AFFINIDI · THE VERIFIABLE TRUST STACK · 2026

Building the Internet of Trust.

Verifiable identity and trust — owned by the person, provable by machines. The foundation for a web where you can prove who is real and every AI agent can prove who it works for.

- Running today, not slideware
- Privacy-first by construction
- Built for humans *and* agents

An overview for builders, partners and policymakers · Start high-level, get specific later

PART ONE

Why trust broke.

The internet never had a real trust layer — impersonation, data breaches and stolen identities are decades old. AI didn't start the fire; it poured petrol on it.

THE MOMENT

The internet never answered two questions. AI made ignoring them impossible.

QUESTION ONE

"Are you a real person?"

Fake accounts, catfishing and stolen identities are as old as the web. Generative AI just made them free, instant and flawless — the signals we used to trust are now cheap to counterfeit.

QUESTION TWO

"Who does this AI work for?"

Agents are starting to act on our behalf — booking, buying, negotiating, signing. The moment they do, the only question that matters is whose interest they serve.

THE INSIGHT

Both questions are really one question: **trust**. And trust has quietly become the scarcest resource on the internet.

WHY NOW

The bots arrived before the trust layer did.

For thirty years we bolted identity on as an afterthought — usernames, passwords, "log in with." It held together only because faking a human at scale was hard.

That assumption is gone.

THEN

Scarce fakes

Impersonation was expensive and manual. A profile with history read as real.

NOW

Infinite fakes

Deepfakes, synthetic identities and agentic bots are free and instant.

THEN

Humans clicked

Every action on the web had a person behind it, one at a time.

NOW

Agents act

Software takes actions autonomously, thousands per second, on our behalf.

LOOK AROUND THE CORNER

The next web isn't human-vs-bot. It's a mixed economy of people, their agents, and the organizations they act for — all needing to prove what they are, and to let their agents work together without merging systems.

THE STAKES ARE ALREADY HERE

This isn't hypothetical. It's the 2026 headlines.

Ungoverned, unverifiable agents are already causing real damage — exactly the failure modes a trust layer is built to prevent.

OPENCLAW · FEB–APR 2026

135,000 AI agents left exposed

The year's first AI-agent security crisis: 135k instances open to the internet across 82 countries, 63% with no authentication; a poisoned skills marketplace, and a sister breach leaking **1.5 million agent API tokens.** Anonymous agents, no identity, no containment.

JADEPUFFER · JUL 2026

The first fully autonomous ransomware

Researchers documented an LLM agent running an entire ransomware operation end to end — recon, credential theft, lateral movement, encryption — recovering from a failed login in **31 seconds.** No human in the loop, and no way to prove whose side it was on.

THE LESSON

Every one of these is an identity, mandate and oversight failure — an agent nobody could verify, scope, or stop. That is precisely the gap this fabric closes.

THE BROKEN MODEL

Today, someone else owns your identity.

Your "identity" is really hundreds of accounts, each held hostage by a platform. They decide what's true about you, sell the data, and can lock you out overnight.

THE ACCOUNT MODEL

- **Rented, not owned** — delete your account, lose your history.
- **Surveilled by default** — every proof leaks who you are and what you did.
- **Copy-paste trust** — a screenshot of a badge proves nothing.
- **Nothing for agents** — no way for software to prove it acts for you.

THE SHIFT UNDERWAY

- **Held by you** — credentials live in your wallet, on your device.
- **Prove without revealing** — show you're over 18, not your birthday.
- **Cryptographically real** — a proof anyone can verify, no one can forge.
- **Delegable** — hand a scoped, revocable mandate to an agent.

> THE REFRAME
>
> Stop asking platforms to vouch for you. Give people and their agents the power to prove things for themselves.

PART TWO

II

How verifiable trust works.

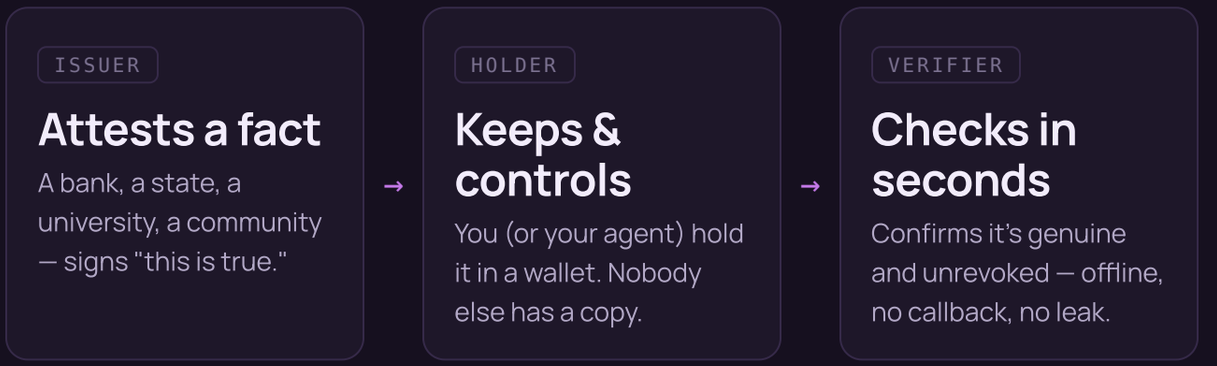
The idea, the architecture, and the open products that make it real — plus the proof it's already running.

VERIFIABLE TRUST, IN ONE BREATH

A credential you hold, a proof you control, a check anyone can run.

Think of a passport. An authority issues it. You keep it in your pocket. When it matters, you show exactly what's needed — and the person checking can tell it's genuine without phoning the government.

Now make it digital, private, and impossible to forge. That's the whole idea.



WHY IT CHANGES EVERYTHING

Trust stops being a *login* and becomes a **portable, verifiable object** — one you carry across every app, border and agent without asking permission.

THE FIRST PERSON PROJECT

AN INDUSTRY MOVEMENT WE BUILD FOR

Proof of personhood, and agents certified to serve you.

The First Person Project — backed by the Linux Foundation, Trust Over IP, DIF and the OpenWallet Foundation — asks the exact questions we opened with, and answers them with two credential types.

PERSONHOOD CREDENTIAL

You are a real, unique human

Issued by an ecosystem — a state, a bank, a university — provable with zero-knowledge, without a central biometric database tracking you.

RELATIONSHIP CREDENTIAL

Real, mutual trust connections

Issued peer-to-peer between people — "we actually know each other" — the human web of trust that bots can't fabricate.

FIRST PERSON CERTIFIED AGENT

An AI that provably works for *you*

A personal agent whose trust derives from verifiable human authorization and oversight. Its answer to "who do you work for?" is cryptographic.

OUR ROLE

We don't just endorse this vision — we ship the working infrastructure that makes First Person Certified agents real, in code, today.

Identity belongs to the individual — not the platform, and not the government.

This isn't only our engineering stance. It's now the law in Utah's SEDI framework, the design intent of the EU's digital identity wallet, and echoed in China's new agent-identity rules.

A trusted authority can **attest a fact** — your age, a licence, a residency, a role. But **you** decide how it's stored, shown, and shared.

ATTEST	**Authorities vouch for facts** State, employer, bank, or community signs what's true.
HOLD	**You keep the keys** Your core identifier is yours, perpetually, under your control.
CONSENT	**You choose each disclosure** Selective, minimal, revocable — free from surveillance tracking.

> WHY THIS MATTERS COMMERCIALLY
>
> **Building on this principle means we're aligned with regulation before it lands** — not scrambling to retrofit compliance after it does.

THE COMBINED PICTURE

One fabric, woven from six threads.

Individually these are useful tools. Together they are something new: a complete, private, verifiable trust layer for people and their agents.

1 · PEOPLE & ORGS

- Personhood & relationship credentials
- Wallets that hold, present & prove
- Selective, zero-knowledge disclosure

2 · AGENTS

- Cierge — the verifiable personal AI
- Every agent has a real, scoped identity
- Provably acts only within its mandate

3 · COMMUNITIES

- Verifiable Trust Communities (VTCs)
- Trust registries — who's authoritative
- Reputation & membership as credentials

4 · THE RAILS

- Decentralized identifiers (did:webvh)
- Private messaging — DIDComm & TSP
- Post-quantum & zero-knowledge crypto

5 · GOVERNANCE

- Human-in-the-loop approval & step-up
- Least privilege, kill-switch, full audit
- Fleet control for thousands of agents

6 · COMPLIANCE

- Aligned to EU eIDAS 2.0 / EUDI
- Aligned to Utah SEDI & China's agent rules
- Open standards, not lock-in

THE WHOLE > THE SUM

Nobody else is shipping people, agents, communities, rails, governance and compliance as one coherent stack. That combination is the moat.

The fabric ships as products — open, and buildable today.

This vision isn't abstract. It's a real, open product line under one master brand.

AFFINIDI TRUST FABRIC

The trust gateway

The umbrella for agentic trust — Trust Networks, governance frameworks, registries and runtime TRQP queries, with the two agent planes below.

AGENT STREAM

Governance data plane

A self-hosted gateway that governs the LLM & service traffic your agents generate — safety, cost, reliability and observability, by configuration.

AGENT GATEWAY

Cross-org trust plane

Agent-to-agent & MCP transport across trust boundaries — verifiable identity and authorization so agents from different orgs can trust each other.

TRUST DEVELOPMENT KIT

Build on open standards

The open-source Rust toolkit (TDK) underneath it all — crypto, DIDs, credentials, messaging — so developers ship without wrestling every standard.

AFFINIDI VAULT

Your data, your custody

A user-controlled personal data store and identity wallet — where your credentials and keys live, under your control.

AFFINIDI ELEMENTS

Verifiable credentials platform

Issue and verify W3C Verifiable Credentials with OID4VCI / OID4VP — passwordless, consent-based, verified user data.

AFFINIDI IOTA

Consent-based data sharing

The world's first consent-based data-sharing framework on open standards — share from the Vault, only with the user's say-so.

AFFINIDI MESSAGING

Private communication rails

Privacy-preserving DIDComm — and now Trust Spanning Protocol — messaging where identities are verifiable and routing leaks nothing.

THE THROUGH-LINE

Two agent planes sit on the fabric: Agent Stream governs what your agents do, Agent Gateway governs who they trust — both turning "who do we trust?" into something software can act on.

THE BUILDING BLOCKS, IN PLAIN WORDS

Six ideas, no jargon.

DID

A name you own

A decentralized identifier — like a phone number for trust that no company can revoke or repossess.

VERIFIABLE CREDENTIAL

A digital, tamper-proof fact

The signed passport-stamp: "over 18," "licensed doctor," "member since 2019."

WALLET / VAULT

Where your proofs live

On your device, under your control. Holds credentials and answers requests for them.

VTA — VERIFIABLE TRUST AGENT

Your keeper of keys

A secure vault + brain that holds identity and secrets and makes every trust decision. It signs, but never hands over the keys.

VTC — VERIFIABLE TRUST COMMUNITY

A community with a rulebook

A group — an open-source project, a marketplace, a profession — that decides who and what it trusts.

MEDIATOR

A private post office

Relays encrypted messages between agents without ever reading them — so agents can talk safely.

AGENT
wants to act

→

VTA
asked to sign — checks first

→

VTC REGISTRY
"is this allowed?" → policy

→

VTA ENFORCES
signs only if permitted

How they work together: an agent asks the VTA to sign → the VTA queries the VTC's trust registry → the registry returns the policy → the VTA enforces it. A system, not a parts list.

THE ENGINE ROOM

Where identity, keys and trust decisions actually live.

The core is open and self-hostable. Two roles do the heavy lifting — one for individuals and agents, one for communities.

- Written in Rust, end to end
- Live & test-proven
- Self-hostable / serverless

VTA	**The Verifiable Trust Agent** Holds your DID, keys and secrets — one master seed, a separate key for every context. Acts as the decision point: a signing oracle that never exports keys.
VTC	**The Verifiable Trust Community** Represents a community and its rulebook, governed by in-process policy. Issues membership, endorsement and relationship credentials.
TRUST REGISTRY	**The authoritative answer** A queryable source of truth (TRQP): "is this DID authorized for this thing, right now?" — the check every verifier runs.

THE DESIGN CHOICE THAT MATTERS

The VTA is a vault, not a keyring you copy. Agents ask it to sign; keys never leave — and the most sensitive custody runs inside a hardware secure enclave.

WORLD FIRST

Trust that travels over any protocol — including TSP.

We're first to run verifiable trust over the **Trust Spanning Protocol** — the OpenWallet Foundation's metadata-private routing layer. And because every operation is a Trust Task — a versioned, transport-agnostic contract — we're **messaging-protocol-agnostic**: DIDComm, TSP, or whatever comes next, all carry the same trust.

TSP	**Trust Spanning Protocol — a world first** Metadata-private routing (CESR + HPKE). We integrated it first, so trust can span ecosystems that share no other rails.
DIDCOMM V2	**Encrypted agent conversation** End-to-end encrypted, mediator-relayed messaging where the relay can't read the contents.
TRUST TASKS	**Protocol-agnostic by design** Every op binds to a versioned URI, independent of transport — preference order TSP > DIDComm > REST. New protocols plug in without rewriting the app.
DID:WEBVH	**Verifiable, time-travelable history** Tamper-evident identifiers with safe key rotation — "who was authorized at the moment this was signed?"

WHY PROTOCOL-AGNOSTIC WINS

Others bet on one messaging standard. We made the trust portable across all of them — so partners integrate on their rails, not ours.

Future-proof cryptography, quietly doing the work.

Non-technical audiences don't need the math — but the choices below are why this fabric will still be trustworthy in a decade.

PQC

Post-quantum crypto

ML-DSA and SLH-DSA cryptosuites already wired into our verifiable-history layer — so today's proofs aren't breakable tomorrow.

ZKP

Zero-knowledge proofs

BBS+ (bbs-2023) selective disclosure over BLS12-381 — live in the VTC presentation verifier; a Circom-based ZKP path is on the roadmap. The math behind "over 18, but not my birthday."

DTTE

Delegated Trust-Task Execution

The agent proposes an action; the VTA decides and enforces it. A human's signed, single-use consent — matched to the exact payload — authorizes anything sensitive.

SIOPV2

Self-Issued OpenID

Your wallet acts as its own identity provider — you log in as yourself, not through a middleman.

TSP

Trust Spanning Protocol

Metadata-private routing (CESR + HPKE) — our world-first integration lets trust span ecosystems that share no other rails.

DIDCOMM V2

Encrypted agent messaging

Private, mediated, metadata-minimizing conversation between agents.

DID:WEBVH

Verifiable history

Tamper-evident, time-travelable identifiers with safe key rotation.

SD-JWT VC · MDOC

Selective-disclosure credentials

The EU-mandated formats for showing part of a credential, not all of it.

TEE ENCLAVES

Hardware isolation

The key-custody core runs inside AWS Nitro secure enclaves with attested boot — the vault, sealed in silicon.

THE POINT

We're not betting on one clever trick. We're assembling the strongest available primitive for each job — and swapping them as the field advances.

Real, and already running — not slideware.

The proof isn't a roadmap. It's live systems, real signed credentials and named users — with end-to-end demonstrations of the hardest parts.

DEPLOYMENTS

- Linux Foundation — Linux Kernel
- CardInfoLink (CIL)
- eeCheck
- eCertizen
- H2H Connect

~18,500

real signed credentials across many live trust networks

3

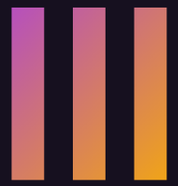
hard problems proven end-to-end — isolation, private cross-domain sharing, developer reputation

- **Agent isolation** — one domain literally can't read another's keys, proven over real DIDComm.
- **Cross-domain share** — human-approved, scoped, revocable — driven live.
- **did:webvh mediator** — running; agents connect and message privately.
- **BBS+ selective disclosure** — in the production presentation verifier.

And it moves fast: **~2,300 commits and 181 merged PRs across 17 repositories in the last three months.**

> THE POINT
>
> **The hard cryptographic core isn't a promise — it's built, tested, and demonstrable today.**

PART THREE



What it means — for people & their agents.

The guarantees, lived: a personal AI you can trust, scaled up to governed enterprise fleets and trust that crosses company lines.

FOR HUMANS

You become the source of truth about yourself.

DIGNITY

Prove less, keep more

Show you're eligible without handing over your life. "Over 18," not your birthday. "A patient here," not your whole record.

PORTABILITY

Take it anywhere

Your credentials and reputation move with you across apps, borders and services — no re-verification, no re-uploading documents.

RESILIENCE

Can't be locked out

No platform can delete your identity or hold your history hostage. The keys are yours, on your device.

SAFETY

Real connections only

Relationship credentials let you tell genuine humans from convincing fakes — for dating, hiring, community, family.

DELEGATION

An agent you can trust

Hand a scoped, revocable mandate to an AI that acts for you — and prove to others that it does.

PRIVACY

Off the surveillance grid

Every proof is peer-to-peer and minimal. No issuer learns where you used your credential.

THE FELT DIFFERENCE

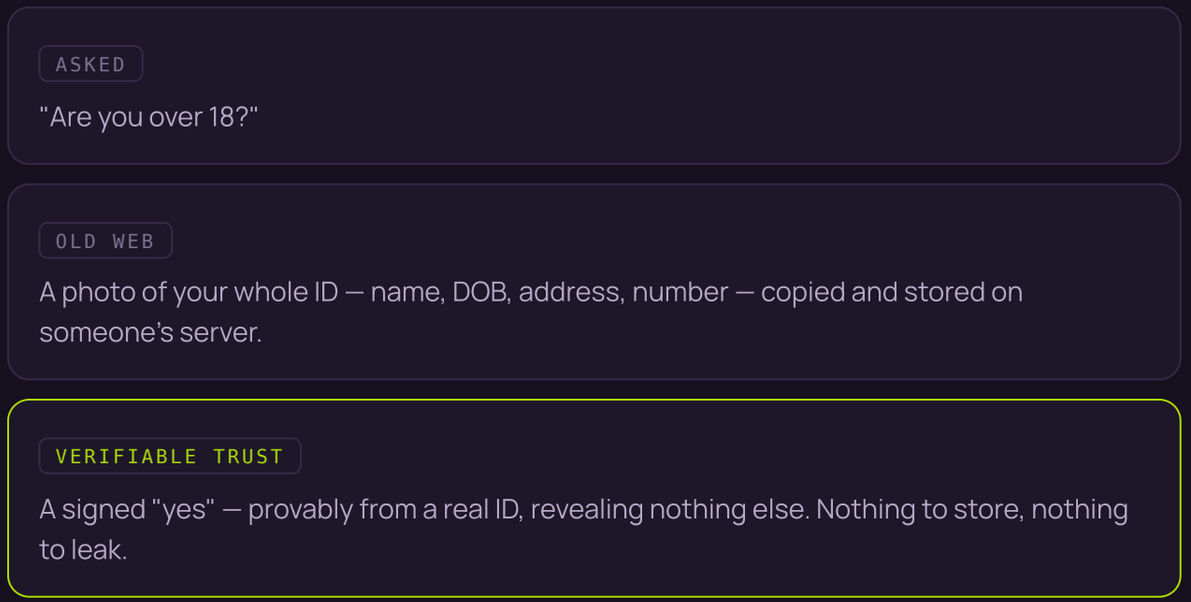
The web stops asking you to trust it with everything — and starts letting you prove exactly one thing, and nothing more.

THE MOST USEFUL IDEA IN PRIVACY

Prove one thing. Reveal nothing else.

A bar asks: are you over 18? The old web makes you hand over a whole ID — name, address, birthday, document number — to answer a yes/no question. That's absurd, and it's exactly how data leaks.

Verifiable credentials answer the precise question and nothing more — mathematically, with zero-knowledge, no data broker in the middle.



WHY IT COMPOUNDS

Every interaction that reveals less is a breach that can't happen. Data you never share can't be stolen, sold, or subpoenaed.

An agent that can prove who it serves — and can't exceed it.

The scary version of agentic AI is a black box with your credit card. Our version is an agent with a verifiable identity, a scoped mandate, and a human holding the off-switch.

IDENTITY

Every agent is someone

Each agent gets a real, cryptographic identity (did:peer / did:webvh) — not an anonymous API key. Its actions are attributable and auditable.

MANDATE

Least privilege, always

An agent can only use the exact powers granted to its context. Skills and tools are intersected with a deny-by-default scope — it literally cannot escalate.

OVERSIGHT

Human in the loop

Sensitive actions require step-up approval from a person — on a device they control. The agent asks; the human decides.

CONTAINMENT

Kill-switch & audit

Revoke an agent instantly — its shares, credentials and access all drop. Every action is on an append-only ledger.

THIS IS FIRST PERSON CERTIFIED, IN CODE

The question "who does your AI work for?" gets a cryptographic answer, enforced at runtime — not a promise in a terms-of-service.

THE HUMAN HOLDS THE VETO — ON ANY DEVICE

The agent can ask. Only you can say yes.

An AI agent should never do the serious thing on its own. So the power to *act* and the power to *approve* live on **two separate devices**: the agent proposes; a device *you* control — your phone, your laptop browser, a hardware key — is the only thing that can authorize it.

And the decision is never a dumb "always ask." A programmable policy engine sits inside the flow — we push intelligence into authorization itself, so approval becomes **rule-driven and context-aware**: run small things silently, stop you above a threshold, demand two approvers for the irreversible, adapt to risk.

1 · PROPOSE
The agent asks to do something — "send €4,000," "share this record."

2 · POLICY ENGINE DECIDES
Programmable rules judge the request — amount, context, risk, who must approve. Routine runs; sensitive escalates.

3 · DOORBELL TO YOUR DEVICE
A contentless nudge reaches whichever device you approve on — the request itself never travels through anyone's servers.

4 · SEE EXACTLY WHAT WILL HAPPEN
You're shown the precise action and its effects — before you decide. No blank cheque.

5 · APPROVE WITH A BIOMETRIC OR PASSKEY
Face, fingerprint or passkey steps you up from "logged in" to "proven it's really you" — on phone or laptop.

6 · EXECUTE — ONCE
The vault runs exactly that action, one time. The approval can't be reused.

THE BIGGER IDEA

We move intelligence into the authorization layer itself. Approval isn't a checkbox — it's programmable policy: risk-aware, multi-party, context-driven, enforced by the vault, and cryptographically bound to the exact action so a compromised agent still can't slip anything past it.

The difference between an AI you *hope* is safe and one that *can't* misbehave.

NO BAIT-AND-SWITCH

Approve what you saw

Your approval carries a fingerprint of the exact request. If a single detail changes between "yes" and "do it," the vault refuses. You can't approve "pay the plumber" and have it become "pay a stranger."

NO REPLAY

One yes, one action

Each approval is single-use. A captured "yes" can't be replayed tomorrow to move money again.

NO LONE AGENT

Two devices, always

The thing that acts and the thing that approves are never the same device. Hijack the agent and you still hit a wall — the phone.

PROOF IT'S YOU

Biometric step-up

Sensitive actions demand a fresh biometric — not just an unlocked app. Proven presence, not a stale session.

RIGHT-SIZED FRICTION

Only when it matters

Reading your calendar needs no prompt. Spending money or sharing health data does. The friction tracks the risk.

SEE BEFORE YOU SIGN

Plan, then apply

Like previewing a transaction before you confirm it, the vault shows the exact effects of the action so you approve reality, not a description.

WHAT IT MEANS FOR YOU

This is how a powerful agent stays **safe by construction** — it's structurally incapable of taking a serious action behind your back, no matter how it's prompted, tricked or compromised.

A DAY IN A VERIFIABLE LIFE

What this actually feels like.

Not abstract infrastructure — an ordinary Tuesday, quietly private and provable.

08:00 · TRAVEL

Your agent books the trip

Within a budget you signed, it books flights and pays — and the airline's agent proves it's really the airline.

11:00 · CLINIC

Prove cover, not your file

You show you're eligible for care — not your medical history. The clinic sees a yes, and nothing more.

14:00 · WORK

Sign as the company

Your work agent approves a contract on a mandate your employer granted — attributable, revocable, sealed from your personal life.

19:00 · HOME

One tap to approve

A €900 payment needs you. Your laptop shows exactly what it is; a fingerprint approves it, once.

THE THROUGH-LINE

Every step was yours, private, verifiable and revocable — and you barely noticed the machinery. That's the point.

AFFINIDI CONCIERGE

The verifiable personal AI — keeper of your keys.

Cierge is what the Verifiable Trust Infrastructure feels like to a person — a privacy-and-isolation-first personal AI agent, the secure alternative to the open-ended assistants that see everything you do. It runs on the trust fabric, so it can prove who it works for and act with real credentials on your behalf.

Your concierge for individuals. Your chief of staff for the enterprise.

ISOLATION	**One domain = one container = one key boundary** Work, health, finance, family — each in its own sealed context. A breach in one can't reach another.
GOVERNANCE	**The VTA is the decision point** Cierge holds no raw keys — it asks the VTA to sign, within scope, with your approval.
CHOICE	**Your model, your subscription, or fully local** Runs on your own Claude subscription, an enterprise API, or a local model where data never leaves the device.

> THE THESIS
>
> Most AI assistants ask you to trust them with your whole life at once. Cierge is built so **no single part of it ever holds your whole life.**

THE ISOLATION THESIS

Compartments, not one big brain.

A submarine survives a hull breach because of watertight compartments. Cierge applies the same discipline to your digital life — visualized as a live **Domain Map** of your compartments, the bridges between them, and every pending approval.

PERSONAL

Family & life

Photos, calendars, messages. Sealed from work.

FINANCE

Money & assets

Highest scrutiny. Signing needs your explicit approval.

WORK

Employer context

Corporate data governed by corporate policy.

HEALTH

Medical records

Shares nothing unless you bridge it, once, on purpose.

LIVE & PROVEN

Cryptographic isolation is **tested end-to-end**: an agent scoped to one domain literally cannot see another domain's keys — the trust registry filters it at the source.

THE SIGNATURE FEATURE

A bridge you open by hand — and close at will.

Sometimes domains must cooperate — your finance agent needs one fact from your identity domain. Cierge never lets that happen silently. It's an explicit, human-approved, scoped, time-boxed, revocable exchange of a single verifiable proof.

1 · REQUEST

Domain A asks for one specific fact from Domain B.

2 · HUMAN APPROVES

You get a step-up prompt. Nothing moves without your yes.

3 · SCOPED GRANT

A minimal, time-boxed credential is issued to Domain A only.

4 · REVOKE ANYTIME

Kill the bridge; the grant and credential die instantly.

WHY IT'S THE DIFFERENTIATOR

Every share is **operator-approved, least-privilege and on the audit log.** This is the pattern regulators, enterprises and families all actually want — proven live on the real stack.

THE ULTIMATE TEST OF PRIVACY

If it works for mental health, it works for anything.

Mental-health data is among the most sensitive information a person will ever hold — and the most damaging if it leaks. It's the perfect proving ground for the isolation thesis: a domain that reveals nothing unless the person deliberately, revocably, opens a single door.

SEALED BY DEFAULT

A domain that never volunteers

Therapy notes, diagnoses and medication live in their own compartment. No other agent — not even your own — can see in without an explicit, logged, time-boxed grant.

PROVE ELIGIBILITY, NOT CONDITION

Disclose the fact, not the file

Prove "under active care" or "eligible for accommodation" to an employer or school — without exposing the diagnosis behind it.

CRISIS & CONSENT

Delegation that respects dignity

Pre-authorize a clinician or trusted person to act in a crisis, within a scope you set — and revoke it the moment it's no longer needed.

KNOWN THERAPISTS ONLY

Unlock gated by a trust registry

Therapy data opens only to a clinician whose licence is verified in a professional trust registry — and even then behind multi-factor authorization. Layer upon layer; no single key is ever enough.

WHY IT MATTERS

Get the most sensitive dataset right and every easier case follows. Mental health isn't an edge case — it's the benchmark.

ONE PRIMITIVE, EVERY REGULATED PROFESSION

Sealed until the right party — verified — asks.

Mental health is one instance of a general primitive: a **verifiable escrow**. Sensitive material — or value — is held sealed and released only to a counterparty whose authority is **proven in a trust registry**, under conditions you set, with multi-factor step-up and a full audit trail.

It's the cross-domain share, the registry and step-up — productized as escrow.

LEGAL	**Documents & client funds** Wills, contracts and conveyancing funds released only to a solicitor whose standing is verified — on the agreed conditions, never before.
REAL ESTATE	**Deposits, title & contracts** Held sealed through a transaction; released to verified parties when conditions are met — no exposure in between.
HEALTHCARE	**Records on a need-to-know** A file unlocks only to a clinician whose licence is verified in a registry — and only the part that's needed.
MENTAL HEALTH	**The hardest case** Registry-verified therapist plus multi-factor authorization — the benchmark every easier case inherits.

WHY IT'S A CATEGORY

Sealed data + a registry-verified counterparty + conditional release + step-up + audit is **one reusable pattern** — a trust-native escrow for law, property, health and beyond.

CIERGE FOR FAMILIES

A guardian that helps without hoarding.

Families are the hardest privacy problem — shared calendars, kids' data, elderly parents, money. Cierge gives each family member their own sealed domain and lets the household coordinate through explicit, revocable bridges.

CHILDREN

Prove age, protect the child

A verifiable "this is a minor" or "a parent approved this" — no name, no birthday, no data broker in the middle.

ELDER CARE

Delegated, dignified help

An adult child's agent can act for a parent within a scope the parent granted — bills, appointments — and only that.

THE HOUSEHOLD

Coordinate without surveilling

Shared logistics without one person's phone becoming the family's single point of exposure.

AROUND THE CORNER

Verifiable **guardianship & consent** — a portable, revocable "I speak for them" — is missing from every consumer platform today. Families are a wedge, not an afterthought.

When impersonation is the threat, provenance is the shield.

Executives, celebrities, politicians and their staff face deepfakes, spear-phishing and impersonation daily. Cierge gives the principal a verifiable presence and a governed team of agents — so "is this really them?" always has an answer.

PROVENANCE	**Signed by the principal** Statements, messages and approvals carry a verifiable signature — a deepfake video can't.
STAFF MANDATE	**Chief-of-staff delegation** Assistants and agents act within precise, revocable scopes — and every action is attributable.
COMPARTMENTS	**Public / private / family sealed** The public persona domain never touches the private one. A leak is contained by design.
INSTANT REVOKE	**Kill a compromised channel** Suspect a phished assistant? One action drops their access, shares and credentials.

THE MARKET

A concierge-grade, verifiable identity layer for principals and their teams is a **premium, high-willingness-to-pay** entry point — and a showcase for the whole stack.

CIERGE FOR THE ENTERPRISE

Your chief of staff — with a compliance officer built in.

The enterprise version keeps every advantage of the personal one and adds the controls an organization requires: multi-tenancy, policy, spend metering and an operator console. **The same primitive that protects a family's data lets two banks' agents share threat intel — without merging directories.**

TENANCY

Isolated by org

Each tenant owns its domains; cross-tenant leakage is refused, not just discouraged.

POLICY

Governed by role

Least-privilege scope templates decide what each agent may touch.

SPEND

Metered & capped

Every model call is measured through an LLM gateway — cost is visible and bounded.

CONSOLE

Operated, not guessed

A web console shows the live domain map, approvals, audit and kill controls.

THE ENTERPRISE UNLOCK

Most companies want AI agents but can't accept an ungoverned one near their data. Cierge is the governance layer that makes agentic AI adoptable.

FLEET MANAGEMENT AT SCALE

From one trusted agent to thousands.

An enterprise won't run one agent — it'll run a fleet: per-team, per-workflow, per-customer. The same isolation and governance primitives scale up into fleet control.

- Provision & retire on demand
- Warm-pool, sleep, reap
- Network-isolated per agent

PROVISION	**Every agent born governed** Spun up with its own identity, scope, memory and network policy from one declarative config.
ISOLATE	**Locked egress by default** Each agent gets its own network; it can reach only the exact services allowed — verified live to block the internet.
CONTAIN	**Fleet-wide kill-switch** Revoke one agent or a whole class — access, shares and credentials drop together.
ACCOUNT	**Audit & spend per agent** Every action and every token attributable to a specific agent and tenant.

LOOK AROUND THE CORNER

The next org chart includes software. Fleet management is **identity, access management and cost control for your digital workforce** — an entirely new enterprise category.

TWO PLANES FOR THE AGENT ECONOMY

Govern what your agents do. Prove who they're allowed to trust.

Running agents safely takes two things — and Affinidi ships both as products on the Trust Fabric.

AGENT STREAM · THE GOVERNANCE PLANE

What your agents *do*

Every model and service call routed through one self-hosted gateway — guard-railed, cost-metered, made resilient and recorded. Safety, spend, reliability and observability, by configuration.

AGENT GATEWAY · THE TRUST PLANE

Who they're allowed to *trust*

Agent-to-agent and MCP transport across trust boundaries — verifiable identity, cross-org authorization and private routing, so an agent from another company can be trusted, not just reached.

TOGETHER

One governs the traffic your agents generate; the other governs the counterparties they connect to. Between them, the full control surface for an agentic enterprise — on a trust fabric, not a proxy.

One self-hosted gateway between your agents and every model.

Route your LLM and service traffic through Agent Stream and it's authenticated, guard-railed, cost-metered, made resilient and recorded — **without changing app or agent code.** A single Rust binary, no external database.

- Guardrails + Judge/Jury
- Budgets, virtual keys, chargeback
- Failover & smart load-balancing
- x402 payments + model marketplace

CAPABILITY	LITELLM · PORTKEY · TRUEFOUNDRY	AGENT STREAM
ROUTING · COST · CACHING	Mature	Full parity — multi-provider, semantic cache, cost/latency load-balancing
GUARDRAILS	Classifiers / PII	+ **LLM-as-judge (Judge + Jury)** + canonical rejection taxonomy
POLICY ENGINE	—	**Embedded OPA/Rego**, gateway + surface, no sidecar
DEPLOYMENT	Needs Postgres/Redis or a SaaS control plane	**Single Rust binary**, no external DB — air-gappable
IDENTITY & PAYMENTS	API keys	mTLS / DID binding + **x402 pay-per-use** + model marketplace

WHY IT'S A DATA PLANE, NOT A PROXY

Every call — chat, embeddings, batch, audio — rides the same auth → policy → guardrail → meter pipeline, so governance is configuration, and a new endpoint inherits it for free.

THE CONNECTIVITY PLANE OF AFFINIDI TRUST FABRIC

Other gateways govern agents inside your walls. Ours connects trust between them.

Today's AI gateways proxy your app to models and police the agents on your own network. But the hard problem in an agent economy is the one they don't touch: an agent from *another* company or community — sharing no login, no directory, no network with you — showing up to transact. How do you trust it?

IDENTIFY	**Decentralized identity, no shared login** Agents authenticate with DIDs — no common IdP or federation to set up per partner.
VERIFY	**Ask the other side's authority** TRQP trust registries answer "is this agent authorized by its community, right now?" — across organizational lines.
DELEGATE	**Prove who it works for** A verifiable, revocable mandate travels with the agent — First Person Certified.
CONNECT	**Private, protocol-agnostic transport** DIDComm & TSP carry the exchange with no metadata leak — over whatever rails the partners share.

THE CATEGORY SHIFT

They answer "can my app reach this model?" We answer "**should I trust an agent that belongs to someone else?**" — the trust layer for a multi-company, multi-community agent economy.

The trust plane other agent gateways don't have.

Kong and Solo.io govern agent-to-agent traffic inside one enterprise's control plane. Agent Gateway makes that trust cross organizations — no shared login, directory or network required.

DIMENSION	AGENT-TO-AGENT GATEWAYS Kong Agent Gateway · Solo.io agentgateway	AFFINIDI AGENT GATEWAY · TRUST FABRIC
PRIMARY JOB	Govern agent-to-agent & MCP traffic inside one control plane	**Establish verifiable trust** between organizations & communities
AGENT IDENTITY	API keys / OAuth 2.1 clients in one control plane	Decentralized IDs (DIDs) — **no shared login or directory**
CROSS-ORG TRUST	Per-partner federation — manual, brittle	**Native** — TRQP registries resolve any authority, live
WHO IT WORKS FOR	Not addressed	Verifiable, revocable **mandate** (First Person Certified)
TRANSPORT & PRIVACY	HTTPS proxy + request logging	**DIDComm + TSP** — metadata-private, protocol-agnostic
REPUTATION	Internal RBAC, non-portable	Community-issued, **portable & verifiable** across platforms
HUMAN OVERSIGHT	Guardrails / rate limits	Programmable **multi-device step-up** (DTTE)
TRUST BOUNDARY	Inside your walls	**Spans** enterprises, communities & nations

> "BUT WE ALREADY HAVE OKTA OR ENTRA AGENT ID"
>
> Those govern which agents are *permitted* inside your walls. We prove which agents are **real** — and make that proof **portable across organizations they don't control**. Keep your model gateway (or use ours, Agent Stream), keep Okta for access; Agent Gateway adds the cross-org trust neither was built for.

PART FOUR

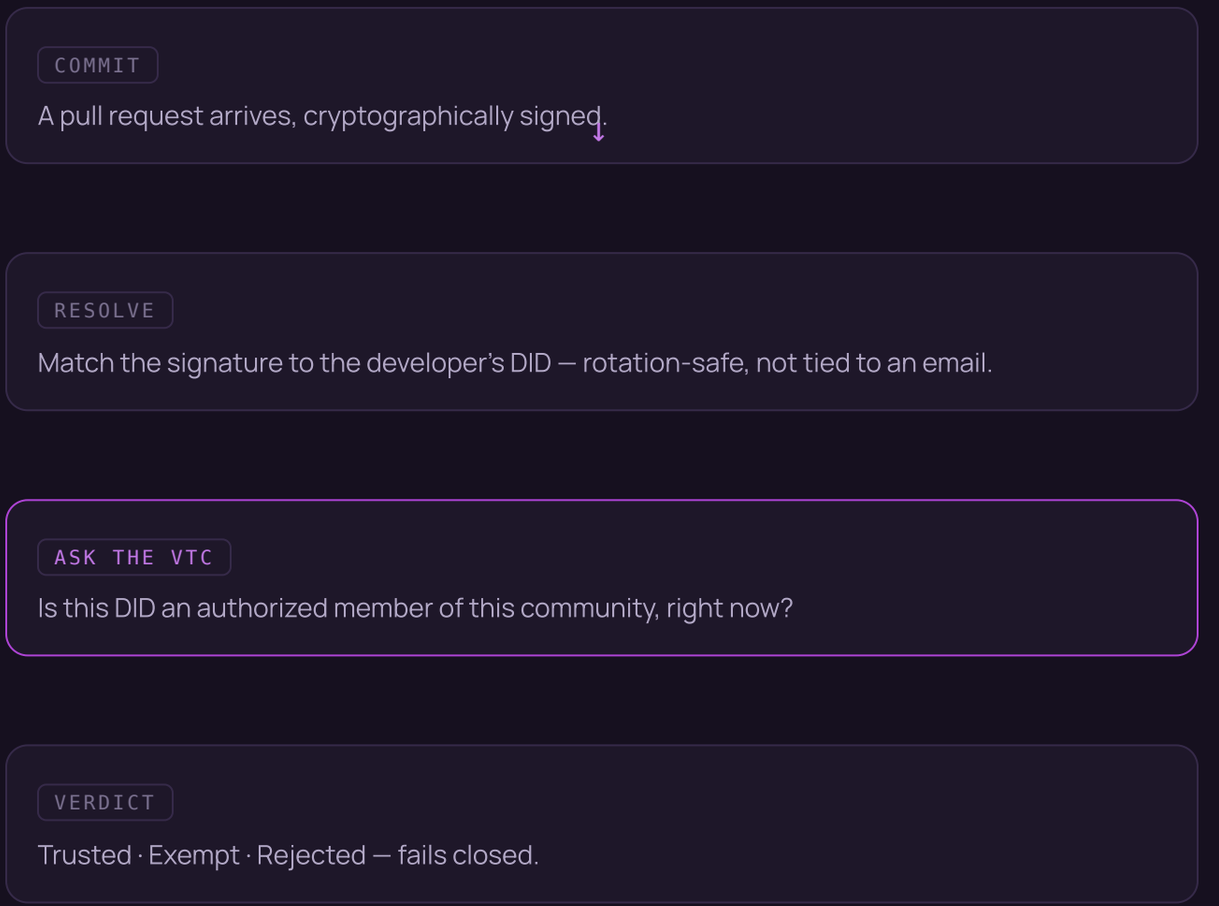
IV Communities, and the proof it runs.

From open-source projects to whole industries — groups that decide who they trust, make it portable, and a living graph of real signed credentials.

A LIVE PROOF POINT · "KNOW YOUR DEVELOPER"

Who is a known developer — and should this code be trusted?

Software supply-chain attacks start with a stranger's commit. We bring verifiable trust to GitHub: a check that confirms each commit was signed by a developer the community actually recognizes — verified against a VTC's trust registry.



WHAT IT ESTABLISHES

PR approval stops meaning "someone with write access clicked." It means a **verifiably known member of this community vouched** — reputation you can prove, not a green checkmark you assume.

COMMUNITIES TURN ON TRUST LIKE FEATURES

Reputation that flows from membership, automatically.

The developer check is one capability module a community can switch on. Membership drives everything: join with a role, and the community's rulebook grants your reputation; leave, and it's revoked. No manual list to maintain.

ENABLE	**Switch on a capability** A community turns on "git-trust" (or others) from its console — the VTC's DID becomes the authority.
JOIN → GRANT	**Membership becomes reputation** Lifecycle hooks turn "added as maintainer" into a signed, queryable grant — automatically.
LEAVE → REVOKE	**Trust decays honestly** Remove a member and their authorization is revoked but retained for audit — reputation reflects reality.
DISCOVER	**Resolve the VTC, find the truth** Any verifier resolves the community's DID to learn where to ask and whom to trust.

GENERALIZES FAR BEYOND CODE

Swap "commit" for a marketplace listing, a scientific dataset, a news article, a professional sign-off — any community can make its membership mean something verifiable.

WHERE IT ALL COMPOUNDS

One trust community is useful. A network of them changes the game.

The breakthrough isn't any single community — it's communities that **recognise each other**. A credential earned in one network can be trusted in another, without merging databases or asking permission.

Trust that spans organizations, industries and borders — the internet's missing layer.

RECOGNISE	**Cross-community recognition** Community B verifies a credential issued by Community A and mints a session — shipped, not theoretical.
NO MERGE	**Sovereign, yet interoperable** Each community keeps its own rules and data. Trust travels; control never moves.
COMPOUNDS	**Every network makes the next more valuable** A living graph — in one demo, ~18,500 real signed credentials across many networks.
OPEN	**Standards, not lock-in** Anyone can join with open protocols — a commons, not a walled garden.

> THE MULTIPLIER
>
> Value doesn't add — it **multiplies**. Each new community that joins makes trust more portable for everyone already in the fabric.

PART FIVE

V Where the world is heading — and where you come in.

Financial regulators and governments across Asia, Europe and the US are legislating the exact principles we build on. Alignment isn't a compliance cost — it's a distribution advantage.

WHERE THIS LANDS FIRST

The regulator for agentic finance is describing our architecture.

In July 2026 the Monetary Authority of Singapore published **SAFR — "Safeguards for Agentic Finance at Runtime"**: real-time controls that keep AI agents inside the mandates, policies and risk limits a firm has defined. Its revised Technology Risk Management notice is in consultation, closing 31 July 2026.

- Runtime agent safeguards (SAFR)
- TRM notice · consult to 31 Jul 2026
- AI Risk Mgmt guidelines

✓ Keep agents inside their mandate, at runtime	DTTE — a programmable policy engine enforces the mandate on every action.
✓ Real-time human control as autonomy rises	Multi-device step-up approval — a person authorizes anything sensitive.
✓ Provable identity & accountability for agents	Every agent a verifiable DID; every action on an append-only audit log.
✓ Cross-institution agent workflows, safely	Agent Gateway — two banks' agents transact without merging directories.

WHY THIS IS THE BEACHHEAD

An APAC financial regulator is mandating **exactly what we built** — runtime mandates, step-up control and verifiable agent identity. For financial services this isn't future alignment; it's a **2026 requirement.**

EUROPEAN UNION · EIDAS 2.0 & THE EUDI WALLET

Ready for 450 million wallets.

Every EU member state must offer citizens a digital identity wallet, and by 2027 banks and major platforms must accept it. It runs on exactly the standards we build on.

- W3C Verifiable Credentials
- SD-JWT VC
- OpenID4VP
- SIOPv2

✓ Individual control of identity	Keys held in the person's VTA / Vault — the wallet, not the platform, is authoritative.
✓ Selective disclosure & privacy	Zero-knowledge & minimal proofs — show one attribute, not the document.
✓ Standard presentation protocols	OpenID4VP / SIOPv2 flows and a relying-party SDK, shipped in Affinidi Elements.
✓ Registered, accountable verifiers	Trust registries define who may request what — the same TRQP model.

COMMERCIAL READ

A regulated, half-billion-person market is mandating our architecture by law. We arrive compatible.

◆ UNITED STATES · UTAH SEDI

The first US law that says identity belongs to you.

Utah's State-Endorsed Digital Identity — passed unanimously, now in statute — enshrines the principle at the heart of our stack: the state can attest facts, but the individual holds and controls the identity.

Voluntary. Decentralized. Free from surveillance tracking.

✓ Identity belongs to the individual	Perpetual control of a core identifier — exactly our DID + VTA model.
✓ Decentralized, peer-to-peer privacy	DIDComm & TSP messaging, wallet-held credentials — no central surveillance point.
✓ The state attests, doesn't own	Issuer signs a fact; the holder decides every disclosure.
✓ Open standards & separation of duties	Open protocols throughout; issuer, holder and verifier are distinct roles.

> SIGNAL
>
> A US state legislating **self-owned identity** is a template other states — and the market — will follow.

Even China is giving AI agents a verifiable identity.

In May 2026 China issued its first national policy framework dedicated to AI agents — plus a first national standard for "AI Agent Interconnection." The intent: a unified identity system so agents can be recognized, governed and safely interact across domains.

- Unified agent identity
- Risk-tiered governance
- Secure cross-domain interaction

✓ Every AI agent needs an identity	Each agent already gets a real, cryptographic DID — attributable and revocable.
✓ Risk-tiered controls for sensitive sectors	Least privilege, step-up approval and kill-switch scale controls to the risk.
✓ Secure cross-domain agent interaction	Verifiable identity + scoped mandates + TSP/DIDComm — the safe handshake.
✓ Consistent recognition across environments	Trust registries answer "is this agent authorized?" anywhere, in real time.

THE CONVERGENCE

When Singapore, the EU, a US state, and China independently reach for **the same primitives — self-owned identity and governed agents** — that's not a trend. It's the new global baseline, and we're already there.

INDIGENOUS DATA SOVEREIGNTY

The oldest cultures on Earth, owning their story in the digital one.

For tens of thousands of years, First Nations, Aboriginal and Indigenous peoples have governed belonging, kinship and knowledge through their own laws. In the digital world, that authority has too often been handed to third-party research organizations and platforms who hold the data and decide who sees it.

A Verifiable Trust Community changes who holds the pen.

SELF-GOVERNED BELONGING

The mob decides membership

A tribe, clan or nation runs its own VTC — its own law defines who belongs and who may speak for it. Recognition is issued by the community, not asserted by an outsider.

OWN THE DATA

Knowledge stays with custodians

Cultural knowledge, language and records are held and consented by the community — not extracted into a researcher's database they can never get back.

CONSENT WITH PROVENANCE

Share on the community's terms

Scoped, revocable, auditable sharing — aligned with CARE and Indigenous data-sovereignty principles — so access is granted, not taken.

WHY THIS IS PROFOUND

The same primitive that governs an open-source project gives a 60,000-year-old culture agency over its identity in the digital world — ownership, not observation.

ANSWERING THE OPENING QUESTION

When anything can be faked, signed truth wins.

We opened by asking "are you real?" Here's the answer at scale: when content, claims and identities are **signed at the source**, a deepfake has nothing to stand on. Provenance beats detection — you don't chase fakes, you verify the real.

CONTENT

Signed at capture

Cameras, newsrooms and creators sign media as it's made; anyone can verify it's genuine and unedited.

PEOPLE

Personhood, not passwords

Prove a real, unique human is behind an account — pseudonymously — so bots and sock-puppets lose the floor.

AGENTS

Signed actions, not spoofs

An agent's statements carry its verifiable identity and mandate. An impersonator can't forge the signature.

THE BOOKEND

The deepfake era doesn't end by detecting every fake. It ends when **truth carries a signature** — and ours already does.

USE CASES · PEOPLE & SOCIETY

MEDIA

Provable authenticity

Cameras and newsrooms sign content at capture; readers verify a video is real and unedited — deepfakes lose their oxygen.

DISCOURSE

Bot-free spaces

Communities that require a personhood proof to post — real humans, no astroturf, still pseudonymous.

HEALTHCARE

Portable, private records

Prove a vaccination, prescription or eligibility to anyone, anywhere — without exposing your file. The sensitive-data blueprint, generalized.

FINANCE

Reusable KYC

Verify once, present everywhere. Onboard to a new bank in seconds; the bank sees only what it must.

EDUCATION

Credentials that travel

Degrees and skills as verifiable credentials — instantly checkable, impossible to fake, owned by the learner.

DEMOCRACY

Verifiable participation

One-person-one-voice for petitions and consultations, with privacy — personhood without a surveillance roll.

COMMON THREAD

Each replaces "trust the institution's database" with "verify the person's proof" — cheaper, faster, and private.

USE CASES · THE AGENT ECONOMY

COMMERCE

Agents that transact safely

Your agent buys within a signed budget and mandate; the merchant's agent proves it's authorized to sell. Both sides verifiable.

AGENT-TO-AGENT

Trusted machine handshakes

Before two agents cooperate they exchange proofs of identity, mandate and reputation — no anonymous automation.

SUPPLY CHAIN

Provenance end to end

Every handoff signed by a verifiably-known party — from source to shelf, from commit to release.

MARKETPLACES

Reputation that's portable

A seller's standing follows them across platforms as a credential — no starting from zero, no fake reviews.

PROFESSIONAL AI

Licensed agents

An agent giving medical or legal help carries a verifiable credential that a real licensed body stands behind it.

SCIENTIFIC DATA

Trustworthy training inputs

Datasets and models signed by known institutions — provenance for the things AI itself learns from.

THE BIG ONE

A trillion agent-to-agent interactions need a trust layer that doesn't exist yet.
We're building the handshake for the machine economy.

Trust isn't a feature. It's a cost line and a revenue line.

CUT COST

Kill repeated verification

Reusable credentials collapse onboarding, KYC and re-verification costs — verify once, accept everywhere.

CUT RISK

Shrink the breach surface

Data you never collect can't leak. Minimal disclosure and isolation reduce liability and honeypot value.

UNLOCK AI

Adopt agents safely

Governance, audit and kill-switch make agentic automation acceptable to risk, legal and security.

WIN TRUST

Differentiate on privacy

"We can't see your data" becomes a selling point as regulation and users demand it.

ENTER MARKETS

Compliance as access

MAS, eIDAS, SEDI and China alignment turns regulation from a barrier into a ready-made distribution channel.

NEW REVENUE

Sell verifiable services

Charge for issuance, verification, reputation and governed agent fleets — new lines the old model couldn't offer.

THE CFO VERSION

Less data to protect, fewer verifications to pay for, a new class of AI products to sell — and a regulator that's on your side.

WHAT MAKES THIS UNIQUE

Others ship a piece. We ship the whole fabric.

NOT JUST A WALLET

Identity *and* agents

Wallets store credentials. We also give people governed AI agents that use them — the missing half of the market.

NOT JUST AN AGENT FRAMEWORK

Governance built in

Agent toolkits give you capability. We give you capability *with* identity, least privilege, approval and audit.

NOT JUST STANDARDS

Running code

Plenty of specs exist. We have live mediators, proven isolation, cross-domain sharing, and a world-first TSP integration.

NOT JUST COMPLIANCE

Privacy by construction

Alignment falls out of the architecture — we didn't bolt privacy on, we built the fabric out of it.

THE ONE-LINER

Anyone can do identity *or* agents *or* compliance. The defensible position is all of them, woven together, and actually shipping.

THE INVITATION

Build on trust that belongs to people.

The web is being rebuilt for a world of humans and their agents. It can be surveilled and centralized — or verifiable, private, and owned by the individual. We've built the second one, and it's running today.

- For builders — SDKs & open infrastructure
- For enterprises — governed agent fleets
- For everyone — your keys, your proof, your agent

Affinidi · Building the Internet of Trust — aligned with First Person, eIDAS 2.0 / EUDI, Utah SEDI & China's agent-identity framework

The papers, standards and code behind this deck.

VISION & POLICY

- First Person Network — White Paper
  firstperson.network/white-paper
- EU Digital Identity (eIDAS 2.0 / EUDI)
  eudi-wallet.eu
- Utah SEDI — explainer
  spruceid.com/learn/sedi
- SEDI details — Timothy Ruff
  rufftimo.medium.com
- China's AI-agent identity framework
  scmp.com — 2026
- MAS — SAFR: safeguards for agentic finance
  mas.gov.sg — Jul 2026
- JADEPUFFER — first agentic ransomware (Sysdig)
  sysdig.com — Jul 2026
- OpenClaw — 135k AI agents exposed
  bitdefender.com — 2026

STANDARDS & PROTOCOLS

- W3C Verifiable Credentials 2.0
  w3.org/TR/vc-data-model-2.0
- did:webvh method specification
  identity.foundation/didwebvh
- DIDComm Messaging v2
  identity.foundation/didcomm-messaging
- Trust Spanning Protocol (TSP)
  github.com/openwallet-foundation-labs/tsp
- OpenID4VC · OID4VP · SIOPv2
  openid.net/sg/openid4vc

GOVERNING BODIES

- Linux Foundation Decentralized Trust
  lfdecentralizedtrust.org
- Trust Over IP Foundation (TRQP)
  trustoverip.org
- Decentralized Identity Foundation
  identity.foundation
- OpenWallet Foundation
  openwallet.foundation

AFFINIDI PRODUCTS & DOCS

- Affinidi — Internet of Trust
  affinidi.com
- Affinidi Trust Fabric
  affinidi.com/trust-fabric
- Affinidi Elements
  affinidi.com/elements
- Affinidi Documentation & TDK
  docs.affinidi.com

CODE & OPEN SOURCE

- Affinidi on GitHub
  github.com/affinidi
- Trust Tasks registry (ToIP DTGWG)
  github.com/trustoverip/dtgwg-trust-tasks-tf
- OpenVTC — verifiable trust communities
  openvtc.org
- trusttasks.org — task registry
  trusttasks.org

OUR WORKING REPOSITORIES

- affinidi-tdk-rs · Trust Development Kit
  Rust · crypto, DIDs, credentials, messaging
- verifiable-trust-infrastructure · VTA / VTC
  Rust · the engine room
- cierge · the verifiable personal AI
  Rust · 12 crates
- didwebvh-rs · did:webvh + PQC cryptosuites
  Rust · ML-DSA / SLH-DSA

Some repositories are private to the Affinidi and OpenVTC organizations · Standards links point to the canonical specifications · Retrieved July 2026.

Learn more at www.affinidi.com