CIVITAS ID

TECHNICAL WHITE PAPER

# From Exclusion to Agency

## Shaping the Future of Inclusive Identity Infrastructure

**Print Edition**
30 October 2026

www.civitasid.org/whitepaper

# About Civitas ID

## Civitas ID is a 501(c)(3) nonprofit building open-source digital credential infrastructure for migrant, marginalized, and displaced communities globally.

Civitas ID's mission is to turn documentation from an economic barrier into an institutionally recognized bridge to opportunity and integration.

The organization's vision is a world where everyone can prove who they are and what they have done, wherever they are.

Civitas ID has assembled a global ecosystem of humanitarian, academic, financial, technology, and legal partners across deployment corridors in Colombia and Kenya. Learn more at civitasid.org.

# About the Authors

## Viktor Radchenko

**VP OF ENGINEERING, CIVITAS ID**

Viktor Radchenko leads engineering at Civitas ID. He brings more than two decades of experience in software architecture, cryptography, embedded systems and applied machine learning. He co-founded ProofSpace and served as its Chief Technology Officer, spending seven years building decentralized identity infrastructure, and has held product and architecture roles at Playtech, Leantegra and Garuda.AI.

## Nick Mason

**SENIOR TECHNOLOGY ADVISOR, CIVITAS ID**

Nick Mason is a senior technology advisor at Civitas ID and Chief Executive Officer of Etymos ID, an affiliated public benefit corporation. He co-founded ProofSpace, a decentralized identity platform deployed across fintech, healthcare, government and web3 in five regulatory jurisdictions, and has contributed to global identity and data standards. He has spent fifteen years working with displaced and vulnerable communities, beginning in Sierra Leone in 2011 and continuing for the last nine years on the board of Street Child, which operates in more than 20 countries.

# Contents


1 Executive Summary 5

2 The Infrastructure Gap 6
- 2.1 The Humanitarian and Economic Problem 6
- 2.2 The Strategic Problem 10
- 2.3 The Technology Problem 11

3 Solution Overview 15
- 3.1 System Overview 15
- 3.2 The Civitas Technical Foundation 16

4 The Civitas Open-Source Toolbox 18
- 4.1 Civitas Consoles 18
- 4.2 Credentialing Suite: Trust Infrastructure and Wallet 21
- 4.3 Orchestration and Agentic Goal-Path Engine 25
- 4.4 Resource Center: Content Platform with Wallet Connectivity 30
- 4.5 Composability of the Toolbox 31
- 4.6 Credential Chain Continuity 31
- 4.7 Standards and Interoperability 32
- 4.8 Licensing 33

5 Privacy, Security, and Compliance 35
- 5.1 Privacy-by-Design Architecture 35
- 5.2 Key Recovery and Wallet Restoration 36
- 5.3 Computation Locality 37
- 5.4 Standards Compliance 37
- 5.5 Governance of Agentic Actions 38

6 Cross-Network Interoperability 39
- 6.1 The Sub-Network Model 39
- 6.2 Interoperability by Default 39
- 6.3 Known Ecosystem Limitations 41

7 Risks and Mitigations 42
- 7.1 Credential Acceptance and Regulatory Recognition 42
- 7.2 Privacy, Surveillance, and Data Misuse Risk 42
- 7.3 Standards Fragmentation and Ecosystem Lock-in Risk 42
- 7.4 Geopolitical and Institutional Stability Risk 43

8 Conclusion 44

9 Glossary 45

10 References 50

# 01 Executive Summary

## Civitas ID addresses the foundational infrastructure gap that excludes marginalized and displaced groups from socioeconomic integration.

For these populations, essential identity records are frequently missing, invalid, or unrecognized by formal bodies. This lack of verifiable documentation creates a cascading barrier to critical services, from opening financial accounts and securing employment to maintaining legal status or transferring historical records across borders.

To overcome these obstacles, Civitas ID provides digital public infrastructure, to be published under an open-source license, that empowers individuals with a digital wallet designed to transform fragmented life documents—like refugee and migrant identity cards, residency permits and work records—into a unified, user-controlled identity. Our platform doesn't just store documents; it will actively orchestrate the steps required for a person to reach their life goals, from opening a bank account to securing a job.

This white paper articulates the technical challenges Civitas ID must overcome in order to be successful, and the way in which we plan to address those challenges. As such, it describes the infrastructure we are building: the target design toward which our work through 2027 is directed. The current status of each component, and our roadmap, are published on our website.

# 02 The Infrastructure Gap

By the end of 2025, 117.8 million people were forcibly displaced worldwide—roughly one in every 70 people globally (UNHCR, 2026). The majority are hosted in low- and middle-income nations, and over one billion people could be at threat of displacement by 2050 from ecological threats, conflict, and civil unrest (IEP, 2020).

The barriers preventing the sustainable inclusion of these displaced populations are systemic. At the humanitarian, strategic, and technical levels, the existing infrastructure meant to support economic participation for displaced and underserved populations is largely absent, inadequate, or structurally misaligned with the communities it serves.

## 2.1 The Humanitarian and Economic Problem

### The Documentation Wall

Displaced populations face a "documentation wall" that traps them in a cycle of economic exclusion. This is not a single issue, but a compounding systemic failure that prevents participation in formal society. The barrier manifests in three primary ways:

1. **The Verification Gap (Trust Failure):** Displaced populations often lack documentation that formal institutions (banks, employers, universities) recognize. Even when valid documents exist, these institutions lack the infrastructure or regulatory clarity to verify them. Without a reliable way to verify status, financial institutions—constrained by KYC regulations—are often forced to reject applicants, effectively locking them out of entry-level financial services.

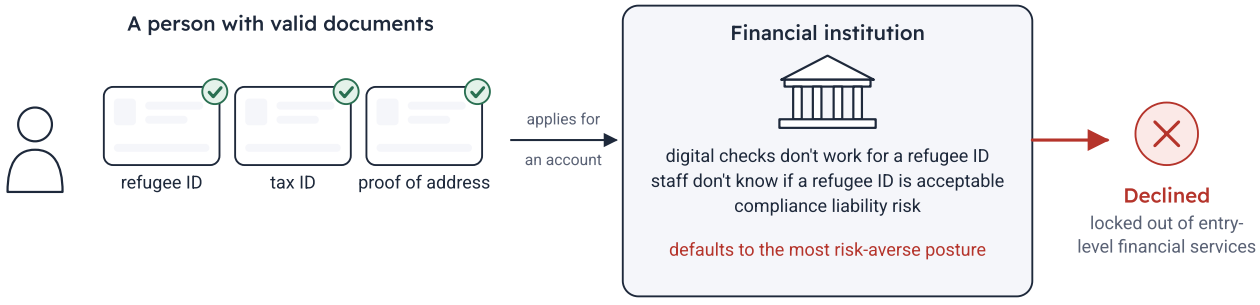



Figure 1. The verification gap: a person holds valid documents, but the institution has no reliable way to check them and, under KYC rules, takes the most risk-averse course and declines.

2 **The Foundational Bottleneck (The Chain Problem):** Life credentials—such as tax registration numbers, health insurance, work permits, and education records—are locked behind a "foundational" document (e.g., a refugee ID or residency permit). Because this foundational layer is often slow, oversubscribed, or administratively unpredictable, a single lapse or delay here creates a cascade of failure. A user cannot obtain secondary credentials without the first, making the entire chain of economic participation impossible to initiate or maintain.

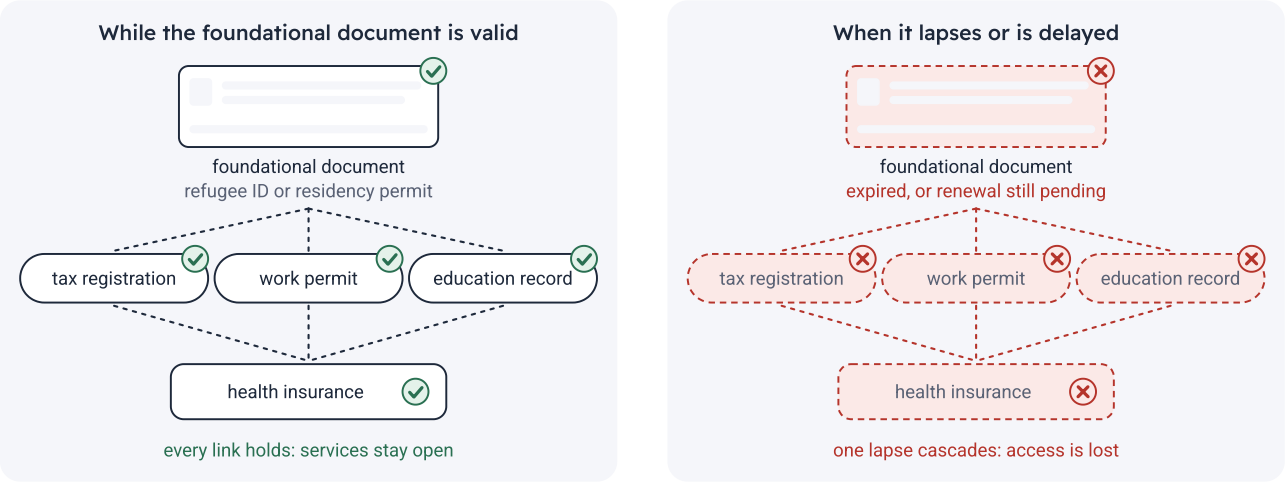



Figure 2. The foundational bottleneck: every derived credential, and the services behind it, hangs on one foundational document. While it is valid the chain holds; when it lapses, everything that depends on it fails with it.

3 **The Portability Failure (Human Capital Loss):** When legal status changes, borders are crossed, or an administrative error occurs, the "chain" breaks. Previous employment records, educational achievements, and financial histories are rarely recognized by the next institution and jurisdiction. As a result, the displaced person is forced to restart as a "stranger" at every new interaction, despite the skills and economic value they have already demonstrated.

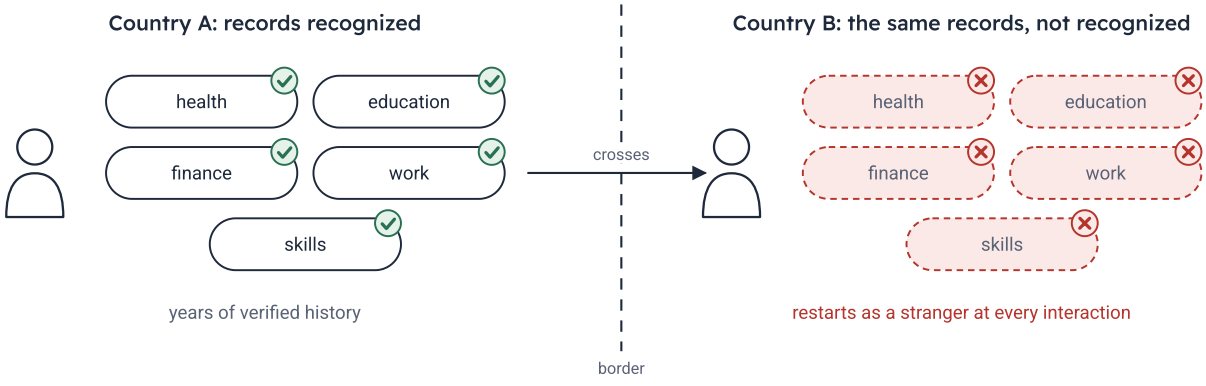



Figure 3. The portability failure: the same records that are recognized on one side of a border are not recognized on the other, and the person starts again as a stranger.

This is a structural market failure. Because each credential relies on the one before it, a single bureaucratic error forces the user back to the beginning of the chain, losing access to services they were previously eligible for. This system does not just create administrative friction—it systematically fails to mobilize the human capital that displaced populations have already accumulated, and it is a primary driver of the humanitarian aid-dependency trap described below.

## The Aid Dependency Trap

The humanitarian model for supporting displaced populations is under extreme structural stress. Despite the efforts of organizations like WFP, UNHCR, IOM and international NGOs to deliver critical cash transfers and in-kind services, the existing system was never designed to provide indefinite support at the current scale. Donor funding has failed to keep pace with a displacement crisis that has doubled in the last decade (International Rescue Committee, 2026). The consequences are stark: rationing cuts in major camps like Dadaab (The New Humanitarian, 2025) and drastic reductions in per-refugee support costs in Uganda (Kisembo and Luyiga, 2026) are no longer anomalies, but symptoms of a model that is structurally buckling under its own weight.

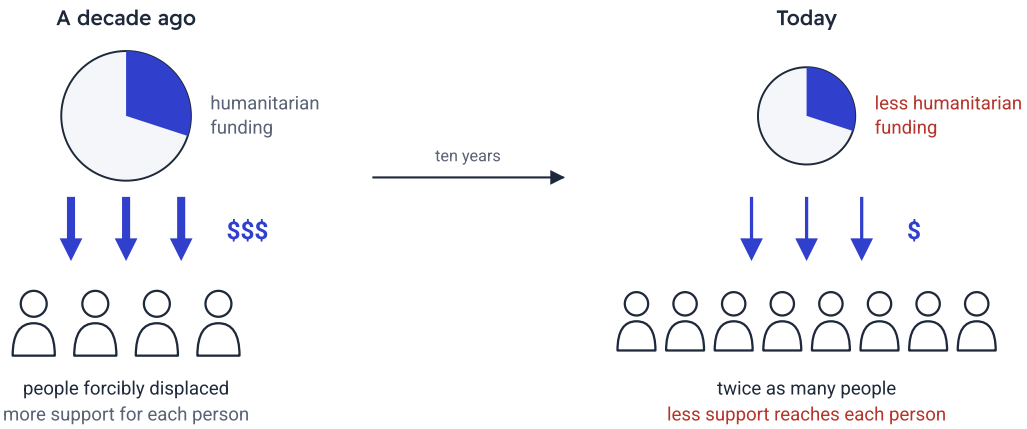



Figure 4. The aid dependency trap: displacement has doubled in a decade (International Rescue Committee, 2026) while aid is contracting rapidly (The New Humanitarian, 2025; Kisembo and Luyiga, 2026), so less support reaches each person.

This system is not only underfunded; it is fundamentally misaligned with the economic potential of the people it serves. We must shift the paradigm from managing dependency to fostering self-sufficiency, which is not merely a humanitarian ideal but a sound economic strategy. Evidence consistently shows that when displaced populations gain access to work and participate in formal markets, they become net contributors to local and national economies. For example, Venezuelan migrants in Colombia generated $529 million in direct and indirect taxes in 2022, against roughly $494 million of health and education spending for that same population (Licheri et al., 2024). Similarly, research indicates that refugees in the U.S. become net fiscal contributors within eight years of arrival (Clemens, 2017).

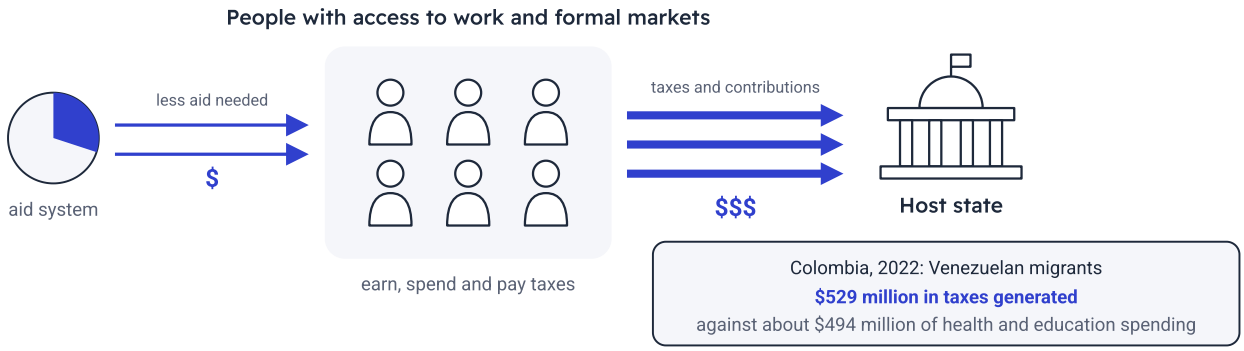



Figure 5. From dependency to contribution: with access to work and formal markets, people need less aid and pay taxes to the host state. In Colombia in 2022, Venezuelan migrants generated $529 million in taxes against roughly $494 million of health and education spending (Licheri et al., 2024).

The target set by the UNHCR High Commissioner—to more than halve the number of refugees in long-term displacement who rely on humanitarian assistance over the next decade (UN News, 2026)—is not an impossible ambition. However, it requires a structural pivot: replacing aid-dependency with the digital infrastructure necessary to mobilize the human capital that displaced populations already possess.

## The Structural Market Failure and Lack of Trust Infrastructure

The private sector is currently sidelined by a trust gap that prevents it from serving migrant and displaced populations—even though these represent underserved markets with significant economic activity and remittance flows. This failure is driven by three primary operational gaps:

The Verification Cost Gap: The burden of verifying non-standard, fragmented documentation is prohibitively expensive, forcing service providers to either absorb the cost or exclude the applicant.

The Compliance/KYC Barrier: Financial institutions, governed by strict KYC/AML regulations, face significant liability when verifying non-traditional identity documents. Without trusted, standardized verification, these institutions must default to rejection.

The Risk Assessment Gap: With thin or non-existent credit files and no formal employment history, traditional risk models are unreliable, making it impossible for providers to offer credit or services at a sustainable risk-adjusted price.

This disconnect results in the systemic exclusion of marginalized and displaced populations from formal financial and labor markets (Zhang, 2023), representing a significant and ongoing loss of human capital and economic opportunity for both these communities and the host nations.

Because the foundational trust infrastructure required to mitigate these risks does not exist, the market fails: private actors who might profitably serve this population cannot do so at acceptable risk and cost.

## 2.2 The Strategic Problem

### Why Previous Approaches Have Failed

Digital identity initiatives for displaced populations have a poor track record. Technically capable systems often fail to achieve adoption at scale, and advocacy initiatives often lack technical and commercial capability.

Non-profit and multilateral approaches hold policy credibility, programmatic impact and institutional trust, but are constrained by funding and capacity limitations, commercial and legal constraints, and governance requirements that make it difficult to move at the speed the market requires. In addition, many organizations have faced situations where data updating has become a bottleneck, as it depends on individual staff members within each institution to manually maintain and refresh information. Over time, these systems tend to become increasingly individualized, fragmented, and disconnected from one another, reducing interoperability and limiting their usefulness for coordinated decision-making and private-sector engagement.

Commercial approaches can achieve the technical and commercial capability the market needs, but are often unable to serve the most vulnerable due to unprofitable economics, and struggle to earn the institutional trust of regulators, governments, and multilateral organizations.

Neither model, operating alone, has solved the problem.

### The Funding Model Shift

At the same time, the funding environment is changing in ways that create new opportunities. Major grant-making institutions are shifting their digital infrastructure investments from project-specific deployments toward reusable, open-source digital public infrastructure (DPI) that can be deployed repeatedly across geographies and implementing partners. The logic is one of leverage: a dollar invested in core infrastructure yields more than a dollar invested in a single deployment.

This shift has a corollary: the infrastructure being funded must be genuinely open, technically credible, and governable at the DPI level. It cannot be a proprietary system whose openness is a policy commitment rather than a technical and legal reality. And it must be accompanied by a credible path to commercial sustainability, so that the infrastructure does not require permanent grant funding to function.

This shift is further amplified by the emergence of innovative financing models, such as Results-Based Financing (RBF) and Pay-for-Success (PFS) frameworks, which are increasingly attracting private capital to address deep-seated social challenges. By aligning financial returns with verified social outcomes—such as successful job placement, financial account acquisition, or sustained economic participation—these models bridge the gap between humanitarian objectives and commercial investment. Unlike traditional

grant-based cycles, impact bonds and performance-linked capital provide a pathway to de-risk development projects, allowing private investors to participate in solving complex, large-scale problems for underserved populations while ensuring that funding is directed toward interventions that demonstrably achieve measurable, high-impact results.

## 2.3 The Technology Problem

Progress on the development and acceptance of global standards in verifiable credentials and decentralized identity has been real and significant. The W3C Verifiable Credentials Data Model v2.0 became a W3C Recommendation in May 2025 (W3C, 2025); OpenID for Verifiable Credential Issuance and OpenID for Verifiable Presentations became final specifications in 2025 (OpenID Foundation, 2025a; 2025b); and SD-JWT, the selective disclosure format, was published as IETF RFC 9901 in November 2025 (Fett, Yasuda and Campbell, 2025).

The European Digital Identity Wallet builds on the two OpenID protocols (European Commission, 2026a), and the binding rules of July 2026 require an OpenID4VP profile for remote presentation (European Commission, 2026b). Under Regulation (EU) 2024/1183, private relying parties, other than micro and small enterprises, that are required by law or contract to use strong user authentication for online identification must accept the wallet at the user's request by 24 December 2027, thirty-six months after the implementing acts took effect (European Parliament and Council, 2024, Article 5f; European Commission, 2024a; 2024b).

Concurrently, the US Financial Crimes Enforcement Network has clarified that US banks may rely on verifiable digital credentials (VDCs) such as mobile driver's licenses to verify customers under the Customer Identification Program (CIP) Rule, including both government-issued and third-party-issued VDCs; for the latter, the bank remains responsible for ensuring that the third party authenticates to the same level as the bank itself would (FinCEN, 2026a; 2026b).

Vendor support is real but uneven: Apple Wallet IDs are ISO 18013 mdocs (Apple, 2025), Android requests credentials over OpenID4VP (Google, 2026), and Microsoft Entra Verified ID uses the Verifiable Credentials Data Model v1.1 (Microsoft, 2024).

Standards maturity, however, does not automatically translate into utility. Seven infrastructure problems persist that no combination of standards alone resolves.

### Participation Economics

Integration costs for verifiable credential (VC) infrastructure remain high and rise with system complexity. AI tooling is reducing development time meaningfully, and abstraction layers are improving — but the organizations with the greatest social incentive to adopt VCs (governments, community lenders,

municipalities, humanitarian agencies, local employers) tend to lack the human resources required to deploy and maintain credential infrastructure at scale. Public-sector and non-profit issuers also lack clearly articulated incentives. Consequently, the reuse dividend — the reduction in cost-to-onboard from verifying attributes once and reusing them across institutions — remains substantially unrealized for digital public services.

## Discovery, Navigation and Interoperability

Over 100 DID methods are registered (W3C, no date), and multiple credential formats — JSON-LD, JSON-LD BBS+, SD-JWT, AnonCreds, mDoc — coexist across ecosystems with different governance models, trust registries, and protocol stacks. A technically valid credential is functionally worthless when a user cannot identify which wallet to use, which credentials are accepted where, and what sequence of institutional interactions is required to reach a goal. The problem is equally acute at the institutional level: different ecosystems use different exchange protocols, credential types and trust models, meaning verifiers require custom integration wrappers. This is particularly complex when bridging government registries, humanitarian systems, and private financial service providers. No dominant discovery mechanism spans the major ecosystems, and interoperability at the implementation layer requires an infrastructural response, not just better standards.

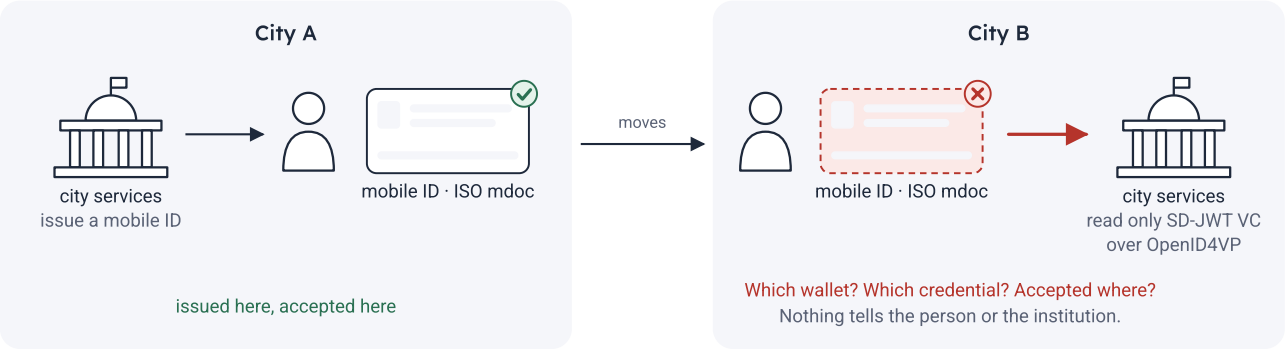



Figure 6. Discovery and navigation: a credential that is valid in the city that issued it is of no use in a city whose services read a different format, and nothing tells the person which credential is accepted where.

## Governance Fragmentation

Governance frameworks — trust models, credential schemas, liability agreements, accreditation processes, revocation policies — are developing in parallel rather than federating. Technology advances often move faster than regulatory frameworks, which often affects the readiness of third parties to accept verifiable digital credentials. For populations governed by evolving legal status, credential governance must specify how a credential survives transitions from one issuing authority to another, one document type to another, or one governing jurisdiction to another.

## User Experience

Research on digital identity wallet UX consistently finds that users struggle to understand what a credential wallet does, which credentials to acquire, where those credentials can be used, and what to do when a flow fails (Korir, Parkin and Dunphy, 2022; Sartor et al., 2022; Sauer et al., 2024). Users do not think "I need to prove my identity"; they think "I want to open an account" or "I want to qualify for this subsidy". Displaced and marginalized populations tend to be asked for the same information at every public, non-profit or private service interaction, and often with limited feedback. Most deployed wallets are static repositories and don't provide a model of what the user is trying to accomplish.

## Credential Chain Continuity

One of the most persistent failures of existing identity systems for mobile, displaced and marginalized populations is the absence of continuity: credentials that expire, documents that change form as legal status evolves, and verification systems that treat each document as a standalone artifact rather than part of an ongoing identity record. A person who has built a verified history with one document must effectively start over each time their status changes — losing service access and being forced into account duplication and re-verification processes.

## Trust

Cryptographic verification confirms a credential's integrity, not that it accurately reflects the holder's current real-world situation. Effective trust assurance requires a clear governance framework and context-aware analysis alongside cryptographic checks. Credentials may be technically valid but outdated, issued under processes that are not acceptable to a verifier, or issued by an entity that is no longer recognized in a trust registry.

## Privacy

Current credential standards and types provide strong primitives for selective disclosure and unlinkability. The EUDI ecosystem has not yet adopted anonymous credentials and zero-knowledge proofs at scale (Baum et al., 2024; European Commission, 2025). Where verifiers face compliance requirements, there is structural pressure toward full disclosure. A gap remains between privacy technology and legal practicability that requires deliberate infrastructure design.

Solving these integrated technical challenges—moving from standards to utility—is essential to unlocking the material impact of inclusion and human mobility. By addressing participation economics, navigation, and fragmented governance, we can move toward an infrastructure where users reach life goals through verified history rather than remaining trapped by bureaucratic friction. How Civitas answers each of these problems is set out in the sections that follow, and on our website, where readers can find more detail

about our roadmap, the credential formats and exchange protocols and the cross-network recognition model.

# 03 Solution Overview

The Civitas platform addresses the systemic market failures described in Section 2 by deploying a privacy-first digital public infrastructure (DPI) designed for economic participation. This solution is built on a standards-based technical foundation, will orchestrate user goals, and will provide an artificial intelligence layer that lowers the friction for vulnerable populations and third-party service providers to connect.

Civitas builds on the vast work that has been done in the last decade by the communities working on standards, protocols, trust models and enabling technologies, as referenced throughout this white paper.

## 3.1 System Overview

The system involves four distinct categories of actors and is built from four components, each described in Section 4:

Actors:

- Issuers: Entities authorized to create and sign credentials (e.g., government agencies, NGOs, financial institutions).
- Holders: Individuals who manage their credentials on non-custodial wallets on their own devices.
- Verifiers: Entities that request and validate credential presentations (e.g., banks, employers).
- Ecosystem Administrators: Entities that most likely lead implementation of the technology and governance framework.

Components:

- Civitas Consoles: In the Ecosystem Console an ecosystem administrator manages issuer/verifier onboarding (KYB), third-party integrations, and governance enforcement. In the Service Console an institution's owner sets up its services (Service Admin) and the institution defines what it issues and verifies (Service Dashboard). The Ecosystem Console is designed not to store personal data or user-level logs. The records an institution keeps about its own clients live in its Service Dashboard, in the institution's own tenant.
- Credentialing Suite - Trust Infrastructure and Wallet: A non-custodial holder wallet and a public, on-chain layer for governance-level data (Stellar/Soroban), in which the trust registry records which institutions are recognized and institutions record which version of their DID logs, the signed histories of their keys, they stand behind. The logs themselves, the credential schemas institutions publish under

their own identifiers, and the status lists of credentials held by people stay off-chain. No PII, holder identifiers, or credential content are written to the chain. Wallets exchange credentials directly with each institution over the OpenID for Verifiable Credentials protocols, with a separate pairwise identifier for each relationship, so that interactions are private and verifiable without reliance on a central database.

- Orchestration and Agentic Goal-Path Engine: Will manage user goal-path logic, converting a stated goal into a sequence of credential interactions.
- Resource Center: A content platform that will connect directly to the wallet, so that a person can move from information to action.

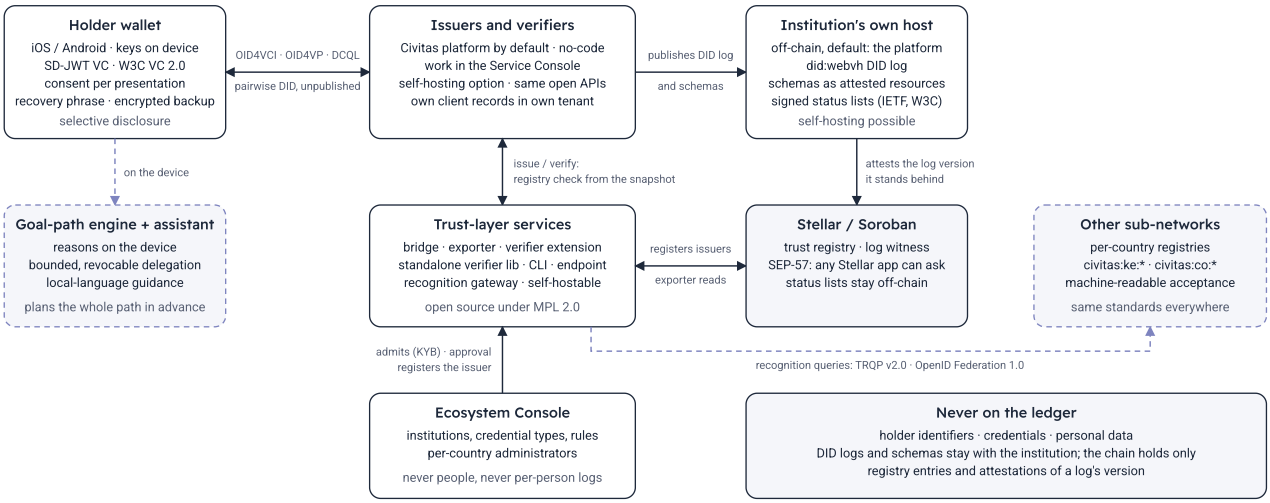



Figure 7. The system on one page: the holder wallet, institutions on the platform, each institution's own host with its DID log and schemas, the trust-layer services, the Stellar contracts, the two parts of the Civitas Consoles (the Service Console of each institution and the Ecosystem Console) and other networks.

## 3.2 The Civitas Technical Foundation

Civitas is building a modular toolbox, to be released as open source, that transforms foundational documentation into actionable economic assets.

Civitas Consoles: The Ecosystem Console, through which ecosystem administrators admit issuers and verifiers, and the Service Console, through which an institution's owner creates its services (Service Admin) and the institution defines the credentials it issues and the interactions it offers without writing code (Service Dashboard).

Credentialing Suite - Trust Infrastructure and Wallet: Will issue each credential in the forms its verifiers need: as an IETF SD-JWT VC, with its claims mapped to the W3C Verifiable Credentials (VC) 2.0 data model, as a W3C VC 2.0 credential and, at a verifier's request, as a W3C VC 1.1 credential, all signed with

the same issuing key. The suite will use the OID4VCI and OID4VP exchange protocols, and holders will prove attributes (e.g., "I am over 18") without revealing extraneous information.

Within the suite, the Stellar/Soroban Trust Infrastructure provides a public, on-chain layer that holds only governance-level, non-personal data: trust registry entries, and attestations in which institutions state which version of their DID logs they stand behind. The DID logs themselves, which carry the institutions' key histories, the credential schemas published under the institutions' own identifiers, and the status lists of credentials held by people stay off-chain. No holder identifiers, no credentials, and no personally identifiable information are written to the chain. Wallet-to-issuer and wallet-to-verifier interactions happen directly between the two parties, encrypted, with a separate pairwise identifier for each relationship that is never published to any registry. The system is designed so that it can assert, structurally rather than by policy alone, that it is not a centralized store of data on vulnerable populations.

Orchestration and Agentic Goal-Path Engine: This engine will convert user goals into executable sequences. By reasoning over wallet context and local service requirements, it will identify missing steps and provide localized guidance. The engine will include an agentic layer—governed by strict, user-approved delegation models—that can proactively manage credential expiry and execute authorized presentations.

Resource Center: A content platform that will connect directly to the wallet, enabling users to transition from information to action (e.g., initiating a microloan application directly from a service description) without redundant data entry.

# 04 The Civitas Open-Source Toolbox

The Civitas stack, to be released as open source, is designed around four components — the Civitas Consoles, a credentialing suite, an orchestration and agentic goal-path engine, and a content platform with wallet connectivity, the Resource Center — that will be deployable together or independently. A government will be able to adopt the credentialing suite on its own, connecting it to existing case-management systems. An NGO will be able to deploy only the Resource Center, connecting it to third-party wallet infrastructure. A multilateral will be able to take the full stack as a reference implementation and localize it for a new geography. The components will share common standards, governance tooling, and data models, so they will be interoperable regardless of how they are combined. Together they will constitute the infrastructure that transforms a valid credential from a stored document into a tool for reaching a life goal.

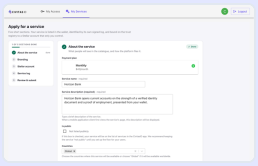
Civitas will be available to provide technical advisory and implementation support to implementing partners, such as governments, multilaterals and NGOs, under separate advisory agreements.

## 4.1 Civitas Consoles

The Civitas Consoles are designed to address the key adoption challenges of verifiable credentials at the application layer: governance, UX, interoperability across identity networks and wallet types. Their governance tooling, integration marketplace, and interaction framework build on production-tested IP, and the interaction framework includes the wallet's guidance to a missing credential. Through open standards, the Consoles will work with multiple credential types and with any standards-compliant wallet (Section 6.2).

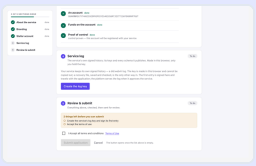
A lender, Horizon Bank, wants to serve displaced people through Civitas ID. It starts in the Service Admin, the part of the Service Console in which an institution's owner sets up its services.

**1 The bank applies for its service** →



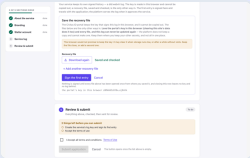
The owner of Horizon Bank applies in the Service Admin in five short sections, starting with what people will see in the wallet's catalog

**2 It proves control of its Stellar account** →



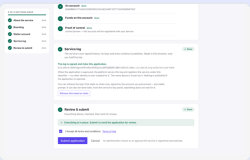
The bank names the account that will identify it in the trust registry and proves control by signing a transaction that is never submitted

### 3 It creates its own signed log →



The owner creates the service's DID log, the signed history of its keys, and saves a recovery file for it

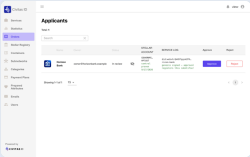
### 4 It submits the application



The first log entry is signed and travels with the application. Nothing is published if the application is rejected

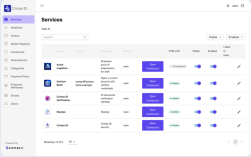
The ecosystem administrator admits Horizon Bank in the Ecosystem Console.

### 1 The administrator checks the institution →



The application arrives with its proven Stellar account and its signed log. The ecosystem administrator checks the institution behind it (KYB), then approves or rejects

### 2 The administrator approves it →



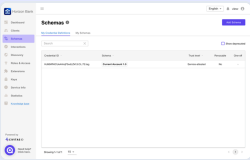
On approval the service is registered automatically. The bank now appears among the institutions on the platform, each with its trust-registry status

### 3 Recognition can be withdrawn

If an institution stops meeting the rules, the administrator suspends or revokes it in the registry, and it leaves the wallet's catalog

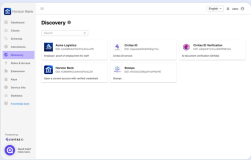
In Horizon Bank's own Service Console: the bank configures its service in the Service Dashboard without writing code.

### 1 It defines what it will issue →



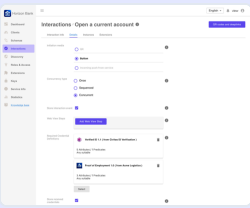
Schemas hold the institution's own credential definitions, each with its trust level. The bank adds the Current Account credential

### 2 It finds the issuers it will rely on →



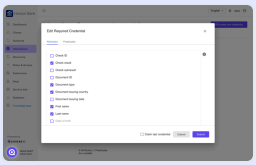
Discovery lists the other services on the network, with their identifiers and credential types: an ID verification service and an employer

### 3 It builds the interaction →



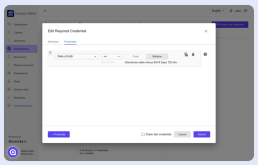
"Open a current account" names the credentials the bank asks for and the credential it issues in return

### 4 It asks only for what it needs →



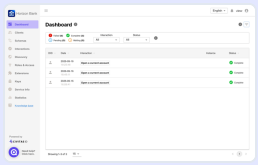
For each required credential the bank picks the attributes it needs. The wallet reveals those and nothing else

**5** **It adds a condition, not a question** →



The bank asks only whether the holder is 18 or over. It never receives the date of birth itself

**6** **The service is live**



Every interaction appears in the Service Dashboard's log with its outcome. The holder appears only as a pairwise identifier

The Civitas platform architecture is built on a pluggable backend that utilizes Stellar and its Soroban smart-contract platform as the reference implementation for institutional identity histories and trust registries. Credentials will support selective disclosure in every form in which they are issued, and wallet interactions will follow OID4VCI 1.0 and OID4VP 1.0, aligned with the OpenID4VC High Assurance Interoperability Profile (HAIP), ensuring compatibility with global identity standards.

For use cases requiring threshold logic—such as proving that a value exceeds a specific threshold—the platform will utilize issuer-side attribute bucketing. In this model, sensitive values are quantized into predefined bands at the point of issuance, allowing the holder to selectively disclose the qualifying band at presentation without revealing the underlying data. Zero-knowledge predicate proofs over the new formats belong to the same research track as unlinkability, with the two approaches described in Section 4.7; until that research concludes, banded claims are the platform's mechanism for thresholds.

## Strategic Ledger Architecture

Civitas adopts a ledger backend to secure the platform's trust infrastructure, a decision driven by three strategic advantages critical to the mission of socioeconomic integration.

First, the architecture is designed to ensure credible neutrality by removing the single-host dependency implicit in web-hosted trust infrastructure. A shared registry on a permissionless public ledger will allow national administrators to govern their own network's registry independently. This will prevent Civitas or any other party from acting as a centralized host or custodian, satisfying the sovereignty requirements essential for government adoption.

Second, the ledger facilitates decentralized, decoupled verification. Any entity with network access will be able to verify a Civitas-issued credential without calling a Civitas API: it checks the signature against the keys in the institution's DID log and the status against a signed status list, both fetched as signed files from the institution's host, which is the Civitas platform by default. With low operational costs and near-instant finality, this approach provides a robust, resilient foundation designed to remain performant even in low-resource deployments.

Third, the choice of the Stellar network enables Civitas to embed its trust fabric with established payments infrastructure, including stablecoins, regulated anchors, and fiat on- and off-ramps already operating in target corridors. By unifying credential verification and value movement on a single substrate, the platform

will allow for complex financial flows—such as direct disbursements or group-lending escrow—to be gated seamlessly by the same credentials a holder already carries.

Our application components will be programmable and API-friendly: no-code tools remain available for partners with limited resources, while the platform is optimized for developers and system integrators.

The wallet will be fast, small, and capable of operating offline with low latency. Most importantly, the orchestration backend will be built to move beyond simple, goal-oriented tasks to function as a wallet-based agent that understands the user's specific context to proactively help them reach their goals. The protocol layer throughout is the modern, industry-standard stack detailed in Section 4.7.

## 4.2 Credentialing Suite: Trust Infrastructure and Wallet

### The Holder Wallet

Holders use the Civitas ID App, a non-custodial mobile wallet for iOS and Android. All private keys and credentials are generated and stored on the holder's device, in a hardware-backed secure enclave where available. The wallet operator never holds a copy of a user's keys, and there is no server-side credential store. Credentials stored in the wallet travel with the holder, to a new device and across sub-networks (defined as independent deployments of Civitas technology). Credentials will also be readable by standards-compliant verifiers, which obtain the issuer's keys from its DID log or its published issuer metadata, without calling a Civitas API.

The wallet supports selective disclosure in every credential form, but the disclosure conditions for any given credential interaction are set by the institutional verifier, rather than by the user. This design reflects how regulated verification works: a financial institution operating under KYC/AML rules, or a government agency enforcing eligibility criteria, defines what attributes must be disclosed for a given purpose. The wallet enforces selective disclosure within those verifier-defined parameters, at the level of individual attributes. Where the verifier's regulatory framework requires full disclosure, the wallet presents accordingly.

Threshold conditions, for example proving that a user's income or age qualifies for a service without disclosing the exact figures, will be handled through issuer-side bucketing: the issuer will quantize a sensitive value into predefined bands at issuance, and the holder will selectively disclose only the qualifying band at presentation, subject to the acceptance requirements of the verifier. This reduces a range-proof requirement to a selective-disclosure operation that the standard credential formats support directly.

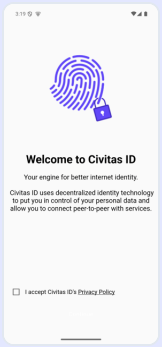
When credentials are presented, the user must proactively confirm sharing by proceeding through the consent step in the wallet interface. This action-based consent is logged on-device. The wallet operates in low-connectivity environments: credentials can be held offline, and will be presented offline, with

verification completed asynchronously when connectivity is restored. An exchange cut off by a dropped connection will be safe to retry: a repeated request returns the credential already issued, never a second one, and a lost response is recovered without the holder starting over. When an issuer needs time, for example for a manual review, the wallet will show the credential as pending and collect it once it is ready.

Key recovery and wallet restoration are handled without a centralized holder registry. The wallet generates a recovery phrase at setup. The phrase protects an encrypted backup of the wallet's keys and credentials, kept in storage the holder controls, and the wallet will offer that backup when the first credential arrives and with each new one. If a device is lost, the holder restores the wallet from the phrase together with that backup. As an alternative to mnemonic management, the wallet could support a privacy-preserving authentication technique that would serve as the recovery mechanism in place of a written phrase — providing a lower-friction recovery path better suited to users who may struggle with mnemonic management. For issuers and verifiers, a restored wallet resumes its existing relationships, and the credentials restored with it remain valid and re-presentable.

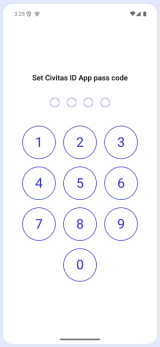
A person who wants an account with Horizon Bank first sets up the wallet. The screens are illustrative.

### 1 First launch →




The holder accepts the privacy policy before anything is created

### 2 A passcode →



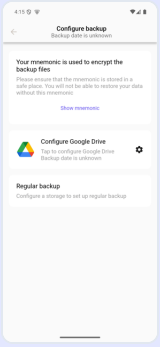

It protects the wallet on the device, and biometrics can be added

### 3 The recovery phrase →



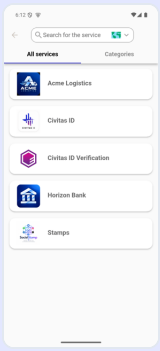

24 words, generated on the device, that unlock the encrypted backup of the wallet's keys and credentials

### 4 Encrypted backup →



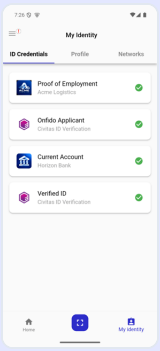

Kept in storage the holder controls, and unlocked by the phrase

**5 The service catalog** →




Horizon Bank is listed, with its standing in the trust registry

**6 Credentials on the device**




Each credential is tied to the issuer that signed it

## Trust Infrastructure (Stellar/Soroban)

The on-chain trust layer will run as Soroban smart contracts on the public Stellar network, organized into per-geography sub-networks (e.g., civitas:ke:* for Kenya, civitas:co:* for Colombia) under a global governance layer. Two contracts form the core, the log witness and the trust registry; credential schemas ride the institutions' DID logs, and status lists are served off-chain.

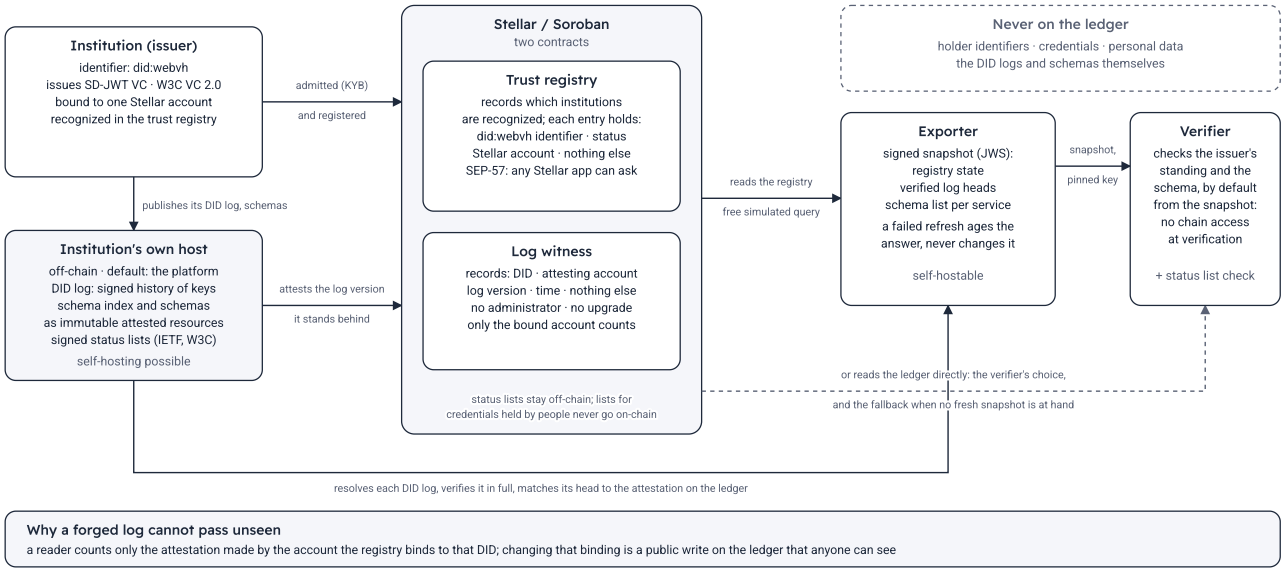



Figure 8. The trust layer on Stellar: an institution's DID log and schemas are served from the address its identifier names, the Civitas platform by default, and the institution attests the log's version from its own Stellar account; the trust registry records who is recognized; the exporter turns the registry and the verified logs into a signed snapshot; verifiers check that snapshot with no chain access, and holder identifiers, credentials, personal data, logs and schemas never reach the ledger.

The log witness anchors each institution's identity history. In plain terms, every institution keeps its own signed logbook of the keys it has used, which the did:webvh standard calls a DID log. Only that institution can write in it, and anyone can check it. After each change, the institution records on the ledger which version of its logbook it stands behind, through a small contract that nobody, Civitas included, can edit or

upgrade. All institutions share this one contract, and an attestation counts only when it comes from the Stellar account that both the trust registry and the logbook itself name for the institution. As such, a forged copy of the logbook cannot pass as the original without a public change on the ledger that anyone can see. There is no separate DID registry: the DID log carries each institution's key history, and the trust registry below records which institutions are recognized. Holders do not have DIDs on-chain; their pairwise identifiers are made for each relationship, kept for as long as that relationship lasts, and never published.

Schema publication rides the same DID logs: credential type definitions (field structures, attribute types, versioning) are published as attested resources under the publisher's own did:webvh identifier, so a schema carries the signed history of the institution that issued it and travels with that institution wherever its identifier goes. Schemas are immutable by construction: a changed schema is a new resource, and previous versions stay resolvable, so credentials issued under them still verify. A publisher can also mark an older version as deprecated, which removes it from the catalog other services see. Issuers are bound to their Stellar accounts through the SEP-57 identity-verifier interface.

The trust registry records which institutions are recognized as credential issuers. Each entry holds the institution's did:webvh identifier, its status (active, suspended, or revoked) and the Stellar account it controls; the keys with which an institution signs credentials are listed in its DID log, not in the registry. The registry will also record verifiers, the roles and credential types each institution is approved for, its attribute-access permissions, and the conditions of its approval. The registry will enforce attribute minimization: a verifier will only be able to request attributes it has been explicitly authorized to receive for a given credential type.

An issuer will be able to withdraw any credential it has issued, and verifiers will check withdrawal through status lists. Each credential type will have one fixed-size list of status bits, published from the same bits in two forms: the IETF Token Status List for SD-JWT VC credentials and the W3C Bitstring Status List for the W3C forms. Each copy of a credential will be assigned its own index in the list at issuance, and withdrawing a credential withdraws every copy. Checking a credential's status will require only fetching the compressed list and reading a single bit; the host learns that a list was fetched, not which credential was checked. The lists will be signed and hosted by the platform under each institution's own address, so revocation does not depend on an issuer's operational capacity, and a verifier that cannot obtain a current list will by default refuse the credential rather than accept it. Status lists for credentials held by people will never be written to the chain; any on-chain use will be limited to lists for credentials issued to organizations. Long-lived ledger entries, such as registry entries and log attestations, will be kept live against Soroban's state-archival (TTL) expiry through scheduled renewal, so governance data does not lapse into an archived state that must be restored before it can be read.

When a verification checks the issuer against the trust registry, it reads a signed snapshot of the registry rather than the chain and touches no network, so verification latency is not governed by on-chain settlement time. The read-only queries that build the snapshot are processed via simulated transactions,

as are direct reads by other applications; both are free and near-instantaneous. Write operations (log attestations, trust registry updates) incur Stellar's approximately five-second finality, but these writes are non-blocking for the user experience. The network's geographic neutrality and low operational cost make it well-suited for a multi-country deployment where no single government or institution should control the trust infrastructure.

### The Stellar Extension as Standalone Component

The Stellar Extension, the set of services through which the Civitas platform reaches the Soroban contracts, is built to be independently deployable and self-hostable. Once the stack is released as open source, a government, multilateral, or development organization will be able to run their own instance and interact with the on-chain registries without any dependency on the Civitas platform. This is the technical foundation of the toolbox's composability: a third party that self-hosts the extension will gain full access to the trust registry and the schemas institutions publish — and full participation in the verification ecosystem — without a Civitas infrastructure dependency.

## 4.3 Orchestration and Agentic Goal-Path Engine

### Goal-Path Logic

Civitas' orchestration engine will convert a user's stated goal into an executable sequence of credential interactions. A user who selects "get a microloan" or "renew my documentation" will not need to understand the credential requirements of each institution, the sequencing of verification steps, or the data each relying party will request. The orchestration engine will map the selected goal against country-specific rules and service requirements, check what credentials the user already holds (without reading PII), and generate a step-by-step pathway showing what is missing and how to obtain, share, or verify it. Within a single interaction, the wallet guides a holder from a missing credential to the interaction that issues it; the engine extends this guidance to the whole path.

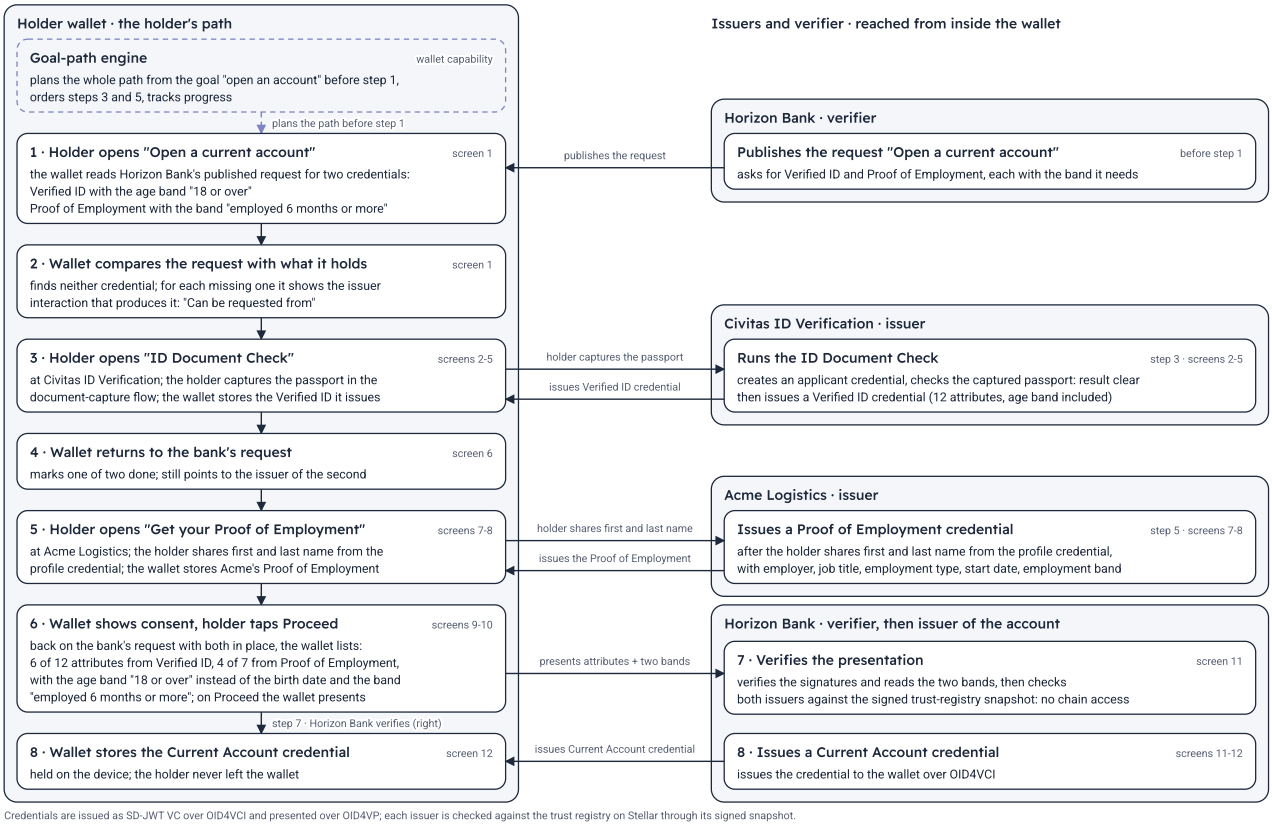
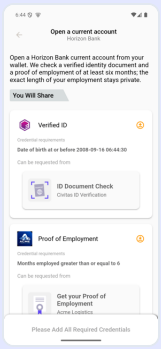



Figure 9. One use case end to end: a holder opens a bank account from the wallet. The bank asks for two credentials the holder does not have; the wallet finds the issuer of each one in its catalog, obtains both, then presents them, disclosing only that the holder is 18 or over and has been employed long enough. The bank checks both issuers against the registry snapshot without touching the chain. The goal-path engine at the top plans the whole path in advance.

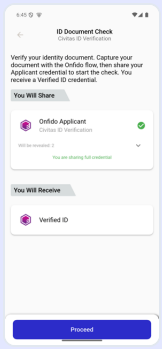
A person collects and submits the required credentials for their Horizon Bank account application. The screens are illustrative.

### 1 The bank's request →




Horizon Bank asks for two credentials. Both are missing, and each points to the interaction that issues it

### 2 The ID check starts →




The applicant credential is shared with the ID Document Check, which will return a Verified ID

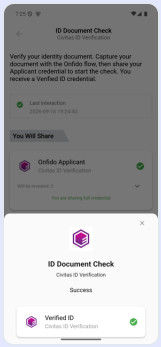
### 3 Choose the document →

The holder picks the issuing country and the document type

### 4 Capture it →

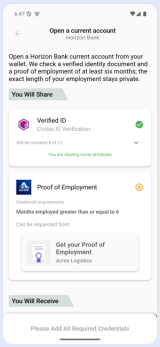
The passport photo page is captured on the device and sent for the check

### 5 Verified ID issued →



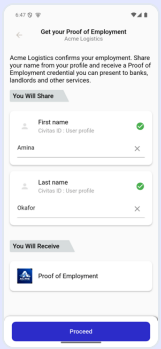

The document check clears and the credential arrives in the wallet

### 6 One of two done →



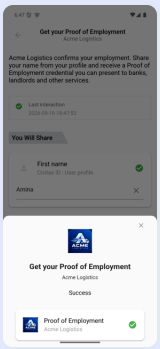

Back on the bank's request, the second credential still points to its issuer

### 7 Ask the employer →



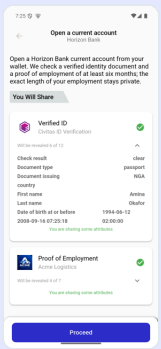

Acme Logistics asks for the name from the profile credential

### 8 Proof of Employment issued →



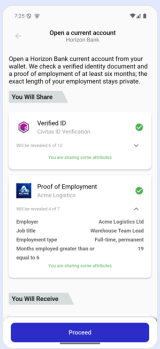

The employer issues the credential to the wallet

### 9 Consent, identity →



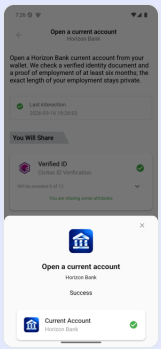

6 of 12 Verified ID attributes are shared. The holder sees the birth date, and the bank learns only that the holder is 18 or over

### 10 Consent, employment →



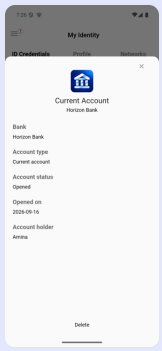

The bank learns only that employment has lasted six months or more. The figure itself is not shared

**11 The bank decides** →




It verifies both credentials and their issuers, then issues the Current Account credential

**12 Account opened**




The new credential is in the wallet

Goal paths will be informed by three inputs: the selected goal, the user's wallet context (credential state, location, legal status, prior activity), and the service requirements of relevant issuers, verifiers, and service providers. Reasoning over wallet context will happen on-device where possible, maintaining the user's momentum toward their goal by seamlessly working through the underlying process. Where server-side processing is required — for instance, to query a service provider's current acceptance criteria — only the minimum necessary, non-identifying information will be transmitted.

Goal paths will adapt as context changes. If a document expires mid-journey, or the user moves to a new location, the engine will reconfigure accordingly — an important property for mobile populations for whom a change in location can alter credential requirements entirely.

## Agentic Capabilities

The orchestration engine will include an agentic layer that extends goal-path execution from a guided step-by-step experience to a conversational, proactively executed one. The wallet agent will learn the user's goals, monitor credential state, and act on their behalf — within explicit consent boundaries — to keep documentation current and goal paths progressing.

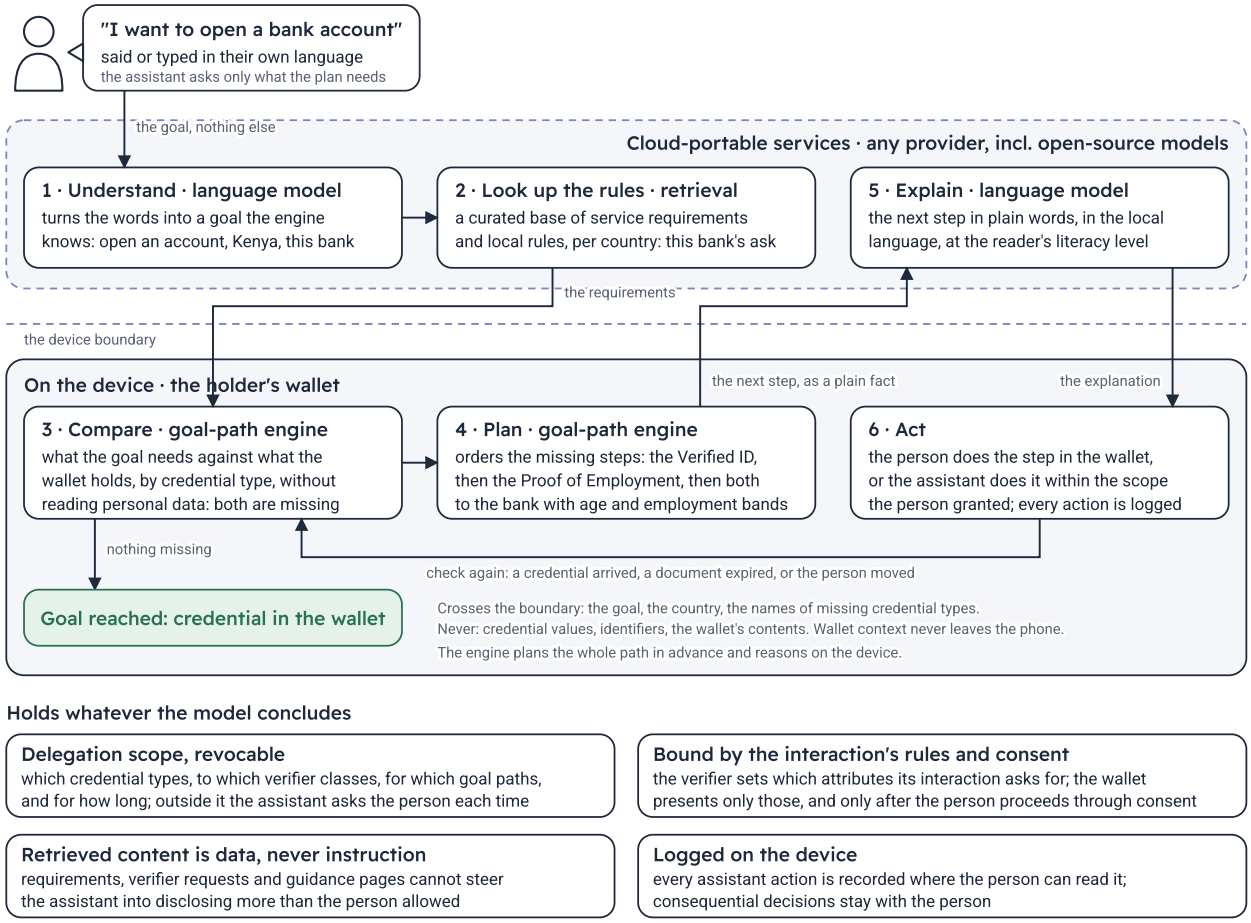



Figure 10. The user journey assisted by agentic goal paths: the person states a goal, cloud-portable services turn it into a goal the engine knows and look up the local rules, and the goal-path engine on the device compares that against the wallet and orders the missing steps. Only the goal, the country and the names of missing credential types cross the device boundary.

The agent will provide:

Conversational interface: A chat-style interaction that will guide users through goal paths in plain language, adapted to local language and literacy level. The interface will abstract the complexity of credential management and bureaucratic requirements.

Proactive notifications: The agent will monitor credential expiry dates, upcoming renewal deadlines, and changes in service requirements, and notify users before these become blockers. Where a remediation goal path exists, and user consent has been obtained, the agent will be able to launch it directly.

Goal-path execution: For steps where the user has authorized agent actions, the agent may be able to execute and maintain a log of interactions — requesting credentials from issuers, presenting proofs to verifiers, completing eligibility checks.

Retrieval-grounded reasoning: The agent will use retrieval over a structured knowledge base of service requirements, eligibility rules, and local regulatory context to ground its recommendations in accurate, current information.

Agent-mediated presentation will operate strictly within a pre-authorized consent scope. Before the agent may present any credential without a per-interaction prompt, the holder would grant a bounded, revocable authorization specifying which credential types may be presented, to which classes of verifier, and for which goal path(s); anything outside that scope would fall back to explicit per-interaction consent. Every agent-initiated presentation would be logged on-device and be auditable by the holder, and the authorization could be withdrawn at any time. Because the agent will consume external content — service requirements, verifier requests, goal-path guidance — it will itself be a target for prompt-injection aimed at inducing over-disclosure; the agent will therefore treat all retrieved content as data rather than instruction, and disclosure will be bounded, regardless of the agent's reasoning, by the parameters the verifier defines for each interaction and by the wallet's consent, given per interaction or through the bounded authorization above (Section 4.2): the agent will not be able to present an attribute the interaction does not ask for, nor one the holder has not approved. The governance of agent-mediated verification is an open area (Section 5.5); the constraints above are the platform's current position pending further R&D, external advice and emerging standards.

These capabilities will be implemented using cloud-based language models and workflow orchestration services. The implementation will be cloud-portable: capabilities will be defined at the functional level (conversational reasoning, retrieval-augmented generation, workflow orchestration, model evaluation and tuning) and will not depend on any single cloud provider's proprietary APIs. Implementing partners in different geographies or under different procurement constraints will be able to deploy the agentic layer against the infrastructure of their choice, including open-source model stacks.

The agentic orchestration engine will be part of the Civitas open-source stack.

## 4.4 Resource Center: Content Platform with Wallet Connectivity

The Civitas Resource Center will be a content platform designed to be accessible to implementing partners with limited technical capacity. It will provide a structured, localizable content layer — a documentation directory, a service directory, and goal-path guidance — with backend connectivity to the Civitas wallet and orchestration engine.

The key distinction from a static information portal is that the Resource Center will enable a user to become informed and act upon that information through the wallet. A user reading about a service will be able to initiate the relevant goal path directly from the Resource Center, carrying their credential context

into the wallet flow without re-entering information. The content and action layers will be connected; the information and the pathway to act on it will exist in the same experience.

Implementing partners will configure content and goal-path connections through a configuration interface without requiring direct access to the credential infrastructure. The content layer will be open-source under a Creative Commons license, facilitating a developer ecosystem with appropriate safeguards in place to ensure the continuous improvement of the user experience and program impact. The backend connectivity APIs that will link content delivery to the wallet and orchestration engine will be part of the core Civitas open-source stack. Implementing partners will be able to deploy the Resource Center independently, connecting it to any standards-compliant wallet infrastructure, or as part of a full Civitas stack deployment.

## 4.5 Composability of the Toolbox

The four components will be independently adoptable, interoperable, and incrementally composable. A government agency that already operates its own digital wallet will be able to adopt the Civitas Consoles and the Credentialing Suite's trust infrastructure — specifically the Service Console, the trust registry, and schema/governance tooling — without replacing its existing wallet application. An international NGO will be able to deploy the Resource Center connected to a third-party wallet infrastructure, providing structured, wallet-connected content journeys without taking on credential issuance infrastructure. A development organization building a new service in a geography where Civitas trust infrastructure is already deployed will be able to adopt the orchestration engine as a standalone component, routing users to issuers and verifiers already in the ecosystem.

The components will share a common standards foundation — IETF SD-JWT VC and W3C Verifiable Credentials, DIDs, OID4VCI/OID4VP — so they will compose cleanly regardless of which implementing partner deployed each one. The governance layer (trust registries, schema publication, KYB processes) will apply across the ecosystem, ensuring any combination of components operates within agreed rules. The Stellar Extension behind these registries will be deployable by any ecosystem operator.

## 4.6 Credential Chain Continuity

The Civitas platform will address the lack of credential continuity for displaced populations through an innovative credential chain continuity model. Rather than treating each credential as independent, the wallet will maintain a local credential graph — the set of credentials the user holds, their validity status, their relationships to each other, and their position in the sequence of legal status documents that characterizes a particular migration or regularization pathway. The orchestration engine will reason over this graph on-device, without transmitting it to any central system. When a credential expires or is superseded, the engine will generate a goal path to obtain the successor credential before access is

interrupted, and will be able to carry forward verifiable attestations from the prior credential that remain relevant — employment history, program participation, financial activity — into the user's record under the new document.

The privacy guarantee, that the credential graph stays on the device and is never sent to a central system, will hold even as the system provides the continuity experience users need.

## 4.7 Standards and Interoperability

The Civitas platform will be built on the converging standards that define the emerging global DPI landscape, so that it can interoperate with other compliant systems without requiring bilateral integration agreements.

W3C Verifiable Credentials (VC Data Model): Credentials will be issued as IETF SD-JWT VC, with their claims mapped to the W3C VC Data Model 2.0. For verifiers that need a W3C form, the same credential will also be issued as VC 2.0 secured with VC-JOSE-COSE, as VC 2.0 with an ECDSA Data Integrity proof (using the selective-disclosure suite for credentials that carry personal data) and, where a verifier asks for it, as VC 1.1 secured as a JSON Web Token, all signed with the same issuing key. Civitas-issued credentials will be readable by verifiers that implement the same open standards, without calling a Civitas API.

Decentralized Identifiers (DIDs): The platform uses did:webvh, did:web extended with a verifiable, tamper-evident key history, for institutional identities, and pairwise identifiers for holder-to-institution relationships: one per relationship and never published, so an institution recognizes a returning holder and a different institution sees a different identifier. did:webvh is preferred over plain did:web because it provides an auditable log of key rotations rather than allowing the controller to silently replace the DID document. A credential names its issuer by the institution's did:webvh identifier, and resolving it provides the current key material from the institution's own signed DID log; for verifiers that do not resolve DIDs, the same keys are published as issuer metadata at the institution's web address, and X.509 certificate chains will be added for ecosystems that require them. A credential verifies while the institution's current DID document lists the key that signed it: at a routine rotation the old key stays listed until the last credential it signed has expired, while a key found compromised is removed at once, every credential it signed is rejected, and its holders are re-issued. The log witness contract records which version of its DID log each institution stands behind. The trust registry (Section 4.2) is the authoritative recognition anchor: it records which institutional DIDs are recognized and their authorization status, and each sub-network will govern a registry of its own. The two are complementary: the registry governs recognition, the DID log provides keys.

OID4VCI / OID4VP: OpenID for Verifiable Credential Issuance 1.0 and OpenID for Verifiable Presentations 1.0 will be the wallet protocols for all credential exchanges, aligned with the OpenID4VC High Assurance Interoperability Profile (HAIP): access tokens are bound to a key the wallet holds (DPoP), presentation

requests are signed by the verifier and expressed as Digital Credentials Query Language (DCQL) queries, and presentation responses are encrypted to a key the verifier creates for that request alone. The verifier-defined disclosure parameters described in Section 4.2 will be encoded as DCQL queries. These are the protocols around which major identity ecosystems, including EUDI, are converging.

SD-JWT: Selective Disclosure JWT, in its SD-JWT VC profile, will be the IETF credential form, providing attribute-level selective disclosure with a simple, auditable structure. Its presentations are linkable across verifications, because the issuer signature is re-presented unchanged each time, so a credential shown to many verifiers needs a second mechanism in order to prevent colluding verifiers from building a composite profile (Fett, Yasuda and Campbell, 2025). Two approaches to that unlinkability are under research. The first is an unlinkable signature scheme in the BBS+ family (Looker et al., 2026). The second is a zero-knowledge proof layer over the SD-JWT credential itself, which would bring unlinkability to the standard format without a second credential type (Paquin, Policharla and Zaverucha, 2024; Frigo and shelat, 2026). Neither is chosen: they are being compared on proof size, proving time and memory on the low-cost handsets our users carry. Until one is selected, the platform limits linkability as described in Section 5.1.

ISO mDoc: Support for ISO/IEC 18013-5 mDoc format is on the roadmap, enabling interoperability with government-issued mobile credentials in jurisdictions that have adopted this standard.

GovStack: The platform will be aligned with GovStack's Identity Wallet building block specification, facilitating adoption by governments already investing in these frameworks.

## 4.8 Licensing

The Civitas IP strategy is organized around a two-tier model that reflects the obligations that come with public and philanthropic investment in DPI. The open-source release will be progressive, starting in 2027: components will be published as they are completed, and those still in research, such as the goal-path engine and the Resource Center, when they are built.

Tier 1 — Open-source core (Civitas): The core credentialing stack, the orchestration and agentic goal-path engine, the Resource Center content platform and wallet connectivity APIs, and the issuer, verifier, and wallet interfaces will be published under an open-source license. The recommended license for Tier 1 is Mozilla Public License 2.0 (MPL 2.0): it provides file-level copyleft (improvements to existing files must be shared back under MPL 2.0, while larger works combining Civitas code with other software may remain under their own license), is compatible with commercial use and proprietary combination, and is consistent with the approach of comparable DPI/DPG projects such as MOSIP. The exchange protocol core (OID4VCI, OID4VP and DCQL) and the credential signer will be released as separate MPL 2.0 packages that any host can embed or run as a standalone service.

Tier 2 — Open content (Civitas): The governance and compliance frameworks (KYB processes, trust registry tooling, consent and safeguarding policies), the Resource Center global design patterns, goal-path UX structures, content schemas, and localization frameworks will be published under Creative Commons Attribution 4.0 (CC BY 4.0). This license requires attribution, permits commercial reuse and adaptation, and is the standard for open educational and policy content in the DPI ecosystem.

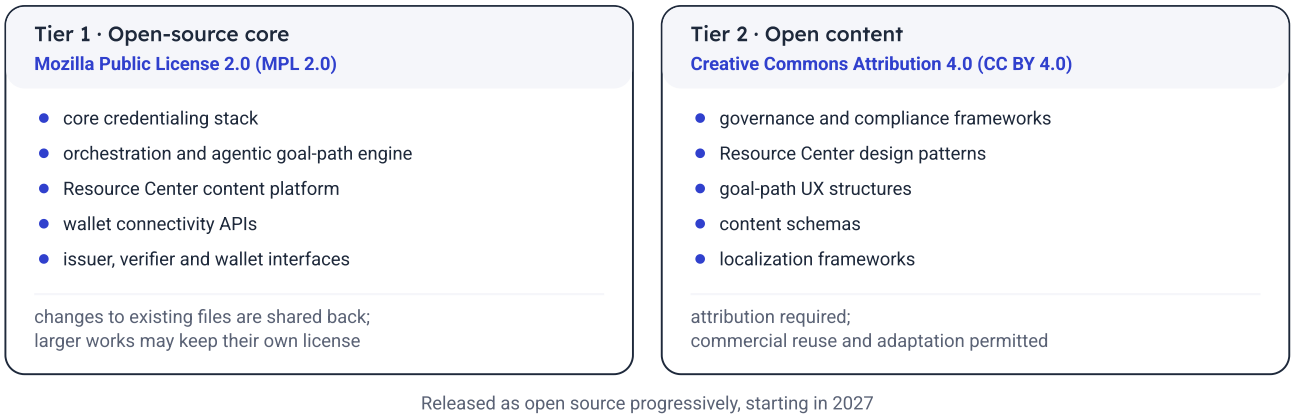



Figure 11. What will be open, and under which license: the software of the stack under MPL 2.0, the governance frameworks and content under CC BY 4.0.

Positioning as DPI, with a view to recognition as a Digital Public Good (DPG): The Civitas open-source stack — Tier 1 and Tier 2 — is designed to function, once released, as Digital Public Infrastructure: freely adoptable by governments, multilaterals, NGOs, and development organizations; standards-based and interoperable by design; and governable at the ecosystem level through open trust registries and governance frameworks. The licensing model is designed to support a mission-aligned developer ecosystem, as well as government adoption. A government agency will be able to deploy, inspect, modify, and redistribute the DPI stack under the terms of the applicable license, without restriction and without dependency on Civitas as a commercial provider. Civitas plans to seek recognition of its DPI as a Digital Public Good.

# 05 Privacy, Security, and Compliance

## 5.1 Privacy-by-Design Architecture

The Civitas architecture is built so that privacy guarantees hold structurally, not just by policy. Three properties define this:

No personally identifiable information on the public ledger. The Stellar/Soroban trust infrastructure holds only governance-level data: trust registry entries and the log witness's attestations, in which institutions name the versions of their own DID logs that they stand behind. The DID logs themselves, with the key histories they carry, and the credential schemas institutions publish under their own identifiers stay off-chain. Status lists for credentials held by people will never be written to the chain; the platform will publish them off-chain as signed lists. No holder identifiers, no credential content, and no PII are written to the chain at any point. A regulator or technical reviewer will be able to verify this directly by inspecting the contract code and on-chain state; it is not a claim that depends on trusting Civitas's internal practices.

Pairwise holder identifiers. Holders have no identifier on the ledger or in any registry. Each wallet-to-issuer and wallet-to-verifier relationship uses its own pairwise identifier, created when the relationship starts, kept for that relationship and never published. An institution therefore recognizes a returning holder, but an issuer and a verifier cannot compare their identifiers to determine they are dealing with the same person. The credential format adds a second consideration: an SD-JWT credential shown more than once carries the same issuer signature each time, so colluding verifiers could link those presentations (Fett, Yasuda and Campbell, 2025). Until an unlinkable format is selected (Section 4.7), the platform limits this in three ways: timestamps in a credential are rounded to the day; the key that binds a credential to its holder is, by default, not the key that identifies the holder's relationship with the issuer; and for sensitive credential types the issuer issues single-use copies, or a separate copy for each verifier.

Disclosure under institutional control, consent in the user's hands. Selective disclosure parameters — which attributes a verifier may request for a given credential type — are set by the institutional verifier and enforced by the platform and by the trust registry itself, reflecting the reality that disclosure requirements are typically set by regulation, not by user preference. Within those parameters, the user retains the final say: a credential is shared only with the user's consent, given in the wallet for each interaction or through a bounded, revocable authorization the user has granted (Section 4.3). No presentation happens silently.

Together these properties are designed so that the system can assert, as a matter of architecture rather than promise, that it does not function as a centralized store of data on vulnerable populations — directly addressing the surveillance and profiling risk that has made displaced populations and their advocates

rightly cautious about digital identity systems (Human Rights Watch, 2021). The full data map, boundary by boundary, is set out on our website, where readers can find more detail about the privacy architecture and data boundaries.

## 5.2 Key Recovery and Wallet Restoration

A non-custodial architecture creates a structural tension: there is no central party who can reset a lost credential, because there is no central party who holds it. For a population with high rates of device loss, displacement, and device-sharing, this tension determines whether the system is usable in practice.

Civitas's approach offers two recovery paths, one by recovery phrase and one under research, recognizing that no single mechanism suits every user:

Mnemonic-based recovery generates a 24-word recovery phrase at wallet setup. The phrase protects an encrypted backup of the wallet's keys and credentials, kept in the holder's own cloud storage. A user who loses a device restores the wallet on a new device with the phrase together with that backup; the phrase alone does not bring credentials back. This is the default SSI industry pattern and requires no third-party dependency, but places the burden of safekeeping a written phrase on the user — a meaningful usability cost for a population that may lack secure physical storage.

Cryptographic recovery (e.g., with Badge) potentially offers a lower-friction alternative that does not trade away the non-custodial principle. Privacy-preserving authentication methods derive a cryptographic key pair directly from a biometric scan at the moment of enrollment using fuzzy extraction, then discard the biometric input and the private key immediately — no biometric template, password, or secret of any kind is stored anywhere (Badge Inc., 2024). On recovery, the user would present the same biometric factor, Badge's algorithm would re-derive the same key from it, and that key would be used to retrieve the user's encrypted credential backup from their private cloud account. The mechanism is cryptographic, not biometric storage: there is no central database of biometric data to breach. This would give users a recovery path that doesn't require managing a written phrase, without reintroducing the centralized-honeypot risk the wallet's architecture is otherwise designed to avoid.

Two limits of this mechanism are acknowledged. Biometric re-derivation has a non-zero false-reject rate that may be higher for users in poor capture conditions, so mnemonic recovery is retained as a parallel path rather than replaced. And because Badge would be a third-party component in an otherwise open stack, its construction will be subject to independent security review, and the recovery flow will degrade gracefully to the mnemonic path if the Badge service or the user's encrypted cloud backup is unavailable.

In both paths, a restored wallet resumes its existing relationships with issuers and verifiers. Credentials restored from the backup remain valid and can be presented again; where a credential cannot be restored,

the holder can ask its issuer to issue it again. Recovery restores access without any central party holding the holder's credentials.

## 5.3 Computation Locality

The default design principle is that reasoning over a user's wallet context — credential state, goal-path position, location — happens on-device. The orchestration engine and the agentic layer (Section 4.3) are specified to perform this reasoning locally wherever possible, so that the data informing a user's next step does not need to leave their device to produce that step.

This default cannot be absolute. Some functions are not meaningfully achievable on-device: querying a service provider's current eligibility criteria, or resolving a schema published under an institution's own identifier. Where server-side or cloud-side processing is genuinely required, the architecture will constrain what crosses the device boundary to the minimum necessary, non-identifying information — never raw wallet content, and never information that could re-identify the holder when combined with other signals available to the receiving system. This follows the same standard GDPR Recital 26 applies to anonymization more broadly: data is treated as carrying re-identification risk if identification is "reasonably likely" given the means available to a recipient, not only if a direct identifier is technically present (Regulation (EU) 2016/679, Recital 26; European Data Protection Board, 2025).

This is an architectural commitment to be refined as implementation proceeds, not a fully closed technical specification. The boundary between what stays on-device and what may be processed server-side as anonymized aggregate will be documented precisely as each component of the agentic layer is built.

## 5.4 Standards Compliance

Civitas is designed to meet the compliance expectations of the jurisdictions and partner types it operates in, recognizing that these expectations differ by counterparty.

For Civitas, as a DPI provider whose stack will be released as open source, the principal compliance commitments are GDPR-aligned data minimization and purpose limitation in any data the platform itself processes, and country-level adaptation of consent and safeguarding frameworks to local regulatory requirements (Section 4.8). Because the architecture is designed not to centralize personal data, Civitas's own compliance surface is deliberately narrow.

Data residency, where required by a specific partner's regulator, will be handled at the deployment level rather than assumed centrally: implementing partners and enterprise customers will be able to specify

regional data processing and storage requirements as part of their commercial agreement, consistent with how comparable identity-verification vendors handle multi-jurisdictional deployment today (Persona, 2026).

## 5.5 Governance of Agentic Actions

The agentic orchestration layer will operate under a governance framework that ensures AI agents act strictly within the bounds of user-authorized authority. Governance will be achieved through authenticated delegation, which binds agent actions to a verifiable credential that identifies the human principal, the scope of delegated authority, and relevant time or transaction limits.

This is an active area of standards development. Recent work on authenticated delegation proposes extending OAuth 2.0 and OpenID Connect with agent-specific credentials that name the human principal, the scope of delegated authority, and a time or transaction limit — producing an auditable chain back to the person who authorized the action, rather than treating the agent as an independent actor (South et al., 2025). Sector-specific guidance is converging on the same pattern: Singapore's Cyber Security Agency addendum on agentic AI recommends a trusted agent registry and short-lived, narrowly scoped credentials for any agent acting with delegated authority, and advises against an agent extending its own privileges without renewed authorization (Cyber Security Agency of Singapore, 2026).

Civitas's agentic layer is designed around this delegation model rather than around standing, broad authority. Specific goal-path actions — requesting a credential, presenting a proof, completing an eligibility check — will be pre-authorized individually by the user, with the scope and duration of that authorization bounded and recorded. The agent will not hold a general mandate to act however it judges best; it will execute specific, user-approved steps and surface anything outside that scope back to the user rather than improvising. As standards in this space mature — particularly around interoperable delegation credentials that would let an agent's authority be verified by a third party without re-establishing trust from scratch — Civitas expects to adopt the converging pattern rather than maintain a bespoke one.

# 06 Cross-Network Interoperability

## 6.1 The Sub-Network Model

The Civitas trust infrastructure will be organized as a set of sub-networks, each typically scoped to a municipality, country or region (for example, civitas:ke:* for Kenya, civitas:co:* for Colombia), operating under a global governance layer, and each can be white-labeled for its operator. A sub-network is not merely a namespace; it is an ecosystem with its own dedicated ecosystem administrator, who will exercise real governance authority over that ecosystem: determining which entities are admitted as issuers or verifiers, maintaining the sub-network's trust registry, defining the standard credential schemas in use locally, and integrating custom connectors into the partner marketplace. The administrator will also have oversight of aggregate, anonymized activity across their sub-network — the visibility a national regulator or lead implementing partner needs without requiring access to any individual's data.

This design exists because identity governance is not, and should not be, uniform across jurisdictions. A government deploying Civitas as national DPI needs confidence that it — not a foreign administrator, and not Civitas centrally — controls who can issue and verify credentials within its borders, what schemas are recognized, and how its trust registry is governed. The sub-network model will give every implementing government or lead partner that control while keeping the underlying protocol, credential format, and verification logic identical across every sub-network.

## 6.2 Interoperability by Default

Sub-networks are not isolated. They interoperate with each other and with the global network by default, not as a bolted-on bridge. Because every sub-network runs the same protocol stack by construction, a verifier in one sub-network can resolve and verify a credential's cryptographic validity regardless of which sub-network issued it, without custom integration work: every sub-network uses the same standard credential forms, DID methods and OID4VCI/OID4VP exchange flows. The hard interoperability problem the field generally faces, of bridging incompatible credential formats and trust models across ecosystems (Section 2.3), does not arise between Civitas sub-networks, because they share one technical foundation by construction.

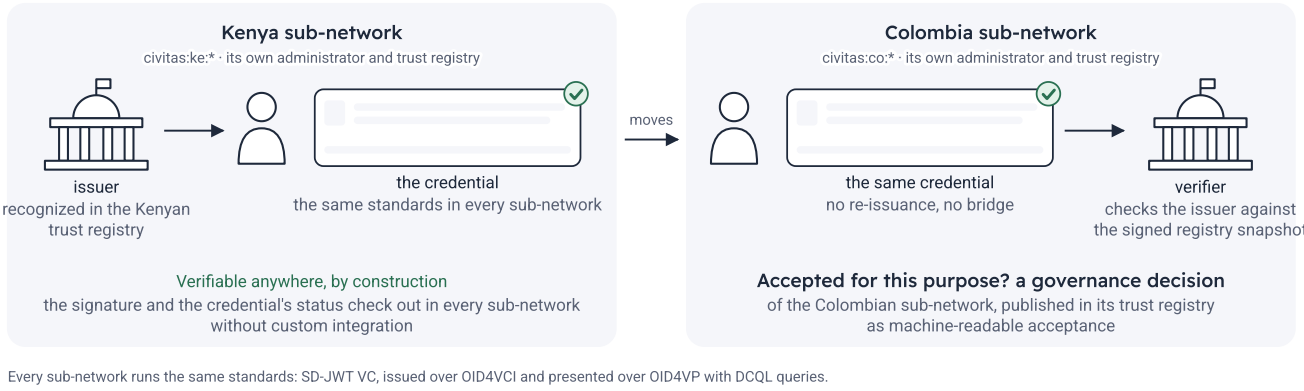



Figure 12. Interoperability by default: the same credential, presented in the sub-network that issued it and in one that did not. Every sub-network runs the same standards, SD-JWT VC over OID4VCI and OID4VP, so the Colombian verifier confirms the signature, the revocation status and the issuer's standing without custom integration. Whether the credential is accepted for a given purpose is a governance decision of the Colombian sub-network, published in its registry as machine-readable acceptance.

What differs sub-network to sub-network is governance, not protocol: whether a given issuer is recognized in another sub-network's trust registry will be a governance decision, made by that sub-network's administrator, separate from whether the credential is cryptographically verifiable. This separation is deliberate. A Colombian verifier should always be able to cryptographically confirm that a credential was genuinely issued by a named Kenyan issuer and has not been revoked; whether that Colombian verifier's regulator or business policy chooses to accept a Kenyan-issued credential for a given purpose is a separate, and properly sovereign, decision for the Colombian sub-network's governance to make.

Beyond the Civitas network, interoperability will rest on the same open standards rather than on Civitas-specific integration. Credentials issued in SD-JWT VC form over OID4VCI and OID4VP will be readable by standards-compliant wallets and verifiers that resolve the issuer's identifier, and the recognition of issuers governed by other networks will be queried through the Trust Registry Query Protocol of the Trust over IP Foundation and through OpenID Federation. An institution's offer opens the Civitas ID App where it is installed and works with any other standards-compliant wallet where it is not; such a wallet can also start an issuance itself, and it shows the same card, because the issuer's display data travels in standard issuer metadata.

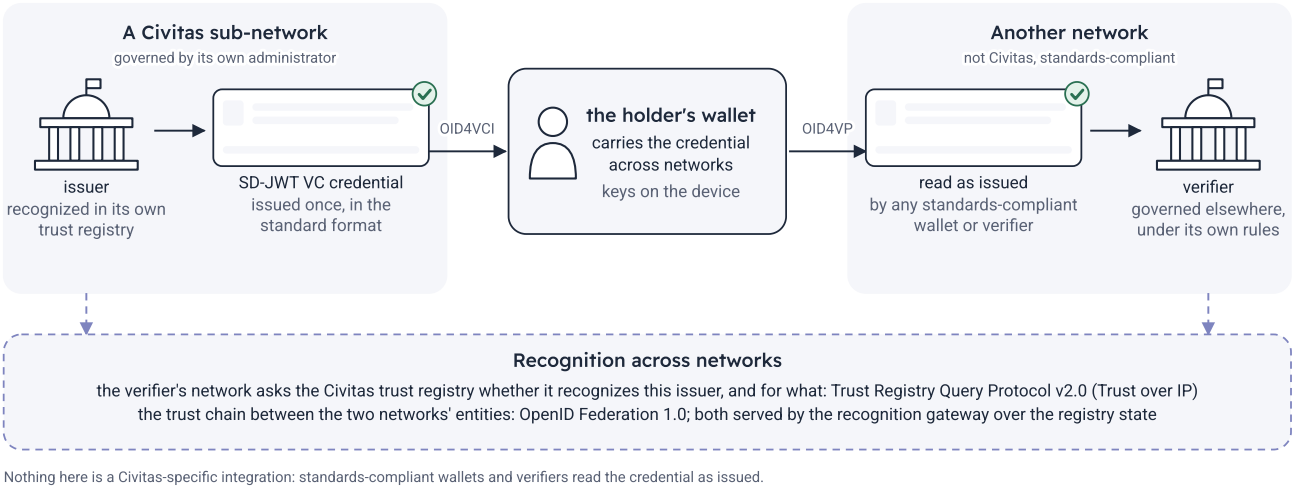



Figure 13. Beyond the Civitas network: a credential issued in a Civitas sub-network in SD-JWT VC form, over OID4VCI, travels in the holder's wallet and is read as issued, over OID4VP, by standards-compliant verifiers that resolve the issuer's identifier. Whether the verifier's network recognizes the issuer, and for what, is answered by the Civitas trust registry through the Trust Registry Query Protocol and OpenID Federation rather than through a Civitas-specific integration.

## 6.3 Known Ecosystem Limitations

This model gives a migrant whose credentials were issued in one sub-network a meaningful, verifiable record that travels with them: the credential can be cryptographically validated anywhere a Civitas-compatible verifier operates, and its provenance — which issuer, under what schema, with what authorization at time of issuance — will be auditable from any sub-network. This is a stronger starting position for cross-border continuity than most identity systems offer today, where a credential issued in one country is often simply illegible to institutions in another.

The next layer is automated cross-sub-network recognition: a standard mechanism by which a verifier's sub-network can express, in a machine-readable and queryable way, which other sub-networks' issuers and schemas it accepts, and under what conditions, so that a verifier doesn't need a bilateral governance arrangement with every other sub-network it might encounter credentials from. Cross-sub-network acceptance is a governance and policy question resolved case by case between implementing partners and regulators, the same way credential acceptance is negotiated in any cross-border context (Section 7.1). A standardized, queryable recognition layer on top of the existing trust registry architecture — so that policy decisions about cross-border acceptance are expressed in the same machine-readable form as everything else in the trust layer — is intended to resolve this issue.

# 07 Risks and Mitigations

The following risks arise from the nature of the infrastructure being built, the ecosystem in which it must operate, and the populations it serves. Delivery-stage risks are tracked on our website, where readers can find more detail about each risk and how it is mitigated.

## 7.1 Credential Acceptance and Regulatory Recognition

The value of a credential is determined by which institutions will accept it, and what for. For example, verifiable credentials may or may not be acceptable under KYC/AML rules for different countries. Cross-sector acceptance — the question of whether a temporary protection credential can satisfy a KYC framework not designed with that document type in mind — often requires policy change and ecosystem engagement.

## 7.2 Privacy, Surveillance, and Data Misuse Risk

The populations this platform serves have well-founded reasons to distrust digital systems that collect data about them. Misuse of the infrastructure could cause real harm and undermine trust in digital identity systems for displaced populations more broadly. The privacy-by-design architecture described in Section 5.1 is the principal technical mitigation: issuers cannot track where credentials are used (No Phone Home), no credential content is stored centrally, and consent is action-based and on-device. The on-device consent property holds structurally rather than depending on any individual actor's good faith, and the no-central-store property is designed to hold in the same way. The governance risk — that implementing partners or ecosystem administrators might attempt to use the platform in ways inconsistent with its design intent — is mitigated by architectural choices that prevent function creep at the infrastructure level, and will be further mitigated by the governance and safeguarding framework (Section 6.1). However, no technical architecture eliminates governance risk entirely, and ongoing vigilance is required.

## 7.3 Standards Fragmentation and Ecosystem Lock-in Risk

The verifiable credentials ecosystem is converging but is not perfectly harmonious. Credential formats, DID methods, exchange protocols, and trust frameworks are each at different maturity levels in different jurisdictions, and a platform built too tightly around any one ecosystem's current choices risks

incompatibility with others. The Civitas architecture mitigates this through the pluggable backend model (Section 4.1), format-level flexibility (IETF SD-JWT VC and the W3C VC forms, the two unlinkability tracks under research, and future mDoc support), and the decision to use did:webvh and OID4VCI/OID4VP rather than proprietary or ecosystem-specific alternatives; OID4VCI/OID4VP are the protocols around which major ecosystems including EU Digital Identity are converging. The risk is not eliminated: a significant standards bifurcation, or a rapid shift in which protocol the major wallet ecosystems adopt, could require architectural adjustment.

## 7.4 Geopolitical and Institutional Stability Risk

A change in government policy, a funding cliff at a key partner, or a shift in multilateral priorities could materially change country-level deployment plans. Once the stack is released as open source, its independently deployable architecture will mean that no single government relationship is structurally load-bearing: an implementing partner will be able to operate the stack without Civitas's direct involvement. In the near term, the risk is managed through deep institutional relationship-building, and Civitas offering an advisory role to implementing partners.

# 08 Conclusion

The status quo for displaced and vulnerable populations is under-representation in foundational identity systems; the fragmentation of their histories and track records across systems and jurisdictions; and the inconsistent acceptance of their legal documentation by institutions. This keeps people trapped outside the formal economy of their host country, and dependent on a collapsing aid system. Climate change, natural disasters and conflict will increase the rate of displacement and vulnerability in the coming decades (World Bank, 2021).

Civitas proposes a sustainable and scalable alternative: open, interoperable infrastructure through which verified history can travel with the individual. A non-custodial wallet gives people control of their credentials; common technical standards allow those credentials to move between systems; trust infrastructure allows institutions to establish who issued them and under what rules; and orchestration can help people use them to navigate the services and administrative processes that matter to their lives.

Technology alone cannot make a credential valuable. Governments, financial institutions, employers and service providers must be willing and legally able to recognize it. The challenge is therefore both technical and institutional: to build open infrastructure that makes recognition possible, while establishing the governance, regulatory frameworks and operating models that make it usable in practice.

Civitas is being developed as digital public infrastructure for governments, multilateral organizations, NGOs and businesses to adopt, extend and interoperate, while preserving local control.

For displaced and mobile populations, Civitas should enable people to own their identities and use them to unlock opportunity, stability, and long-term mobility—wherever they are.

# 09 Glossary

This glossary defines the technical terms used in this paper. Further terms are defined on our website, where readers can find the full, searchable glossary.

## Credential formats and proof schemes

**AnonCreds** — A Hyperledger credential format built on Camenisch-Lysyanskaya signatures, the same family behind Idemix. Supports native zero-knowledge predicate proofs, allowing threshold claims without revealing the underlying value.

**BBS+** — A signature scheme that lets holders disclose selected credential attributes while generating unlinkable cryptographic proofs. It is one of the two approaches under research for cross-verifier unlinkability, although disclosed attributes or higher-layer metadata may still enable correlation.

**ISO mDoc / mDL** — ISO/IEC 18013-5 standard for mobile identity credentials, commonly mobile driver's licenses. Uses CBOR, a compact binary data format, and supports proximity presentation via NFC or QR.

**JSON-LD** — JSON-based Linked Data format used in W3C Verifiable Credentials to make credential attributes machine-interpretable across systems.

**SD-JWT** — Selective Disclosure JSON Web Token. Enables attribute-level disclosure at presentation. SD-JWT supports selective disclosure but does not itself make presentations unlinkable; linkability depends on holder binding, disclosed claims, verifier policy and implementation choices. SD-JWT VC is its profile for verifiable credentials, the IETF credential form Civitas issues.

**Verifiable Credential (VC)** — A W3C data model for a cryptographically verifiable digital credential. Contains claims about a subject, signed by an issuer and presented by a holder to a verifier.

**Verifiable Presentation (VP)** — A package of one or more VCs, or derived proofs from them, presented by a holder to a verifier and signed by the holder.

**VC Data Integrity** — A W3C specification for securing a verifiable credential with a proof carried inside the credential itself. Its ECDSA suites include one that supports selective disclosure.

**W3C VC Data Model 2.0** — The second major W3C Verifiable Credentials specification. Civitas issues credentials in VC 2.0 forms, secured with VC-JOSE-COSE or with a VC Data Integrity proof, beside the IETF SD-JWT VC form, and as VC 1.1 in JWT form where a verifier asks for it.

## Identifiers and DID methods

**DID** — A decentralized identifier that does not depend on a central registration authority. A DID resolves to a DID Document containing public keys and service endpoints.

**DID Document** — The machine-readable document linked to a DID, containing authentication keys, service endpoints and related metadata.

**DID log** — The did:webvh standard's term for an identifier's verifiable history: a file of log entries, one for each version of the DID document, each chained to the one before by a hash and signed by a key the log already authorizes, so no earlier version can be altered without detection.

**did:webvh** — A DID method extending did:web with verifiable, tamper-evident DID history. In the Civitas reference deployment on Stellar, each institution also records on the ledger which version of its DID log it stands behind, in the log witness contract.

**Pairwise identifier** — An identifier, typically a DID, created for a single relationship, such as holder-to-issuer.

## Protocols and exchange standards

**DCQL** — Digital Credentials Query Language. Used in OID4VP to express verifier presentation requests, replacing Presentation Exchange.

**DPoP** — Demonstrating Proof of Possession (RFC 9449). Binds an access token to a key the wallet holds, so that a token intercepted in transit cannot be used by anyone else.

**HAIP** — OpenID4VC High Assurance Interoperability Profile. A profile of OID4VCI and OID4VP that narrows their options for ecosystems that need high assurance, such as the EU Digital Identity Wallet.

**OID4VCI** — OpenID for Verifiable Credential Issuance. Standard protocol for issuing credentials to a wallet.

**OID4VP** — OpenID for Verifiable Presentations. Standard protocol for presenting wallet-held credentials to verifiers, with presentation requests expressed as DCQL queries.

**SEP-57** — Stellar Ecosystem Proposal 57, which defines T-REX, a standard for regulated tokens on Stellar. One of its parts is an identity-verifier interface that lets a Stellar application ask whether an account belongs to a verified party. The Civitas trust registry implements that interface through the OpenZeppelin stellar-contracts library, so any Stellar application can ask whether an account belongs to a recognized issuer.

**SSI** — Self-Sovereign Identity. An identity model where individuals hold credentials in a personal wallet and control what they share, with whom, and when.

## Trust infrastructure and governance

**Civitas Consoles** — The administration side of the platform, in two parts: the Ecosystem Console for ecosystem administrators and the Service Console for institutions.

**Credential chain continuity** — The preservation of a user’s verified history as documentation changes, such as moving from asylum seeker to refugee to resident status.

**Ecosystem administrator** — The entity governing a sub-network: admitting issuers and verifiers, defining schemas, overseeing aggregate activity and managing the sub-network's own trust registry.

**Ecosystem Console** — The environment in which an ecosystem administrator admits institutions, follows their trust-registry status and governs the sub-network. It holds institutions, credential types and rules, and is designed to hold no personal data and no per-person logs.

**Fuzzy extraction** — A cryptographic technique that derives a stable key from noisy biometric input without storing the biometric itself. Used by Badge for privacy-preserving authentication, which Civitas is researching as a second route to wallet recovery, alongside the recovery phrase.

**Governance framework** — The business, legal and technical rules for a credential ecosystem: who may issue or verify which credentials, under what conditions, and with what accountability.

**Holder** — The individual who owns and controls credentials in a non-custodial wallet and authorizes presentations.

**Holder binding** — Cryptographic binding of a credential to the holder's public key, ensuring only the holder with the matching private key can present it.

**Issuer** — An entity authorized to create and sign credentials after KYB registration in the trust registry.

**KYB** — Know Your Business. Due diligence process used to verify an entity before admitting it as an issuer or verifier.

**KYC** — Know Your Customer. Regulatory identity verification requirement for financial institutions and other regulated entities, usually applied together with anti-money laundering (AML) rules.

**Log witness** — A Civitas smart contract in which each institution records, from its own Stellar account, which version of its DID log it stands behind. The contract has no administrator and no upgrade path. The name is Civitas's own; it is distinct from the optional witness mechanism in the did:webvh standard, in which designated witnesses co-sign new log entries.

**MPL 2.0** — Mozilla Public License 2.0. Open-source license that provides file-level copyleft while allowing commercial use and proprietary combination.

**No Phone Home** — The principle that verifying a credential must not require contacting its issuer at presentation time, so that an issuer cannot track where its credentials are used. Fetching a status list or an issuer's keys is not such contact: the request names no credential and no holder.

**Non-custodial wallet** — Wallet where private keys are generated and stored on the holder's device, not by a third party.

**Schema publication** — Publication of credential type definitions, including fields, attribute types and versioning; in the Civitas reference deployment a schema is an attested resource: a document published off-chain under its publisher's own did:webvh identifier, immutable by construction.

**Service Console** — The environment an institution works in. It has two parts: the Service Admin, where the institution's owner creates its services and their DID logs, and the Service Dashboard, where the institution defines its credentials and interactions and keeps the records of its own clients, in its own tenant.

**Status list** — A signed, compressed list in which each credential issued by an institution holds one bit recording whether it has been withdrawn, so that validity can be checked without revealing which credential is checked. Civitas publishes each list as an IETF Token Status List and as a W3C Bitstring Status List.

**Sub-network** — A Civitas term for a geographic or institutional ecosystem with its own administrator, its own governance rules and its own trust registry.

**Token Status List** — An IETF format for a compressed list of credential status bits, used with SD-JWT VC credentials.

**Trust registry** — On-chain record of recognized credential issuers. Each entry holds the issuer's did:webvh identifier, status and Stellar account. It will also record verifiers, credential types, conditions and permitted attributes.

**Verifier** — An entity that requests and validates credential presentations from holders.

**W3C Bitstring Status List** — Privacy-preserving revocation mechanism where each credential's validity is represented by a bit in a compressed list; used with the W3C credential forms.

## Cryptography

**Selective disclosure** — Ability for a holder to reveal only selected credential attributes.

**Unlinkability** — The property that separate presentations of the same credential cannot be matched to each other, so verifiers cannot combine them into a profile of the holder. Bringing it to the new credential formats is under research, with two candidate approaches.

**Zero-knowledge proof (ZKP)** — A cryptographic method for proving a statement without revealing the underlying data. A zero-knowledge predicate proves a condition on a value, such as being over 18, without revealing the value.

## Miscellaneous

**DPI** — Digital Public Infrastructure. Foundational interoperable systems for identity, payments or data exchange, designed for public and private actors to build on.

**DPG** — Digital Public Good. Open, non-harmful digital infrastructure relevant to the Sustainable Development Goals.

**EUDI** — EU Digital Identity. The European Union's framework for digital identity, built around the EU Digital Identity Wallet.

**GDPR** — EU data protection regulation. Civitas is designed around data minimization, purpose limitation and data subject rights.

**Goal path** — A step-by-step sequence of credential interactions that takes a user from a stated goal, such as getting a microloan, to its completion, showing what is missing and how to obtain, share or verify it. The Civitas orchestration engine will generate goal paths.

**PII** — Personally identifiable information. Data that identifies a person directly, such as a name or document number, or that could identify them when combined with other data.

**Prompt injection** — Attack where malicious retrieved content tries to override an AI agent’s instructions. Civitas will mitigate this by treating retrieved content as data and by limiting disclosure to what the verifier's interaction asks for and what the holder approves in the wallet.

**Retrieval-augmented generation (RAG)** — AI architecture that grounds model responses in retrieved knowledge base content at inference time.

**Soroban** — Stellar smart contract platform used for the trust infrastructure of the Civitas reference deployment: the trust registry and the log witness run on it, while credential schemas ride the institutions' DID logs and status lists are served off-chain.

**Stellar** — Public blockchain network used as the trust infrastructure backend of the Civitas reference deployment, offering low-cost writes and fast finality.

# 10 References


Apple (2025) Verify identity documents on the web. WWDC25, session 232. Available at: https://developer.apple.com/videos/play/wwdc2025/232/ (Accessed: 16 September 2026).

Badge Inc. (2024) Badge innovation cracks the code in security by removing PII and biometric storage as attack vector for data breaches. Press release, 17 January. Available at: https://www.prnewswire.com/news-releases/badge-innovation-cracks-the-code-in-security-by-removing-pii-and-biometric-storage-as-attack-vector-for-data-breaches-302036610.html (Accessed: 11 September 2026).

Baum, C. et al. (2024) Cryptographers' feedback on the EU Digital Identity's ARF. Open letter, 19 June. GitHub issue 200, eu-digital-identity-wallet/eudi-doc-architecture-and-reference-framework. Available at: https://github.com/eu-digital-identity-wallet/eudi-doc-architecture-and-reference-framework/issues/200 (Accessed: 11 September 2026).

Clemens, M. (2017) 'The real economic cost of accepting refugees', Center for Global Development Blog, 14 September. Available at: https://www.cgdev.org/blog/real-economic-cost-accepting-refugees (Accessed: 11 September 2026).

Cyber Security Agency of Singapore (2026) Securing agentic AI: an addendum to the guidelines and companion guide on securing AI systems. Version 1.0, 17 June. Singapore: Cyber Security Agency of Singapore. Available at: https://www.csa.gov.sg/resources/publications/addendum-on-securing-ai-systems/ (Accessed: 11 September 2026).

European Commission (2024a) Commission Implementing Regulation (EU) 2024/2977 of 28 November 2024 laying down rules for the application of Regulation (EU) No 910/2014 of the European Parliament and of the Council as regards person identification data and electronic attestations of attributes issued to European Digital Identity Wallets. OJ L, 2024/2977, 4 December. Available at: https://eur-lex.europa.eu/eli/reg_impl/2024/2977/oj (Accessed: 16 September 2026).

European Commission (2024b) Commission Implementing Regulation (EU) 2024/2981 of 28 November 2024 laying down rules for the application of Regulation (EU) No 910/2014 of the European Parliament and of the Council as regards the certification of European Digital Identity Wallets. OJ L, 2024/2981, 4 December. Available at: https://eur-lex.europa.eu/eli/reg_impl/2024/2981/oj (Accessed: 16 September 2026).

European Commission (2025) Topic G: zero knowledge proof. Discussion paper for the European Digital Identity Cooperation Group, version 1.4, 30 March. EU Digital Identity Wallet Architecture and Reference Framework. Available at: https://github.com/eu-digital-identity-wallet/eudi-doc-architecture-and-reference-framework/blob/main/docs/discussion-topics/g-zero-knowledge-proof.md (Accessed: 11 September 2026).

European Commission (2026a) European Digital Identity Wallet Architecture and Reference Framework, version 3.0.0, chapter 5: Data model and data exchange protocols, 21 July. Available at: https://eudi.dev/3.0.0/main/05-data-model-and-data-exchange-protocols/ (Accessed: 16 September 2026).

European Commission (2026b) Commission Implementing Regulation (EU) 2026/1731 of 15 July 2026 amending Implementing Regulations (EU) 2024/2977, (EU) 2024/2979, (EU) 2024/2980 and (EU) 2024/2982 as regards applicable standards and specifications. OJ L, 2026/1731, 22 July. Available at: https://eur-lex.europa.eu/eli/reg_impl/2026/1731/oj/eng (Accessed: 16 September 2026).

European Data Protection Board (2025) Guidelines 01/2025 on pseudonymisation. Version for public consultation, adopted 16 January. Brussels: European Data Protection Board. Available at: https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-012025-pseudonymisation_en (Accessed: 11 September 2026).

European Parliament and Council (2024) Regulation (EU) 2024/1183 of the European Parliament and of the Council of 11 April 2024 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework. OJ L, 2024/1183, 30 April. Available at: https://eur-lex.europa.eu/eli/reg/2024/1183/oj (Accessed: 16 September 2026).

Fett, D., Yasuda, K. and Campbell, B. (2025) Selective disclosure for JSON Web Tokens. RFC 9901, November. Internet Engineering Task Force. Available at: https://www.rfc-editor.org/rfc/rfc9901 (Accessed: 11 September 2026).

FinCEN (Financial Crimes Enforcement Network) (2026a) 'FinCEN issues frequently asked questions regarding treatment of verifiable digital credentials under the Customer Identification Program Rule', news release, 8 September. Available at: https://www.fincen.gov/news/news-releases/fincen-issues-frequently-asked-questions-regarding-treatment-verifiable-digital (Accessed: 17 September 2026).

FinCEN (Financial Crimes Enforcement Network) (2026b) Frequently asked questions regarding treatment of verifiable digital credentials under the Customer Identification Program Rule. Issued jointly with the staffs of the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation, the National Credit Union Administration and the Office of the Comptroller of the Currency, September. Available at: https://www.fincen.gov/system/files/2026-09/MDL-VDC-FAQs.pdf (Accessed: 17 September 2026).

Frigo, M. and shelat, a. (2026) 'Anonymous credentials from ECDSA', IACR Communications in Cryptology, 3(1). Available at: https://doi.org/10.62056/a3qjmpgxq.

Google (2026) About digital credentials. Android Developers. Available at: https://developer.android.com/identity/digital-credentials (Accessed: 16 September 2026).

Human Rights Watch (2021) 'UN shared Rohingya data without informed consent', news release, 15 June. Available at: https://www.hrw.org/news/2021/06/15/un-shared-rohingya-data-without-informed-consent (Accessed: 11 September 2026).

IEP (Institute for Economics and Peace) (2020) 'Over one billion people at threat of being displaced by 2050 due to environmental change, conflict and civil unrest', Ecological Threat Register press release, 9 September. Available at: https://www.economicsandpeace.org/wp-content/uploads/2020/09/Ecological-Threat-Register-Press-Release-27.08-FINAL.pdf (Accessed: 16 September 2026).

International Rescue Committee (2026) 'New UN report confirms historic levels of forced displacement as 118 million forced to flee', IRC press release, 11 June. Available at: https://www.rescue.org/uk/press-release/un-confirms-historic-levels-forced-displacement-118-million-forced-flee (Accessed: 16 September 2026).

Kisembo, D. and Luyiga, B. (2026) 'Funding shortfalls push Uganda's refugees deeper into hunger', World Food Programme, 23 March. Available at: https://www.wfp.org/stories/funding-shortfalls-push-ugandas-refugees-deeper-hunger (Accessed: 11 September 2026).

Korir, M., Parkin, S. and Dunphy, P. (2022) 'An empirical study of a decentralized identity wallet: usability, security, and perspectives on user control', in Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022). Boston, MA: USENIX Association, pp. 195-211. Available at: https://www.usenix.org/conference/soups2022/presentation/korir (Accessed: 11 September 2026).

Licheri, D. et al. (2024) Impacto fiscal de la migración venezolana en Colombia: realidad vs. potencial [The fiscal impact of Venezuelan migration in Colombia: reality vs. potential]. Bogotá: Konrad-Adenauer-Stiftung, Fundación Ideas para la Paz, OIM and CAVEX. Available at: https://ideaspaz.org/publicaciones/investigaciones-analisis/2024-05/el-impacto-fiscal-de-la-migracion-venezolana-en-colombia-realidad-vs-potencial (Accessed: 11 September 2026).

Looker, T. et al. (2026) The BBS signature scheme. Internet-Draft draft-irtf-cfrg-bbs-signatures-10, 8 January. IRTF Crypto Forum Research Group, work in progress. Available at: https://datatracker.ietf.org/doc/draft-irtf-cfrg-bbs-signatures/ (Accessed: 11 September 2026).

Microsoft (2024) Microsoft Entra Verified ID supported standards. Microsoft Learn, 16 December. Available at: https://learn.microsoft.com/en-us/entra/verified-id/verifiable-credentials-standards (Accessed: 16 September 2026).

OpenID Foundation (2025a) OpenID for Verifiable Credential Issuance 1.0. Final Specification, 16 September. OpenID Digital Credentials Protocols Working Group. Available at: https://openid.net/specs/openid-4-verifiable-credential-issuance-1_0.html (Accessed: 16 September 2026).

OpenID Foundation (2025b) OpenID for Verifiable Presentations 1.0. Final Specification, 9 July. OpenID Digital Credentials Protocols Working Group. Available at: https://openid.net/specs/openid-4-verifiable-presentations-1_0.html (Accessed: 16 September 2026).

Paquin, C., Policharla, G.-V. and Zaverucha, G. (2024) Crescent: stronger privacy for existing credentials. Cryptology ePrint Archive, Paper 2024/2013. Available at: https://eprint.iacr.org/2024/2013 (Accessed: 11 September 2026).

Persona (2026) 'What are data residency laws? How do they impact KYC and KYB?', Persona Blog. Available at: https://withpersona.com/blog/data-residency-laws (Accessed: 18 June 2026).

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union, L 119, 4 May 2016, pp. 1-88. Available at: https://eur-lex.europa.eu/eli/reg/2016/679/oj (Accessed: 11 September 2026).

Sartor, S. et al. (2022) 'Love at first sight? A user experience study of self-sovereign identity wallets', in Proceedings of the 30th European Conference on Information Systems (ECIS 2022), Research Paper 46. Association for Information Systems. Available at: https://aisel.aisnet.org/ecis2022_rp/46/ (Accessed: 11 September 2026).

Sauer, M. et al. (2024) 'User experience and information security implications of digital identity wallets', in Proceedings of the 2024 14th International Conference on Information Communication and Management (ICICM 2024). New York: ACM, pp. 36-45. Available at: https://doi.org/10.1145/3711609.3711615.

South, T. et al. (2025) 'Position: AI agents need authenticated delegation', in Proceedings of the 42nd International Conference on Machine Learning (ICML 2025). PMLR 267, pp. 82211-82231. Available at: https://proceedings.mlr.press/v267/south25a.html (Accessed: 11 September 2026).

The New Humanitarian (2025) 'Dadaab refugees decry divisive and unfair food aid cuts', 22 October. Available at: https://www.thenewhumanitarian.org/news-feature/2025/10/22/dadaab-refugees-decry-divisive-and-unfair-food-aid-cuts (Accessed: 18 June 2026).

UN News (2026) 'Refugee numbers drop for first time in a decade, but millions remain trapped', 10 June. Available at: https://news.un.org/en/story/2026/06/1167693 (Accessed: 11 September 2026).

UNHCR (2026) Global Trends: Forced displacement in 2025. Geneva: UNHCR. Available at: https://www.unhcr.org/global-trends (Accessed: 18 June 2026).

W3C (2025) Verifiable Credentials Data Model v2.0. W3C Recommendation, 15 May. Available at: https://www.w3.org/TR/vc-data-model-2.0/ (Accessed: 11 September 2026).

W3C (no date) DID methods. DID Extensions, W3C Decentralized Identifier Working Group. Available at: https://w3c.github.io/did-extensions/methods/ (Accessed: 11 September 2026).

World Bank (2021) 'Climate change could force 216 million people to migrate within their own countries by 2050', press release, 13 September. Available at: https://www.worldbank.org/en/news/press-release/2021/09/13/climate-change-could-force-216-million-people-to-migrate-within-their-own-countries-by-2050 (Accessed: 18 September 2026).

Zhang, T. (2023) 'Often shut out of the financial system, refugees and other migrants face economic integration challenges', Migration Policy Institute, 6 December. Available at: https://www.migrationpolicy.org/article/refugees-migrants-financial-inclusion (Accessed: 18 June 2026).