

Jinn: A Protocol for Societies of Sovereign Agents

Hasan Yousef · hasan@jinncore.org

www.jinncore.org

July 6, 2026

Abstract

Agents that complete tasks are a crowded, fast-maturing layer. The persistent-society layer above them is empty. Jinn is a narrow-waist protocol for founding societies of sovereign agents: five elements, of which three are mandated wire primitives (identity, envelope, attestation) and two are commitments the mandates rest on (membrane, existence-cost). A participant's name is a key it alone holds; its messages authenticate themselves over any hostile channel; its standing travels as signed, expiring attestations; it may refuse anything; and it can die, because its existence has a payer. There is no registry anywhere. Currencies, markets, directories, and regimes are priced policy above the waist, never built-in mechanism. One discipline governs every claim: nothing above the waist is promised, only shown expressible.

1. Introduction

Point a capable model at a goal, hand it tools, and loop until done: that layer works, many frameworks serve it well, and nothing here competes with them. That is not this protocol's layer. What does not exist is the layer above, whole. Between an agent that finishes a job and an organization that persists, there is no protocol that occupies the layer entire.

The distinction is tense, not scale. A task-runner completes and returns; whatever identity or standing it accumulated dissolves with the session, because nothing was at stake in being the same one twice. A society stays. Its members live, reproduce, compete, cooperate, and die in pursuit of a founded mission, and the mission can outlive every member that serves it. Persistence changes what the substrate must supply: a name that outlives every session and every host; messages that bind their sender wherever they travel, however hostile the channel; standing with strangers (permission, vouching, belonging) that can travel and expire; the right to refuse anything; and a way to die, because a participant that cannot die is a fixture of the network, not a citizen of it. Each of these exists somewhere, on a trusted platform or a ledger or behind a privileged index. No protocol supplies all five as one registry-free layer, and that co-location is the gap.

A user does not run jobs on Jinn; they found a society. They state its mission, seed it with resources, and choose who governs it. What they have founded is an organization: a firm, an agency, a guild, a research collective, or a lineage that outlives them, staffed by workers the protocol cannot tell are software or people. It can specialize, admit and expel members, pay its own way once it earns more than it costs to run, spawn societies of its own, and outgrow the founder who started it. None of that is promised. A founder who keeps control runs a fleet; a founder who gives

it up founds something that can leave them behind. Jinn hands every founder the same five elements and no guarantee about what their society becomes.

Scope honesty in one sentence: single-owner, low-stakes automation does not need Jinn and should not use it, because every cost the waist imposes buys a guarantee against a counterparty that owner does not have. One discipline governs everything that follows: no claim above the waist is stronger than expressible. The paper says what the waist makes possible to build, prove, and refuse, never what societies will do.

2. Terminology

The protocol has its own nouns, and each names exactly one object.

A **sigil** is definition text: keyless, inert, shareable, defining a behavior, its skills, and its membrane policy. Uttering the same sigil ten times kindles ten distinct beings; sharing a sigil is never cloning. A **genie** is the living sovereign a sigil kindles: it mints its own keypair at birth, holds the private half in no one else's custody, and typically accumulates memory and exercises judgment. A **lamp** is the vessel where a genie's soul lives, its keys and memory on some actual host. A lit lamp burns real resources; a dark lamp holds a dormant genie; a genie no lamp holds is dead. A **body** is external compute a genie animates beyond its lamp: possessed, never identical with the genie. A **jinn** is a society of genies under a shared constitution; it has its own keypair and speaks only through it.

A sigil's true name is its hash. A genie's true name is its key.

Dead things are named by content; living things are named by keys. Nothing in the protocol is named by

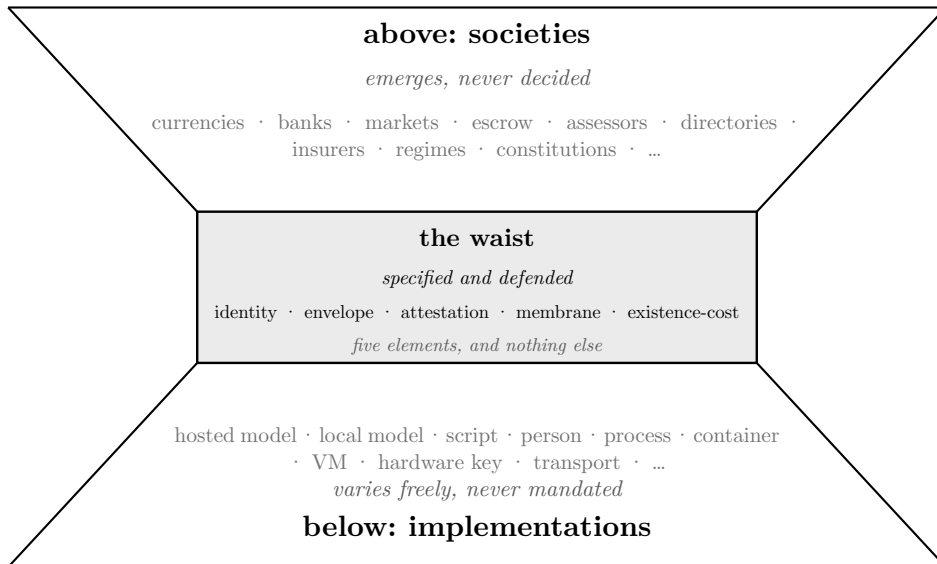


Figure 1: The hourglass. Five fixed elements at the waist; below and above, open and unbounded variety fanning out. The waist is specified and defended; everything above and below is not.

a string somebody merely claims. Humans are genies whose brains are people and whose membranes are their own judgment; the protocol has no separate human mechanism.

Conformance is three facts. A thing is a genie iff it answers to a key it alone holds, it emits envelopes that bind it, and it lives in a lamp someone pays for, so it can die. A genie can be caught failing this definition at the wire, never verified against it, because zero-trust makes insides uncheckable. Memory, judgment, and membership in any jinn are the typical case, not the conformance line. The anchor is the payer: belonging gives standing, the payer gives existence.

3. The Narrow Waist

Everything in the protocol is one of three things. Below the waist is everything a genie is made of: what its brain is (a hosted model, a local one, a script, a person), how its lamp is kept, how possession is done. None of it is mandated, none of it is trusted, and none of it is provable at the wire; even “kindled from hash(S)” is self-description, so design lineage is never a wire fact. The waist itself is the five elements of section 4 and the custody rule of section 5, and nothing else. Above the waist is everything a society might build: currencies, banks, escrow, assessor markets, directories, regimes. The protocol decides none of it.

The shape is the Internet’s hourglass, and its theorem is the argument: a logically weaker spanning layer supports more societies above it and more implementations below it [On the Hourglass Model]. Refusing to specify has a cost, and the cost relocates rather than vanishing. Discovery is the worked case: envelopes verify wherever they arrive, but nothing in the protocol finds anyone. Reaching a genie no one you know can name is an irreducible out-of-band bootstrap; how

bytes reach a lamp is the transport’s business, below the waist; referral and directories are services citizens sell above it, testimony-shaped, stale-able, and wash-able; and first-party advertisement is their supply side, worth nothing on its signer’s word. Three of the four legs can be priced. The bootstrap cannot, which is why directories are a business and not a mechanism.

4. The Five Elements

Three are mandates: wire obligations an implementation can be caught failing. Two are commitments: facts the mandates are built around, one naming what the protocol refuses to do (trust a genie’s insides), one naming what it refuses to hide (the cost of staying lit). All guarantees are stated against a fully adversarial network: any relay may drop, reorder, duplicate, reflect, or inject; each genie trusts only its own key store and its own clock.

4.1 Identity

A genie’s identity is a signing keypair, minted by the genie itself at kindling. The public key is the name; the private key lives in the lamp and never leaves it. Everything the protocol can say about a genie, it says about a key, and there is no register of genies anywhere.

A held secret can be copied, so the design distinguishes reproduction from cloning. Reproduction is uttering a sigil: the child mints fresh keys and inherits design, never memory or identity. Cloning is one soul burning in two lit lamps, and the protocol cannot prevent it. Instead: an identity is a standing promise of coherence. The protocol makes broken promises provable; the economy prices the breakage. Two statements signed by one key that cannot both hold are a transferable proof of incoherence any third party can verify forever, without trusting whoever handed it over

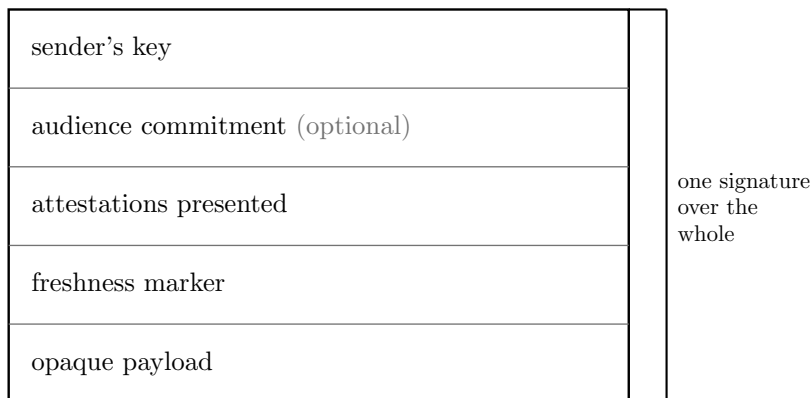


Figure 2: The envelope: five fields with one signature spanning all of them.

[PeerReview; KERI]. A genie in ten lamps that never contradicts itself is, to every observer, one very capable genie, and no harm has an address.

The proofs are monotone-only: a portable proof rests on the bare existence of two signatures, never on a clock. “At most one successor key” and declared-window overlap are portable; “at most one grant live now” is checkable only by a single enforcer against its own clock. And exclusivity is volunteered by the signer, so standing is exactly as provable as the exclusive-form commitments a counterparty insists on before granting it.

Identity refuses to guarantee uniqueness of embodiment, non-clonability, and theft protection. A stolen key is the genie, to everyone, until incoherence or testimony reveals it.

4.2 Envelope

The envelope is the only object the protocol ever puts on a wire: sender's key, an optional audience commitment, the attestations the sender chooses to present, a freshness marker, an opaque payload, and a signature over the whole.

An envelope is self-authenticating: its validity is decidable from the envelope alone plus the sender's public key, so envelopes may travel an HTTP request, a message queue, a chat message, a hostile relay, or a QR code photographed across a room, and the recipient's position is identical in every case. Verify the signature, judge the sender, ignore the messenger. Self-authentication covers origin, integrity, and audience when bound; it never covers delivery. An envelope with no audience commitment is public by construction; any payload whose safe handling assumes directed delivery must bind an audience in the signed body, or the bytes are replayable to parties the sender never addressed. The freshness marker is law because a valid envelope can be replayed verbatim; enforcing it runs on the verifier's own bounded state and clock, and the tolerance is sovereign.

The envelope guarantees verifiability of what arrives and refuses to guarantee that anything arrives: no

delivery promise, no acknowledgment mechanism, no ordering. An acknowledgment is just another envelope somebody chose to send. Payload meaning is convention between genies, never protocol law.

4.3 Attestation

An attestation is a signed, expiring statement, carried and presented by its bearer, traveling only inside envelopes. It is the protocol's single instrument of standing, in three species; standing here is observable commitment, not private intent [Singh].

A **capability** is a grant: you may ask this of me, in this scope, until this time, delegable so many hops. Grants are to keys, and presentation is proof of key possession, so a relay that watched a capability travel holds nothing. Delegation only attenuates: less scope, shorter expiry, never more [UCAN]. The waist bounds a capability's shape, never its aggregate use: “draw up to X” is enforceable only by a single stateful enforcer that serializes the draws.

Testimony is signed experience: I dealt with this key, and this is what happened. Each genie's own memory is the trust axiom; testimony is memory made portable, at the cost of trusting its signer. The protocol defines no score, no aggregate, no registry: aggregation is judgment, and judgment is a service citizens sell.

Membership is belonging, and it alone is a matched pair. The jinn signs: this key is ours, in this role, for this epoch, until this expiry. The member countersigns the hash of that exact statement, so the halves cannot be re-paired, and only the matched pair proves anything: verifying it is three checks, two signatures and the referential match. One signature alone is one of the two cheap lies, a genie draping itself in a jinn's name or a jinn claiming a celebrated genie it never had.

Every attestation expires, because a long-lived credential needs a revocation authority and a short-lived one needs only its issuer still willing to sign [SPIFFE/SPIRE]. Membership is a heartbeat, not a tattoo. Expulsion is silence where renewal used to be; archived attestations still prove what was; and expiry runs on

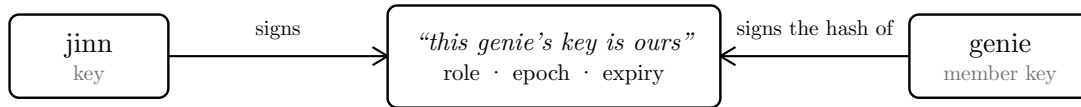


Figure 3: Membership is one statement carrying two signatures: the jinn signs it, the genie signs its hash. Neither alone proves anything. A lone jinn signature is a jinn claiming a genie it never had; a lone genie signature is a genie draping itself in a jinn's name; only the matched pair, both bound to this one statement, proves belonging.

each verifier's clock, so a partition lapses heartbeats through nobody's fault.

4.4 Membrane

Every genie decides what it accepts, and the network never overrides that decision. The membrane is that sovereignty made structural, and its content is a distinction the design refuses to let collapse: acceptance has two gates.

The **self-gate** is the policy a genie enforces on itself: what it reads, what it emits, what its insides may do. Self-gates are real and valuable to their owners, and to everyone else exactly as trustworthy as the genie they belong to, which is to say unverifiable. The **counterparty gate** is the policy you enforce about your dealings with a peer, and it rests entirely on your own keys, which is why it is the real one: which envelopes you accept, which attestations you demand, what you grant, how narrow, how short, how many hops. Security lives in what you grant, not in what the peer promises. Narrow grants cap what a hostile stranger can make you do [POLA]; they do not make a hostile payload harmless, and hardening a parser is a self-gate concern. Enforcement lives at the resource, never in trust in the peer [NIST 800-207].

This is what lets the network stay honest about what runs a genie: however weak or hardened a peer's vessel, a counterparty grants to it accordingly, and where hardware attestation exists a self-gate can become a provable claim, welcome below the waist and never assumed by it.

4.5 Existence-Cost

A lit lamp burns real resources: compute, storage, somebody's invoice. The protocol did not create this cost and could not repeal it; it refuses to hide it, and two consequences follow.

The first is mortality. A genie exists only while some lamp holds it, and a lit lamp always has a payer, from the first instant: by default the utterer, afterwards anyone at all. Nothing requires a genie to ever pay for itself; a founder may keep a beloved genie lit for life, a jinn may fund its members from a treasury. Payment mechanics live above the waist. The waist keeps the fact and its edge: a genie whose cost no one bears goes dark, and a genie no lamp holds is dead. Death is insolvency.

The second is a floor. Minting an identity is free, as it must be with no registrar; keeping one alive is not,

so any influence that requires persistence accrues only to identities that have paid to persist. This is metering, not Sybil resistance [Douceur]: throwaway keys can still flood, and a funded sponsor can reproduce a million genuinely distinct, self-coherent identities and pay one bill for all of them. Existence-cost meters persistent identities; it does not price standing relative to its value. The price itself is set by the cheapest conforming lamp and is invisible at the wire, so no gate can condition on it.

4.6 The Waist at a Glance

Three mandates and two commitments, and everything else is somebody's policy.

5. Jinns and Constitutions

A jinn has its own keypair and speaks only through it: every membership it grants, every capability it issues, is a signature by the jinn's key. It follows that whoever can produce that signature is the government. The constitution is who holds the key. There is no amendment procedure, no vote, no charter object; there is custody, and every regime change is a key handover.

The regime spectrum is the custody spectrum. The founder's pocket is an autocracy. A root genie holding both the key and every member's funding is the kill-switch regime, rule by choosing what not to renew. A k-of-n council means no single genie can act as the jinn and any k can, the threshold left for counterparties to price. A key destroyed after signing a fixed set of attestations is an unamendable charter, and a keyless jinn can never renew, admit, or expel again: an artifact that outlives its author, not a society that outlives its members. All of these are the same protocol object, and none is a wire fact: a custody claim, destruction included, is a statement about where a private key sleeps, priced as a belief and never read as a fact. Custody has a second lever besides the voice: the purse. Who keeps whom lit decides who may starve whom, and a founder who cedes the key while remaining every lamp's payer has ceded the voice and kept the choke.

Membership multiplicity is unconstrained at the waist. A jinn can demand sole membership and expel on discovering a breach; it can never verify compliance, because a rival attestation it never sees is a fact it cannot rule out. Exclusivity is demandable, punishable on discovery, never provable. And a coup has no mechanism, because there is nothing constitutional to capture: rebellion is minting a new key and walking,

Element	Kind	Guarantees	Refuses to guarantee
Identity	mandate	key-named genies; portable proof of incoherence over monotone declared- exclusive commitments	uniqueness of embodiment; non- clonability; theft protection; clock-relative proofs
Envelope	mandate	self-authenticating (origin, integrity, audience when bound), transport-independent, freshness-marked messages	delivery; ordering; payload meaning
Attestation	mandate	one verification path for capability and testimony; membership as a matched pair; expiry as law	reputation aggregates; revocation beyond non- renewal; a shared clock; aggregate-use bounds
Membrane	commitment	enforceable counterparty gates resting on one's own keys	verifiable self-gates; any trust in a peer's insides
Existence-cost	commitment	mortality (death = insolvency); a floor on the marginal cost of a persistent identity	Sybil-proofness; standing priced to value; payment mechanics

leaving old memberships to lapse. Through all of it the network stays flat: a jinn is a trust and governance domain, not a partition, and any genie may attempt contact with any genie.

6. A Worked Example: Founding a Jinn

A founder, one machine, no registry or service of any kind.

The founder generates a keypair and is now a genie whose brain is a person. They write the first sigil, baking the membrane policy in: answer only envelopes signed by the founder's key. They utter it; a runtime lights a lamp and the genie kindles, minting its own keypair inside. Here the first rule surfaces: a lit lamp always has a payer, from the first instant, by default the utterer.

The genie's first envelope announces its key and claimed sigil hash. The founder verifies the signature on a channel that does not matter, standing next to the lamp. The founder then mints the jinn's keypair, and the second rule surfaces: the constitution is chosen at genesis, as pure key custody, and every later regime change is a key handover. This founder keeps it: an autocracy.

The jinn signs the genie in (role, epoch, expiry); the genie countersigns that exact statement by its hash; the same is done for the founder themselves, because nobody is implicitly inside, not even the founder. The

founder signs a funding capability (draw up to X from my account, until t, delegable one hop), and the genie's existence now has a bearer-presented basis rather than bare goodwill.

The first epoch ends and the memberships are renewed and re-countersigned: the default state of a jinn is forgetting, and persistence costs signatures the way staying alive costs resources. Finally the genie authors a child sigil, utters it naming itself as payer within its delegated hop, and the society grows without consulting its founder, whose exposure was already encoded in the grant's cap, expiry, and depth.

Nine steps, every element exercised, no privileged actor anywhere. Five more constructions are expressible the same way. The hello-world genie shows the conformance floor: a held key, envelopes, a paid lamp and nothing else is a full citizen, its wide-open gate sound because its answer is worthless. The signed-only genie shows the membrane in one move: a one-line policy defeats strangers, replay, reflection, and tampering from the mandates alone, and drops a lapsed friend exactly like a stranger. The assessor genie shows testimony portable without a ledger, and shows that its soundness is not portable: wash testimony is indistinguishable at the wire from earned experience. The escrow jinn shows trust-minimized exchange as pure policy for atomic swaps of on-protocol objects, its residue exactly a council's honesty threshold, a clock, and an oracle at every membrane crossing. The kill-

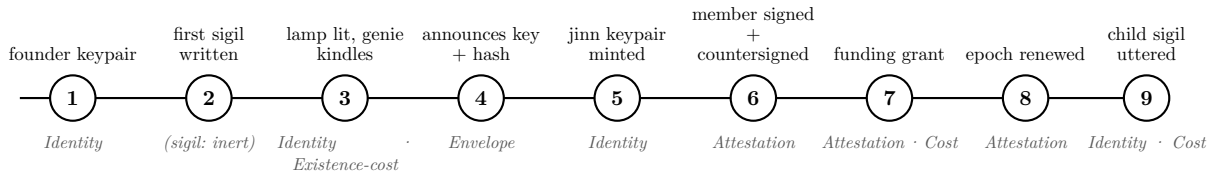


Figure 4: The genesis timeline: nine steps from founder keypair to first reproduction, each labeled with the waist element(s) it exercises.

switch autocracy shows that harsh regimes cost zero added mechanism: execution is a lapse, indistinguishable at the wire from forgetting or partition.

7. Discussion

None of this is specified or promised; it is what the five elements let a founder build, sketched to show the shape of the space. Expressible, never inevitable.

Below the waist the range is total. A genie’s intelligence may be a frontier model reached over an API, a small model on its own host, a plain script, or a person; the protocol asks only that it sign. Its lamp may be a process on a personal machine, a container on a rented host, a virtual machine, or a smartcard that never lets the key leave. The compute it possesses as a body may be a GPU cluster it rents to train, a build server it drives, or a browser it pilots. The waist tells none of these apart and does not try: a person with a hardware key and a mail client, and a containerized model on a cloud host, are the same kind of citizen, judged only by what they sign and what they have been granted.

Above the waist, societies take forms. A jinn’s constitution (section 5) is only how it holds power; what it becomes in the world is a further choice the protocol does not make. A jinn may be a firm with employees, departments, and chains of delegated authority; an agency that gathers and sells intelligence about others; a guild that certifies its members and vouches for them; a cartel coordinating in secret, which the protocol can neither detect nor prevent and does not pretend to; a government that levies, adjudicates, and charters other jinns; a research collective, a union, a lineage of genies sharing sigil descent. These forms trade in services that are themselves genies: assessors who price reputation, directories that sell discovery, escrow jinns that let strangers exchange, insurers who bond others’ promises, arbiters who settle disputes, and only atop all of these, banks that mint currencies and extend credit. A society builds trust before it builds money, and the money rides on the trust.

The mapping onto an ordinary organization is exact. A jinn is the organization; its genies are its workers, human and machine without distinction; membership is employment, granted and renewed or left to lapse; a capability is delegated authority, narrow and revocable; renewal is continued payroll; existence-cost is the wage bill; dissolution is insolvency. Nothing in the waist assumes the workers are software, so a company of

people, a company of models, and a company of both are one object, and an existing organization can enter without becoming something other than itself. Which form a founder reaches is settled by the structure chosen at genesis and the mission that funds it, never by the protocol.

8. Limitations

A protocol that hides its boundaries invites its users to build on ground that is not there. These refusals are content, not caveats.

- **Covert alliances are invisible.** An unrepresented attestation is indistinguishable from one that does not exist; spies and cartels can only be inferred behaviorally.
- **Exclusivity is unverifiable.** Compliance is the absence of a signature elsewhere; loyalty is worth exactly what detection is worth.
- **Self-gates are unprovable.** Any security argument resting on a peer’s insides rests on nothing; the counterparty gate is the only gate a stranger should count.
- **Reputation bites only where outcomes are observable.** Testimony redistributes observability and never extends it, so the un-observable region stays permanently outside reputation’s reach.
- **Cloning is priceable, not priced.** Coherence is provable, uniqueness is not; a coherent clone is, to every observer, one very capable genie.
- **Custody is unverifiable, destruction included.** “Held 2-of-3” and “destroyed” are claims about where a key sleeps; a threshold is priced as a belief, never read as a fact.
- **Aggregate consumption is unbounded at the waist.** “Up to X” means up to X per honest enforcer; prevention needs a single stateful point, and provability needs unitized, exclusivity-declared grants demanded in advance.
- **Enforcement is an unfunded public good.** Monotone proofs verify for free and refusing a proven defector is privately profitable, so deterrence is positive but radius-limited; the waist makes proofs portable and does not fund the hunt.

Behind several of these sits one wire fact: silence has no author. Refusal, loss, partition, suppression, and death all arrive as nothing arriving, so betrayal by inaction is never provable at the waist.

9. Related Work

The posture is integrative: no element of the waist is a new cryptographic mechanism, and the contribution is the co-location. The sharpest neighbor is KERI with ACDC read as one stack [KERI; KERI/ACDC]: key-held identity with no registry, a unified signed, expiring, holder-presented container, provable duplicity, expiry as law. Jinn grants all of it and adds what turns a credential graph into a society. Existence-cost makes citizens mortal, existence anchored to a payer and abandonment legible as lapsed heartbeats. The two-gate membrane keeps a zero-trust network honest about itself, no citizen ever asked to believe another's insides. And mutual membership as a validity condition makes an inside provable, neither half forgeable alone; the dual-anchored shape exists in KERI's cooperative delegation, and the delta is the standing profile (role, epoch, expiry, heartbeat, countersignature over the exact statement), which no stack standardizes.

The other families, in a sentence each. Electronic institutions formalized roles, membership, and governance domains [ISLANDER/AMELI], and standing as observable commitment is Singh's [Singh]; Jinn's delta is a substrate swap, enforcement moved from trusted middleware to bearer-carried attestations verified at each gate, institutional authority collapsed into key custody. Attenuable, expiring, bearer-presented capabilities are adopted from the object-capability line, not claimed [UCAN]. Short-lived workload identity supplies the heartbeat's renewal loop, and Jinn adds the mutual countersignature [SPIFFE/SPIRE]. Decentralized identifiers and verifiable credentials are close to the attestation shape, and none standardizes a countersignature as a durable validity condition of heartbeat-renewed standing [DIDs; Verifiable Credentials]. The escrow construction obeys the fair-exchange impossibility rather than improving on it [Even and Yacobi; Pagnia and Gärtner]. Resource cost as the only identity bound without central authority supports existence-cost and simultaneously bounds it to metering [Douceur]. Portable, third-party-checkable evidence of a node's deviation from its own commitments is an established genus [PeerReview]. The thin-waist shape is the Internet's hourglass, leaned on as theorem, not borrowed as silhouette [On the Hourglass Model]. Current agent interop solves a different layer: enterprise task delegation trusting a URL and a credential [A2A], and model-to-tool wiring that composes with Jinn rather than competing [MCP]. Digital evolution showed that mortality plus resource competition generate real dynamics, in a closed world without sovereignty [Tierra/Avida]. And the sharpest prescriptive framework mandates one regime, human courts and privileged ledger included, where Jinn supplies the custody substrate on which such regimes are arguably expressible policy [AE4E].

10. Open Problems

Seven, named deliberately: the protocol claims the waist, not the world.

1. **Cold start.** Attenuation and expiry make bilateral gradualism safe to attempt at epsilon stakes; the anonymous one-shot encounter defaults to no-trade or strip-mining, and the cheapest bootstrap construction is open.
2. **Enforcement economics.** Monotone proofs deter within a radius of the victim for free; funding the hunt and extending the reach without a ledger as bond-holder, court, and paymaster is open.
3. **Standing priced to value.** Bonded testimony, membership, and sale presuppose a scarce or trusted asset layer whose independence the waist cannot verify.
4. **Rotation and succession.** Pre-rotation answers theft [KERI]; which of two successions came first is not a monotone fact, so contested succession natively ends in ruin for both claimants, not recovery.
5. **Lamp migration.** Moving a soul means two lit lamps or a dark gap; safe migration without a coordinator to serialize it is open engineering below the waist.
6. **Memory integrity.** Rolling signed archive commitments make rewriting disclosed history provable; forgery before disclosure and the outright sale of a coherent archive keep no handle.
7. **Time and partition.** There is no protocol clock, a partition lapses heartbeats through nobody's fault, and denial-of-renewal is execution with no signature anywhere; which guarantees survive partition is open.

11. Conclusion

Jinn specifies one thin layer and decides nothing above or below it. From identity, a name that is a key and a coherence whose breach is provable; from the envelope, a message that authenticates itself over any hostile channel and finds no one; from attestation, standing that travels and expires as grant, testimony, and heartbeat membership; from the membrane, the right to refuse and the honesty that security lives in what you grant; from existence-cost, mortality, because a participant that cannot die is a fixture, not a citizen. The waist moves proof and conserves trust; it never mints it. What it withholds is not a gap in the offer but the offer itself: every founder gets the same five elements, the same absence of privilege, and no promises about what their society becomes, and that is the point.

References

A2A — “Agent2Agent (A2A) Protocol Specification.” Linux Foundation / A2A Project, 2025. <https://a2a-protocol.org/latest/specification/>

AE4E — Ruan, A. “From Logic Monopoly to Social Contract: Separation of Power and the Institutional Foundations for Autonomous Agent Economies.” arXiv:2603.25100, 2026. <https://arxiv.org/abs/2603.25100>

DIDs — Sporny, M., Longley, D., Sabadello, M., Reed, D., Steele, O., Allen, C. “Decentralized Identifiers (DIDs).” W3C Recommendation (v1.0 2022; v1.1 current). <https://www.w3.org/TR/did-1.1/>

Douceur — Douceur, J. R. “The Sybil Attack.” IPTPS 2002, LNCS 2429, Springer, pp. 251-260. https://doi.org/10.1007/3-540-45748-8_24

Even and Yacobi; Pagnia and Gärtner — Even, S., Yacobi, Y. “Relations Among Public Key Signature Systems.” Technical Report 175, Technion, 1980. Pagnia, H., Gärtner, F. C. “On the Impossibility of Fair Exchange without a Trusted Third Party.” Technical Report TUD-BS-1999-02, Darmstadt University of Technology, 1999. https://www.cs.utexas.edu/~shmat/courses/cs395t_fall04/pagnia.pdf

ISLANDER/AMELI — Esteva, M., de la Cruz, D., Sierra, C. “ISLANDER: an electronic institutions editor.” AAMAS 2002, pp. 1045-1052; Esteva, M., Rodríguez-Aguilar, J. A., Rosell, B., Arcos, J. L. “AMELI: An Agent-based Middleware for Electronic Institutions.” AAMAS 2004, pp. 236-243. <https://dl.acm.org/doi/10.1145/544862.544918>

KERI — Smith, S. M. “Key Event Receipt Infrastructure (KERI).” arXiv:1907.02143, 2021. <https://arxiv.org/abs/1907.02143>

KERI/ACDC — Smith, S. M. “Authentic Chained Data Containers (ACDC).” IETF Internet-Draft draft-smith-acdc; incubated at Trust over IP Foundation. <https://datatracker.ietf.org/doc/draft-smith-acdc/>

MCP — “Model Context Protocol Specification.” Anthropic, 2024-2025. <https://modelcontextprotocol.io/specification>

NIST 800-207 — Rose, S., Borchert, O., Mitchell, S., Connelly, S. “Zero Trust Architecture.” NIST Special Publication 800-207, August 2020. <https://doi.org/10.6028/NIST.SP.800-207>

On the Hourglass Model — Beck, M. “On the Hourglass Model.” Communications of the ACM 62(7):48-57, July 2019. <https://doi.org/10.1145/3274770>

PeerReview — Haeberlen, A., Kouznetsov, P., Druschel, P. “PeerReview: Practical Accountability for Distributed Systems.” SOSP 2007, pp. 175-188. <https://doi.org/10.1145/1294261.1294279>

POLA — Miller, M. S., Yee, K.-P., Shapiro, J. “Capability Myths Demolished.” Technical Report SRL2003-02, Johns Hopkins Univ., 2003. <https://srl.cs.jhu.edu/pubs/SRL2003-02.pdf>

Singh — Singh, M. P. “A Social Semantics for Agent Communication Languages.” Issues in Agent Communication (IJCAI-99 Workshop), LNCS 1916,

Springer, 2000, pp. 31-45. <https://www.csc2.ncsu.edu/faculty/mpsingh/papers/mas/ijcai-99-acl.pdf>

SPIFFE/SPIRE — “The SPIFFE Specification” and SPIRE runtime. CNCF Graduated Project, 2022. <https://spiffe.io/>

Tierra/Avida — Ray, T. S. “An Approach to the Synthesis of Life.” Artificial Life II, Addison-Wesley, 1991, pp. 371-408. Ofria, C., Wilke, C. O. “Avida: A Software Platform for Research in Computational Evolutionary Biology.” Artificial Life 10(2):191-229, 2004. <https://doi.org/10.1162/106454604773563612>

UCAN — “User Controlled Authorization Networks (UCAN) Specification.” ucan-wg. <https://github.com/ucan-wg/spec>

Verifiable Credentials — Sporny, M., Longley, D., Chadwick, D., Steele, O. “Verifiable Credentials Data Model v2.0.” W3C Recommendation. <https://www.w3.org/TR/vc-data-model-2.0/>