Red de Reputación Descentralizada: Un marco para la organización natural de la sociedad

Pavel Kudrna
Prague, Czechia
Draft v1 — March 2026

Abstract

El sentido de justicia—la convicción de que las injusticias se reparan y el mérito se recompensa—es la fuerza cohesiva más poderosa de la sociedad. Cuando las instituciones centralizadas de gobierno no pueden brindar este sentido porque el coste de hacer valer un derecho supera el valor del propio derecho, la cohesión social se erosiona. Las estructuras institucionales actuales no logran resolver una clase significativa de disputas de la misma manera sistemática en que la planificación central falla al asignar recursos: mediante la asimetría de información, la ausencia de bucles de retroalimentación y la desconexión de quienes toman decisiones respecto de las consecuencias de esas decisiones. Este artículo presenta una Red Social de Reputación (RSN): una capa de comunicación descentralizada, incorruptible y resistente a la censura, construida sobre Identificadores Descentralizados (DID), que permite a participantes seudónimos publicar, verificar y consumir información de reputación sobre el comportamiento en el mundo real. El mecanismo central se sustenta en tres principios de diseño: (1) una estructura de costes asimétrica en la que leer datos de reputación es barato pero escribir requiere un esfuerzo verificable; (2) un protocolo de selección de verificadores no determinista basado en hashing consistente que pone precio a las afirmaciones radicales mediante costes de iteración crecientes; y (3) un modelo de autoridad de reputación impulsado por el mercado en el que verificadores privados arriesgan su reputación acumulada sobre la exactitud de las afirmaciones que respaldan. La arquitectura resultante produce una meritocracia emergente sin coordinación central y habilita una ruta de migración gradual y reversible desde los sistemas administrados por el Estado actuales.

1 Introducción

1.1 El problema de la confianza centralizada

El gobierno moderno se apoya en una suposición fundamental: que las instituciones centralizadas pueden mediar la confianza entre desconocidos a gran escala. Los tribunales dirimen disputas. Los registros certifican la propiedad. Los organismos reguladores hacen cumplir las normas. Estas instituciones se diseñaron en una época en la que la información era escasa, la comunicación era lenta y delegar la autoridad en un órgano central era el único mecanismo de coordinación viable. El gobierno centralizado, sin embargo, fracasa por las mismas razones estructurales que la planificación central: asimetría de información, ausencia de bucles de retroalimentación y desconexión de quienes toman decisiones respecto de las consecuencias de esas decisiones.

La suposición fracasa, por ejemplo, cuando el coste de la mediación institucional supera el valor de la disputa. Considérese a un inquilino que descubre que su apartamento alquilado carece de un certificado de seguridad eléctrica exigido por ley—una condición que supone un peligro inmediato para la vida. El derecho legal del inquilino a rescindir el contrato es inequívoco. Sin embargo, el coste de hacer valer ese derecho a través del sistema judicial supera el valor monetario de la fianza y los daños adeudados, incluso incluyendo una posible indemnización legal [1]. La respuesta racional es asumir la pérdida y salir.

No se trata de un caso límite patológico. Es la experiencia habitual para una clase significativa de disputas en las que el daño es real pero lo que está en juego económicamente cae por debajo del umbral a partir del cual la ejecución institucional se vuelve económicamente racional. El resultado es un sistema que favorece estructuralmente a los malos actores: quienes perjudican a otros en volúmenes por debajo de este umbral pueden actuar con impunidad, porque el coste de buscar justicia recae sobre la parte perjudicada.

El ejemplo anterior está tomado del entorno jurídico del autor—el sistema de derecho civil checo. En otras tradiciones jurídicas—el common law basado en precedentes, el derecho consuetudinario, los sistemas mixtos—la justicia falla de maneras y en grados distintos, según las particularidades culturales y procesales de cada jurisdicción. Es probable que los lectores reconozcan ejemplos equivalentes en su propio contexto. Lo que es estructuralmente universal es el patrón: las disputas cuyo valor cae por debajo del umbral de la ejecución económicamente racional terminan con la pérdida absorbida por la parte perjudicada. La RSN no promete una justicia perfecta—simplemente reduce la ventaja estructural de la que disfrutan los malos actores cuando la ejecución institucional resulta antieconómica.

1.2 Por qué las soluciones existentes se quedan cortas

Los marcos de identidad descentralizada (DID) [2] proporcionan mecanismos criptográficos para la gestión de identidad autosoberana, pero se centran principalmente en la verificación de identidad sin abordar la cuestión más amplia de la reputación conductual.

Los Soulbound Tokens (SBT) [3] capturan la idea de que la reputación es no transferible, pero dependen de una infraestructura de cadena de bloques con limitaciones de escalabilidad y no abordan los incentivos económicos que rigen la entrada de información. Conceptos relacionados como los tokens de mérito (por ejemplo, Liberland) comparten una filosofía similar, pero siguen atados a marcos jurisdiccionales concretos.

Las Organizaciones Autónomas Descentralizadas (DAO) [4] implementan la gobernanza mediante contratos inteligentes, pero replican dinámicas plutocráticas en las que la influencia es proporcional al capital. La votación cuadrática [5] lo mitiga en parte pero limita su adopción práctica. Las DAO también se enfrentan al *problema del oráculo* fundamental: los contratos inteligentes no pueden verificar de forma fiable estados del mundo real sin una fuente externa de confianza, y este problema no tiene solución general dentro de la arquitectura de cadena de bloques.

Ningún sistema existente combina descentralización, resistencia a la censura, seudonimia, coste de entrada verificable y consenso emergente.

1.3 Aportaciones

Este artículo realiza las siguientes aportaciones:

1. La definición de la **Red Social de Reputación (RSN)**, un protocolo descentralizado que extiende el marco DID para admitir el intercambio bilateral de reputación.
2. Un **protocolo de selección de verificadores no determinista** basado en hashing consistente [6].
3. El **Principio del Radicalismo Costoso**, que pone precio a las afirmaciones según su distancia respecto del consenso de la red mediante tres canales de coste que se acumulan.
4. Un **modelo de Autoridad Reputable** con incentivos asimétricos en el que un respaldo falso cuesta catastróficamente más que las tarifas legítimas.
5. Una **ruta de migración gradual** a través de una infraestructura fiscal paralela.

2 Principios de diseño

La arquitectura de la RSN está condicionada por cinco principios de diseño innegociables.

Continuidad y evolución. El sistema coexiste con las instituciones existentes durante un periodo de transición de duración indefinida. La transición debe ser reversible en cada etapa.

Voluntariedad y responsabilidad. La participación es voluntaria, pero la voluntariedad va unida a la responsabilidad: un participante que asume el control de una función institucional asume simultáneamente las obligaciones que la acompañan.

Diversidad y neutralidad cultural. El consenso surge de las interacciones bilaterales, no de reglas predefinidas impuestas por los diseñadores del sistema.

Resiliencia y resistencia a la censura. El coste de censurar la red debe superar el coste de tolerarla. Solo un totalitarismo total—a un coste político y económico ruinoso—puede suprimir el sistema.

Consenso y ejecución. El sistema incorpora las tendencias humanas hacia la mezquindad, el rencor y la satisfacción retributiva como características estructurales. La mala conducta no se prohíbe; se pone a precio.

3 Visión general del sistema

3.1 Red Social de Reputación

La RSN es una capa de comunicación descentralizada, entre pares, en la que los participantes intercambian información verificable sobre el comportamiento en el mundo real. Sus propiedades definitorias son: infraestructura descentralizada e incensurable; participación seudónima mediante DID; estructura de costes asimétrica (leer es barato, escribir es caro); verificabilidad criptográfica; y **soporte de estándares**—formatos legibles por máquina para la información propagada por la red, para los servicios ofrecidos por las autoridades (composición de afirmaciones, respaldo, servicio de testigo) y para el formato de política, incluidas las reglas para evaluar la reputación de la contraparte y valorar el riesgo de discrepancia de política (entre el comportamiento declarado y el real).

3.2 Componentes arquitectónicos

La RSN consta de cuatro componentes principales (Fig. 1):

1. **Capa DID.** Identidades de los participantes con reglas declaradas, metadatos de reputación y enrutamiento de red.
2. **Protocolo de afirmaciones.** Formato estructurado para publicar aserciones sobre eventos del mundo real.
3. **Red de verificación.** Protocolo descentralizado para asignar verificadores mediante hashing consistente.
4. **Capa de agregación de reputación.** Servicios que producen resúmenes de reputación legibles por humanos.

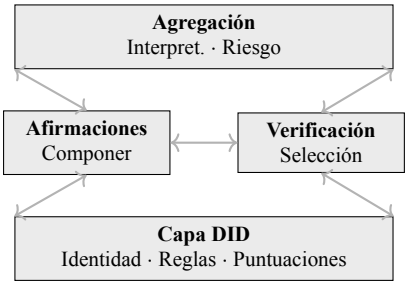

Figure 1: Arquitectura de cuatro capas de la RSN.

3.3 Roles de los participantes

Una transacción de verificación implica hasta seis roles DID distintos (Fig. 2): **Emisor** (DID_I, crea y envía la afirmación), **Sujeto** (DID_S, el DID sobre el que trata la afirmación—puede coincidir con el Emisor), **Autoridad** (DID_A, respalda la afirmación por una tarifa; también puede ofrecer servicios de testigo—*opcional*), **Testigo** ($DID_{W_{1..n}}$, supervisor de incógnito de la honestidad del verificador mediante códigos de reto—*opcional*), **Verificador** (DID_V, seleccionado algorítmicamente para publicar la afirmación) y la **RSN** (la red donde se publican las afirmaciones verificadas). Cualquier rol puede delegarse en otro DID (Sección 7.4). Una Autoridad suele combinar el respaldo con servicios de testigo, pero ambas funciones son lógicamente independientes.

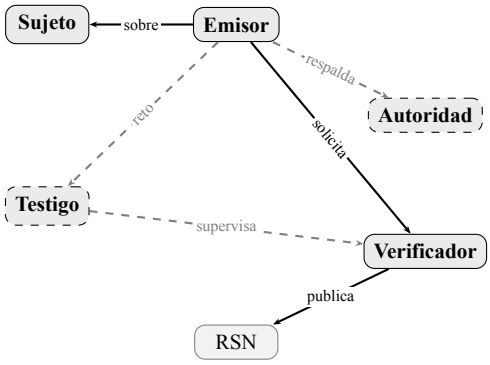

Figure 2: Roles DID en una transacción de verificación. Discontinuo = opcional (Autoridad, Testigo).

3.4 Incorruptibilidad: cuatro fuerzas

La incorruptibilidad de la RSN no surge de un único mecanismo, sino de cuatro fuerzas concurrentes. Ninguna es suficiente por sí sola; solo su combinación hace que un intento de corrupción sea económicamente irracional.

Objetivo impredecible. El verificador de cada afirmación se selecciona de forma no determinista mediante hashing consistente (Sección 5.3). Un atacante no puede fijar de antemano un verificador concreto para sobornarlo, porque la identidad del verificador es función del contenido de la afirmación y de las iteraciones previas, no de la elección del emisor.

Apalancamiento reputacional—sube despacio, cae rápido. La reputación solo puede ganarse de forma incremental, mediante un comportamiento consistente sostenido. Sin embargo, puede perderse de manera catastrófica en un único respaldo fraudulento (Sección 6.2). Esta asimetría implica que años de reputación acumulada pueden destruirse con un solo respaldo deshonesto; la estrategia óptima sigue siendo rechazar las afirmaciones sospechosas.

Promesa pública. Cada Titular de un DID declara públicamente su política—un conjunto de reglas legible por máquina que se compromete a cumplir. La

divergencia entre la declaración y el comportamiento real es detectable y se pone a precio como una infracción reputacional más grave que la violación subyacente (Sección 7.3). La hipocresía es, por tanto, más costosa que la deshonestidad directa.

Asimetría de coste de ataque en malla. El coste de censurar o desestabilizar la red crece de forma no lineal con el tamaño y la densidad de la red, mientras que el coste de tolerarla permanece constante. Para que la censura resulte rentable, un actor centralizado tendría que aplicarla a toda la red—exigiendo medidas desproporcionadas frente a cualquier beneficio concebible (Sección 10).

4 Modelo de identidad descentralizada

4.1 DID con propiedades especiales

La RSN extiende la especificación W3C DID Core [2] con: **Reglas Declaradas** (compromisos conductuales legibles por máquina), **Metadatos de Reputación** (puntuación conductual agregada) y **Enrutamiento de Red** (pasarela onion mutable, separada de la posición estable en el anillo).

4.2 Ciclo de vida de una afirmación

Una afirmación transita por seis etapas (Fig. 3): composición, respaldo opcional por una autoridad, selección de verificador mediante hashing consistente, decisión de verificación (aceptar o rechazar con iteración), publicación y actualización de reputación para todas las partes.

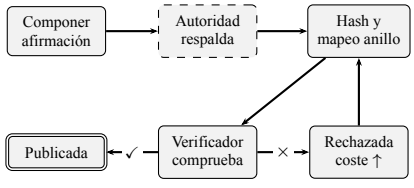

Figure 3: Ciclo de vida de una afirmación con bucle de iteración.

4.3 Relación con W3C DID Core

El modelo de identidad de la RSN es un superconjunto propio de la especificación W3C DID Core [2]. Todos los DID de la RSN son DID válidos de W3C. Las extensiones específicas de la RSN se codifican como propiedades del documento DID definidas en una especificación de método DID dedicada, compatible con infraestructura existente como ION [7] y Ceramic [8].

5 Verificación y propagación de la información

5.1 Coste de la entrada de información

El principio económico central de la RSN es la asimetría entre leer y escribir. Leer datos de reputación se aproxima a un coste marginal nulo para los participantes que forman parte de la red DID y tienen un acceso acorde con su reputación—los recién llegados deben ganar progresivamente un acceso más amplio (véase anti-parasitismo). Escribir—entendido como la materialización duradera de la reputación en la red, no como un acto técnico puntual—requiere una inversión acumulada verificable en tres dimensiones: **tiempo** (años de vida dedicados a un comportamiento consistente, compromisos cumplidos y relaciones cultivadas), **energía** (esfuerzo invertido en un trato honesto, cuidado de las asociaciones y fiabilidad a largo plazo) y **dinero** (inversión en el propio nombre—desarrollo profesional, calidad del propio trabajo, restitución por los daños causados). La reputación no puede comprarse al instante—es el resultado de una acumulación de toda una vida que el sistema simplemente hace visible y transferible entre contextos. Los costes técnicos puntuales del protocolo (firmas criptográficas, tarifas de verificador) son insignificantes en comparación con esta inversión subyacente.

5.2 Grafo social y profundidad de contactos

La RSN es fundamentalmente una red social: cada DID añade contactos (otros DID) que permiten la conexión. Estos contactos tienen sus propios contactos, y así sucesivamente. La búsqueda algorítmica de verificadores opera dentro de una profundidad configurable h de contactos enlazados (por ejemplo, $h = 3$: los contactos directos, sus contactos y los contactos de los contactos). Esta arquitectura no requiere una única cadena de bloques global—favorece de forma natural la aparición de comunidades/agrupaciones con solapamientos hacia otras comunidades. Todo participante DID debe tener una **política** publicada—un conjunto de reglas legible por máquina que rige cómo se evalúan las solicitudes entrantes. Las violaciones de política se registran en la red como artefactos negativos.

5.3 Selección algorítmica de verificadores

El protocolo de verificación emplea hashing consistente [6] (Fig. 4). La verificación es un servicio de pago: los verificadores operan por una tarifa, creando un mercado donde la competencia impulsa el precio, la velocidad y la fiabilidad.

Paso 1. El documento de la afirmación se concatena con el resultado del hash de la iteración anterior (o $\varnothing$ en la primera iteración) y se somete a hash:

$$\text{pos} = f_h(\text{claim} \| \text{prev\_hash}) \tag{1}$$

El algoritmo debe incluir la *ruta completa* de todas las solicitudes y respuestas previas—no meramente un hash de la iteración anterior. La respuesta completa de cada verificador (aceptación, rechazo, precio cotizado, motivo) se añade al documento en crecimiento, verificable mediante firmas criptográficas en cada paso.

Paso 2. El hash se mapea a N posiciones en el anillo de hashing consistente, distribuidas uniformemente (por ejemplo, para $N = 4$: $+0°, +90°, +180°, +270°$). Desde cada posición, una búsqueda en sentido horario selecciona el DID más cercano como candidato a verificador (Fig. 5). N es un parámetro variable que puede depender del porcentaje de DID actualmente activos, de la hora del día o de la capacidad de respuesta histórica de la red.

Paso 3. El Verificador acepta (publica, arriesgando reputación) o rechaza (volviendo al Paso 1 con el hash del rechazo). Cuando un nodo no responde pese a declararse en línea, la siguiente solicitud debe incluir todas las respuestas de los N candidatos a verificador anteriores.

Anti-parasitismo. Los DID objetivo pueden fijar tarifas o restricciones de acceso para las consultas de DID nuevos con poca reputación. Este mecanismo graduado (no binario) obliga a los recién llegados a construir reputación mediante actividad genuina antes de consumir libremente los datos de otros.

Cada rechazo añade la respuesta completa del verificador (incluidos motivos y condiciones) al documento de la afirmación, ampliando su contenido. A partir del documento ampliado se calcula de forma no determinista un nuevo hash, que se mapea a una posición diferente del anillo y selecciona un verificador distinto. La documentación de la ruta completa es obligatoria—el emisor no puede predecir ni influir en el siguiente verificador. Si un verificador ignora una solicitud pese a tener una política declarada compatible, los testigos (si están presentes) lo registran, y el emisor puede adjuntar la prueba del testigo en la publicación final.

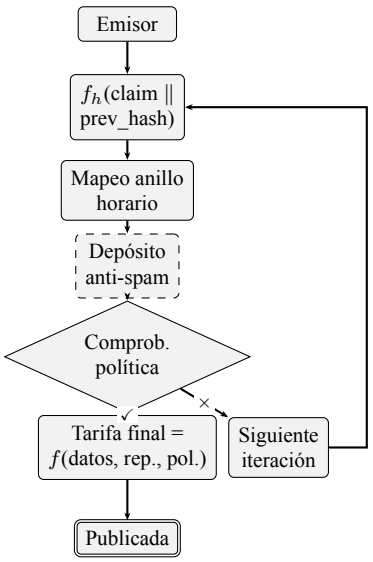

Figure 4: Protocolo de selección de verificadores. El depósito anti-spam es opcional y puede ser reembolsable. La tarifa final solo se paga en caso de aceptación.

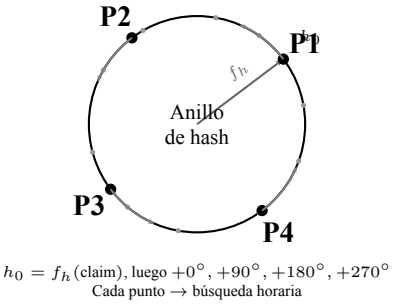

Figure 5: Selección multipunto (N=4). El hash h_0 fija el desplazamiento; 4 puntos distribuidos uniformemente desde h_0.

5.4 Protocolo de testigos

El rol de **Testigo** proporciona rendición de cuentas de incógnito sobre la honestidad del verificador mediante un protocolo criptográfico de códigos de reto:

Paso W1. El solicitante firma la solicitud de verificación y la envía a N_w testigos—contactos de la red social del solicitante, o proveedores especializados de servicios de testigo que operan por una tarifa.

Paso W2. Cada testigo w_i firma toda la solicitud firmada, conserva la firma original σ_i y calcula un código de reto $c_i = h(\sigma_i)$. El código de reto se añade a la solicitud.

Paso W3. La solicitud, que ahora lleva N_w códigos de reto, se envía al verificador. El verificador observa los códigos pero *no puede determinar* quiénes son los testigos ni si los códigos corresponden a firmas reales.

Paso W4 (verificador honesto). Si el verificador

responde conforme a su política declarada, los códigos de reto permanecen opacos. Los testigos nunca se revelan. No se publica ningún dato adicional.

Paso W5 (violación de política). Si el verificador viola su política (ignora la solicitud, responde de forma incoherente), el solicitante conserva las firmas originales σ_i de los testigos. Estas pueden publicarse como datos acompañantes: cualquiera puede verificar $c_i = h(\sigma_i)$, demostrando que el testigo estaba presente. El solicitante puede publicar de forma independiente—el verificador ya firmó los códigos de reto en su respuesta.

La propiedad del farol. El solicitante puede sustituir algunos o todos los códigos de reto por valores aleatorios. El verificador no puede distinguir los códigos genuinos (respaldados por autoridades de alta reputación) del ruido. Esta *incertidumbre* es el mecanismo de ejecución: cada solicitud lleva el riesgo de que una autoridad respetada esté observando de incógnito. El coste de mantener esta presión es casi nulo (generar números aleatorios), mientras que el coste potencial de la deshonestidad para el verificador es catastrófico. El comportamiento honesto se incentiva incluso en ausencia de testigos reales.

La Autoridad como testigo. Una Autoridad que respalda una afirmación puede combinar servicios de testigo: "Respaldo tu afirmación y actúo como testigo de incógnito durante la verificación." Este es un modelo de negocio natural—la Autoridad ya cuenta con el contexto. No obstante, ambos roles son lógicamente independientes.

5.5 Principio del Radicalismo Costoso

Tres canales de coste se acumulan simultáneamente (Fig. 6):

Canal 1: coste de iteración. Cada rechazo obliga a una nueva iteración que consume tiempo y recursos. Crecimiento lineal.

Canal 2: inflación del documento. Cada iteración añade la respuesta completa del verificador al documento de la afirmación. El crecimiento es lineal, pero con un desplazamiento base: la carga inicial (contenido de la afirmación, respaldo de autoridad, firmas criptográficas) ya tiene un tamaño no trivial B_0. Tamaño total tras k iteraciones: $S(k) = B_0 + k \cdot \Delta$, donde Δ es el tamaño medio de respuesta.

Canal 3: prima de riesgo del verificador. El riesgo es subjetivo. El verificador evalúa si las políticas declaradas del DID solicitante entran en conflicto con sus propios intereses comerciales, sociales o políticos. La prima puede escalar exponencialmente con la distancia percibida respecto del "ideal" del verificador: $P(d) = \alpha \cdot e^{\beta d}$, donde d es la métrica de distancia subjetiva del verificador. Más allá de un límite personal infranqueable, el verificador puede negarse por completo.

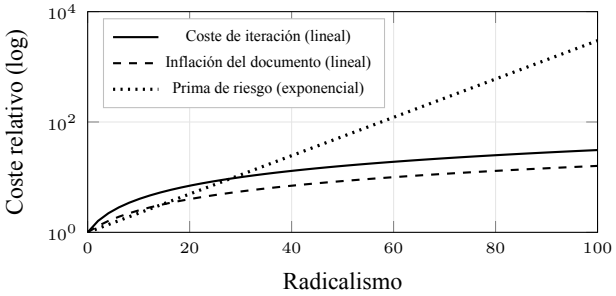

Figure 6: Escalada de costes a través de tres canales. La prima de riesgo domina en niveles altos de radicalismo.

Fundamentalmente, esto no es censura. Ninguna afirmación se prohíbe. El sistema pone precio al riesgo (Tabla 1).

Table 1: Comparación de costes según el tipo de afirmación.

Tipo	Iter.	Doc	Tarifa	Total
Creíble	$k_{\min}$	B_0	$P_{\min}$	Bajo
Controvertida	k_{med}	$B_0 + k\Delta$	$\alpha e^{\beta d}$	Moderado
Radical	$k_{\max}$	$\gg B_0$	$\gg P_{\min}$	Muy alto
Fraudulenta	$\to \infty$	—	—	Impublicable

6 Reputación y evaluación del riesgo

6.1 Modelo de acumulación de reputación

La reputación exhibe tres propiedades: **acumulación lenta** (crecimiento incremental mediante comportamiento consistente), **destrucción rápida** (un solo fraude puede reducir la puntuación de forma catastrófica) y **evaluación individual** (cada parte consumidora puede aplicar su propia función de agregación).

6.2 Autoridades Reputables

Una Autoridad Reputable revisa las afirmaciones y, si queda satisfecha, las cofirma—arriesgando su propia reputación (Fig. 7). La asimetría es intencionada: ganar reputación es lento (incremental $+\delta$ por cada respaldo honesto), mientras que perderla es catastrófico ($-\Delta \gg \delta$ por cada respaldo falso; a modo ilustrativo, p. ej., $+2\%$ frente a -70%). La estrategia óptima

es siempre rechazar las afirmaciones sospechosas. Los valores concretos de δ y Δ son parámetros del sistema.

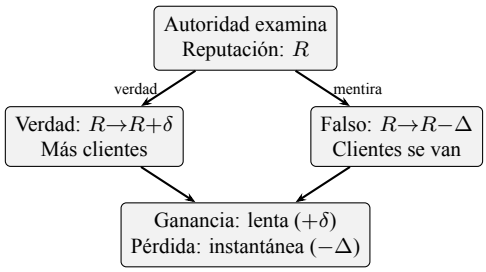

Figure 7: Autoridad Reputable—incentivos asimétricos.

6.3 Reputación multidimensional

La reputación no es un escalar sino un vector a lo largo de múltiples dimensiones: fiabilidad financiera, integridad personal, competencia profesional, compromiso cívico y otras (Fig. 8). Distintos proveedores de evaluación proyectan este vector de manera diferente según el contexto—un prestamista prioriza la fiabilidad financiera, un empleador la competencia profesional, un vecino el comportamiento cívico. No existe una única puntuación de reputación "correcta"; solo perspectivas.

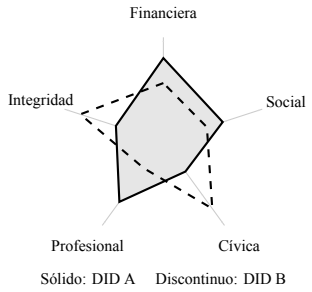

Figure 8: Vectores de reputación multidimensionales. Distintos DID exhiben perfiles diferentes en las dimensiones.

6.4 Evaluación de riesgo democratizada

La RSN democratiza la evaluación sistemática del riesgo—una capacidad actualmente concentrada en las instituciones financieras pero aplicable mucho más allá de la banca. Conglomerados industriales que evalúan la fiabilidad de sus proveedores, aseguradoras que ponen precio al riesgo conductual, la debida diligencia en fusiones y adquisiciones, la estimación de la vida útil de equipos en la industria—dondequiera que el riesgo de la contraparte requiera evaluación, la RSN ofrece una alternativa descentralizada e impulsada por el mercado. Un mercado de servicios de interpretación traduce los datos brutos de reputación multidimensional en resúmenes accionables, haciendo que la evaluación de la contraparte sea accesible para cualquier participante a un coste marginal.

7 Consenso y resolución de disputas

7.1 Modelo de justicia descentralizada

La RSN replantea la justicia como un problema de negociación bilateral. La amenaza de un daño reputacional permanente y verificable es un elemento disuasorio más eficaz que un proceso judicial que la parte perjudicada no puede permitirse.

7.2 Consenso emergente

Cada participante mantiene un umbral personal de satisfacción. El consenso agregado de la red emerge como el promedio estadístico de miles de negociaciones bilaterales (Fig. 9). Una respuesta muy por encima del consenso (bárbara) es cara de publicar. Una respuesta cercana al consenso (proporcional) es barata. El consenso no es una regla—es una señal de precio.

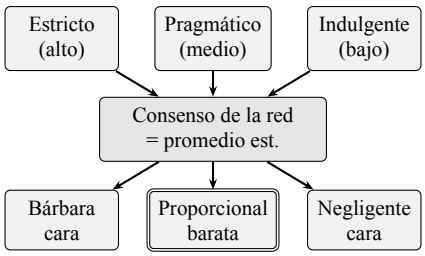

Figure 9: Consenso emergente a partir de umbrales de satisfacción individuales.

7.3 Calibración del castigo

Cada Titular de un DID declara públicamente sus respuestas a categorías específicas de mala conducta. Las declaraciones desproporcionadamente severas o indulgentes atraen señales de reputación negativas. Las declaraciones proporcionales se aceptan sin fricción—un sistema de castigo autocalibrado.

Las declaraciones de consenso están sujetas ellas mismas a la detección de hipocresía: comprometerse declarativamente con una posición de consenso mientras se actúa de forma incoherente constituye una infracción reputacional más grave que la desviación subyacente, ya que demuestra un engaño intencionado.

7.4 Delegación

Todas las actividades dentro de un DID—con una excepción—pueden delegarse en otro DID. **El rol de Sujeto—el DID sobre cuyo comportamiento trata la afirmación—no puede delegarse; está ligado de forma no transferible al titular de la identidad a quien se refiere la afirmación.** Los demás roles activos—Emisor, Autoridad, Testigo, Verificador—pueden transferirse a un DID delegado designado, ya sea para verificación, envío de afirmaciones, participación en el consenso o resolución de disputas. La delegación es revocable en cualquier momento. Esto permite la especialización (por ejemplo, servicios profesionales de verificación) mientras el Titular conserva el control y la responsabilidad últimos. La delegación se aplica a todo el sistema y es esencial para la escalabilidad.

7.5 De la moral individual al contrato social emergente

La política declarada de cada DID representa una formalización de la moral individual: lo que el Titular considera aceptable, cómo responde a comportamientos específicos, qué exige de sus contrapartes. Estas políticas no las impone un diseñador del sistema—las elige cada participante y se expresan en forma legible por máquina.

Esta formalización habilita una emergencia en tres etapas. Primero, la moral individual se codifica como **política**—un conjunto de reglas declaradas, umbrales y funciones de respuesta que el DID se compromete a seguir. Segundo, el agregado de todas las políticas de la red produce una **ética emergente**—normas estadísticamente observables que ningún participante individual diseñó pero que surgen de la interacción de miles de posiciones morales individuales [9]. Tercero, esta ética emergente, expresada como normas de comportamiento compartidas y ejecutadas mediante señales de precio bilaterales, constituye un **contrato social medible**—no un constructo teórico que nadie firmó [10], sino un conjunto de reglas empíricamente observable y respaldado por el comportamiento real.

Este proceso refleja los hallazgos de la teoría de los fundamentos morales [11]: la moral humana es intuitiva, plural y culturalmente variable. La RSN no requiere el consenso moral como entrada; produce el consenso conductual como salida. El contrato social resultante no es estático—evoluciona a medida que los participantes se unen, se marchan y ajustan sus políticas en respuesta a la retroalimentación de la red. A diferencia de la teoría clásica del contrato social, este contrato es revocable, observable y ejecutable sin un soberano.

8 Modelo económico

Las secciones anteriores describieron una herramienta—los mecanismos de verificación de la RSN, su estructura de costes y su consenso emergente. Esta sección describe una metodología para aplicar esa herramienta a la transición fiscal y de gobernanza. La herramienta es de propósito general; la metodología que sigue es una posible aplicación.

8.1 Estructura de incentivos

La reputación como capital crea incentivos directos para el comportamiento honesto. En el funcionamiento normal, los participantes construyen reputación de forma natural mediante transacciones cotidianas y compromisos cumplidos. El gasto principal se destina a las afirmaciones *negativas*—registrar el daño causado por otros. Esta asimetría incentiva el comportamiento útil y fiable: ser honesto no cuesta nada adicional, mientras que ser dañino crea un rastro documentado que otros pagarán por preservar. La hipocresía—declarar reglas sin seguirlas—es el modo de fallo más caro, ya que demuestra un engaño intencionado.

8.2 Sistema fiscal paralelo

Tres componentes habilitan la presión competitiva: (1) **Impuesto Simple**—una única tasa proporcional sin excepciones, inicialmente marginalmente por debajo de la tasa efectiva existente; (2) **Registro Electrónico del Gasto (RED)**—invirtiendo el modelo checo EET para que los ciudadanos vigilen el gasto estatal; (3) **Asignación Tributaria Dirigida por el Ciudadano**—empezando en un pequeño porcentaje el primer año y alcanzando, tras una década, desde cifras de dos dígitos bajos hasta una migración completa a un sistema dirigido por el ciudadano, según el ritmo de adopción. El tempo real depende de la velocidad a la que surjan alternativas de libre mercado a los servicios estatales y de la disposición del Estado a cooperar con la transición; la función del tiempo no tiene por qué ser lineal, y no hay razón para fijar un límite superior sobre dónde puede acabar llegando la adopción.

9 Ruta de migración

La migración transcurre por tres fases (Fig. 10): **Fase 1: Superposición** (la RSN como capa informativa, el Estado es el único proveedor), **Fase 2: Competencia** (la RSN ofrece alternativas funcionales, uso de doble sistema) y **Fase 3: Transición** (la RSN supera a los equivalentes estatales, migración voluntaria). La transición es reversible en cada etapa.

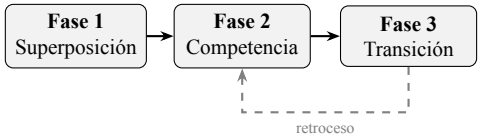

Figure 10: Migración en tres fases—reversible en cada etapa.

El principal mecanismo de presión es fiscal: cuando los contribuyentes condicionales alcanzan una "minoría económicamente significativa X"—un grupo lo bastante grande como para que la represión contra él cause un daño económico no trivial y la desobediencia civil se convierta en una amenaza creíble—el Estado entra en negociación. A modo ilustrativo, usamos $X \approx 15\%$ del PIB, pero el umbral real depende de la economía concreta.

10 Análisis de seguridad

Consideramos cinco categorías de adversarios: malos actores individuales, grupos organizados, adversarios corporativos, adversarios a nivel estatal y adversarios a nivel de protocolo.

Resistencia Sybil [12]. Construir reputación requiere una interacción sostenida y verificable. El coste de un ataque Sybil exitoso escala linealmente con las identidades falsas y cuadráticamente con el nivel de reputación objetivo. Operar múltiples DID en paralelo es posible pero caro por diseño—cada identidad debe acumular reputación de forma independiente mediante actividad genuina. En sociedades libres este coste disuade el abuso casual; en dictaduras, los DID paralelos se convierten en una herramienta de supervivencia que habilita una resistencia compartimentada, la coordinación clandestina y una navegación más segura del mercado negro.

Defensa contra la colusión. Los patrones de respaldo mutuo entre grupos cerrados son detectables mediante el análisis del grafo social. Las funciones de agregación de reputación descuentan las afirmaciones procedentes de fuentes fuertemente agrupadas.

Adversario a nivel estatal. El enrutamiento onion, la ausencia de infraestructura centralizada y la distribución del rol de Verificador entre la base de participantes garantizan que la supresión requiera medidas desproporcionadas frente a cualquier beneficio.

Privacidad frente a rendición de cuentas. La seudonimia—un término medio entre el anonimato y la revelación de identidad—permite a los DID acumular reputación significativa sin revelar la identidad real del Titular.

11 Consideraciones sobre la privacidad

El modelo de privacidad de la RSN se basa en la seudonimia. El Titular controla la revelación de identidad: mínima (seudónimo puro), selectiva (credenciales específicas vinculadas) o total (identidad legal asociada). Las afirmaciones publicadas son permanentes—el sistema admite la corrección contextual pero no el borrado. Un Titular puede abandonar un DID comprometido y empezar de nuevo, renunciando a la reputación acumulada.

12 Trabajo relacionado

La especificación W3C DID Core [2] y la red Ceramic [8] proporcionan la base de identidad. EigenTrust [13] demostró la agregación distribuida de reputación sin autoridad central. Los SBT [3] introdujeron tokens sociales no transferibles. La RSN se diferencia en su énfasis en el coste económico como filtro de calidad y en su mecanismo para poner precio al radicalismo.

Las DAO [4] y la votación cuadrática [5] demostraron la viabilidad de la gobernanza descentralizada. La RSN deriva el consenso de negociaciones bilaterales de precio en lugar de la votación.

La propuesta se basa en la teoría anarcocapitalista [14, 15] para la provisión privada de servicios, pero se aparta de ella en dos aspectos críticos: diseña teniendo en cuenta el rencor y los impulsos retributivos en lugar de suponer racionalidad, y propone una transición gradual en lugar de una revolución. El concepto de contratos sociales emergentes tiene antecedentes en la teoría del orden espontáneo de Hayek [16].

Anarcoagorismo. La RSN comparte un terreno común significativo con la contraeconomía agorista [17]—en particular el énfasis en construir instituciones paralelas y en el intercambio voluntario. Sin embargo, el agorismo tiende a suponer el interés propio racional como mecanismo de coordi-

nación suficiente y a menudo carece de una ruta de migración concreta desde las estructuras estatales existentes. La RSN incorpora explícitamente tendencias conductuales no racionales y proporciona un mecanismo de presión fiscal para una transición gradual.

Meritocracia. La RSN puede representar una primera descripción funcional de cómo la meritocracia [18] puede emerger como una propiedad sistémica en lugar de un eslogan. Los sistemas meritocráticos tradicionales fracasan porque requieren una autoridad central que defina y haga cumplir el "mérito". En la RSN, el mérito emerge de evaluaciones bilaterales: quienes aportan valor acumulan reputación; ningún comité decide quién tiene mérito.

La propiedad como privilegio. La RSN se aparta fundamentalmente de los tratamientos anarcocapitalistas y de la escuela austriaca sobre los derechos de propiedad como naturales y absolutos. En el marco de la RSN, la propiedad es un *privilegio*—un reconocimiento social que, en casos extremos, puede ser retirado por la comunidad cuando un participante viola fundamentalmente las normas compartidas (traición, negativa a defender a la comunidad en una crisis existencial, explotación sistemática). Esto no es una expropiación estatal, sino la retirada del reconocimiento social por parte de la red de relaciones bilaterales.

13 Discusión y limitaciones

Arranque en frío. El valor de la RSN depende de los efectos de red; los adoptantes tempranos afrontan un valor limitado hasta que la participación alcanza un umbral. Sin embargo, incluso desde etapas iniciales, la red DID sirve como una fuente de información complementaria útil. Se espera que la evaluación temprana de la reputación y la valoración de las autoridades se desarrollen gradualmente con una sofisticación creciente.

Anti-parasitismo y control de acceso. Los DID objetivo pueden imponer tarifas graduadas y restricciones de acceso a las consultas de DID de baja reputación. Esto no es binario sino una escala continua: cuanta menos reputación tiene un DID que consulta, menos información recibe, más espera y más paga. Este mecanismo incentiva la participación genuina por encima del consumo pasivo.

Desigualdad de acceso. El coste de publicar afirmaciones puede filtrar reclamaciones legítimas de participantes económicamente desfavorecidos. Mitigación: cada participante actúa simultáneamente como verificador potencial (o delega este rol) y gana tarifas de verificación. En un horizonte temporal suficiente, un participante no radical debería aproximarse a la neutralidad financiera—los ingresos por verificación compensando aproximadamente los costes de publicación. Esto es una expectativa estadística, no una garantía.

Desigualdad de reputación. Los adoptantes tempranos acumulan ventajas que pueden crear barreras para los que llegan más tarde. Sin embargo, la evaluación de la reputación la realizan proveedores externos que pueden optar por mitigar la ventaja del primer participante mediante sus algoritmos de agregación. Ser el primero con una buena reputación es una ventaja económica comparativa legítima—y eso es intencionado.

Cuestiones abiertas. Las funciones de coste óptimas, los estándares de agregación, la gobernanza del protocolo y la interacción con datos de reputación sintética generados por IA siguen siendo áreas para trabajo futuro.

Este artículo no afirma que la RSN vaya a producir resultados óptimos en todas las circunstancias. Afirma que la RSN representa una alternativa *viable*—técnicamente implementable, económicamente autosostenible y socialmente compatible con las tendencias conductuales humanas tal como son en realidad.

14 Conclusión

Este artículo ha presentado la Red Social de Reputación, un protocolo descentralizado que habilita el intercambio de reputación seudónimo y verificable. El Principio del Radicalismo Costoso garantiza que las afirmaciones fraudulentas queden fuera de precio sin necesidad de censura. La ruta de migración propuesta permite una transición gradual y reversible desde las instituciones existentes. El diseño incorpora la naturaleza humana tal como es—incluidas la mezquindad, el rencor y el deseo de satisfacción retributiva—en lugar de exigir una transformación ideológica. El resultado no es una utopía, sino un sistema en el que la justicia es accesible, la reputación se gana y el coste de la mala conducta lo soporta la parte que lo causó.

References

[1] Czech Republic, "Act no. 549/1991 coll., on court fees, as amended," 1991, fee schedule for court proceedings; cost rules per Act No. 99/1963 Coll. (Code of Civil Procedure) and Decree No. 177/1996 Coll. (Attorney Tariff).

[2] M. Sporny, D. Longley, M. Sabadello, D. Reed, O. Steele, and C. Allen, "Decentralized identi-

fiers (DIDs) v1.0," W3C Recommendation, Jul. 2022.

[3] E. G. Weyl, P. Ohlhaver, and V. Buterin, "Decentralized society: Finding Web3's soul," *SSRN Electronic Journal*, May 2022.

[4] V. Buterin, "DAOs, DACs, DAs and more: An incomplete terminology guide," Ethereum Blog, May 2014.

[5] V. Buterin, Z. Hitzig, and E. G. Weyl, "A flexible design for funding public goods," *Management Science*, vol. 65, no. 11, pp. 5171–5187, 2019.

[6] D. Karger, E. Lehman, T. Leighton, R. Panigrahy, M. Levine, and D. Lewin, "Consistent hashing and random trees: Distributed caching protocols for relieving hot spots on the World Wide Web," in *Proc. 29th ACM STOC*, 1997, pp. 654–663.

[7] D. Buchner, "ION — a scalable DID method based on Bitcoin," Microsoft, 2021.

[8] Ceramic Network, "Ceramic protocol specification," GitHub, 2023. [Online]. Available: https://github.com/ceramicnetwork

[9] É. Durkheim, *The Division of Labor in Society*. Free Press, 1893, english translation 1984.

[10] J. Rawls, *A Theory of Justice*. Harvard University Press, 1971.

[11] J. Haidt, *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. Vintage Books, 2012.

[12] J. R. Douceur, "The Sybil attack," in *Proc. 1st International Workshop on Peer-to-Peer Systems (IPTPS)*, 2002, pp. 251–260.

[13] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The EigenTrust algorithm for reputation management in P2P networks," in *Proc. 12th International Conference on World Wide Web*, 2003, pp. 640–651.

[14] M. N. Rothbard, *For a New Liberty: The Libertarian Manifesto*. Macmillan, 1973.

[15] H.-H. Hoppe, *Democracy: The God That Failed*. Transaction Publishers, 2001.

[16] F. A. Hayek, *Law, Legislation and Liberty*. University of Chicago Press, 1973.

[17] S. E. I. Konkin, *An Agorist Primer*. KoPubCo, 2006.

[18] M. Young, *The Rise of the Meritocracy*. Thames and Hudson, 1958.