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

Pavel Kudrna
Prague, Czechia
Rascunho v1 — Março de 2026

Abstract

O sentido de justiça—a convicção de que os agravos são reparados e o mérito é recompensado—é a mais forte força de coesão da sociedade. Quando as instituições centralizadas de governação não conseguem proporcionar este sentido, porque o custo de fazer valer um direito excede o próprio valor desse direito, a coesão social erode-se. As estruturas institucionais actuais falham na resolução de uma classe significativa de litígios sistematicamente da mesma forma que o planeamento central falha na afectação de recursos: através da assimetria de informação, da ausência de mecanismos de retroacção e da desconexão entre os decisores e as consequências das 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 pseudónimos publicar, verificar e consumir informação de reputação sobre comportamento no mundo real. O mecanismo central assenta em três princípios de conceção: (1) uma estrutura de custos assimétrica em que ler dados de reputação é barato mas escrever exige um esforço verificável; (2) um protocolo não determinístico de seleção de verificadores baseado em hashing consistente, que fixa o preço de alegações radicais através de custos de iteração crescentes; e (3) um modelo de autoridade de reputação orientado pelo mercado, em que verificadores privados apostam a reputação acumulada na exatidão das alegações que endossam. A arquitetura resultante produz uma meritocracia emergente sem coordenação central e possibilita um caminho de migração gradual e reversível a partir dos sistemas actualmente administrados pelo Estado.

1 Introdução

1.1 O Problema da Confiança Centralizada

A governação moderna assenta num pressuposto fundamental: o de que as instituições centralizadas conseguem mediar a confiança entre desconhecidos em larga escala. Os tribunais julgam litígios. Os registos certificam a propriedade. Os organismos reguladores fazem cumprir normas. Estas instituições foram concebidas numa era em que a informação era escassa, a comunicação era lenta e a delegação de autoridade num corpo central era o único mecanismo de coordenação viável. A governação centralizada, contudo, falha pelas mesmas razões estruturais que o planeamento central: assimetria de informação, ausência de mecanismos de retroacção e desconexão entre os decisores e as consequências das suas decisões.

O pressuposto falha, por exemplo, quando o custo da mediação institucional excede o valor do litígio. Considere-se um inquilino que descobre que o apartamento arrendado carece de um certificado de segurança elétrica legalmente exigido—uma condição que representa perigo imediato para a vida. O direito legal do inquilino a rescindir o contrato é inequívoco. Contudo, o custo de fazer valer esse direito através do sistema judicial excede o valor monetário da caução e dos danos devidos, mesmo incluindo a eventual compensação legal [1]. A resposta racional é absorver a perda e sair.

Este não é um caso patológico e excecional. É a experiência habitual para uma classe significativa de litígios em que o dano é real mas os valores monetários em jogo ficam 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: quem prejudica os outros em montantes abaixo deste limiar pode agir impunemente, porque o custo de procurar justiça recai sobre a parte lesada.

O exemplo acima é retirado do enquadramento jurídico do autor—o sistema checo de direito civil.

Noutras tradições jurídicas—direito comum baseado em precedente, direito consuetudinário, sistemas mistos—a justiça falha de formas diferentes e em graus diferentes, consoante as especificidades culturais e processuais da jurisdição. Os leitores reconhecerão provavelmente exemplos equivalentes no seu próprio contexto. O que é estruturalmente universal é o padrão: litígios cujo valor fica abaixo do limiar de execução economicamente racional terminam com a perda a ser absorvida pela parte lesada. A RSN não promete uma justiça perfeita—limita-se a reduzir a vantagem estrutural de que os maus atores gozam quando a execução institucional não é economicamente viável.

1.2 Porque as Soluções Existentes Ficam Aquém

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

Os Soulbound Tokens (SBTs) [3] captam a ideia de que a reputação é intransferível, mas dependem de infraestrutura blockchain com limitações de escalabilidade e não abordam os incentivos económicos que regem a entrada de informação. Conceitos relacionados, como os tokens de mérito (p. ex., Liberland), partilham uma filosofia semelhante mas permanecem vinculados a enquadramentos jurisdicionais específicos.

As Organizações Autónomas Descentralizadas (DAOs) [4] implementam governação através de contratos inteligentes, mas replicam dinâmicas plutocráticas em que a influência é proporcional ao capital. O voto quadrático [5] atenua parcialmente este problema mas limita a adoção prática. As DAOs enfrentam também o *problema do oráculo* fundamental: os contratos inteligentes não conseguem verificar de forma fiável estados do mundo real sem uma fonte externa de confiança, e este problema não tem solução geral no âmbito da arquitetura blockchain.

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

1.3 Contribuições

Este artigo apresenta as seguintes contribuições:

1. A definição da **Rede Social de Reputação (RSN)**, um protocolo descentralizado que estende o modelo 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 Dispendioso**, que fixa o preço das alegações consoante a sua distância ao consenso da rede através 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 os emolumentos legítimos.
5. Um **caminho de migração gradual** através de uma infraestrutura fiscal paralela.

2 Princípios de Conceção

A arquitetura da RSN é condicionada por cinco princípios de conceção inegociáveis.

Continuidade e Evolução. O sistema coexiste com as instituições existentes durante um período de transição de duração indefinida. A transição tem de ser reversível em todas as fases.

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

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

Resiliência e Resistência à Censura. O custo de censurar a rede tem de exceder o custo de a tolerar. Só um totalitarismo pleno—a um custo político e económico ruinoso—pode suprimir o sistema.

Consenso e Execução. O sistema incorpora as tendências humanas para a mesquinhez, o rancor e a satisfação retributiva como elementos estruturais. A má conduta não é proibida; tem um preço.

3 Visão Geral do Sistema

3.1 Rede Social de Reputação

A RSN é uma camada de comunicação descentralizada e ponto a ponto na qual os participantes trocam informação verificável sobre comportamento no mundo real. As suas propriedades definidoras são: infraestrutura descentralizada e incensurável; participação pseudónima através de DIDs; estrutura de custos assimétrica (ler é barato, escrever é dispendioso);

verificabilidade criptográfica; e **suporte a normas**—formatos legíveis por máquina para a informação propagada através da rede, para os serviços oferecidos pelas autoridades (composição de alegações, endosso, serviço de testemunho) e para o formato de política, incluindo as regras de avaliação da reputação da contraparte e a estimativa do risco de discrepância de política (entre o comportamento declarado e o real).

3.2 Componentes Arquiteturais

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

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

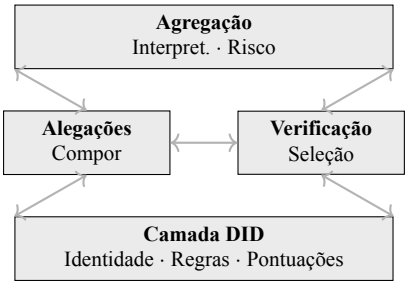

Figure 1: Arquitetura da RSN em quatro camadas.

3.3 Papéis dos Participantes

Uma transação de verificação envolve até seis papéis DID distintos (Fig. 2): **Emissor** (DID_I, cria e submete a alegação), **Sujeito** (DID_S, o DID a que a alegação diz respeito—pode coincidir com o Emissor), **Autoridade** (DID_A, endossa a alegação mediante emolumentos; pode também oferecer serviços de testemunho—*opcional*), **Testemunha** ($DID_{W_{1..n}}$, monitor incógnito da honestidade do verificador através 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 noutro DID (Secção 7.4). Uma Autoridade combina frequentemente o endosso com serviços de testemunho, mas as duas funções são logicamente independentes.

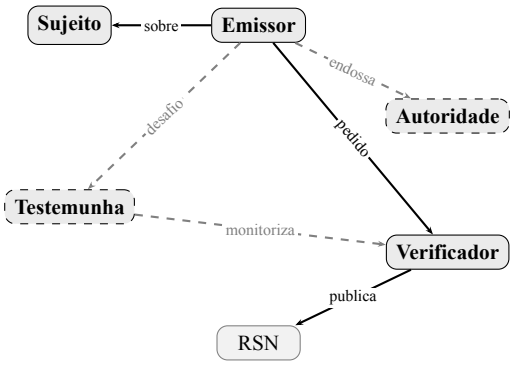

Figure 2: Papéis DID numa transação de verificação. Tracejado = opcional (Autoridade, Testemunha).

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; só a sua combinação torna economicamente irracional qualquer tentativa de corrupção.

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

Alavancagem reputacional—subida lenta, queda rápida. A reputação só pode ser adquirida de forma incremental, através de um comportamento consistente e sustentado. Pode, contudo, ser perdida de forma catastrófica num único endosso fraudulento (Secção 6.2). Esta assimetria significa que anos de reputação acumulada podem ser destruídos por um único endosso desonesto; a estratégia ótima continua a ser rejeitar alegações suspeitas.

Promessa pública. Cada Titular de DID declara publicamente a sua política—um conjunto de regras legíveis por máquina que se compromete a cumprir. A divergência entre a declaração e o comportamento real é detetável e tem um preço, como uma infração reputacional mais grave do que a violação subjacente (Secção 7.3). A hipocrisia é, por isso, mais dispendiosa do que a desonestidade direta.

Assimetria do custo de ataque em malha. O custo de censurar ou desestabilizar a rede cresce de forma não linear com a dimensão e a densidade da rede, ao passo que o custo de a tolerar permanece constante.

Para que a censura compensasse, um ator centralizado teria de a aplicar a toda a rede—exigindo medidas desproporcionadas face a qualquer benefício concebível (Secção 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** (pontuação comportamental agregada) e **Encaminhamento na Rede** (gateway onion mutável, separado da posição estável no anel).

4.2 Ciclo de Vida da Alegação

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

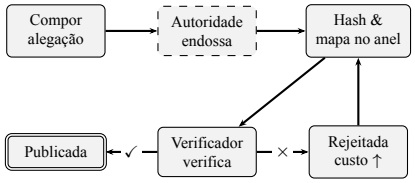

Figure 3: Ciclo de vida da alegação com laço de iteração.

4.3 Relação com a 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 numa especificação de método DID dedicada, compatível com infraestruturas existentes como a ION [7] e a Ceramic [8].

5 Verificação e Propagação da Informação

5.1 Custo de Entrada da Informação

O princípio económico central da RSN é a assimetria entre ler e escrever. Ler dados de reputação aproxima-se de um custo marginal nulo para os participantes que fazem parte da rede DID e têm acesso proporcional à sua reputação—os recém-chegados têm de conquistar gradualmente um acesso mais amplo (ver anti-parasitagem). Escrever—entendido como a materialização durável da reputação na rede, e não como um ato técnico pontual—exige um investimento cumulativo verificável em três dimensões: **tempo** (anos de vida dedicados a um comportamento consistente, a compromissos cumpridos e a relações cultivadas), **energia** (esforço investido em negociar com honestidade, cuidar das parcerias e manter uma fiabilidade a longo prazo) e **dinheiro** (investimento no próprio nome—desenvolvimento profissional, qualidade do trabalho, reparação dos danos causados). A reputação não pode ser comprada instantaneamente—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, emolumentos dos verificadores) são negligenciáveis face a este investimento subjacente.

5.2 Grafo Social e Profundidade de Contactos

A RSN é fundamentalmente uma rede social: cada DID adiciona contactos (outros DIDs) que permitem a ligação. Esses contactos têm os seus próprios contactos, e assim sucessivamente. A pesquisa algorítmica de verificadores opera dentro de uma profundidade configurável h de contactos ligados (p. ex., $h = 3$: os contactos diretos, os contactos destes e os contactos dos contactos). Esta arquitetura não exige uma única blockchain global—favorece naturalmente o surgimento de comunidades/agrupamentos com sobreposições noutras comunidades. Cada participante DID tem de ter uma **política** publicada—um conjunto de regras legível por máquina que rege a forma como os pedidos recebidos são avaliados. As violações de política são registadas na rede como artefactos negativos.

5.3 Seleção Algorítmica de Verificadores

O protocolo de verificação recorre a hashing consistente [6] (Fig. 4). A verificação é um serviço remunerado: os verificadores atuam mediante emolumentos, criando um mercado em que a concorrência determina o preço, a rapidez e a fiabilidade.

Etapa 1. O documento da alegação é concatenado com o resultado do hash da iteração anterior (ou $\varnothing$ na primeira iteração) e sujeito a hash:

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

O algoritmo tem de incluir o *caminho completo* de todos os pedidos e respostas anteriores—e não apenas um hash da iteração anterior. A resposta completa de cada verificador (aceitação, rejeição, preço proposto, motivo) é acrescentada ao documento em crescimento, verificável através de assinaturas criptográficas em cada etapa.

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

Etapa 3. O Verificador aceita (publica, apostando reputação) ou rejeita (voltando à Etapa 1 com o hash da rejeição). Quando um nó não responde apesar de declarar estar em linha, o pedido seguinte tem de incluir todas as respostas de todos os N candidatos a verificador anteriores.

Anti-parasitagem. Os DIDs-alvo podem definir emolumentos ou restrições de acesso para consultas provenientes de DIDs novos, com pouca reputação. Este mecanismo gradual (não binário) obriga os recém-chegados a construir reputação através de atividade genuína antes de poderem consumir livremente os dados dos outros.

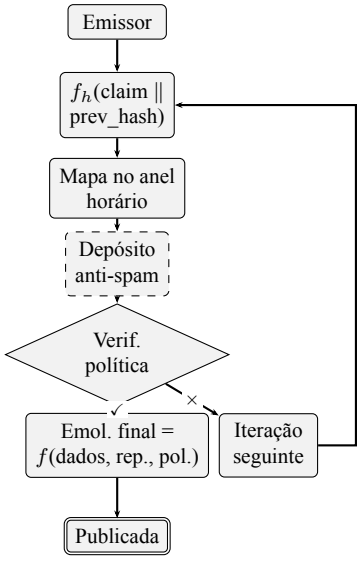

Figure 4: Protocolo de seleção de verificadores. O depósito anti-spam é opcional e pode ser reembolsável. Os emolumentos finais só são pagos em caso de aceitação.

Cada rejeição acrescenta a resposta completa do verificador (incluindo motivos e condições) ao documento da alegação, ampliando o seu conteúdo. A partir do documento ampliado, é calculado de forma não determinística um novo hash, que mapeia numa posição diferente do anel e seleciona um verificador diferente. A documentação do caminho completo é obrigatória—o emissor não pode prever nem influenciar o verificador seguinte. Se um verificador ignorar um pedido apesar de uma política declarada compatível, as testemunhas (se presentes) registam-no, e o emissor pode anexar a prova testemunhal aquando da publicação final.

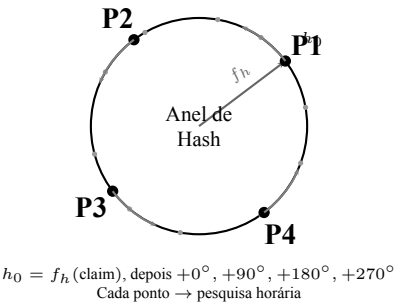

Figure 5: Seleção multiponto (N=4). O hash h_0 define o desvio; 4 pontos distribuídos uniformemente a partir de h_0.

5.4 Protocolo de Testemunho

O papel de **Testemunha** confere responsabilização incógnita para a honestidade do verificador através de um protocolo criptográfico de códigos de desafio:

Etapa W1. O requerente assina o pedido de verificação e envia-o a N_w testemunhas—contactos da rede social do requerente ou prestadores especializados de serviços de testemunho que atuam mediante emolumentos.

Etapa W2. Cada testemunha w_i assina todo o pedido assinado, guarda a assinatura original σ_i e calcula um código de desafio $c_i = h(\sigma_i)$. O código de desafio é acrescentado ao pedido.

Etapa W3. O pedido, agora transportando N_w códigos de desafio, é enviado ao verificador. O verificador observa os códigos mas *não consegue determinar* quem são as testemunhas nem se os códigos correspondem a assinaturas reais.

Etapa W4 (verificador honesto). Se o verificador responder de acordo com a sua política declarada, os códigos de desafio permanecem opacos. As testemunhas nunca são reveladas. Não é publicado qualquer dado adicional.

Etapa W5 (violação de política). Se o verificador violar a sua política (ignorar o pedido, responder de

forma inconsistente), o requerente detém as assinaturas originais σ_i das testemunhas. Estas podem ser publicadas como dados complementares: qualquer pessoa pode verificar $c_i = h(\sigma_i)$, provando que a testemunha estava presente. O requerente pode publicar de forma independente—o verificador já assinou os códigos de desafio na sua resposta.

A propriedade do bluff. O requerente pode substituir alguns ou todos os códigos de desafio por valores aleatórios. O verificador não consegue distinguir os códigos genuínos (respaldados por autoridades de elevada reputação) do ruído. Esta *incerteza* é o mecanismo de execução: cada pedido comporta o risco de que uma autoridade respeitada esteja a observar de forma incógnita. O custo de manter esta pressão é praticamente nulo (gerar números aleatórios), ao passo que 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 combinar serviços de testemunho: "Endosso a sua alegação e sirvo de testemunha incógnita durante a verificação." Trata-se de um modelo de negócio natural—a Autoridade já dispõe do contexto. Contudo, os dois papéis são logicamente independentes.

5.5 Princípio do Radicalismo Dispendioso

Três canais de custo somam-se em simultâneo (Fig. 6):

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

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

Canal 3: Prémio de Risco do Verificador. O risco é subjetivo. O verificador avalia se as políticas declaradas do DID requerente entram em conflito com os seus próprios interesses comerciais, sociais ou políticos. O prémio pode escalar exponencialmente com a distância percecionada face ao "ideal" do verificador: $P(d) = \alpha \cdot e^{\beta d}$, onde d é a métrica de distância subjetiva do verificador. Além de uma fronteira pessoal intransponível, o verificador pode recusar por completo.

Fundamentalmente, isto não é censura. Nenhuma alegação é proibida. O sistema fixa o preço do risco (Tabela 1).

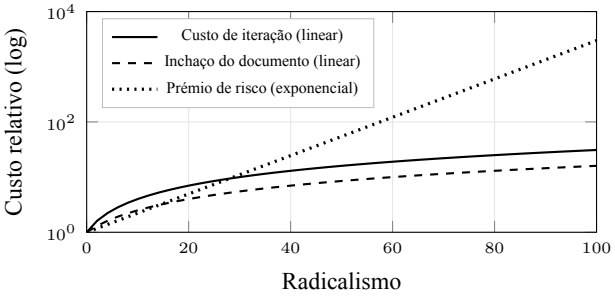

Figure 6: Escalada de custos ao longo de três canais. O prémio de risco domina em níveis elevados de radicalismo.

Table 1: Comparação de custos por tipo de alegação.

Tipo	Iter.	Doc	Emol.	Total
Credível	$k_{\min}$	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 Avaliação de Reputação e de Risco

6.1 Modelo de Acumulação de Reputação

A reputação exibe três propriedades: **acumulação lenta** (crescimento incremental através de comportamento consistente), **destruição rápida** (uma única fraude pode reduzir a pontuação de forma catastrófica) e **avaliação individual** (cada parte consumidora pode aplicar a sua própria função de agregação).

6.2 Autoridades Reputáveis

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

6.3 Reputação Multidimensional

A reputação não é um escalar, mas um vetor ao longo de múltiplas dimensões: fiabilidade financeira, integridade pessoal, competência profissional, envolvimento

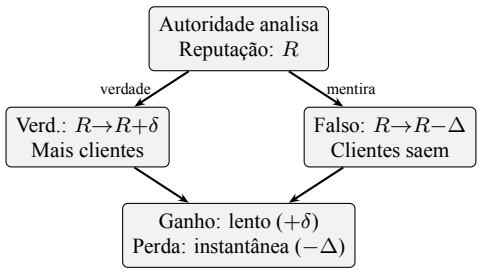

Figure 7: Autoridade Reputável—incentivos assimétricos.

cívico e outras (Fig. 8). Diferentes prestadores de avaliação projetam este vetor de forma distinta consoante o contexto—um credor prioriza a fiabilidade financeira, um empregador a competência profissional, um vizinho o comportamento cívico. Não existe uma única pontuação de reputação "correta"; apenas perspetivas.

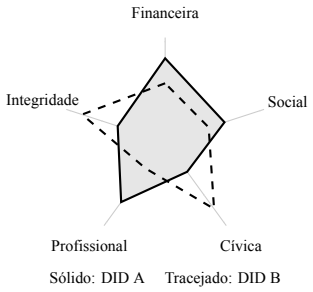

Figure 8: Vetores de reputação multidimensionais. Diferentes DIDs exibem perfis distintos ao longo das dimensões.

6.4 Avaliação de Risco Democratizada

A RSN democratiza a avaliação sistemática de risco—uma capacidade atualmente concentrada nas instituições financeiras mas aplicável muito para além da banca. Conglomerados industriais que avaliam a fiabilidade de fornecedores, seguradoras que fixam o preço do risco comportamental, diligências prévias para fusões e aquisições, estimativa do tempo de vida útil de equipamentos na indústria transformadora—onde quer que o risco de contraparte exija avaliação, a RSN oferece uma alternativa descentralizada e orientada pelo mercado. Um mercado de serviços de interpretação traduz os dados brutos de reputação multidimensional em resumos acionáveis, tornando a avaliação de contrapartes acessível a qualquer participante a custo marginal.

7 Consenso e Resolução de Litígios

7.1 Modelo de Justiça Descentralizada

A RSN reformula a justiça como um problema de negociação bilateral. A ameaça de dano reputacional permanente e verificável é um dissuasor mais eficaz do que um processo judicial 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 dispendiosa de publicar. Uma resposta próxima do consenso (proporcional) é barata. O consenso não é uma regra—é um sinal de preço.

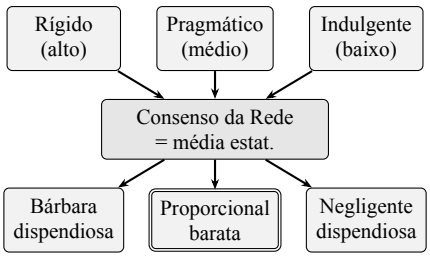

Figure 9: Consenso emergente a partir de limiares individuais de satisfação.

7.3 Calibração da Punição

Cada Titular de DID declara publicamente as respostas a categorias específicas de má conduta. Declarações desproporcionadamente severas ou brandas atraem sinais reputacionais negativos. Declarações proporcionais são aceites sem fricção—um sistema de punição autocalibrado.

As próprias declarações de consenso estão sujeitas à deteção de hipocrisia: comprometer-se declarativamente com uma posição de consenso mas comportar-se de forma inconsistente constitui uma infração reputacional mais grave do que o desvio subjacente, pois demonstra um engano intencional.

7.4 Delegação

Todas as atividades no âmbito de um DID—com uma exceção—podem ser delegadas noutro DID. **O papel de Sujeito—o DID cujo comportamento a alegação descreve—não pode ser delegado; está intransferivelmente vinculado ao titular de identi-**

dade a que a alegação se refere. Os restantes papéis ativos—Emissor, Autoridade, Testemunha, Verificador—podem ser transferidos para um DID delegado designado, seja para verificação, submissão de alegações, participação no consenso ou resolução de litígios. A delegação é revogável a qualquer momento. Isto permite a especialização (p. ex., serviços profissionais de verificação) enquanto o Titular retém o controlo e a responsabilidade últimos. 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. Estas políticas não são impostas por um conceptor do sistema—são escolhidas por cada participante e expressas em formato legível por máquina.

Esta formalização possibilita uma emergência em três fases. Primeiro, a moralidade individual é codificada como **política**—um conjunto de regras declaradas, limiares e funções de resposta que o DID se compromete a cumprir. Segundo, o agregado de todas as políticas da rede produz uma **ética emergente**—normas estatisticamente observáveis que nenhum participante isolado concebeu, mas que resultam da interação de milhares de posições morais individuais [9]. Terceiro, esta ética emergente, expressa como normas comportamentais partilhadas e reforçada por sinais de preço bilaterais, constitui um **contrato social mensurável**—não uma construção teórica que ninguém assinou [10], mas um conjunto de regras empiricamente observável e respaldado por comportamento efetivo.

Este processo espelha as conclusões da teoria dos fundamentos morais [11]: a moralidade humana é intuitiva, pluralista e culturalmente variável. A RSN não exige consenso moral como entrada; produz consenso comportamental como saída. O contrato social resultante não é estático—evolui à medida que os participantes entram, saem e ajustam as suas políticas em resposta à retroação da rede. Ao contrário da teoria clássica do contrato social, este contrato é revogável, observável e executável sem um soberano.

8 Modelo Económico

As secções anteriores descreveram uma ferramenta—os mecanismos de verificação da RSN, a estrutura de custos e o consenso emergente. Esta secção descreve uma metodologia para aplicar essa ferramenta à transição fiscal e de governação. A ferramenta é de uso geral; a metodologia que se segue é uma das aplicações possíveis.

8.1 Estrutura de Incentivos

A reputação como capital cria incentivos diretos ao comportamento honesto. Em funcionamento normal, os participantes constroem reputação naturalmente através de transações quotidianas e compromissos cumpridos. A despesa principal recai sobre as alegações *negativas*—o registo de danos causados por outros. Esta assimetria incentiva um comportamento útil e fiável: ser honesto não acarreta custo adicional, ao passo que ser prejudicial cria um rasto documentado que outros pagarão para preservar. A hipocrisia—declarar regras sem as cumprir—é o modo de falha mais dispendioso, pois demonstra um engano intencional.

8.2 Sistema Fiscal Paralelo

Três componentes possibilitam a pressão concorrencial: (1) **Imposto Simples**—uma única taxa proporcional sem exceções, inicialmente ligeiramente abaixo da taxa efetiva existente; (2) **Registo Eletrónico de Despesa (RED)**—invertendo o modelo checo EET para que os cidadãos fiscalizem a despesa do Estado; (3) **Afetação de Impostos Dirigida pelo Cidadão**—começando em alguns pontos percentuais no primeiro ano e alcançando, ao fim de uma década, algo entre as dezenas baixas de por cento e uma migração completa para um sistema dirigido pelo cidadão, consoante o ritmo de adoção. O tempo efetivo depende da rapidez com que surgem alternativas de mercado livre aos serviços do Estado e da disposição do Estado para cooperar com a transição; a função do tempo não tem de ser linear, e não há razão para fixar um limite superior ao ponto onde a adoção pode acabar por chegar.

9 Caminho de Migração

A migração processa-se em três fases (Fig. 10): **Fase 1: Sobreposição** (a RSN como camada informacional, o Estado é o único prestador), **Fase 2: Concorrência** (a RSN oferece alternativas funcionais, uso de duplo sistema) e **Fase 3: Transição** (a RSN supera os equivalentes estatais, migração voluntária). A transição é reversível em todas as fases.

O principal mecanismo de pressão é fiscal: quando os contribuintes condicionais atingem uma "minoria

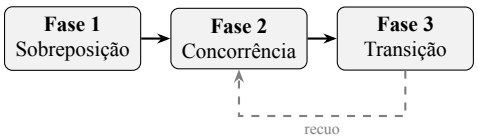

Figure 10: Migração em três fases—reversível em todas as fases.

economicamente significativa X"—um grupo suficientemente grande para que a repressão contra ele cause danos económicos não triviais e a desobediência civil se torne uma ameaça credível—o Estado entra em negociação. A título ilustrativo, 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 empresariais, adversários ao nível do Estado e adversários ao nível do protocolo.

Resistência a Sybil [12]. Construir reputação exige uma interação sustentada e verificável. O custo de um ataque Sybil bem-sucedido cresce linearmente com o número de identidades falsas e quadraticamente com o nível de reputação-alvo. Operar múltiplos DIDs em paralelo é possível mas dispendioso por conceção—cada identidade tem de acumular reputação de forma independente através de atividade genuína. Nas sociedades livres, este custo dissuade o abuso ocasional; nas ditaduras, os DIDs paralelos tornam-se uma ferramenta de sobrevivência que possibilita uma resistência compartimentada, a coordenação clandestina e uma navegação mais segura no mercado negro.

Defesa contra conluio. Os padrões de endosso mútuo entre grupos fechados são detetáveis através da análise do grafo social. As funções de agregação de reputação descontam as alegações provenientes de fontes densamente agrupadas.

Adversário ao nível do Estado. O encaminhamento onion, a ausência de infraestrutura centralizada e a distribuição do papel de Verificador por toda a base de participantes garantem que a supressão exige medidas desproporcionadas face a qualquer benefício.

Privacidade vs. responsabilização. A pseudonímia—um meio-termo entre o anonimato e a divulgação de identidade—permite que os DIDs acumulem reputação significativa sem revelar a identidade real do Titular.

11 Considerações de Privacidade

O modelo de privacidade da RSN assenta na pseudonímia. O Titular controla a divulgação de identidade: mínima (pseudónimo puro), seletiva (credenciais específicas associadas) ou total (identidade legal associada). As alegações publicadas são permanentes—o sistema suporta a correção contextual mas não o apagamento. Um Titular pode abandonar um DID comprometido e recomeçar, abdicando da reputação acumulada.

12 Trabalho Relacionado

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 distingue-se pela ênfase no custo económico como filtro de qualidade e pelo seu mecanismo de fixação do preço do radicalismo.

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

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

Anarco-agorismo. A RSN partilha um terreno comum significativo com a contraeconomia agorista [17]—em particular a ênfase na construção de instituições paralelas e na troca voluntária. Contudo, o agorismo tende a pressupor o autointeresse racional como mecanismo de coordenação suficiente e carece frequentemente de um caminho de migração concreto 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 propriedade sistémica, 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: quem entrega valor

acumula reputação; nenhum comité decide quem tem mérito.

Propriedade como privilégio. A RSN afasta-se fundamentalmente do tratamento anarcocapitalista e da escola austríaca dos direitos de propriedade como naturais e absolutos. No quadro da RSN, a propriedade é um *privilégio*—um reconhecimento social que pode, em casos extremos, ser retirado pela comunidade quando um participante viola fundamentalmente as normas partilhadas (traição, recusa de defender a comunidade numa crise existencial, exploração sistemática). Não se trata de expropriação estatal, mas da retirada do reconhecimento social pela rede de relações bilaterais.

13 Discussão e Limitações

Arranque a frio. O valor da RSN depende de efeitos de rede; os primeiros aderentes deparam-se com um valor limitado até que a participação atinja um limiar. Contudo, mesmo em fases iniciais, a rede DID serve de fonte de informação complementar útil. Espera-se que a avaliação inicial de reputação e a avaliação de autoridades se desenvolvam gradualmente, com sofisticação crescente.

Anti-parasitagem e controlo de acesso. Os DIDs-alvo podem impor emolumentos graduais e restrições de acesso às consultas de DIDs de baixa reputação. Não se trata de uma escala binária, mas contínua: quanto menos reputação um DID consulente tiver, menos informação recebe, mais tempo espera e mais paga. Este mecanismo incentiva a participação genuína em detrimento do consumo passivo.

Desigualdade no acesso. O custo de publicar alegações pode filtrar queixas legítimas de participantes economicamente desfavorecidos. Mitigação: cada participante serve simultaneamente de potencial verificador (ou delega esse papel) e aufere emolumentos de verificação. Ao longo de um horizonte temporal suficiente, um participante não radical deverá aproximar-se da neutralidade financeira—o rendimento de verificação a compensar aproximadamente os custos de publicação. Trata-se de uma expectativa estatística, não de uma garantia.

Desigualdade de reputação. Os primeiros aderentes acumulam vantagens que podem criar barreiras para os que chegam mais tarde. Contudo, a avaliação de reputação é realizada por prestadores terceiros que podem optar por atenuar a vantagem do primeiro entrante através dos seus algoritmos de agregação. Ser o primeiro com boa reputação é uma vantagem económica comparativa legítima—e isso é intencional.

Questões em aberto. As funções de custo ótimas, as normas de agregação, a governação do protocolo e a interação com dados de reputação sintética gerados por IA continuam a ser áreas para trabalho futuro.

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

14 Conclusão

Este artigo apresentou a Rede Social de Reputação, um protocolo descentralizado que possibilita a troca de reputação pseudónima e verificável. O Princípio do Radicalismo Dispendioso garante que as alegações fraudulentas ficam com um preço proibitivo sem censura. O caminho de migração proposto possibilita uma transição gradual e reversível a partir das instituições existentes. A conceção 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 uma utopia, mas um sistema em que a justiça é acessível, a reputação é conquistada e o custo da má conduta é suportado pela parte que a 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.