
- Alignment with global trust models
 - Even if a country does not use eIDAS as such, the system is compatible with similar models:
 - separation between "simple" and "high-assurance" signatures,
 - use of accredited trust service providers where required,
 - and clear, long-term verifiability of certificates and timestamp evidence.

This means that when a document is presented as evidence, the State can show cryptographically and organisationally that it was issued by the right authority, with the right signing process, at the right time.

4.4.2 Proving integrity and non-repudiation over time

Beyond origin, the key evidentiary question is: *"Has this document changed since it was archived?"*

- Immutable references via IPFS and hashes
 - Each archived document is referenced by a CID (content hash) and, if needed, additional hash fields stored on Fabric.
 - To verify integrity, a court expert can re-hash the content and compare it to the stored values. Any change in the bits breaks the match.
- Ledger-backed history (Fabric)
 - Fabric keeps an append-only history of the Archive:
 - when it was created,
 - which CIDs were associated,
 - whether any technical migrations happened (e.g. re-encryption, format conversion),
 - and who triggered them.
 - This history is collectively validated by peers from multiple institutions, making silent, unilateral tampering extremely hard.
- Non-repudiation of operations



- Each key operation (creation, access, policy change, key destruction) is recorded as a Fabric transaction, signed by the acting entity's credentials.
- Combined with MPC + ZKP for key operations, this provides strong evidence that:
 - the action was initiated by specific roles,
 - it followed a defined protocol,
 - and it was accepted by the consortium at a specific time.

In a dispute, the State can thus provide not only the document, but also a cryptographically verifiable timeline of its life inside the archive.

4.4.3 Chain of custody and proof packages

In practice, when a digital record is used as evidence, what is presented is often a "proof package" rather than just a file:

- Contents of a proof package might include:
 - the archived document itself (in its evidentiary format, e.g. PDF/A),
 - its Archive metadata (issuing authority, dates, legal basis, document type),
 - hashes and CIDs,
 - references to signatures, seals and timestamps,
 - a selection of ledger events showing creation and key lifecycle steps,
 - and, when relevant, a summary of access history or legal holds.
- Reconstruction from the system
 - Because all these elements live in Fabric, IPFS, and Indy, an authorised official can generate this proof package on demand, for a given archived.
 - The generation itself becomes a logged event, ensuring that even the process of preparing evidence is traceable.
- Global alignment, local rules
 - One country may require compliance with a strict national trust list and certified time-stamping authorities; another may focus on court-approved expert verification.
 - The architecture supports both:
 - trust frameworks plug in at the level of signature types, certificate policies and timestamp providers,
 - while the underlying integrity, origin and chain-of-custody evidence remains the same.

By designing the archive so that every evidentiary requirement corresponds to a concrete technical artefact—signature, hash, ledger entry, ZK proof—the State can adapt to different national and international standards while keeping a single, coherent technical core. The result is a digital archiving platform whose records are not only well stored, but legally defensible wherever they need to be presented.

5. Governance, organisation and operating model at State scale

Technology alone cannot guarantee trustworthy archiving. For a State, the real guarantee comes from how the system is governed: who has the right to change rules, who operates which nodes, how conflicts are resolved, and how new institutions are onboarded. The proposed IPFS + Hyperledger (Fabric + Indy) architecture is therefore coupled with a consortium governance model, where ministries, agencies, and (optionally) regional or departmental entities share responsibility in a structured way.



5.1 State-level consortium model

The archiving platform is operated as a multi-institution consortium, not as a monolithic system owned by a single ministry or vendor. The idea is simple: every institution that has legal responsibility over archives should also have a controlled technical role in the platform, proportional to its mandate and capacity.

A practical model includes four layers of governance:

1. Strategic council (policy and direction)
2. Technical steering committee (architecture and evolution)
3. Security & compliance board (risk and conformity)
4. Operational consortium (node operators and service owners)

Each layer has clear responsibilities and is represented in Fabric/Indy as specific roles and credentials.

5.1.1 Strategic council – “why and what”

The Strategic council is the political and strategic brain of the consortium. It typically includes:

- the ministry or agency in charge of digital transformation or IT,
- the national archives authority,
- key line ministries (justice, interior, finance, health, education),
- possibly the data protection authority and other independent regulators.

Its main tasks are to:

- define the overall mandate and scope of the platform (which archives, which sectors, which phases of the archival chain),
- approve major policy frameworks (who can join the consortium, general access principles, long-term funding),
- validate high-impact changes (e.g. adding a new class of critical archives, cross-border data exchange agreements),
- arbitrate conflicts that cannot be resolved at lower levels (e.g. competing priorities between ministries).

Decisions at this level are infrequent but structural. They are reflected in Fabric as updates to global parameters and policies, often requiring multi-signature approvals from several council members' credentials.

5.1.2 Technical steering committee – “how and when”

The Technical steering committee is responsible for the architecture and technical roadmap:

- defining the reference architecture and the “golden patterns” for new use cases,
- approving new Fabric channels, chaincodes, and Indy schemas,
- deciding when and how to adopt new technologies (new hash algorithms, new encryption schemes, post-quantum transitions),
- supervising performance, scalability, and interoperability with other national or regional platforms.



Its members are senior architects and engineers from the national IT agency, leading ministries, and possibly regional IT providers. They work closely with the operational teams and with external experts (universities, industry).

In the platform, their authority is expressed through:

- code-signing keys for approved chaincode and configuration packages,
- privileged roles that can propose (but not unilaterally enforce) configuration changes,
- participation in MPC key operations when cryptographic policies are updated.

Fabric endorsement policies can require that certain configuration transactions (e.g. deploying a new chaincode that affects lifecycle rules) are endorsed by peers operated under the responsibility of this committee.

5.1.3 Security & compliance board – “safe and lawful”

The Security & compliance board focuses on risk management and legal conformity:

- mapping security controls in 4.x to national cybersecurity and critical infrastructure requirements,
- performing or supervising risk assessments, penetration tests, and audits,
- reviewing incidents, near-misses, and major policy violations,
- ensuring alignment with archive law, evidence law, sector-specific rules, and data protection principles.

It includes representatives from:

- the national cybersecurity agency or equivalent,
- data protection authority,
- national archives and justice,
- security officers from key ministries.

This board does not manage day-to-day operations but can:

- approve or block certain high-risk features (e.g. emergency access modes, cross-border connectors),
- require additional logging, ZK proofs or controls for sensitive operations,
- issue formal recommendations that must be translated into Fabric/Indy policies and operational procedures.

5.1.4 Operational consortium – “who runs which nodes”

The Operational consortium is where technology and institutions meet daily. Its members are the organisations that actually run nodes and services:

- Fabric peers for each major institution (ministries, departments, possibly large municipalities or agencies),
- IPFS nodes at national, regional, and departmental levels,
- Indy validators and agents,
- application gateways, monitoring, and MPC key services.

Each participating organisation signs a consortium operating agreement that defines:

- which nodes it runs (peer, IPFS, Indy, gateway, MPC share holder, etc.),
- its service level commitments (uptime, maintenance windows, incident response),
- its duties (logging, backup, configuration management, participation in upgrades),



- and the procedures for joining, upgrading, or leaving the consortium.

Technically, each operator has:

- a set of organisation DIDs and Fabric MSP identities,
- MPC shares if it participates in key management,
- and specific roles in chaincode policies (e.g. "at least one departmental peer must endorse archives in domain X").

Operations are coordinated through:

- regular consortium meetings (weekly/bi-weekly),
- a shared change management process (change tickets, approvals, test environments, rollback plans),
- and a formal onboarding process for new organisations and nodes:
 - security baseline assessment,
 - provisioning of certificates, DIDs, and MPC shares,
 - connection to monitoring and incident channels,
 - controlled deployment of peers, IPFS nodes or agents.

POINT OF ATTENTION: Shared decision-making and vetoes

A central idea of this consortium model is that no single actor can unilaterally do dangerous things. For example:

- Changing lifecycle rules for a class of archives may require:
 - a proposal signed by a line ministry,
 - approval signatures from the national archives authority and the data protection authority,
 - and a Fabric transaction endorsed by peers from at least N different organisations.
- Executing a key destruction operation via MPC may require:
 - explicit approvals from the Technical steering committee and Security board (as Fabric transactions),
 - participation of a threshold of MPC key holders from different institutions.

In other words, governance rules are encoded in the technology: Fabric endorsement policies, Indy credential issuance policies, and MPC approval thresholds are concrete expressions of institutional balances of power.

5.2 Roles and responsibilities

To keep the system understandable and governable, each institution plays clear, named roles. The idea is that, for any important decision or incident, it is obvious who proposes, who validates, who operates, and who audits.

Below, we structure roles along two axes:

- Governance roles – who decides the rules.
- Operational roles – who runs the nodes and day-to-day services.

5.2.1 Governance roles

These roles correspond to the governance layers described in 5.1, but in a more concrete way.



Role	Level	Main responsibilities	Typical holder
<i>Political Sponsor</i>	Strategic council	Overall mandate, political support, arbitration of major priorities and disputes	Minister or senior official (Digital, Justice, Interior)
<i>Archival Authority</i>	Strategic + Compliance	Defines archival policies, retention rules, transfer to historical archives	National archives directorate
<i>Data Protection / Privacy Lead</i>	Compliance board	Ensures data protection principles are embedded in design and operations	Data protection authority or State DPO
<i>Cybersecurity Authority</i>	Compliance board	Defines security baseline, validates risk management and incident response	National cybersecurity agency
<i>Domain Owner</i>	Strategic + Technical	Owns requirements and policies for a given domain (justice, tax, health, civil status, etc.)	Line ministry or regulator
<i>Technical Architect (Core)</i>	Technical committee	Defines reference architecture, validates changes to Fabric/Indy/IPFS design	National IT agency / central digital team
<i>Domain Architect</i>	Technical committee	Adapts the core architecture to one domain's specific needs	IT architect from each domain ministry
<i>Legal / Evidence Advisor</i>	Compliance board	Ensures evidentiary value, signature/timestamp practices, and alignment with procedural law	Justice ministry, courts, legal experts

These governance roles do not operate servers; they set rules that are then encoded as:

- chaincode policies on Fabric,
- credential schemas and issuer rights on Indy,
- MPC approval thresholds and procedures for key operations,
- security baselines and configuration standards.

5.2.2 Operational and technical roles

These roles exist inside ministries, agencies, departments, and sometimes larger municipalities. They translate high-level rules into day-to-day operations.

Role	Scope	Responsibilities
<i>Node Operator – Fabric Peer</i>	Organisation / Department	Runs and maintains Fabric peers (upgrades, monitoring, backups), applies security baseline, participates in endorsement
<i>Node Operator – IPFS</i>	Organisation / Region/Dept	Runs IPFS nodes, manages pinning policies, monitors storage, ensures replication and restore tests
<i>Node Operator – Indy Validator / Agent</i>	Organisation / National/Dept	Runs Indy validator nodes or edge agents, manages connectivity, keeps software up to date
<i>Archiving Service Owner</i>	Organisation / Domain	Owns the archiving APIs and backend services, coordinates releases, handles integration with business applications
<i>Business Application Owner</i>	Domain application	Integrates archiving functions into line-of-business systems (justice, tax, civil status, health, etc.)
<i>Security Officer (Local)</i>	Organisation / Department	Ensures local compliance with security baseline, handles incidents, coordinates with national cybersecurity authority



<i>Data Protection Officer (Local)</i>	Organisation	Reviews new uses of the archive, checks privacy impact, supports DPIA and data subject requests
<i>SSI Credential Issuer Admin</i>	Organisation / HR / Authority	Manages issuance and revocation of staff credentials (judges, officers, clerks, etc.) under defined schemas and rules
<i>MPC Key Custodian</i>	Institution holding a key share	Participates in multi-party key operations (derive, rotate, destroy) according to thresholds and policies
<i>Audit & Reporting Analyst</i>	Cross-cutting	Uses ledger events and logs to produce compliance reports, usage statistics, and supports investigations/audits

In many institutions, one person may hold several of these roles, but they are conceptually distinct. For example:

- A departmental IT team may be both *Fabric peer operator* and *IPFS node operator*.
- A ministry's CISO may be both *local security officer* and member of the central Security & compliance board.

5.2.3 RACI-style view for key activities

For clarity, the table below summarises who is Responsible (R), Accountable (A), Consulted (C), and Informed (I) for a few critical activities. (Exact mapping can be adapted per country.)

Activity	Strategic Council	Tech Steering	Security & Compliance	Domain Owner	Node Operators	MPC Custodians	DPO (central/local)
<i>Define new archive class & retention policy</i>	A	C	C	R	I	I	C
<i>Change lifecycle rules (chaincode update)</i>	A	R	C	R	I	I	C
<i>Onboard new ministry/agency into consortium</i>	A	R	C	R	R	C	I
<i>Run day-to-day Fabric/IPFS/Indy operations</i>	I	C	C	I	R	I	I
<i>Execute key rotation/destruction via MPC</i>	I	C	A	C	I	R	I
<i>Major security incident handling</i>	I	C	A	C	R	C	C
<i>Approve design for high-risk new use case</i>	A	R	A	R	I	C	C

This RACI view ensures that:



- No critical action (e.g. changing lifecycle rules, destroying keys) is done by a single actor without oversight.
- Operational teams know who to call for decisions, and decision-makers understand which teams they depend on.

5.2.4 Interaction patterns (day-to-day life of the platform)

In everyday operation:

- Business teams (justice, tax, civil status...) talk primarily to the Business Application Owner and Archiving Service Owner to evolve their workflows.
- Those owners, together with Domain Architects, translate requirements into changes on Fabric/Indy/IPFS, reviewed by the Technical Steering Committee.
- Node Operators implement and deploy these changes following standard change management, under the supervision of local Security Officers.
- MPC Custodians and DPOs step in whenever lifecycle, destruction, or particularly sensitive data is involved.
- The Security & Compliance board reviews incidents, audit findings, and high-risk changes periodically; the Strategic council intervenes only when strategic choices or conflicts arise.

This division of roles and responsibilities guarantees that the archiving platform is not just technically sound, but also organisationally credible: each institution knows its responsibilities, and external auditors can see clearly who does what, with what authority, and under which controls.

5.3 Admission and onboarding of administrations

Bringing a new ministry, agency, department or (large) municipality into the archiving consortium is not just a technical plug-in. It's a controlled process that touches legal mandates, governance, security posture, and operational capacity. The goal is to make onboarding repeatable and predictable, so that the platform can grow without losing control.

We can structure onboarding into five main phases:

1. Eligibility & scoping
2. Legal & governance alignment
3. Technical readiness & security baseline
4. Node provisioning & integration
5. Validation, go-live & continuous compliance

A high-level view:

Phase	Main question	Key outputs
1	Should this entity join, and how?	Scope, roles, priority archives
2	Under which rules?	Signed agreements, governance representation
3	Is it technically and securely ready?	Security baseline check, capacity plan
4	How does it connect?	Nodes, credentials, integration completed
5	Can we trust it in production?	Go-live decision, monitoring, audit hooks active

Onboarding a New Member to Archiving Consortium

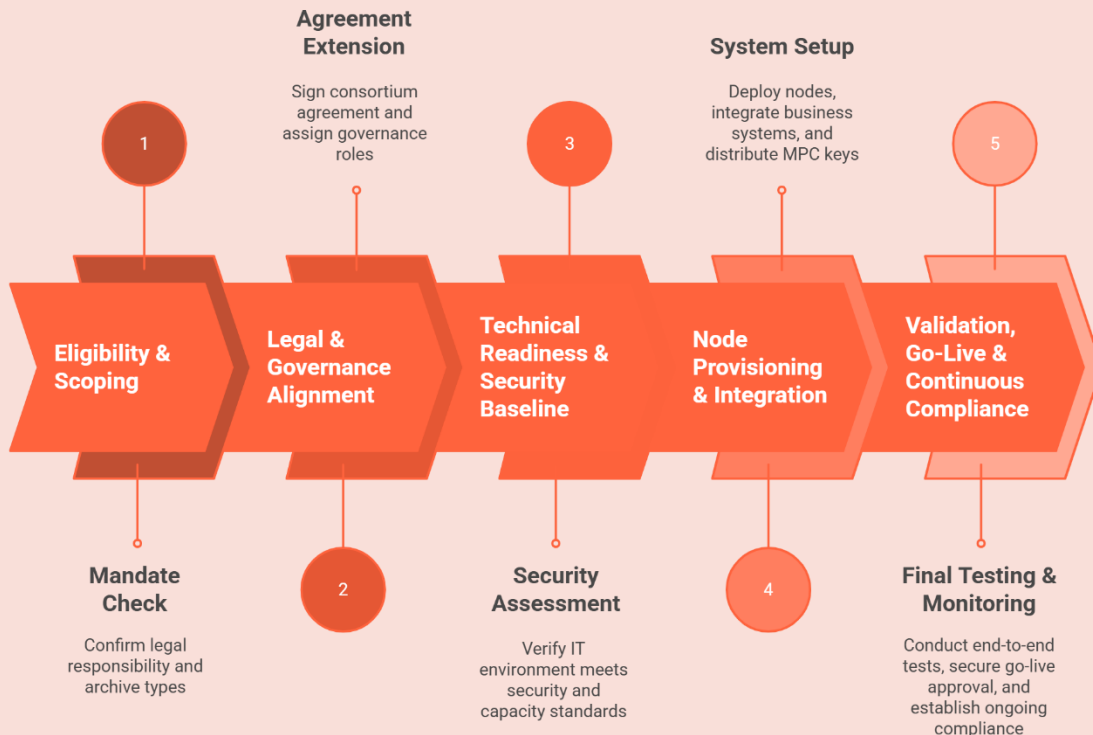


Figure 4: Structure onboarding phases

5.3.1 Phase 1 – Eligibility and scoping

Objective: clarify *why* and *how* a given administration should join.

Key activities:

- Mandate check
 - Confirm the institution’s legal responsibility for archives (e.g. sector law, archival law, internal regulations).
 - Classify the types of archives concerned (justice, taxation, civil status, health, education, municipal records, etc.).
- Scope definition
 - Decide which archive classes are in scope for the first wave (e.g. only new archives, or also historical backfiles).
 - Define the intended role:
 - full consortium member (running Fabric/IPFS/Indy nodes),
 - or edge participant (only gateways and wallets, no ledger/storage nodes).
- Priority and roadmap fit
 - Align with the overall deployment roadmap (e.g. “wave 2: all departments in justice”, “wave 3: selected municipalities”).
 - Rough sizing (volume of documents, users, latency needs) to feed technical planning.

Output: a short onboarding dossier summarising mandate, scope, expected role, and priority.



5.3.2 Phase 2 – Legal and governance alignment

Objective: make sure the new member is integrated into the rules of the game, not just the network.

Key activities:

- Consortium agreement extension
 - Add the institution as a signatory to the consortium operating agreement (or an appendix).
 - Specify:
 - its responsibilities (archives, data protection, security),
 - its rights (governance participation, access to logs, reporting),
 - applicable sanctions or remedial actions in case of breach.
- Governance roles
 - Decide whether the institution obtains:
 - a seat (or observer status) in the Operational consortium,
 - eventual participation in Technical steering or Security & compliance if it's a large or critical player.
- Legal & compliance mapping
 - Map the institution's own legal constraints (sector laws, professional secrecy) to the policy framework of the platform.
 - Identify any specific clauses needed (e.g. stricter access audit for sensitive sectors).

Output: signed legal documents, assigned governance roles, and initial compliance mapping.

5.3.3 Phase 3 – Technical readiness and security baseline

Objective: verify that the administration's IT environment can host nodes or gateways without weakening the platform.

Key activities:

- Security baseline assessment
 - Check that the institution meets minimum controls (patching, hardening, backup, incident response, access management).
 - If gaps exist, define a remediation plan and make onboarding conditional on its completion.
- Network and connectivity design
 - Design the connectivity to national/regional data centres: VPN, government backbone, segmentation.
 - Define where local components sit: DMZ, application zone, infrastructure zone.
- Capacity and resilience planning
 - Estimate required resources (CPU, RAM, storage, bandwidth) for:
 - Fabric peer(s),
 - IPFS node(s),
 - Indy agent(s),
 - gateway services.
 - Decide on high-availability (HA) setup where relevant.
- Operational role assignment
 - Nominate local Node Operators, Security Officer, local DPO, and (if applicable) MPC key custodian.
 - Ensure they are trained on procedures, tooling, and escalation paths.



Output: a technical readiness report with “green/yellow/red” status and agreed remediation actions.

5.3.4 Phase 4 – Node provisioning and integration

Objective: actually connect the administration to the platform and plug in its business systems.

Key activities:

- Identity and credential setup
 - Issue organisation DIDs and Fabric MSP identities.
 - Configure SSI credential issuance flows (e.g. staff roles: civil status officer, judge, tax officer).
- Node provisioning (for full members)
 - Deploy and configure:
 - Fabric peer(s) for relevant channels,
 - IPFS node(s) with pinning policies (local vs central replication),
 - Indy edge agent(s) for local credential/verification flows.
 - Connect nodes to central monitoring and logging.
- MPC key share distribution (if applicable)
 - For institutions participating in key management:
 - generate/distribute MPC shares according to threshold policies,
 - register them in the MPC service, tested with non-production keys first.
- Business integration
 - Integrate key line-of-business systems with the archiving API:
 - map business events to ArchiveDocument / GetDocument calls,
 - configure initial metadata and policy templates per document type,
 - implement front-end features (e.g. “send to archive”, “view archived file”).

Output: nodes up and connected, applications integrated in a controlled test environment.

5.3.5 Phase 5 – Validation, go-live and continuous compliance

Objective: ensure the new member behaves correctly before and after going live.

Key activities:

- End-to-end functional testing
 - Test typical flows with real users: archiving, consultation, legal hold, destruction requests.
 - Verify that policies behave as intended (e.g. municipal officer cannot see justice archives, etc.).
- Security and performance tests
 - Run targeted penetration tests or vulnerability scans (especially on exposed gateways).
 - Measure end-to-end latency and throughput for typical scenarios; adjust pinning and routing if needed.
- Go-live decision
 - Prepare a short go-live dossier summarising:
 - readiness status,
 - residual risks,
 - mitigation measures.
 - Get formal approval from:
 - Domain Owner,



- Technical steering representative,
- Security & compliance representative.
- Continuous compliance hooks
 - Ensure that logs from the new member's nodes feed the central SIEM.
 - Schedule periodic reviews (e.g. annual) of security posture, credential issuance, and use of the archive.
 - Define indicators: number of archives created, access patterns, error/incident rate.

Output: formal onboarding completion, with the institution fully part of the consortium and monitored like any existing member.

With this onboarding model, the platform is not just technically scalable, but also institutionally scalable: each new administration is integrated through a controlled process that preserves security, legal alignment, and the overall quality of the archiving system.

5.4 System evolution and maintenance

A state-scale archive must live for decades. That means the platform must be easy to change without breaking trust. Evolutions (new use cases, new policies, new cryptography) and day-to-day maintenance (patches, upgrades, configuration changes) are managed through a deliberate, repeatable process rather than ad-hoc interventions.

5.4.1 Change management lifecycle

Every change, from a small configuration tweak to a new chaincode version, follows the same high-level lifecycle:

Step	Purpose	Main actors
<i>Proposal</i>	Describe what needs to change and why	Domain Owner, Technical Architect
<i>Impact analysis</i>	Assess security, legal, performance implications	Tech Steering, Security & Compliance, DPO
<i>Design & implementation</i>	Produce code/config, tests, rollback plan	Dev teams, Node Operators
<i>Review & approval</i>	Validate design against policies and standards	Tech Steering, Security & Compliance, Domain Owner
<i>Staged deployment</i>	Deploy to test/pre-prod, then pilot, then full prod	Node Operators, Archiving Service Owners
<i>Post-deployment review</i>	Check behaviour, metrics, incidents, user feedback	Tech Steering, Ops, Domain Owner

This lifecycle is encoded in tools and in the platform itself:

- Fabric:
 - Chaincode upgrades require a formal approval process (chaincode definition update) endorsed by several organisations.
 - Network configuration changes (e.g. adding peers, changing policies) are also channel configuration transactions with multi-party signatures.
- Indy:



- New schemas, credential definitions or revocation registries must be registered by authorised issuers and can be reviewed by the Technical steering committee before use.
- MPC key management:
 - Cryptographic policy changes (e.g. key rotation frequency, algorithm changes) are expressed as MPC policy updates and require threshold approval from key custodians and security authorities.

5.4.2 Release and upgrade strategy (IPFS, Fabric, Indy, services)

To avoid “big bang” upgrades, components follow a rolling, decoupled release strategy:

- Archiving services and APIs
 - Versioned APIs (e.g. /v1/archiveDocument, /v2/archiveDocument) allow business applications to migrate gradually.
 - Backwards compatibility is maintained as much as possible; deprecations are announced with clear timelines.
- Fabric
 - Peers and orderers are upgraded one subset at a time to avoid downtime, following vendor/community LTS cycles.
 - Chaincodes are versioned explicitly (e.g. archive_cc_v1, archive_cc_v2), with a migration plan for state if needed.
 - Endorsement policies can temporarily accept both old and new chaincode versions, easing migration.
- IPFS
 - Upgrades are performed node by node; adding new nodes with the new version before decommissioning old ones supports horizontal migration.
 - New features (e.g. new CID formats, new pinning strategies) are introduced with opt-in configuration flags.
- Indy and SSI components
 - Validator nodes are upgraded in waves, ensuring consensus is never lost.
 - Wallets and agents are upgraded with attention to backwards proof compatibility (old credentials can still be proved against new proof formats where possible).
- Infrastructure and OS
 - Underlying OS and container runtimes are patched through standard infrastructure-as-code pipelines, with blue/green or canary deployments where feasible.

The guiding principle: no single change should require a full system stop, and rollback paths must exist for each release.

5.4.3 Algorithm agility and long-term cryptography

Over the lifetime of the archive, cryptographic algorithms will need to change (deprecation, vulnerabilities, post-quantum).

The architecture anticipates this by:

- Abstracting algorithms in metadata
 - Archive records always store which algorithms and parameters were used:
 - hash function (for CIDs and additional hashes),
 - encryption algorithm and key size,
 - signature scheme and curve,
 - proof system for ZKPs.
- Supporting multiple algorithms in parallel



- Chaincodes and services are designed to accept and verify multiple algorithm families at the same time (e.g. SHA-256 and a post-quantum hash; ECDSA and a PQ signature).
- New archives can be created with new algorithms while old ones remain verifiable.
- Planned re-protection and migration
 - For highly sensitive archives, a policy may require periodic re-encryption or re-signing when algorithms are deprecated.
 - This is done as a controlled batch process:
 - content is re-encrypted with new keys,
 - new CIDs and hashes are associated to the Archive,
 - Fabric records the link between old and new representations,
 - MPC+ZKP proves that re-protection was done under approved procedures.
- Post-quantum roadmap
 - The Technical steering committee maintains a roadmap for post-quantum readiness:
 - evaluation of PQ algorithms,
 - test deployments on non-critical channels,
 - eventual switch for new archives while keeping mixed environments verifiable.

This ensures cryptographic continuity: an archive created today can still be trusted and verifiable thirty years from now.

5.4.4 Operational maintenance and health management

Day-to-day maintenance focuses on keeping the system healthy, predictable and auditable:

- Monitoring and SLOs
 - Clear Service Level Objectives for key operations (archive, consult, search, lifecycle evaluation).
 - Dashboards track latency, error rates, peer health, IPFS replication lag, proof verification times, MPC queue lengths.
 - SLO breaches trigger investigation and, if needed, infrastructure scaling or configuration tuning.
- Patch and vulnerability management
 - Regular vulnerability scans for nodes, containers and code dependencies.
 - Critical patches for Fabric, IPFS, Indy, and cryptographic libraries are applied under accelerated change procedures, with risk-based prioritisation.
- Capacity and data growth management
 - Periodic review of storage growth, metadata size on Fabric, and Indy ledger size.
 - Use of pruning, channel partitioning, and off-chain indexing to keep ledgers manageable without losing auditability.
- Documentation and knowledge management
 - Living documentation for:
 - runbooks,
 - change procedures,
 - incident playbooks,
 - cryptographic policies.
 - Training plans for new operators, auditors, and architects, so that the system does not become dependent on a few individuals.
- Continuous improvement loop
 - Regular post-mortems for incidents and major changes feed into improved policies, better automation, and refined governance rules.



- Lessons learned from one domain (e.g. justice) are shared across others (e.g. taxation, health).

5.5 Economic model and cost sharing

A state-scale digital archiving platform is a public infrastructure, not a one-off IT project. Its economic model must therefore cover:

- Initial investment: design, build, first domain onboarding, migration of priority archives.
- Recurrent costs: hosting, operations, security, governance, and support.
- Long-term evolution: new use cases, cryptography updates, legal/regulatory changes.

Because multiple ministries, agencies, departments, and municipalities use and benefit from the platform, costs must be shared in a transparent and predictable way, with mechanisms that avoid both free-riding and overly complex billing.

For concreteness, the figures below assume a medium-sized West African country (5–10 million inhabitants, 7 departments, 26 municipalities, 4 initial domains: justice, civil status, taxation, education; ~1 PB hot/warm + 4 PB cold archives after 5 years). All amounts are indicative 2025 USD.

5.5.1 Main cost components (with indicative ranges)

A. One-off / initial (years 0–2)

Cost category	What it covers	Typical range (USD)	Notes
<i>Core design & build</i>	Global architecture, core IPFS/Fabric/Indy setup, chaincodes, SSI schemas, security, base APIs	2–4 M	Funded centrally as “nationwide infrastructure”
<i>Domain onboarding (per domain)</i>	Domain modelling (justice, tax, civil status, etc.), integration with 2–3 key apps	0.5–1.2 M / domain	Often co-funded by line ministries
<i>Data migration (legacy archives)</i>	Cleaning, transforming, loading prior archives	\$500–1,500 per TB	Strongly depends on data quality and preparation

Example for 4 domains + 400 TB migrated in the first wave:

- Core design & build: ≈ \$3 M
- 4 domains × ≈ \$0.8 M: ≈ \$3.2 M
- 400 TB × ≈ \$1,000/TB: ≈ \$0.4 M

→ Total initial investment over 2–3 years: ≈ \$6.5–7.5 M

B. Recurring / annual (steady state)

Cost category	What it covers	Typical annual range (USD)
---------------	----------------	----------------------------



<i>Infrastructure – compute & network</i>	Servers/VMs for IPFS, Fabric, Indy, archiving services, gateways, monitoring	\$250–450 k
<i>Infrastructure – storage</i>	≈1 PB hot/warm + 4 PB cold (including replication, backup, overhead)	\$350–700 k
<i>Core operations team</i>	8–12 FTE (SREs, platform engineers, security, architects, support)	\$800 k–1.2 M
<i>Security & compliance</i>	SIEM, audits, penetration tests, DPIAs, certifications, training	\$150–300 k
<i>Governance & consortium running</i>	Committees, documentation, onboarding support, training, workshops	\$80–150 k
<i>Evolution & innovation</i>	New use cases, UX, post-quantum pilots, new domains, analytics, cross-border features	\$250–500 k

→ Rough annual run cost: ≈ \$1.9–3.3 M/year, shared by all ministries, 7 departments, and 26 municipalities.

5.5.2 Cost allocation principles (core vs domain vs local)

Rather than precise per-GB or per-transaction billing (which discourages usage), the platform uses simple, rule-based allocation:

- Core layer (national)
 - Includes: Fabric ordering service, core IPFS backbone, Indy validators, MPC+ZKP key services, central security tooling, and consortium governance.
 - Funding: central budget (digital transformation / national IT agency / finance ministry).
- Domain layer (ministries)
 - Includes: domain chaincodes, schemas, archiving services for justice, taxation, health, education, etc., plus domain-specific integrations.
 - Funding: each line ministry co-funds its domain costs, proportional to usage and legal mandate.
- Local layer (departments / municipalities)
 - Includes: departmental IPFS and Fabric peers, local gateways, integration with local systems, user training and support.
 - Funding: departmental and municipal budgets, often with central subsidies for smaller entities.

A workable split for a \$2.3 M/year budget (mid-range) is:

- Core / national: ≈ 60% ⇒ ≈ \$1.4 M/year
- Domains / ministries: ≈ 30% ⇒ ≈ \$690 k/year (e.g. Justice ≈ \$250 k, Interior/civil status ≈ \$200 k, Finance/tax ≈ \$150 k, Education ≈ \$90 k)
- Local (7 departments + 26 municipalities): ≈ 10% ⇒ ≈ \$230 k/year
 - 7 departments as node operators: 7 × ≈ \$20 k ≈ \$140 k/year
 - 26 municipalities as edge participants: 26 × ≈ \$3.5 k ≈ \$90 k/year (with possible full subsidy for the smallest ones)

In perception terms:

- A big ministry pays roughly the cost of a modest IT project per year (~\$200–250 k).
- A department pays about the cost of 0.5–1 engineer per year (~\$20 k co-funding).



- A municipality contributes a symbolic but real amount (a few thousand dollars/year), or is fully subsidised.

5.5.3 Five-year TCO and comparison with the “silo status quo”

On a 5-year horizon for the same reference scenario:

- Initial build + first waves: \approx \$7 M (one-off)
- 5 years of run (mid-range, \approx \$2.4 M/year): \approx \$12 M

→ Total 5-year TCO: \approx \$19 M for a country-wide, multi-domain archival platform.

By contrast, in a typical “silo” situation:

- 4 major ministries each operate their own ECM/archiving stack:
 - \$800 k–1.5 M/year per ministry (infrastructure, licences, ops)
 - 4 ministries \Rightarrow \$3.2–6 M/year just for central silos
 - plus departmental and agency mini-solutions, backups, and ad-hoc “scan & archive” projects
- Over 5 years, total spending easily exceeds \$20–25 M, with:
 - weaker security and resilience,
 - duplicated vendor contracts and skills,
 - no shared investment in cryptography and compliance,
 - limited interoperability and evidence consistency.

The shared IPFS + Hyperledger platform sits in the same cost bracket but:

- replaces several legacy stacks,
- is designed for long-term legal and cryptographic robustness,
- and offers a single sovereign infrastructure that any ministry, department or municipality can join at predictable, modest marginal cost.

5.5.4 ROI in a West African context

In a medium-sized West African country (GDP \$15–50 B, State budget \approx 20–35% of GDP, IT \approx 1–3% of the budget), allocating \$2–3 M/year to a shared archiving and trust infrastructure represents:

- \approx 0.1–0.3% of the central State budget,
- less than the budget of one large sectoral project (e.g. a new tax system),
- but with benefits that cut across justice, taxation, civil registry, education, health, and more.

Using conservative assumptions, once core sectors are live, annual quantifiable benefits typically fall in the range \$2.8–7.4 M/year, broken down as follows:

- Decommissioning legacy silos
 - Fragmented archiving systems across 4–6 big ministries \approx \$2 M/year in total.
 - Replacing 50–70% of that with the shared platform \Rightarrow \$1–1.4 M/year saving.
- Improved tax and customs performance
 - Tax revenue often \$2–4 B/year.
 - A 0.5–1% improvement in effective collection (better evidence, fewer “lost” cases) \Rightarrow \$10–40 M/year in extra revenue.
 - If only 10–20% of that improvement is attributed to the archiving backbone \Rightarrow \$1–8 M/year.
- Justice efficiency



- Example: 20,000 cases/year, average cost \$300–500.
- 5–10% reduction in wasted time (adjournments, missing files) ⇒ \$300 k–1 M/year in efficiency gains.
- Civil registry and social programme targeting
 - Social/education spending often \$200–500 M/year.
 - 0.5–1% better targeting thanks to reliable civil status archives ⇒ \$1–5 M/year; a fraction attributable to the platform ⇒ \$0.5–2 M/year.
- Risk reduction and crisis resilience
 - Avoiding catastrophic loss of court, tax, or civil-status archives can prevent tens of millions of dollars in downstream damage over a decade.

Against an annual run cost of \$2–2.5 M/year, this gives a steady-state ROI roughly in the range:

- Low scenario: $2.8 / 2.5 \approx 1.1\times$
- Mid scenario: $4.5 / 2.3 \approx 2\times$
- High scenario: $7.4 / 2.3 \approx 3.2\times$

This is before counting softer but real benefits:

- regional and cross-border trust (e.g. ECOWAS-wide diploma or identity verification),
- improved investor confidence from stronger rule of law and records,
- synergies with digital ID, payments, and trade platforms.

5.5.5 Aligning incentives and service-level benefits

The economic model should drive the behaviours the State wants:

- Encourage consolidation, discourage silos
 - Ministries are not penalised for migrating from legacy systems; central programmes can co-fund migration.
 - Savings from decommissioned solutions are progressively reallocated to sustain the shared platform.
- Reward good data hygiene and standardisation
 - Programmes that clean data and adopt common schemas receive targeted funding or lower marginal integration costs, reflecting the lower long-term TCO of well-structured archives.
- Stabilise multi-year funding
 - Because archives and cryptography outlive political cycles, the platform budget is planned on a multi-year horizon (e.g. 5 years), not year-to-year.
 - This supports serious investments (post-quantum crypto, infrastructure renewal) without constant renegotiation.
- Ensure transparency
 - High-level cost and benefit indicators (number of archives preserved, legacy systems retired, major incidents avoided, revenue gains, justice backlog reductions) are published to auditors and, where appropriate, the public.
 - This frames the platform as a long-term public good, not a niche IT expense.

Finally, the same economic backbone supports multiple high-value services particularly relevant in West Africa:

- National digital case file for justice (trusted, tamper-evident case records).
- Secure e-Civil Registry (birth/marriage/death archives reused by ID, tax, social security, banks).
- Trusted tax & customs archive (evidence for audits and trade facilitation).
- Diploma and professional certificate verification (verifiable credentials for education and professions).



- Health & social protection evidence store (secure, audited records underpinning targeted programmes).

All these use cases share the same IPFS + Hyperledger + SSI infrastructure and thus amortise the investment over many sectors, reinforcing the economic case for a single, sovereign, state-scale digital archiving platform.

6. Risks, limitations and success factors

The proposed architecture is technically ambitious but its real difficulty is institutional, not purely technical. Deploying a state-scale digital archiving platform based on IPFS and Hyperledger requires:

- long-term political and financial commitment,
- coordination across ministries and levels of government,
- legal and regulatory adaptations,
- and sustained capacity-building for civil servants and technical teams.

If these conditions are not met, the platform risks becoming just another pilot or a siloed “blockchain-for-show” project disconnected from core public services. This section outlines the main risks and limitations and identifies the critical success factors needed for real impact.

6.1 Strategic preconditions and leadership commitment

The first and most important risk is insufficient high-level support. A sovereign, cross-cutting infrastructure of this kind touches justice, finance, interior, civil service, archives, cybersecurity, and digital transformation. Without clear backing from “top heads” (ministers, cabinet-level leadership, sometimes the head of state or government), it will be extremely hard to align priorities, budgets, and timelines.

Several strategic preconditions must therefore be in place:

A. Clear political mandate and ownership

The platform needs a formal mandate, expressed in policy documents or legal texts, that states:

- why the State is investing in a shared digital archiving infrastructure,
- which institutions are responsible for leading it,
- and how it links to national digital strategies, justice reforms, tax reforms, civil registry reforms, and data protection policies.

Ideally, a lead ministry or national digital agency is explicitly tasked with steering the project, with a strong political sponsor who can arbitrate conflicts and secure budget over multiple years.

B. Cross-ministerial governance with real authority

Because archives and records cut across sectors, no single ministry can impose the platform on others without resistance. The consortium governance model described in section 5 must be backed by:

- formal participation of key ministries (justice, interior, finance, education, health, national archives, cybersecurity agency, data protection authority),



- decision rights clearly allocated to the Strategic council, Technical steering committee, and Security & compliance board,
- mechanisms for joint approvals (e.g. changes to lifecycle rules, key destruction) so that critical decisions are shared and defensible.

Without this, each ministry may continue to invest in its own solutions, and the shared backbone will remain under-used.

C. Stable multi-year funding and prioritisation

A state-scale archive cannot be built on project-by-project, one-year budgets. To be credible, the State must:

- commit a multi-year funding envelope (e.g. 5-year horizon) covering both build and run phases,
- make the platform a priority dependency for major reforms (e.g. digital justice, e-tax, civil registration, digital ID),
- and ensure that development partners and donors understand that sectoral projects should build on, not bypass, the shared infrastructure.

If funding is uncertain or fragmented, ministries will hesitate to migrate critical data and will keep legacy systems "just in case", reducing ROI and increasing complexity.

D. Legal and regulatory readiness

Even if the architecture is designed to respect data protection and archival law, some adjustments are usually needed:

- explicit recognition of digital archives and electronic signatures/timestamps as having evidentiary value,
- provisions for cryptographic erasure and lifecycle rules that balance retention and rights (e.g. right to erasure, sealing),
- mandates for public bodies to use the shared archive for specific classes of documents, at least for new data.

Top-level political support is essential to drive these legal changes through parliaments or regulatory authorities and to align sectoral laws (justice, tax, health, education) with the new infrastructure.

E. Institutional change management and capacity building

A digital archiving backbone changes how administrations work with documents and evidence. It implies:

- new responsibilities for IT teams (running peers, IPFS nodes, SSI agents),
- new routines for clerks, officers, judges, and archivists (digital-first workflows, proof-based access),
- and new skills in cybersecurity, cryptography, and data protection.

Leadership must therefore commit to:

- change management programmes (training, communication, support),
- embedding the platform into standard operating procedures and job descriptions,



- and incentivising administrations to actually use the platform instead of reverting to old practices.

Without this human and organisational investment, the technology risks remaining underutilised or misused.

F. Alignment with development partners and regional agendas

In many West African contexts, major digital systems are co-financed by donors and development banks. Leadership needs to:

- position the archiving platform as a shared foundation for these sectoral projects,
- encourage partners to reuse it for evidence, audit trails, and data retention instead of funding isolated, proprietary solutions,
- and align the platform with regional initiatives (e.g. ECOWAS/UEMOA digital ID, diplomas, trade facilitation), which increases its strategic value.

Top-level support is crucial here: without clear messaging and negotiation from ministers and senior officials, donors may continue to fund siloed, short-lived systems.

G. Summary: technology as an instrument of political will

In practice, the most important success factor is that senior leadership sees the platform as a strategic instrument:

- to strengthen rule of law and trust in public institutions,
- to protect the State's legal memory and digital sovereignty,
- and to make major reforms (justice, tax, social protection, civil registry) sustainable and auditable.

If this vision is shared and translated into mandates, governance structures, legal changes, budgets, and capacity-building, the technical challenges of IPFS, Hyperledger Fabric, and Indy are manageable. If not, the architecture described in this paper will remain a theoretical reference rather than a living, national infrastructure.

6.2 Technical and operational risks and constraints

Even with strong political backing, the proposed architecture raises non-trivial technical and operational risks. These do not invalidate the approach, but they must be recognised early and actively managed.

6.2.1 Complexity of a multi-layer, multi-technology stack

The platform deliberately combines several advanced components: IPFS, Hyperledger Fabric, Hyperledger Indy, SSI wallets/agents, and MPC+ZKP key management. This increases the architectural surface compared to a classic "database + storage + application server" model. Risks include:

- Teams underestimating the learning curve, leading to incorrect configurations or insecure shortcuts.
- Overly complex initial designs that are hard to operate or explain to auditors.
- Dependencies between components (e.g. Fabric relying on identity infrastructure, IPFS relying on network topology) creating failure modes that are not obvious.



Mitigations:

- Start with a minimal viable stack (few channels, simple policies, limited MPC workflows) and grow progressively.
- Invest in reference implementations and templates (for chaincodes, metadata, policies, IPFS configs) that new domains can reuse.
- Use infrastructure-as-code and automated testing to detect configuration drift early.

6.2.2 Operational maturity and 24/7 service reliability

A state-scale archive is not a lab environment; it must support continuous operations, often 24/7, across multiple data centres. Risks include:

- Lack of SRE skills (Site Reliability Engineering) and modern DevOps practices in public IT teams.
- Inadequate monitoring and alerting, leading to late detection of problems (e.g. broken replication, failed nodes, chaincode misbehaviour).
- Difficulty coordinating maintenance windows across ministries and regions.

Mitigations:

- Build a central platform operations team with explicit SRE skills and clear SLAs/SLOs, and gradually train departmental teams.
- Standardise monitoring: one shared observability stack (metrics, logs, traces, dashboards) for all Fabric/IPFS/Indy nodes.
- Adopt phased rollout and blue/green deployments for upgrades, with rollback procedures tested in pre-production.

6.2.3 Network, power and infrastructure constraints

In many countries (including West African contexts), network and power infrastructures are less stable than in wealthy regions:

- Departments and municipalities may experience link instability, low bandwidth, or frequent outages.
- Power cuts can impact local nodes and corrupt poorly configured systems.

Mitigations:

- Design for graceful degradation: when departmental nodes are offline, regional or national nodes can temporarily serve read requests.
- Prioritise bandwidth-efficient designs (caching at regional/department level, careful tuning of IPFS replication and Fabric block sizes).
- Provide reference hardware and configuration profiles tailored to constrained environments (e.g. small servers for departmental IPFS/peers, local UPS requirements).

6.2.4 Performance and user experience expectations

If the system feels slow or unreliable compared to local, ad-hoc solutions, users may resist adoption. Risks include:

- High latency for document consultation if IPFS content is not locally cached or if network routing is suboptimal.
- Perceived complexity (e.g. SSI proofs, multi-factor authentication) leading users to circumvent security controls.



Mitigations:

- Place IPFS and Fabric peers close to heavy users (departments, large courts, tax centres) and use targeted pinning strategies.
- Optimise the critical paths: archiving and consultation APIs should be simple and fast; heavy processes (analytics, batch lifecycle evaluation) can run off-path.
- Invest in good front-end UX integrated into line-of-business applications, so that SSI and proof-based access feel natural, not like an extra barrier.

6.2.5 Skill shortages and dependency on a few experts

IPFS, Fabric, Indy, and MPC+ZKP are still niche skills in most public sectors.

Risks:

- Over-reliance on a small number of internal champions or external consultants.
- Difficulty maintaining the platform if key individuals leave.
- Slow adoption in ministries that lack technically confident staff.

Mitigations:

- Plan a capacity-building programme from the outset: training tracks for architects, developers, operators, archivists, and legal teams.
- Encourage regional collaboration (e.g. between West African countries) to share expertise, patterns, and even common components.
- Use widely adopted, well-documented open-source distributions and frameworks to reduce vendor lock-in and simplify hiring.

6.2.6 Governance encoded in software – rigidity vs. agility

A strength of this architecture is that governance rules (policies, lifecycles, approvals, MPC thresholds) are codified in chaincodes and configurations. The risk is:

- If rules are encoded too rigidly, adapting to legal changes or exceptional situations becomes slow and bureaucratic.
- If rules are encoded too loosely, security and probative value are weakened.

Mitigations:

- Design policy abstraction layers: high-level rules (e.g. "destroy after 30 years unless legal hold") are parameterised rather than hard-coded, so updates do not require low-level code changes.
- Maintain test channels / sandboxes where new policies and workflows can be trialled with realistic data before production deployment.
- Embed legal and compliance experts into the change management process, so that adjustments to code and configuration reflect genuine regulatory needs.

6.2.7 Integration burden and legacy coexistence

In practice, the platform will coexist for years with legacy systems.

Risks:

- Integration is underestimated, leading to partially connected systems and double entry.
- Projects treat the archive as "one more external service" instead of the primary place for long-term records.



Mitigations:

- Make the platform an explicit dependency for major digital projects (justice, tax, civil registry, social protection) in national digital strategies and donor agreements.
- Provide standardised SDKs and adapters for common stacks (Java/.NET back-ends, typical case management systems, document scanners, etc.).
- Define clear sunset plans for legacy archiving components at each ministry, with milestones and shared monitoring.

Taken together, these technical and operational risks are manageable but real. The key is to treat the platform as critical national infrastructure with professional operations, disciplined change management, and continuous capacity-building-not as an experimental “blockchain project” sitting at the margins of government IT.

6.3 Legal, institutional and political risks

Even if the technology is solid and operations are well run, the project can still fail for legal, institutional or political reasons. These risks must be made explicit, especially in contexts where laws are evolving and institutions have overlapping mandates.

6.3.1 Misalignment with existing laws and judicial practice

Courts, archives, and regulators often rely on legal frameworks that were written for paper records and centralised systems. Risks include:

- Laws that do not clearly recognise digital archives, distributed ledgers, or electronic signatures as having the same evidentiary value as paper.
- Procedural codes that assume physical files (e.g. “the original file kept in the registry of court X”) and are silent on shared digital records.
- Judges and lawyers who are unfamiliar or uncomfortable with cryptographic evidence (hashes, CIDs, ZK proofs, ledger logs), leading to inconsistent acceptance in practice.

If these gaps are not addressed, ministries and courts may keep parallel paper or legacy systems “just in case”, undermining the value of the new platform.

Mitigations:

- Early legal gap analysis involving justice, national archives, data protection, and sector regulators.
- Targeted legal reforms that explicitly define digital archives, recognise cryptographic evidence, and set rules for their use.
- Training and guidance documents for courts on how to interpret and weigh evidence produced by the platform.

6.3.2 Mandate conflicts and institutional resistance

Public archives touch many powerful actors: ministries, national archives, supreme audit institutions, courts, security forces. Risks include:

- Perceived loss of control by ministries over “their” archives when a shared platform is introduced.
- Overlapping mandates between national archives, line ministries, and regulators, leading to turf wars over who sets rules or runs nodes.



- Resistance from IT departments or vendors who have invested in existing systems and contracts and may oppose migration.

Such resistance can result in:

- partial adoption (some ministries opt out),
- duplication (ministries keep their own systems and treat the shared platform as a peripheral service),
- or political blockages when critical design or governance decisions are needed.

Mitigations:

- A clear political mandate that defines roles and responsibilities (as in section 5), supported at cabinet level.
- Transparent governance where major institutions have a formal voice (strategic council, technical steering, compliance board).
- Transition plans that recognise existing investments and provide realistic migration paths and incentives, not just obligations.

6.3.3 Data protection authority concerns and public trust

Strong identity, cryptography and cross-sector usage can trigger legitimate concerns from:

- Data protection authorities (DPAs) worried about excessive centralisation of sensitive data, function creep, and mass access.
- Civil society and media, if the platform is perceived as a “surveillance” tool or a way to consolidate political control over information.

Risks:

- DPAs may block or restrict key features (cross-domain access, long-term retention, analytics),
- public controversies may damage trust and slow adoption, especially for sensitive domains (health, criminal justice, tax).

Mitigations:

- Involve DPAs from the very beginning so that data protection is designed in (section 4.3) rather than bolted on.
- Make data minimisation and strict, auditable access control central to the architecture (SSI, ZKP, policies on Fabric).
- Communicate clearly that the archive is not a central data warehouse, but a controlled memory layer with strong safeguards and traceability.

6.3.4 Donor/partner fragmentation and political cycles

In many countries, digital projects are co-financed by donors, banks and regional programmes, each with its own timelines and priorities. Risks:

- Sectoral projects (tax, justice, ID, health) may be designed and funded without using the shared archive, creating new silos.
- Changes in government or ministerial leadership may reset priorities, leaving the platform half-implemented, frozen in pilot mode.

Mitigations:



- Position the platform as a foundational “public good” in national digital strategies and in discussions with partners.
- Include the use of the shared archive as an explicit requirement or strong recommendation in donor project documents and funding agreements.
- Secure multi-year commitments that survive individual political cycles, e.g. by anchoring the platform in mid-term national plans or regional frameworks.

In short, the main non-technical risk is that the platform remains politically and legally under-specified: no-one is clearly forced or incentivised to use it, laws do not fully recognise it, and each institution continues to solve its problems alone. Avoiding this outcome requires deliberate legal work, stable leadership, and inclusive governance.

6.4 Limitations of the approach and realistic scope

Even if successfully implemented, the IPFS + Hyperledger (Fabric + Indy) architecture does not solve everything. Being clear about its limitations helps manage expectations and keeps the project focused.

6.4.1 It does not fix bad source data or weak processes

The platform can guarantee integrity, traceability, and controlled access to what it is given; it cannot guarantee that:

- the original decision was fair or lawful,
- the classification of a document was correct,
- the metadata is accurate or complete,
- or that the document was actually the one that should have been issued.

If ministries continue to produce poorly structured, incomplete or erroneous records, the archive will preserve those errors more securely-nothing more.

Implication: the project must be accompanied by records management improvements and data quality efforts in source systems; otherwise, legal and operational gains will be limited.

6.4.2 It is not a real-time transactional platform for all data

The architecture is optimised for archival records, not for:

- ultra-low latency transactional processing,
- massive streaming workloads (e.g. raw sensor data, clickstreams),
- or heavy analytical queries directly on Fabric or IPFS.

While off-chain indexing and analytics can be added, the platform is not meant to be the “database of everything” for the State. It should hold canonical, legally relevant versions of documents and key metadata, not every transient internal state.

Implication: transactional systems (tax calculation engines, case management workflows, health HIS, etc.) continue to exist, and the archive is integrated as the trusted memory layer, not a replacement for all back-end databases.

6.4.3 It cannot eliminate all political or organisational misuse

The platform makes it technically hard to:



- silently alter records without detection,
- destroy archives without audit trails,
- or give access without leaving traces.

However, it cannot fully prevent:

- misuse of legal powers to mark records as secret or to impose broad legal holds,
- politically motivated over-classification or under-classification,
- selective enforcement (e.g. choosing when to “discover” existing records).

These are governance and culture problems, not technical ones. The platform can provide better evidence and transparency, but it cannot by itself guarantee impartiality or good faith.

Implication: expectations should focus on reducing opportunities for tampering and loss, not on magically solving corruption or politicisation.

6.4.4 Dependence on infrastructure, skills, and vendors

The platform reduces lock-in compared to proprietary archiving solutions, but it still depends on:

- reliable electricity and connectivity, especially at departments and regional data centres;
- the ability to hire and retain people with skills in distributed systems, security, and cryptography;
- a healthy open-source ecosystem and/or trusted vendors for IPFS, Fabric, Indy, and SSI components.

Implication: countries with very constrained infrastructure or severe skill shortages may need simpler initial deployments, more reliance on regional/shared expertise, and a gradual path towards the full architecture.

Cryptographic and technological uncertainty over decades

The design is “algorithm agile”, but there is always residual uncertainty:

- future vulnerabilities in current algorithms,
- the timing and impact of post-quantum cryptography,
- the evolution of standards for SSI, verifiable credentials, and distributed ledgers.

Implication: part of the long-term cost is continuous adaptation. The State must accept that cryptographic migration and platform evolution are recurring tasks, not one-time activities.

Scope: start with the most valuable records, not everything

Finally, trying to “archive everything for everyone” from day one is unrealistic. The architecture is best applied first to:

- archives with high legal and economic value (judgements, tax decisions, civil status, key administrative acts),
- domains where loss or tampering would be catastrophic (justice, tax, civil registry),
- and contexts where cross-institution access and evidentiary value are crucial.

Less critical records, or purely operational documents with low long-term value, may remain in simpler systems or be integrated later.

Implication: a phased scope-starting with a limited set of high-impact document types and gradually expanding-is both technically and institutionally more realistic than a “big bang”.

Recognising these limitations is not a weakness; it is essential to framing the platform as what it really is: a powerful, sovereign, cryptographically robust foundation for the State’s digital

memory—one that must coexist with other systems, be supported by better processes and laws, and be deployed progressively where it brings the greatest value.

7. Implementation roadmap and recommendations

A state-scale archive cannot be deployed “in one shot.” To keep risk under control and build trust, the platform should start small, prove value quickly, and then expand. This section proposes a pragmatic roadmap and a set of first use cases aligned with the architecture and the West African context described earlier.

7.1 Phased rollout roadmap

A realistic path is to move in four main phases, over roughly 4–6 years. Timelines are indicative; the key point is the sequence and the learning loop.

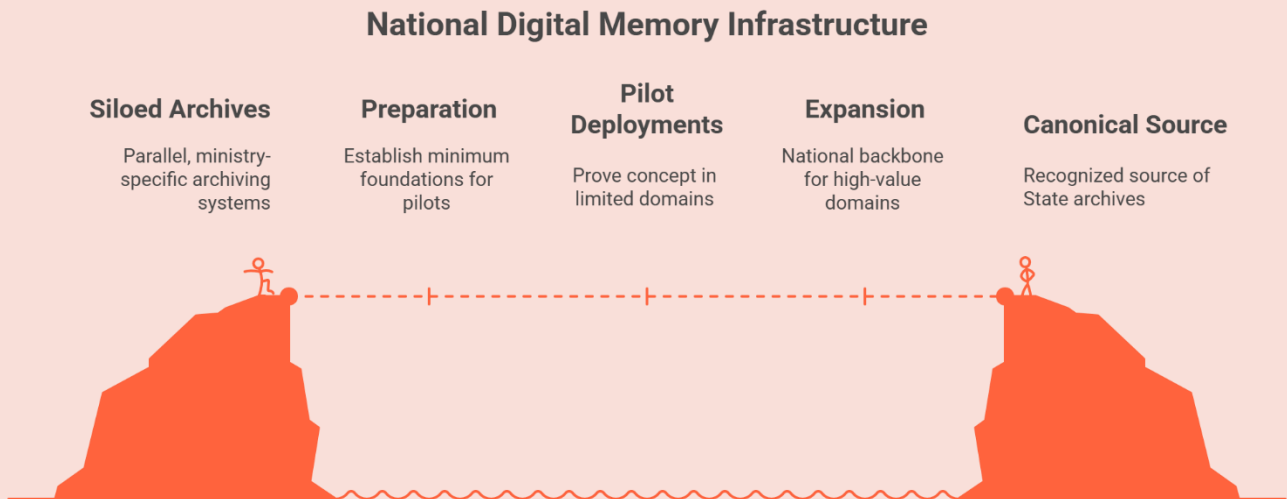


Figure 5: Road for the national memory



Table 1: The roadmap for archiving transformation at state level

Phase	Years	Objective	Scope	Key actions	Success criteria
Phase 0 – Preparation and foundation	Year 0–1	Put in place the minimum foundations so that pilots are credible and auditable.	Foundational (national) <ul style="list-style-type: none"> • Core decision to create a shared platform. • Initial governance bodies (Strategic council, Technical steering, Security & compliance). • First legal and architectural baselines. 	Political & legal preparation <ul style="list-style-type: none"> • Formal decision to create the shared platform. • Nomination of lead institution and initial governance bodies. • First legal gap analysis (archives law, evidence law, data protection, sectoral laws). Core technical setup <ul style="list-style-type: none"> • Design and deploy a minimal backbone: <ul style="list-style-type: none"> – 2 national data centres + 1 regional site, – Fabric ordering service + a few peers, – private IPFS cluster (small but resilient), – initial Indy network (validators + 1–2 agents), – basic MPC+ZKP key service for content keys. Reference architecture & templates <ul style="list-style-type: none"> • Define standard Archive model, metadata templates, initial chaincodes (archive creation, access control, lifecycle skeleton). • Publish SDKs and API specifications for integration. Pilot governance & operations team • Assemble the core platform operations team (8–12 people over time, starting smaller). • Define monitoring, change management, basic incident procedures. 	<ul style="list-style-type: none"> • Running “version 0” of the platform on test networks. • Draft legal and governance documents produced. • Short list of candidate pilot domains and regions agreed.



<p>Phase 1 – Pilot deployments in 1–2 domains and limited geography</p>	<p>Year 1–2</p>	<p>Prove the concept on a limited but visible perimeter, without trying to cover everything.</p>	<p>Domains</p> <ul style="list-style-type: none"> • Civil status (birth, marriage, death certificates). • Justice decisions (high court or one court of appeal). <p>Geography</p> <ul style="list-style-type: none"> • Capital region + 1–2 departments (out of 7). • Small number of municipalities in those departments. 	<ul style="list-style-type: none"> • Integrate 2–3 line-of-business systems per domain (civil status software, court case management, etc.). • Implement end-to-end flows: archive creation, consultation, legal hold, basic lifecycle. • Train local users (civil status officers, judges, clerks, archivists) and iterate UX based on feedback. • Start using SSI credentials for a limited group of staff and for citizen access in controlled scenarios. • Run a first limited data migration (e.g. last 2–5 years of civil status and judgements in pilot areas). 	<ul style="list-style-type: none"> • System is stable and meets agreed SLOs (e.g. open archived document in < 3–5 seconds for typical cases). • Local teams can operate their nodes with support from the central platform team. • First audits and legal reviews confirm evidentiary value and compliance. • At least one real case (court decision, complex civil status situation) is successfully handled using the archive as source of truth.
<p>Phase 2 – Expansion to core domains and all departments</p>	<p>Year 2–4</p>	<p>Move from “pilot” to national backbone for a small number of high-value domains.</p>	<p>Domains</p> <ul style="list-style-type: none"> • Civil status. • Justice decisions. • Taxation & customs (selected processes and documents). • Education diplomas (higher education + key professional schools). <p>Geography</p> <ul style="list-style-type: none"> • All 7 departments (each with ≥ 1 Fabric peer + IPFS node + Indy agent). • 20–26 municipalities as edge participants (gateways, optional local IPFS caching). 	<ul style="list-style-type: none"> • Roll out departmental nodes and connect them via the governmental backbone. • Standardise metadata and policy templates per document type and domain. • Plan and execute progressive migration from ministry-specific archiving systems for targeted documents. • Strengthen MPC+ZKP key management and incident response procedures. • Begin using the platform as the official archive for selected categories of acts (e.g. all new birth certificates, all new appellate decisions, all new high-value tax assessments). 	<ul style="list-style-type: none"> • For in-scope documents, ministries and departments no longer rely on parallel, siloed archives for legal or evidentiary purposes. • Data protection authority and national archives confirm that retention, lifecycle and access rules align with law. • First economic and governance indicators show value: fewer lost files, reduced adjournments, improved auditability, initial tax/registry gains.



Phase 3 – Consolidation, additional domains and regional integration	Year 4–6	Stabilise and scale the platform as a national digital memory infrastructure, and connect to regional initiatives where relevant.	<p>Additional / deeper domains</p> <ul style="list-style-type: none">• Health (carefully scoped data: major medical decisions, eligibility for programmes).• Social protection.• Public finance and procurement.• Regulatory decisions (telecom, energy, competition). <p>Regional / cross-border</p> <ul style="list-style-type: none">• Verification of diplomas or professional licences within ECOWAS/UEMOA.• Support to regional digital ID and trade platforms.	<ul style="list-style-type: none">• Decommission a growing number of legacy archiving systems, freeing budget.• Introduce more advanced features:<ul style="list-style-type: none">– analytics and dashboards over archival metadata (with privacy safeguards),– more sophisticated lifecycle and classification policies,– post-quantum cryptography pilots on selected channels.• Strengthen regional partnerships and mutual recognition of trust frameworks.	<ul style="list-style-type: none">• Platform is recognised in law and practice as the canonical source of State archives for in-scope domains.• Clear net economic benefit: annual savings and gains exceed total run costs by a comfortable margin.• Donors and partners use the platform as the default archival and evidence layer for new digital investments.
--	----------	---	---	--	--

7.2 Suggested first use cases (high value, low regret)

Choosing the right first use cases is critical: they must be politically visible, legally important, but also technically manageable. Below are recommended candidates for Phase 1 and early Phase 2, particularly in a West African context.

#	Use case	Why	Scope (Phase)	Benefits
1	Civil status: birth, marriage and death certificates	<ul style="list-style-type: none"> • Foundation for identity, voting, education, social protection, and banking. • Often fragile today: paper registers in municipalities, partial digitisation, frequent loss or damage. 	Phase 1–2 <ul style="list-style-type: none"> • New registrations in selected municipalities, then all municipalities. • Priority digitisation and archiving of recent years (e.g. last 5–10 years) and high-risk registers (urban centres). 	<ul style="list-style-type: none"> • Reliable, nationally accessible registry; easier issuance of certified copies anywhere. • Stronger basis for national ID and for targeting social and education programmes. • Immediate, understandable impact for citizens and local officials.
2	Justice decisions: high courts and courts of appeal	<ul style="list-style-type: none"> • Judgements are high-value legal assets; losing them undermines rule of law. • Volume is significant but still manageable compared to lower-level courts. 	Phase 1–2 <ul style="list-style-type: none"> • All new decisions from the Supreme Court / Court of Cassation and Courts of Appeal. • Later extension to selected first-instance courts (e.g. commercial or anti-corruption jurisdictions). 	<ul style="list-style-type: none"> • Guarantees long-term availability and integrity of key precedents. • Eases access for lawyers, lower courts, academia, and sometimes the public (for non-sensitive decisions). • Creates a strong “proof of value” for judges and justice leadership.
3	Taxation and customs: high-value assessments and decisions	<ul style="list-style-type: none"> • Direct link to domestic resource mobilisation and trust in tax administration. • Many disputes hinge on the existence and integrity of specific notices, assessments, and rulings. 	Early Phase 2 <ul style="list-style-type: none"> • Archive all new: <ul style="list-style-type: none"> – large taxpayer assessments, – major audit reports and decisions, – key customs rulings and exemptions. 	<ul style="list-style-type: none"> • Stronger evidentiary base for audits and appeals. • Reduces revenue loss due to “missing files” or weak evidence. • Improves trust of donors and investors in the robustness of the tax system.
4	Education diplomas and professional certificates	<ul style="list-style-type: none"> • Diploma fraud is a known issue; verification is slow and manual. • Youth mobility and regional labour markets (ECOWAS/UEMOA) need trusted, portable education evidence. 	Phase 2 <ul style="list-style-type: none"> • Archive and issue verifiable credentials for: <ul style="list-style-type: none"> – university diplomas, – professional school certificates (teachers, nurses, engineers, etc.), – key professional licences (lawyers, doctors, accountants). 	<ul style="list-style-type: none"> • Faster, more reliable verification for employers and foreign institutions. • Reduced fraud and reputational risk for national education systems. • A positive, citizen-facing use case for SSL credentials.
5	Public procurement and high-value	<ul style="list-style-type: none"> • Central to transparency, anti-corruption, and public financial management. 	Phase 2–3 <ul style="list-style-type: none"> • Archive all key documents relating to tenders above a certain threshold: tender 	<ul style="list-style-type: none"> • Easier audits and investigations.



administrative decisions	<ul style="list-style-type: none"> • Often scrutinised by supreme audit institutions and donors. 	notices, bids, evaluation reports, award decisions, contract amendments.	<ul style="list-style-type: none"> • Helps demonstrate good governance to citizens and partners. • Reuses the same architecture (IPFS content, Fabric metadata, SSI identities) with relatively limited additional complexity.
--------------------------	---	--	--

Starting with them allows the State to prove the platform’s usefulness quickly, build institutional confidence, and then extend the same backbone to other domains and regions with a much stronger narrative and operational experience.

7.3 Key milestones and decision points

Phase	ID	Type	Milestone / Decision	When / Goal
Phase 0 – Preparation & foundation	M0.1	Decision	Political mandate confirmed – formal government decision designating lead institution and core partners.	Early Phase 0
	M0.2	Decision	Governance and scope baseline – approval of Strategic council, Technical steering, Security & compliance, and high-level scope.	Phase 0
	M0.3	Deliverable / Decision	Legal gap analysis validated – report on archives law, evidence law, data protection, sector laws, plus decision on which changes are addressed now vs later.	Phase 0
	M0.4	Deliverable / Decision	“Version 0” platform live on test network – minimal IPFS + Fabric + Indy + MPC backbone; decision to move to pilots.	End of Phase 0 → start Phase 1
Phase 1 – Pilot deployments	M1.1	Decision	Pilot design approved – selection of pilot domains (e.g. civil status + justice) and regions; detailed pilot plan and budget validated.	Start of Phase 1
	M1.2	Deliverable / Decision	First production pilot go-live – real archives flowing for limited perimeter; decision that performance/reliability are acceptable (or need corrections).	Mid-Phase 1
	M1.3	Decision	Legal/archival recognition for pilot scope – national archives + justice recognise digital archive as official record for pilot scope.	After first stable months of pilots
	M1.4	Deliverable / Decision	Pilot evaluation and scale-up decision – evaluation report (technical, legal, organisational, economic); go / no-go / adjust for Phase 2 expansion.	End of Phase 1
Phase 2 – Expansion to core domains	M2.1	Decision	National architecture & standards frozen (v1) – adoption of v1 standards (Archive model, metadata, policies, SSI schemas) for core domains.	Early Phase 2



& all departments	M2.2	Deliverable / Decision	Departmental readiness certification – checklists confirming each department meets infra and security baseline; authorisation to join as node operators.	During rollout across 7 departments
	M2.3	Decision	Official “no new silos” rule – circular/decision that new major digital projects must use the shared archive for in-scope documents (unless formally exempted).	Mid-Phase 2
	M2.4	Decision	First legacy decommissioning – formal retirement of at least one ministry-specific archiving system with records transferred to the platform.	Late Phase 2
Phase 3 – Consolidation & regional integration	M3.1	Decision	National recognition as canonical archive – legal texts and sectoral rules updated so the platform is the canonical archival system for defined domains.	Early Phase 3
	M3.2	Decision	Regional / cross-border interoperability decision – agreements (e.g. ECOWAS/UEMOA) on verification of records (diplomas, licences, etc.) via the platform.	Mid-Phase 3
	M3.3	Decision	Long-term funding and upgrade plan (5–10 years) – adoption of a mid-term investment and funding plan (crypto migration, infra renewal, new domains).	Late Phase 3 / start of next cycle

7.4 Indicators and KPIs

Category (Level 1)	KPI ID	KPI (Level 2)	Description / Example measure
Usage & adoption	U1	Volume of archived records	Number of new archives per month, by domain and department.
	U2	Coverage ratio per domain	% of in-scope documents actually archived (e.g. % of new birth certificates, % of appellate decisions).
	U3	Active institutions	Number of ministries, departments, municipalities with ≥ X operations/month on the platform.
	U4	User base	Number of active professional users (judges, officers, clerks, archivists) and citizen interactions.
Integrity, security & compliance	S1	Integrity verification success rate	% of sampled documents where hashes/CIDs, signatures and audit trail are correct and complete.
	S2	Access control violations / incidents	Number of confirmed access breaches or policy misconfigurations per year.
	S3	Time to detect and respond	Mean time to detect (MTTD) and mean time to respond (MTTR) for security incidents affecting archived data.
	S4	Legal / compliance findings	Number and severity of issues raised by DPA, national archives, audit bodies specifically about the platform.
Performance & reliability	P1	Response time for core operations	Median and 95th percentile latency for archive creation, document consultation, integrity checks.



	P2	Availability	Uptime of central services and departmental nodes (e.g. % of time SLOs are met).
	P3	Replication lag	Time between initial archiving and confirmation that required replicas (national + regional + departmental) exist.
	P4	Node health	Number of nodes out of service, average time to restore, frequency of serious incidents per node type.
Economic & governance	E1	Legacy systems decommissioned	Number of ministry/agency archiving solutions retired; estimated budget freed or avoided.
	E2	Efficiency gains in justice and tax	Indicators such as reduced adjournments due to missing files, number of disputes resolved using archive evidence, incremental revenue attributable to better evidence.
	E3	Social / registry outcomes	% of births registered on time; % of acts retrievable within SLA; reduction in duplicate/fraudulent identities detected.
	E4	Cost vs benefit (ROI trend)	Annual platform run cost vs quantified savings and gains (from 5.5), tracked over time to show ROI evolution.

7.5 Recommendations to policymakers and donors

This final subsection distils the paper into actionable recommendations, with a particular focus on West African governments and their partners.

7.5.1 Treat the platform as national infrastructure, not a sectoral project

Policymakers should:

- Position the IPFS + Hyperledger + SSI archive as a horizontal, sovereign infrastructure, like the national backbone network or ID system.
- Anchor it in national digital strategies and mid-term plans, with explicit links to justice, tax, civil registry, education, and social protection reforms.

Donors and partners should:

- Recognise the platform as a public good and avoid funding new siloed archiving components in sectoral projects.
- Where possible, channel resources (TA, hardware, integration work) through or around the shared backbone.

7.5.2 Start small but with high-value, high-visibility use cases

Policymakers should:

- Focus initial phases on civil status, justice decisions, and key tax documents, in a limited set of departments and municipalities.
- Set clear success criteria and timelines for pilots, including at least one or two real cases where the archive is used to resolve a dispute or prove a right.

Donors should: Support these Phase 1 pilots with targeted funding for integration, training, and evaluation, rather than pushing for broad but shallow coverage.



7.5.3 Secure multi-year funding and cross-ministerial governance

Policymakers should:

- Commit to a 5-year funding envelope covering both build and run, rather than relying on fragmented annual budgets.
- Formally establish and empower the Strategic council, Technical steering committee, and Security & compliance board, with representation from justice, finance, interior, national archives, DPA, and cybersecurity agency.

Donors should: Align their support to reinforce these governance structures (e.g. funding secretariats, expert support, training), not bypass them.

7.5.4 Align laws and practices early, especially for evidence and archives

Policymakers should:

- Mandate a legal review to adapt archives law, evidence law, and sectoral regulations so that digital archives on the platform have clear probative value.
- Involve judges, national archives, and DPAs in drafting guidelines for using the platform as evidence and as official archive.

Donors should: Support legal and institutional reform components (consultations, drafting, capacity building for courts and archives) alongside technical work.

7.5.5 Invest in people and regional collaboration, not only in software

Policymakers should:

- Create training pathways for platform operators, developers, archivists, judges, tax officials, and DPOs, with incentives for staff to specialise and remain in public service.
- Encourage South-South and regional collaboration (e.g. between UEMOA/ECOWAS countries) to share patterns, tools, and even parts of the SSI and Indy ecosystems.

Donors should: Fund capacity-building programmes (regional academies, communities of practice, joint pilots) and support local universities and companies in building expertise around these technologies.

7.5.6 Make the economic case explicit and track ROI

Policymakers should:

- Use the economic model and ROI analysis (section 5.5) to justify the investment:
 - decommissioning silos,
 - improved tax collection,
 - justice efficiency,
 - better social targeting.
- Require the platform team to publish annual summaries showing costs, savings, and gains.
- Donors should:
- Include the platform's KPIs and ROI in their own monitoring frameworks, so that success is visible and can justify continued or expanded support.



7.5.7 Be honest about limitations and scope

Finally, both policymakers and donors must remember:

- The platform does not replace political will, fair procedures, or good governance; it supports them.
- It should be used first where the combination of legal value, economic impact, and technical feasibility is strongest, and expanded progressively.

Framed this way, the proposed IPFS + Hyperledger (Fabric + Indy) architecture is not a speculative experiment but a pragmatic, future-oriented way to safeguard the State's digital memory, strengthen rule of law, and support economic and social development-especially in regions where institutional trust and resilience are under pressure.

8. Conclusion – MindStack view: building a sovereign digital memory

This paper has described how IPFS, Hyperledger Fabric and Hyperledger Indy can be combined to build a state-scale digital archiving platform: a backbone for integrity, identity, and long-term evidence. Technically, it is about content-addressed storage, distributed ledgers, SSI credentials and MPC+ZKP key management. Politically and institutionally, it is about something deeper: how a State organises its memory, and who is allowed to write, read, and prove that memory over time.

From a MindStack perspective, the question is not only *"Which technologies do we use?"* but *"What kind of thinking do these technologies encode for the State?"*. A paper-based archive encodes scarcity, opacity, and locality. A siloed digital archive encodes fragmentation and short-termism. A shared, cryptographically verifiable archive encodes continuity, verifiability, and shared responsibility. It makes it technically difficult to lose, secretly alter, or selectively reveal records-without pretending to solve, on its own, the political or cultural problems behind those behaviours.

For West African governments and their partners, the stakes are high but clear. The same infrastructure that preserves civil status, court decisions, tax records, diplomas and procurement files can also underpin regional trust, better revenue, more effective social policies, and stronger rule of law. The investment is modest compared to national budgets and to the costs of repeated failures and lost archives. The real challenge is not money or even technology; it is coordination, legal clarity, and sustained leadership over a decade or more.

MindStack's core argument is that such an infrastructure should be treated as a strategic cognitive asset of the State: a way for public institutions to remember consistently, prove honestly, and decide transparently. The combination of IPFS, Hyperledger Fabric and Indy is one concrete path to that goal-open, auditable, and compatible with regional and global trust frameworks-provided that governance, law, and capacity-building keep pace with the code.

If that alignment is achieved, the result is more than a technical platform. It is a sovereign digital memory that future governments, courts, citizens and partners can rely on-long after the current generation of systems and political cycles has passed.



MindStack closing quote

You can reuse this as a pull-quote or visual caption:

"A State's real digital transformation begins the day its memory becomes sovereign, verifiable, and shared."

End of White Paper

State-Scale Digital Archiving with IPFS and Hyperledger: A Sovereign, Secure Infrastructure for Public Records.



Executive Business Case for West African Governments

State-Scale Digital Archiving with IPFS + Hyperledger (Fabric & Indy)

1. Purpose of this annex

This annex is written for ministers, permanent secretaries, DGs, and senior advisors in West African governments and regional institutions (UEMOA/ECOWAS). It summarises, in non-technical terms, why investing in a sovereign digital archiving and trust infrastructure makes economic and strategic sense, and how it directly supports priority reforms in taxation, justice, civil registration, education, and social protection.

2. Strategic context

West African states are rapidly digitising:

- Tax and customs systems to increase domestic revenue mobilisation.
- Civil registration and ID to support social programmes, financial inclusion, and elections.
- Justice and public finance processes under pressure from citizens, investors, and partners for greater transparency and efficiency.

At the same time, many countries still rely on:

- fragile paper archives (risk of fire, flood, loss, manipulation);
- fragmented, ministry-by-ministry document systems;
- foreign cloud or proprietary platforms that weaken digital sovereignty and make long-term evidence management difficult.

A shared national digital archiving backbone, built on open technologies (IPFS, Hyperledger Fabric and Indy), directly addresses these issues: it creates a trusted “memory layer” for the State that all sectors can use, with clear rules and shared costs.

3. Investment overview (order of magnitude)

For a typical medium-sized West African country (5–10 million inhabitants, 7 departments, ~25–30 municipalities onboarded in the first waves), a realistic financial envelope is:

- Initial design & build (2–3 years)
 - Core platform, first 3–4 key domains (justice, tax, civil status, education), initial migration of priority archives.
 - Order of magnitude: USD 6.5–7.5 million one-off.
- Annual run cost (steady state)



- Data centres / sovereign cloud (compute, storage, network), operations team, security and compliance, governance, and evolution budget.
- Order of magnitude: USD 2–2.5 million per year.

Compared to typical public budgets, this is:

- a small fraction of the national IT/digital spend,
- comparable to one medium-sized sectoral project,
- but serving all ministries and levels of government.

4. Expected benefits and ROI (5–10 year horizon)

Even with conservative assumptions, the annual benefits once the platform is used by core sectors can reasonably reach USD 3–7+ million per year, i.e.:

- 1.1× to 3× ROI in steady state,
- plus strong protection against rare but very costly events (loss or destruction of archives, major corruption or fraud cases collapsing due to weak evidence).

Key benefit drivers:

1. Decommissioning legacy silos
 - Today, 4–6 large ministries each pay for their own archiving/ECM stack.
 - Consolidating into a shared backbone can realistically save or avoid USD 1–1.4 M/year in duplicated infrastructure, licences, and small teams keeping “zombie” systems alive.
2. Improved tax and customs performance
 - Better integrity and accessibility of declarations, assessments, and audit trails increase effective collection, especially for large taxpayers and high-risk sectors.
 - Even a very small improvement in collection efficiency (for example, a fraction of a percent of tax revenue) translates into several million dollars per year; allocating only a modest portion of this effect to the archiving platform already covers a significant part of its cost.
3. Faster and more reliable justice
 - Fewer lost files and adjournments; easier sharing between courts, prosecutors, police, and anti-corruption bodies.
 - Gains of even 5–10% in productivity on thousands of cases per year represent hundreds of thousands to around one million dollars in equivalent value, plus strong political gains (visible reduction of backlog, more predictable decisions).
4. Stronger civil registry and social programme targeting
 - Reliable, tamper-evident archives for birth, marriage, and death certificates; better interoperability with national ID, education, health, and social protection systems.
 - Reducing mis-targeting and fraud in scholarships or social transfers by even a small percentage can free hundreds of thousands to several million dollars per year, while improving fairness.
5. Risk reduction and resilience
 - A single fire, flood, or political crisis that destroys court, tax, or registry archives can have tens of millions of dollars in downstream impact.
 - A sovereign, replicated, cryptographically protected archive is effectively an insurance policy for the State’s memory.



5. Key public services enabled

The platform is not an abstract IT project; it directly supports concrete services that matter for West African citizens and businesses:

1. Digital case files for justice
 - Secure, searchable, and shareable electronic case files;
 - Trusted digital copies of judgements and procedural documents;
 - Easier cooperation between anti-corruption bodies, financial intelligence units, and courts.
2. e-Civil Registry & identity backbone
 - Long-term preservation of civil status events;
 - Easier issuance of certified copies anywhere in the country;
 - Direct reuse by national ID projects, election management, and financial sector KYC.
3. Trusted tax/customs archive
 - Digital continuity for all major tax and customs processes;
 - Better auditability for domestic control and partners (IMF, World Bank, regional bodies);
 - Stronger basis for electronic invoicing and regional trade facilitation.
4. Diploma and professional licence verification
 - Ministries of education and professional orders issuing verifiable credentials for diplomas and licences;
 - Employers and foreign partners verifying them online, reducing fraud and improving labour mobility within ECOWAS/UEMOA.
5. Accountability for public spending
 - Archiving key documents related to procurement, budget execution, and audit findings;
 - Easier work for supreme audit institutions and anti-corruption agencies;
 - Improved confidence of citizens and external partners in public financial management.

6. Why now?

Several factors make this the right moment for West African governments to invest:

- Major digital ID, tax, justice, and social protection reforms are already under way; they need a trusted archival layer to be sustainable.
- Regional initiatives (e.g. for digital trade, mutual recognition of diplomas, or cross-border payments) require trusted, verifiable records above the level of individual applications.
- Cryptographic and blockchain technologies have matured enough to be used as serious infrastructure, with open-source stacks and local capacity building.
- Donors and development partners increasingly seek evidence of good governance, transparency, and data protection before funding or recognising digital systems.

The cost of waiting is not neutral: every year spent rolling out new siloed systems multiplies future migration costs and keeps legal risk high.



7. Recommendation to decision-makers

From a policy and investment perspective, the case can be summarised as follows:

1. Treat digital archiving and trust as a national infrastructure, not as a side effect of individual projects.
2. Fund a shared IPFS + Hyperledger-based platform centrally, with clear cost-sharing rules, instead of multiplying small, fragile systems.
3. Start with 3–4 high-impact domains (justice, tax/customs, civil status, education) and one or two pilot regions/departments, then scale up.
4. Use the platform to consolidate legacy solutions over time, freeing budget and staff for higher-value services.
5. Leverage the architecture to negotiate better with vendors and cloud providers, keeping keys, ledgers, and archives under sovereign control.

In economic terms, this is a moderate, predictable investment (a few million dollars per year) with:

- clear direct savings (decommissioning, efficiency),
- strong indirect gains (revenue, rule of law, investor confidence),
- and a crucial role in protecting the legal memory and digital sovereignty of West African States.