Rede de Reputação Descentralizada: Um Arcabouço para a Organização Natural da Sociedade

Pavel Kudrna
Prague, Czechia
Draft v1 — March 2026

Abstract

O senso de justiça—a convicção de que os males são reparados e o mérito é recompensado—é a mais forte força de coesão da sociedade. Quando as instituições centralizadas de governança não conseguem oferecer esse senso porque o custo de fazer valer um direito excede o próprio valor do direito, a coesão social se corrói. As estruturas institucionais atuais não conseguem resolver uma classe significativa de disputas sistematicamente pela mesma razão pela qual o planejamento central falha em alocar recursos: por meio da assimetria de informação, de ciclos de retroalimentação ausentes e da desconexão entre os tomadores de decisão e as consequências de suas decisões. Este artigo apresenta uma Rede Social de Reputação (RSN): uma camada de comunicação descentralizada, incorruptível e resistente à censura, construída sobre Identificadores Descentralizados (DIDs), que permite a participantes pseudônimos publicar, verificar e consumir informações de reputação sobre comportamentos do mundo real. O mecanismo central repousa sobre três princípios de projeto: (1) uma estrutura de custo assimétrica em que ler dados de reputação é barato, mas escrever exige esforço verificável; (2) um protocolo não determinístico de seleção de verificadores baseado em hashing consistente, que precifica alegações radicais por meio de custos de iteração crescentes; e (3) um modelo de autoridade de reputação orientado pelo mercado, no qual verificadores privados apostam sua reputação acumulada na exatidão das alegações que endossam. A arquitetura resultante produz meritocracia emergente sem coordenação central e viabiliza um caminho de migração gradual e reversível a partir dos sistemas atuais administrados pelo Estado.

1 Introdução

1.1 O Problema da Confiança Centralizada

A governança moderna repousa sobre um pressuposto fundamental: o de que instituições centralizadas podem mediar a confiança entre estranhos em larga escala. Tribunais julgam disputas. Cartórios certificam a propriedade. Órgãos reguladores fazem cumprir padrões. Essas instituições foram concebidas numa era em que a informação era escassa, a comunicação era lenta e a delegação de autoridade a um órgão central era o único mecanismo de coordenação viável. A governança centralizada, no entanto, falha pelas mesmas razões estruturais que o planejamento central: assimetria de informação, ciclos de retroalimentação ausentes e desconexão entre os tomadores de decisão e as consequências de suas decisões.

O pressuposto falha, por exemplo, quando o custo da mediação institucional excede o valor da disputa. Considere um inquilino que descobre que o apartamento alugado carece de um certificado de segurança elétrica legalmente exigido—uma condição que representa perigo imediato à vida. O direito legal do inquilino de rescindir o contrato é inequívoco. Ainda assim, o custo de fazer valer esse direito pelo sistema judiciário excede o valor monetário do depósito e dos danos devidos, mesmo incluindo a eventual indenização prevista em lei [1]. A resposta racional é absorver a perda e sair.

Isso não é um caso patológico e excepcional. É a experiência padrão para uma classe significativa de disputas em que o dano é real, mas o valor monetário em jogo fica abaixo do limiar a partir do qual a execução institucional se torna economicamente racional. O resultado é um sistema que favorece estruturalmente os maus atores: aqueles que prejudicam outros em volumes abaixo desse limiar podem agir com impunidade, porque o custo de buscar justiça recai sobre a parte lesada.

O exemplo acima é extraído do ambiente jurídico

do autor—o sistema de direito civil tcheco. Em outras tradições jurídicas—o direito consuetudinário baseado em precedentes (common law), o direito costumeiro, os sistemas mistos—a justiça falha de maneiras diferentes e em graus diferentes, dependendo das especificidades culturais e processuais da jurisdição. É provável que os leitores reconheçam exemplos equivalentes em seu próprio contexto. O que é estruturalmente universal é o padrão: disputas cujo valor fica abaixo do limiar de execução economicamente racional terminam com a perda sendo absorvida pela parte lesada. A RSN não promete justiça perfeita—ela apenas reduz a vantagem estrutural de que os maus atores desfrutam quando a execução institucional é antieconômica.

1.2 Por que as Soluções Existentes Ficam Aquém

Arcabouços de Identidade Descentralizada (DID) [2] fornecem mecanismos criptográficos para a gestão de identidade autossoberana, mas concentram-se sobretudo na verificação de identidade sem abordar a questão mais ampla da reputação comportamental.

Tokens Soulbound (SBTs) [3] captam a percepção de que a reputação é intransferível, mas dependem de uma infraestrutura de blockchain com restrições de escalabilidade e não tratam dos incentivos econômicos que governam a inserção de informação. Conceitos correlatos, como tokens de mérito (por exemplo, Liberland), compartilham uma filosofia semelhante, mas permanecem atrelados a arcabouços jurisdicionais específicos.

Organizações Autônomas Descentralizadas (DAOs) [4] implementam governança por meio de contratos inteligentes, mas reproduzem dinâmicas plutocráticas em que a influência é proporcional ao capital. O voto quadrático [5] mitiga isso parcialmente, mas limita a adoção prática. As DAOs também enfrentam o *problema do oráculo* fundamental: contratos inteligentes não conseguem verificar de forma confiável estados do mundo real sem uma fonte externa confiável, e esse problema não tem solução geral dentro da arquitetura de blockchain.

Nenhum sistema existente combina descentralização, resistência à censura, pseudonimato, custo de entrada verificável e consenso emergente.

1.3 Contribuições

Este artigo apresenta as seguintes contribuições:

1. Definição da **Rede Social de Reputação (RSN)**, um protocolo descentralizado que estende o arcabouço DID para suportar a troca bilateral de reputação.
2. Um **protocolo não determinístico de seleção de verificadores** baseado em hashing consistente [6].
3. O **Princípio do Radicalismo Custoso**, que precifica alegações de acordo com sua distância em relação ao consenso da rede por meio de três canais de custo cumulativos.
4. Um **modelo de Autoridade Reputável** com incentivos assimétricos, em que o endosso falso custa catastroficamente mais do que as taxas legítimas.
5. Um **caminho de migração gradual** por meio de uma infraestrutura fiscal paralela.

2 Princípios de Projeto

A arquitetura da RSN é restringida por cinco princípios de projeto inegociáveis.

Continuidade e Evolução. O sistema coexiste com as instituições existentes durante um período de transição de duração indefinida. A transição deve ser reversível em cada etapa.

Voluntariedade e Responsabilidade. A participação é voluntária, mas a voluntariedade está acoplada à responsabilidade: um participante que assume o controle sobre uma função institucional assume simultaneamente as obrigações que a acompanham.

Diversidade e Neutralidade Cultural. O consenso emerge de interações bilaterais, não de regras predefinidas impostas pelos projetistas do sistema.

Resiliência e Resistência à Censura. O custo de censurar a rede deve exceder o custo de tolerá-la. Somente o totalitarismo pleno—a um custo político e econômico ruinoso—pode suprimir o sistema.

Consenso e Execução. O sistema incorpora as tendências humanas à mesquinhez, ao rancor e à satisfação retributiva como características estruturais. A má conduta não é proibida; ela é precificada.

3 Visão Geral do Sistema

3.1 Rede Social de Reputação

A RSN é uma camada de comunicação descentralizada, ponto a ponto, na qual os participantes trocam informações verificáveis sobre comportamentos do mundo real. Suas propriedades definidoras são: infraestrutura descentralizada e incensurável; partic-

ipação pseudônima por meio de DIDs; estrutura de custo assimétrica (ler é barato, escrever é caro); verificabilidade criptográfica; e **suporte a padrões**—formatos legíveis por máquina para as informações propagadas pela rede, para os serviços oferecidos pelas autoridades (composição de alegações, endosso, serviço de testemunho) e para o formato de política, incluindo regras para avaliar a reputação da contraparte e aferir o risco de descompasso de política (entre o comportamento declarado e o real).

3.2 Componentes Arquiteturais

A RSN é composta por quatro componentes principais (Fig. 1):

1. **Camada DID.** Identidades dos participantes com regras declaradas, metadados de reputação e roteamento de rede.
2. **Protocolo de Alegações.** Formato estruturado para publicar asserções sobre eventos do mundo real.
3. **Rede de Verificação.** Protocolo descentralizado para designar verificadores usando hashing consistente.
4. **Camada de Agregação de Reputação.** Serviços que produzem resumos de reputação legíveis por humanos.

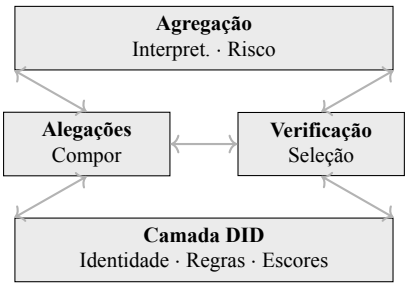

Figure 1: Arquitetura de quatro camadas da RSN.

3.3 Papéis dos Participantes

Uma transação de verificação envolve até seis papéis distintos de DID (Fig. 2): **Emissor** (DID_I, cria e submete a alegação), **Sujeito** (DID_S, o DID sobre o qual a alegação versa—pode coincidir com o Emissor), **Autoridade** (DID_A, endossa a alegação mediante taxa; pode também oferecer serviços de testemunho—*opcional*), **Testemunha** ($DID_{W_{1..n}}$, monitor incógnito da honestidade do verificador por meio de códigos de desafio—*opcional*), **Verificador** (DID_V, selecionado algoritmicamente para publicar a alegação) e a **RSN** (a rede onde as alegações verificadas são publicadas). Qualquer papel pode ser delegado a outro DID (Seção 7.4). Uma Autoridade comumente agrupa o endosso com serviços de testemunho, mas as duas funções são logicamente independentes.

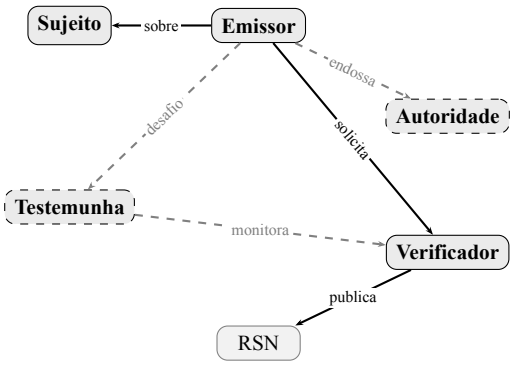

Figure 2: Papéis de DID em uma transação de verificação. Tracejado = opcional (Autoridade, Testemunha).

3.4 Incorruptibilidade: Quatro Forças

A incorruptibilidade da RSN não decorre de um único mecanismo, mas de quatro forças concorrentes. Nenhuma é suficiente por si só; apenas sua combinação torna economicamente irracional uma tentativa de corrupção.

Alvo imprevisível. O verificador de cada alegação é selecionado de forma não determinística por meio de hashing consistente (Seção 5.3). Um atacante não pode escolher previamente um verificador específico para suborno, porque a identidade do verificador é uma função do conteúdo da alegação e das iterações anteriores, não da escolha do emissor.

Alavancagem reputacional—sobe devagar, cai depressa. A reputação só pode ser ganha incrementalmente, por meio de comportamento consistente e sustentado. Ela pode, porém, ser perdida catastroficamente em um único endosso fraudulento (Seção 6.2). Essa assimetria significa que anos de reputação acumulada podem ser destruídos por um único endosso desonesto; a estratégia ótima continua sendo rejeitar alegações suspeitas.

Promessa pública. Todo Titular de DID declara publicamente sua política—um conjunto de regras legível por máquina que ele se compromete a seguir. A divergência entre a declaração e o comportamento real é detectável e precificada como uma infração reputacional mais grave do que a violação subjacente (Seção 7.3). A hipocrisia é, portanto, mais custosa do que a desonestidade direta.

Assimetria de custo de ataque da malha. O custo de censurar ou desestabilizar a rede cresce de forma

não linear com o tamanho e a densidade da rede, ao passo que o custo de tolerá-la permanece constante. Para que a censura compensasse, um ator centralizado teria de aplicá-la sobre toda a rede—exigindo medidas desproporcionais a qualquer benefício concebível (Seção 10).

4 Modelo de Identidade Descentralizada

4.1 DID com Propriedades Especiais

A RSN estende a especificação W3C DID Core [2] com: **Regras Declaradas** (compromissos comportamentais legíveis por máquina), **Metadados de Reputação** (escore comportamental agregado) e **Roteamento de Rede** (gateway onion mutável, separado da posição estável no anel).

4.2 Ciclo de Vida da Alegação

Uma Alegação percorre seis estágios (Fig. 3): composição, endosso opcional por autoridade, seleção de verificador via hashing consistente, decisão de verificação (aceitar ou rejeitar com iteração), publicação e atualização de reputação para todas as partes.

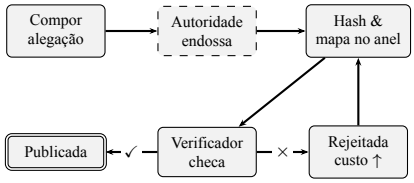

Figure 3: Ciclo de vida da alegação com laço de iteração.

4.3 Relação com o W3C DID Core

O modelo de identidade da RSN é um superconjunto próprio da especificação W3C DID Core [2]. Todos os DIDs da RSN são DIDs W3C válidos. As extensões específicas da RSN são codificadas como propriedades do documento DID definidas em uma especificação de método DID dedicada, compatível com infraestruturas existentes como ION [7] e Ceramic [8].

5 Verificação e Propagação da Informação

5.1 Custo de Inserção de Informação

O princípio econômico central da RSN é a assimetria entre ler e escrever. Ler dados de reputação aproxima-se de custo marginal zero para os participantes que fazem parte da rede DID e têm acesso proporcional à sua reputação—os recém-chegados devem conquistar gradualmente um acesso mais amplo (ver anti-parasitismo). Escrever—entendido como a materialização durável da reputação na rede, não como um ato técnico pontual—exige investimento cumulativo verificável em três dimensões: **tempo** (anos de vida despendidos em comportamento consistente, compromissos cumpridos e relacionamentos cultivados), **energia** (esforço investido em negociar com honestidade, no cuidado com as parcerias e na confiabilidade de longo prazo) e **dinheiro** (investimento no próprio nome—desenvolvimento profissional, qualidade do próprio trabalho, reparação dos danos causados). A reputação não pode ser comprada instantaneamente—ela é o resultado de uma acumulação ao longo da vida que o sistema apenas torna visível e transferível entre contextos. Os custos técnicos pontuais do protocolo (assinaturas criptográficas, taxas de verificador) são desprezíveis em comparação com esse investimento subjacente.

5.2 Grafo Social e Profundidade de Contatos

A RSN é fundamentalmente uma rede social: cada DID adiciona contatos (outros DIDs) que permitem a conexão. Esses contatos têm seus próprios contatos, e assim por diante. A busca algorítmica de verificadores opera dentro de uma profundidade configurável h de contatos ligados (por exemplo, $h = 3$: os contatos diretos de alguém, os contatos deles e os contatos dos contatos). Essa arquitetura não requer uma única blockchain global—ela favorece naturalmente o surgimento de comunidades/clusters com sobreposições em outras comunidades. Todo participante DID deve ter uma **política** publicada—um conjunto de regras legível por máquina que governa como as solicitações recebidas são avaliadas. As violações de política são registradas na rede como artefatos negativos.

5.3 Seleção Algorítmica de Verificadores

O protocolo de verificação emprega hashing consistente [6] (Fig. 4). A verificação é um serviço pago: os

verificadores operam mediante taxa, criando um mercado em que a concorrência impulsiona preço, velocidade e confiabilidade.

Passo 1. O documento da alegação é concatenado com o resultado do hash da iteração anterior (ou ∅ na primeira iteração) e submetido a hash:

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

O algoritmo deve incluir o *caminho completo* de todas as solicitações e respostas anteriores—não apenas um hash da iteração anterior. A resposta completa de cada verificador (aceitação, rejeição, preço cotado, motivo) é anexada ao documento em crescimento, verificável por meio de assinaturas criptográficas em cada etapa.

Passo 2. O hash é mapeado para N posições no anel de hash consistente, distribuídas uniformemente (por exemplo, para $N = 4$: $+0°, +90°, +180°, +270°$). A partir de cada posição, uma busca no sentido horário seleciona o DID mais próximo como candidato a verificador (Fig. 5). N é um parâmetro variável que pode depender do percentual de DIDs atualmente ativos, da hora do dia ou da capacidade de resposta histórica da rede.

Passo 3. O Verificador aceita (publica, apostando reputação) ou rejeita (retornando ao Passo 1 com o hash de rejeição). Quando um nó não responde apesar de declarar estar on-line, a solicitação seguinte deve incluir todas as respostas de todos os N candidatos a verificador anteriores.

Anti-parasitismo. Os DIDs de destino podem estabelecer taxas ou restrições de acesso para consultas provenientes de DIDs novos com pouca reputação. Esse mecanismo gradual (não binário) obriga os recém-chegados a construir reputação por meio de atividade genuína antes de consumir livremente os dados alheios.

Cada rejeição anexa a resposta completa do verificador (incluindo motivos e condições) ao documento da alegação, expandindo seu conteúdo. A partir do documento expandido, um novo hash é computado de forma não determinística, mapeando para uma posição diferente no anel e selecionando um verificador diferente. A documentação do caminho completo é obrigatória—o emissor não pode prever nem influenciar o próximo verificador. Se um verificador ignora uma solicitação apesar de uma política declarada compatível, as testemunhas (se presentes) registram isso, e o emissor pode anexar a evidência das testemunhas na publicação final.

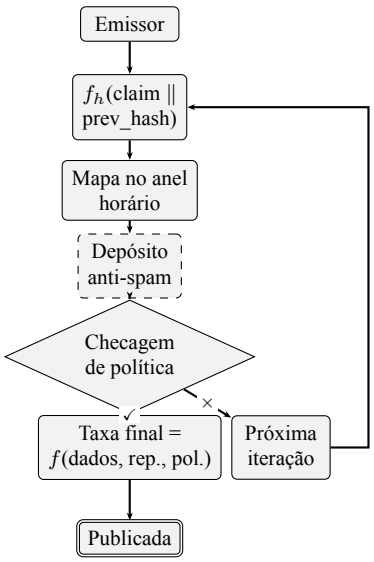

Figure 4: Protocolo de seleção de verificador. O depósito anti-spam é opcional e pode ser reembolsável. A taxa final é paga somente na aceitação.

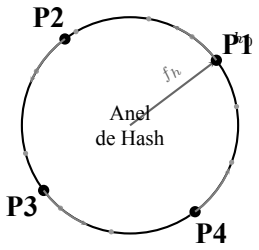

Figure 5: Seleção multiponto (N=4). O hash h_0 define o deslocamento; 4 pontos distribuídos uniformemente a partir de h_0.

5.4 Protocolo de Testemunho

O papel de **Testemunha** fornece responsabilização incógnita para a honestidade do verificador por meio de um protocolo criptográfico de código de desafio:

Passo W1. O solicitante assina a solicitação de verificação e a envia a N_w testemunhas—contatos da rede social do solicitante ou provedores especializados de serviço de testemunho que operam mediante taxa.

Passo W2. Cada testemunha w_i assina toda a solicitação assinada, retém a assinatura original σ_i e computa um código de desafio $c_i = h(\sigma_i)$. O código de desafio é anexado à solicitação.

Passo W3. A solicitação, agora carregando N_w códigos de desafio, é enviada ao verificador. O verificador observa os códigos, mas *não consegue determinar* quem são as testemunhas nem se os códigos correspondem a assinaturas reais.

Passo W4 (verificador honesto). Se o verificador

responde de acordo com sua política declarada, os códigos de desafio permanecem opacos. As testemunhas nunca são reveladas. Nenhum dado adicional é publicado.

Passo W5 (violação de política). Se o verificador viola sua política (ignora a solicitação, responde de forma inconsistente), o solicitante detém as assinaturas originais das testemunhas σ_i. Estas podem ser publicadas como dados complementares: qualquer um pode verificar $c_i = h(\sigma_i)$, provando que a testemunha estava presente. O solicitante pode publicar de forma independente—o verificador já assinou os códigos de desafio em sua resposta.

A propriedade do blefe. O solicitante pode substituir alguns ou todos os códigos de desafio por valores aleatórios. O verificador não consegue distinguir códigos genuínos (respaldados por autoridades de alta reputação) de ruído. Essa *incerteza* é o mecanismo de execução: toda solicitação carrega o risco de que uma autoridade respeitada esteja observando de forma incógnita. O custo de manter essa pressão é quase nulo (gerar números aleatórios), enquanto o custo potencial da desonestidade para o verificador é catastrófico. O comportamento honesto é incentivado mesmo na ausência de testemunhas reais.

Autoridade como testemunha. Uma Autoridade que endossa uma alegação pode agrupar serviços de testemunho: "Endosso sua alegação e sirvo como testemunha incógnita durante a verificação." Esse é um modelo de negócio natural—a Autoridade já dispõe do contexto. No entanto, os dois papéis são logicamente independentes.

5.5 Princípio do Radicalismo Custoso

Três canais de custo se acumulam simultaneamente (Fig. 6):

Canal 1: Custo de Iteração. Cada rejeição força uma nova iteração que consome tempo e recursos. Crescimento linear.

Canal 2: Inchaço do Documento. Cada iteração adiciona a resposta completa do verificador ao documento da alegação. O crescimento é linear, mas com um deslocamento de base: a carga inicial (conteúdo da alegação, endosso de autoridade, assinaturas criptográficas) já tem tamanho não trivial B_0. Tamanho total após k iterações: $S(k) = B_0 + k \cdot \Delta$, onde Δ é o tamanho médio da resposta.

Canal 3: Prêmio de Risco do Verificador. O risco é subjetivo. O verificador avalia se as políticas declaradas do DID solicitante conflitam com os interesses comerciais, sociais ou políticos do próprio verificador. O prêmio pode escalar exponencialmente com a distância percebida em relação ao "ideal" do verificador: $P(d) = \alpha \cdot e^{\beta d}$, onde d é a métrica de distância subjetiva do verificador. Além de um limite pessoal intransponível, o verificador pode recusar por completo.

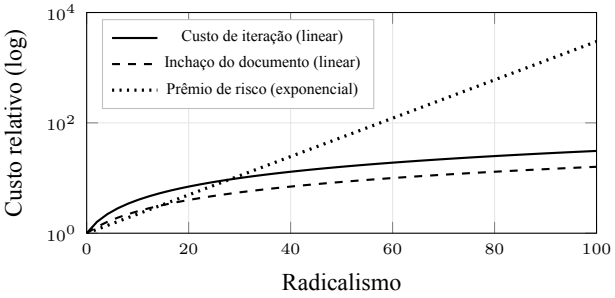

Figure 6: Escalada de custo ao longo de três canais. O prêmio de risco domina em alto radicalismo.

Crucialmente, isso não é censura. Nenhuma alegação é proibida. O sistema precifica o risco (Tabela 1).

Table 1: Comparação de custos entre tipos de alegação.

Tipo	Iter.	Doc	Taxa	Total
Crível	$k_{\min}$	B_0	$P_{\min}$	Baixo
Controversa	k_{med}	$B_0 + k\Delta$	$\alpha e^{\beta d}$	Moderado
Radical	$k_{\max}$	$\gg B_0$	$\gg P_{\min}$	Muito alto
Fraudulenta	$\to \infty$	—	—	Impublicável

6 Reputação e Avaliação de Risco

6.1 Modelo de Acumulação de Reputação

A reputação exibe três propriedades: **acumulação lenta** (crescimento incremental por meio de comportamento consistente), **destruição rápida** (uma única fraude pode reduzir o escore catastroficamente) e **avaliação individual** (cada parte consumidora pode aplicar sua própria função de agregação).

6.2 Autoridades Reputáveis

Uma Autoridade Reputável revisa alegações e, se satisfeita, coassina-as—apostando sua própria reputação (Fig. 7). A assimetria é intencional: ganhar reputação é lento (incremento $+\delta$ por endosso honesto), enquanto perdê-la é catastrófico ($-\Delta \gg \delta$ por endosso falso; a título ilustrativo, por exemplo, $+2\%$ contra -70%). A estratégia ótima é sempre rejeitar alegações suspeitas. Os valores específicos de δ e Δ são parâmetros do sistema.

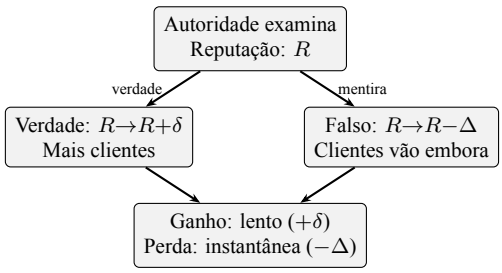

Figure 7: Autoridade Reputável—incentivos assimétricos.

6.3 Reputação Multidimensional

A reputação não é um escalar, mas um vetor ao longo de múltiplas dimensões: confiabilidade financeira, integridade pessoal, competência profissional, engajamento cívico e outras (Fig. 8). Diferentes provedores de avaliação projetam esse vetor de maneira distinta conforme o contexto—um credor prioriza a confiabilidade financeira, um empregador a competência profissional, um vizinho o comportamento cívico. Não existe um único escore de reputação "correto"; apenas perspectivas.

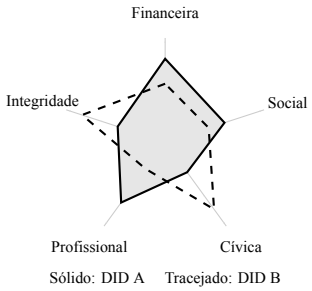

Figure 8: Vetores de reputação multidimensionais. Diferentes DIDs exibem perfis diferentes ao longo das dimensões.

6.4 Avaliação de Risco Democratizada

A RSN democratiza a avaliação sistemática de risco—uma capacidade hoje concentrada nas instituições financeiras, mas aplicável muito além do setor bancário. Conglomerados industriais avaliando a confiabilidade de fornecedores, seguradoras precificando o risco comportamental, due diligence para fusões e aquisições, estimativa de vida útil de equipamentos na indústria—em qualquer lugar em que o risco da contraparte precise ser avaliado, a RSN oferece uma alternativa descentralizada e orientada pelo mercado. Um mercado de serviços de interpretação traduz dados brutos de reputação multidimensional em resumos acionáveis, tornando a avaliação da contraparte acessível a qualquer participante a custo marginal.

7 Consenso e Resolução de Disputas

7.1 Modelo de Justiça Descentralizada

A RSN reformula a justiça como um problema de negociação bilateral. A ameaça de um dano reputacional permanente e verificável é um dissuasor mais eficaz do que ações judiciais que a parte lesada não pode custear.

7.2 Consenso Emergente

Cada participante mantém um limiar pessoal de satisfação. O consenso agregado da rede emerge como a média estatística de milhares de negociações bilaterais (Fig. 9). Uma resposta muito acima do consenso (bárbara) é cara de publicar. Uma resposta próxima ao consenso (proporcional) é barata. O consenso não é uma regra—é um sinal de preço.

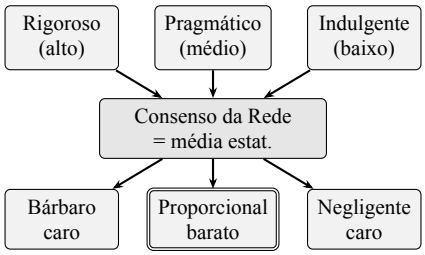

Figure 9: Consenso emergente a partir dos limiares individuais de satisfação.

7.3 Calibração da Punição

Cada Titular de DID declara publicamente respostas a categorias específicas de má conduta. Declarações desproporcionalmente severas ou brandas atraem sinais reputacionais negativos. Declarações proporcionais são aceitas sem atrito—um sistema de punição autocalibrável.

As declarações de consenso estão elas próprias sujeitas à detecção de hipocrisia: comprometer-se declarativamente com uma posição de consenso enquanto se comporta de forma inconsistente constitui uma infração reputacional mais grave do que o próprio desvio subjacente, pois demonstra engano intencional.

7.4 Delegação

Todas as atividades dentro de um DID—com uma exceção—podem ser delegadas a outro DID. **O papel de Sujeito—o DID cujo comportamento é objeto da alegação—não pode ser delegado; ele está intransferivelmente vinculado ao titular da identidade a quem a alegação se refere.** Os

demais papéis ativos—Emissor, Autoridade, Testemunha, Verificador—podem ser transferidos a um DID delegado designado, seja para verificação, submissão de alegação, participação no consenso ou resolução de disputas. A delegação é revogável a qualquer momento. Isso possibilita a especialização (por exemplo, serviços profissionais de verificação) enquanto o Titular mantém o controle e a responsabilidade em última instância. A delegação aplica-se a todo o sistema e é essencial para a escalabilidade.

7.5 Da Moralidade Individual ao Contrato Social Emergente

A política declarada de cada DID representa uma formalização da moralidade individual: o que o Titular considera aceitável, como responde a comportamentos específicos, o que exige das contrapartes. Essas políticas não são impostas por um projetista do sistema—são escolhidas por cada participante e expressas em forma legível por máquina.

Essa formalização possibilita uma emergência em três estágios. Primeiro, a moralidade individual é codificada como **política**—um conjunto de regras declaradas, limiares e funções de resposta que o DID se compromete a seguir. Segundo, o agregado de todas as políticas ao longo da rede produz **ética emergente**—normas estatisticamente observáveis que nenhum participante isolado projetou, mas que surgem da interação de milhares de posições morais individuais [9]. Terceiro, essa ética emergente, expressa como normas comportamentais compartilhadas e executada por meio de sinais de preço bilaterais, constitui um **contrato social mensurável**—não um construto teórico que ninguém assinou [10], mas um conjunto de regras empiricamente observável e respaldado por comportamento efetivo.

Esse processo espelha achados da teoria das fundações morais [11]: a moralidade humana é intuitiva, pluralista e culturalmente variável. A RSN não requer consenso moral como entrada; ela produz consenso comportamental como saída. O contrato social resultante não é estático—ele evolui à medida que os participantes entram, saem e ajustam suas políticas em resposta à retroalimentação da rede. Diferentemente da teoria clássica do contrato social, esse contrato é revogável, observável e executável sem um soberano.

8 Modelo Econômico

As seções anteriores descreveram uma ferramenta—os mecanismos de verificação, a estrutura de custo e o consenso emergente da RSN. Esta seção descreve uma metodologia para aplicar essa ferramenta à transição fiscal e de governança. A ferramenta é de propósito geral; a metodologia a seguir é uma aplicação possível.

8.1 Estrutura de Incentivos

A reputação como capital cria incentivos diretos ao comportamento honesto. Em operação normal, os participantes constroem reputação naturalmente por meio de transações cotidianas e compromissos cumpridos. O principal dispêndio ocorre com alegações *negativas*—o registro de danos causados por outros. Essa assimetria incentiva o comportamento útil e confiável: ser honesto não custa nada extra, enquanto ser danoso cria um rastro documentado que outros pagarão para preservar. A hipocrisia—declarar regras sem seguilas—é o modo de falha mais caro, pois demonstra engano intencional.

8.2 Sistema Fiscal Paralelo

Três componentes permitem a pressão competitiva: (1) **Imposto Simples**—uma alíquota proporcional única, sem exceções, inicialmente marginalmente abaixo da alíquota efetiva vigente; (2) **Registro Eletrônico de Gastos (REG)**—invertendo o modelo tcheco EET para que os cidadãos fiscalizem os gastos do Estado; (3) **Destinação Tributária Dirigida pelo Cidadão**—começando com alguns pontos percentuais no primeiro ano e alcançando, após uma década, algo entre os dois dígitos baixos e uma migração completa para um sistema dirigido pelo cidadão, conforme o ritmo de adoção. O tempo efetivo depende da velocidade com que surgem alternativas de livre mercado aos serviços estatais e da disposição do Estado em cooperar com a transição; a função do tempo não precisa ser linear, e não há razão para estabelecer um limite superior para onde a adoção pode, em última instância, chegar.

9 Caminho de Migração

A migração procede por três fases (Fig. 10): **Fase 1: Sobreposição** (a RSN como camada informacional, o Estado é o provedor exclusivo), **Fase 2: Concorrência** (a RSN oferece alternativas funcionais, uso de sistema duplo) e **Fase 3: Transição** (a RSN supera os equivalentes estatais, migração voluntária). A transição é reversível em cada etapa.

O principal mecanismo de pressão é fiscal: quando os contribuintes condicionais atingem uma "minoria economicamente significativa X"—um grupo grande

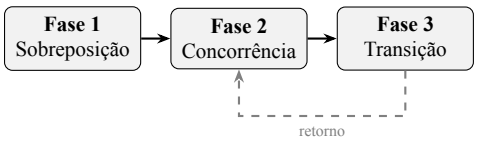

Figure 10: Migração em três fases—reversível em cada etapa.

o bastante para que a repressão contra ele cause dano econômico não trivial e a desobediência civil se torne uma ameaça crível—o Estado entra em negociação. Para fins de ilustração, usamos $X \approx 15\%$ do PIB, mas o limiar efetivo depende da economia específica.

10 Análise de Segurança

Consideramos cinco categorias de adversários: maus atores individuais, grupos organizados, adversários corporativos, adversários de nível estatal e adversários de nível de protocolo.

Resistência a Sybil [12]. Construir reputação exige interação sustentada e verificável. O custo de um ataque Sybil bem-sucedido escala linearmente com o número de identidades falsas e quadraticamente com o nível de reputação almejado. Operar múltiplos DIDs em paralelo é possível, mas caro por projeto—cada identidade deve acumular reputação de forma independente por meio de atividade genuína. Em sociedades livres, esse custo dissuade o abuso casual; em ditaduras, os DIDs paralelos tornam-se uma ferramenta de sobrevivência que possibilita resistência compartimentada, coordenação clandestina e navegação mais segura no mercado negro.

Defesa contra conluio. Padrões de endosso mútuo entre grupos fechados são detectáveis por meio da análise do grafo social. As funções de agregação de reputação descontam alegações provenientes de fontes densamente aglomeradas.

Adversário de nível estatal. O roteamento onion, a ausência de infraestrutura centralizada e a distribuição do papel de Verificador por toda a base de participantes garantem que a supressão exija medidas desproporcionais a qualquer benefício.

Privacidade vs. responsabilização. O pseudonimato—um meio-termo entre o anonimato e a divulgação da identidade—permite que os DIDs acumulem reputação significativa sem revelar a identidade real do Titular no mundo.

11 Considerações sobre Privacidade

O modelo de privacidade da RSN é construído sobre o pseudonimato. O Titular controla a divulgação da identidade: mínima (pseudônimo puro), seletiva (credenciais específicas vinculadas) ou plena (identidade legal associada). As alegações publicadas são permanentes—o sistema suporta correção contextual, mas não apagamento. Um Titular pode abandonar um DID comprometido e recomeçar, abrindo mão da reputação acumulada.

12 Trabalhos Relacionados

A especificação W3C DID Core [2] e a Ceramic Network [8] fornecem a base de identidade. O EigenTrust [13] demonstrou a agregação distribuída de reputação sem autoridade central. Os SBTs [3] introduziram tokens sociais intransferíveis. A RSN difere na sua ênfase no custo econômico como filtro de qualidade e no seu mecanismo de precificação do radicalismo.

As DAOs [4] e o voto quadrático [5] demonstraram a viabilidade da governança descentralizada. A RSN deriva o consenso de negociações bilaterais de preço, em vez de votação.

A proposta baseia-se na teoria anarcocapitalista [14, 15] para a provisão privada de serviços, mas dela se afasta em dois aspectos críticos: projeta para o rancor e os impulsos retributivos em vez de supor racionalidade, e propõe uma transição gradual em vez de revolução. O conceito de contratos sociais emergentes tem antecedentes na teoria da ordem espontânea de Hayek [16].

Anarco-agorismo. A RSN compartilha um significativo terreno comum com a contraeconomia agorista [17]—particularmente a ênfase na construção de instituições paralelas e na troca voluntária. No entanto, o agorismo tende a supor o interesse próprio racional como mecanismo de coordenação suficiente e frequentemente carece de um caminho concreto de migração a partir das estruturas estatais existentes. A RSN incorpora explicitamente tendências comportamentais não racionais e fornece um mecanismo de pressão fiscal para uma transição gradual.

Meritocracia. A RSN pode representar uma primeira descrição funcional de como a meritocracia [18] pode emergir como uma propriedade sistêmica, e não como um slogan. Os sistemas meritocráticos tradicionais falham porque exigem uma autoridade central para definir e fazer valer o “mérito”. Na RSN, o mérito emerge de avaliações bilaterais: aqueles que entregam valor acumulam reputação; nen-

hum comitê decide quem tem mérito.

Propriedade como privilégio. A RSN afasta-se fundamentalmente dos tratamentos anarcocapitalistas e da escola austríaca sobre os direitos de propriedade como naturais e absolutos. No arcabouço da RSN, a propriedade é um *privilégio*—um reconhecimento social que pode, em casos extremos, ser retirado pela comunidade quando um participante viola fundamentalmente normas compartilhadas (traição, recusa em defender a comunidade em crise existencial, exploração sistemática). Isso não é expropriação estatal, mas a retirada do reconhecimento social pela rede de relações bilaterais.

13 Discussão e Limitações

Partida a frio. O valor da RSN depende de efeitos de rede; os primeiros a adotá-la enfrentam valor limitado até que a participação atinja um limiar. Contudo, mesmo desde os estágios iniciais, a rede DID serve como uma útil fonte suplementar de informação. Espera-se que a avaliação inicial de reputação e a aferição de autoridade se desenvolvam gradualmente, com sofisticação crescente.

Anti-parasitismo e controle de acesso. Os DIDs de destino podem impor taxas graduadas e restrições de acesso a consultas provenientes de DIDs de baixa reputação. Isso não é binário, mas uma escala contínua: quanto menos reputação um DID consultante tem, menos informação recebe, mais tempo espera e mais paga. Esse mecanismo incentiva a participação genuína em vez do consumo passivo.

Desigualdade de acesso. O custo de publicar alegações pode filtrar queixas legítimas de participantes economicamente desfavorecidos. Mitigação: todo participante serve simultaneamente como verificador em potencial (ou delega esse papel) e aufere taxas de verificação. Ao longo de um horizonte de tempo suficiente, um participante não radical deveria aproximar-se da neutralidade financeira—a renda de verificação compensando aproximadamente os custos de publicação. Trata-se de uma expectativa estatística, não de uma garantia.

Desigualdade de reputação. Os primeiros a adotar acumulam vantagens que podem criar barreiras para os retardatários. No entanto, a avaliação de reputação é realizada por provedores terceiros, que podem optar por mitigar a vantagem do pioneiro por meio de seus algoritmos de agregação. Ser o primeiro com boa reputação é uma vantagem econômica comparativa legítima—e isso é intencional.

Questões em aberto. Funções de custo ótimas, padrões de agregação, governança do protocolo e a interação com dados de reputação sintética gerados por IA permanecem áreas para trabalho futuro.

Este artigo não afirma que a RSN produzirá resultados ótimos em todas as circunstâncias. Ele afirma que a RSN representa uma alternativa *viável*—tecnicamente implementável, economicamente autossustentável e socialmente compatível com as tendências comportamentais humanas tais como realmente são.

14 Conclusão

Este artigo apresentou a Rede Social de Reputação, um protocolo descentralizado que possibilita a troca pseudônima e verificável de reputação. O Princípio do Radicalismo Custoso garante que alegações fraudulentas sejam precificadas para fora do sistema, sem censura. O caminho de migração proposto viabiliza uma transição gradual e reversível a partir das instituições existentes. O projeto incorpora a natureza humana tal como ela é—incluindo a mesquinhez, o rancor e o desejo de satisfação retributiva—em vez de exigir uma transformação ideológica. O resultado não é a utopia, mas um sistema em que a justiça é acessível, a reputação é conquistada e o custo da má conduta é arcado pela parte que o causou.

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 identifiers (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.