Zdecentralizowana sieć reputacji: Ramy naturalnej organizacji społecznej

Pavel Kudrna
Prague, Czechia
Draft v1 — March 2026

Abstract

Poczucie sprawiedliwości—przekonanie, że krzywdy zostają naprawione, a zasługi nagrodzone—jest najsilniejszą siłą spajającą społeczeństwo. Gdy scentralizowane instytucje władzy nie potrafią zapewnić tego poczucia, ponieważ koszt wyegzekwowania prawa przewyższa wartość samego prawa, spójność społeczna ulega erozji. Obecne struktury instytucjonalne nie potrafią rozstrzygnąć znaczącej klasy sporów w sposób systematycznie taki sam, w jaki centralne planowanie zawodzi przy alokacji zasobów: poprzez asymetrię informacji, brak pętli sprzężenia zwrotnego oraz oderwanie decydentów od konsekwencji ich decyzji. Niniejsza praca przedstawia Reputacyjną Sieć Społeczną (RSN): zdecentralizowaną, niepodatną na korupcję i odporną na cenzurę warstwę komunikacyjną zbudowaną na Zdecentralizowanych Identyfikatorach (DID), która umożliwia pseudonimowym uczestnikom publikowanie, weryfikowanie i konsumowanie informacji reputacyjnych o zachowaniach w świecie rzeczywistym. Podstawowy mechanizm opiera się na trzech zasadach projektowych: (1) asymetrycznej strukturze kosztów, w której odczyt danych reputacyjnych jest tani, ale zapis wymaga weryfikowalnego wysiłku; (2) niedeterministycznym protokole wyboru weryfikatora opartym na spójnym haszowaniu, który wycenia radykalne twierdzenia poprzez rosnące koszty iteracji; oraz (3) rynkowym modelu autorytetu reputacyjnego, w którym prywatni weryfikatorzy stawiają zgromadzoną reputację na dokładność twierdzeń, które popierają. Powstała architektura wytwarza emergentną merytokrację bez centralnej koordynacji i umożliwia stopniową, odwracalną ścieżkę migracji z istniejących systemów administrowanych przez państwo.

1 Wprowadzenie

1.1 Problem scentralizowanego zaufania

Współczesne rządzenie opiera się na fundamentalnym założeniu: że scentralizowane instytucje potrafią pośredniczyć w zaufaniu między obcymi na dużą skalę. Sądy rozstrzygają spory. Rejestry poświadczają własność. Organy regulacyjne egzekwują standardy. Instytucje te zaprojektowano w epoce, gdy informacja była rzadka, komunikacja powolna, a delegowanie władzy do centralnego organu było jedynym wykonalnym mechanizmem koordynacji. Scentralizowane rządzenie zawodzi jednak z tych samych strukturalnych powodów co centralne planowanie: asymetrii informacji, braku pętli sprzężenia zwrotnego oraz oderwania decydentów od konsekwencji ich decyzji.

Założenie to zawodzi na przykład wtedy, gdy koszt instytucjonalnego pośrednictwa przewyższa wartość sporu. Rozważmy najemcę, który odkrywa, że jego wynajmowane mieszkanie nie posiada prawnie wymaganego certyfikatu bezpieczeństwa instalacji elektrycznej—stanu stwarzającego bezpośrednie zagrożenie życia. Prawo najemcy do odstąpienia od umowy jest jednoznaczne. Jednak koszt wyegzekwowania tego prawa na drodze sądowej przewyższa wartość pieniężną kaucji i należnych odszkodowań, nawet wliczając potencjalne ustawowe zadośćuczynienie [1]. Racjonalną reakcją jest pogodzenie się ze stratą i wycofanie się.

Nie jest to patologiczny przypadek skrajny. To domyślne doświadczenie znaczącej klasy sporów, w których szkoda jest realna, ale stawki pieniężne spadają poniżej progu, przy którym instytucjonalne egzekwowanie staje się ekonomicznie racjonalne. Efektem jest system strukturalnie sprzyjający złym aktorom: ci, którzy krzywdzą innych w wolumenach poniżej tego progu, mogą działać bezkarnie, ponieważ koszt dochodzenia sprawiedliwości spada na stronę poszkodowaną.

Powyższy przykład zaczerpnięto ze środowiska

prawnego autora—czeskiego systemu prawa cywilnego. W innych tradycjach prawnych—prawie precedensowym common law, prawie zwyczajowym, systemach mieszanych—sprawiedliwość zawodzi w różny sposób i w różnym stopniu, w zależności od kulturowej i proceduralnej specyfiki danej jurysdykcji. Czytelnicy zapewne rozpoznają równoważne przykłady z własnego kontekstu. Strukturalnie uniwersalny jest wzorzec: spory, których wartość spada poniżej progu ekonomicznie racjonalnego egzekwowania, kończą się przejęciem straty przez stronę poszkodowaną. RSN nie obiecuje doskonałej sprawiedliwości—jedynie zmniejsza strukturalną przewagę, jaką cieszą się źli aktorzy, gdy instytucjonalne egzekwowanie jest nieopłacalne.

1.2 Dlaczego istniejące rozwiązania są niewystarczające

Ramy Zdecentralizowanej Tożsamości (DID) [2] zapewniają mechanizmy kryptograficzne do samodzielnego zarządzania tożsamością, lecz koncentrują się głównie na weryfikacji tożsamości, nie odnosząc się do szerszej kwestii reputacji behawioralnej.

Soulbound Tokens (SBT) [3] ujmują spostrzeżenie, że reputacja jest nieprzenoszalna, ale opierają się na infrastrukturze blockchain z ograniczeniami skalowalności i nie odnoszą się do bodźców ekonomicznych regulujących wprowadzanie informacji. Pokrewne koncepcje, takie jak tokeny zasług (np. Liberland), podzielają podobną filozofię, lecz pozostają związane z konkretnymi ramami jurysdykcyjnymi.

Zdecentralizowane Organizacje Autonomiczne (DAO) [4] realizują zarządzanie poprzez inteligentne kontrakty, ale odtwarzają plutokratyczną dynamikę, w której wpływ jest proporcjonalny do kapitału. Głosowanie kwadratowe [5] częściowo to łagodzi, lecz ogranicza praktyczną adopcję. DAO stają również w obliczu fundamentalnego *problemu wyroczni*: inteligentne kontrakty nie potrafią wiarygodnie weryfikować stanów świata rzeczywistego bez zaufanego zewnętrznego źródła, a problem ten nie ma ogólnego rozwiązania w ramach architektury blockchain.

Żaden istniejący system nie łączy decentralizacji, odporności na cenzurę, pseudonimowości, weryfikowalnego kosztu wejścia i emergentnego konsensusu.

1.3 Wkład

Niniejsza praca wnosi następujący wkład:

1. Definicję **Reputacyjnej Sieci Społecznej (RSN)**, zdecentralizowanego protokołu rozszerzającego ramy DID o wsparcie dwustronnej wymiany reputacji.
2. **Niedeterministyczny protokół wyboru weryfikatora** oparty na spójnym haszowaniu [6].
3. **Zasadę Kosztownego Radykalizmu**, wyceniającą twierdzenia według ich odległości od konsensusu sieci przez trzy nakładające się kanały kosztowe.
4. **Model Reputowanego Autorytetu** z asymetrycznymi bodźcami, gdzie fałszywe poparcie kosztuje katastrofalnie więcej niż uczciwe opłaty.
5. **Stopniową ścieżkę migracji** poprzez równoległą infrastrukturę fiskalną.

2 Zasady projektowe

Architektura RSN jest ograniczona pięcioma nienegocjowalnymi zasadami projektowymi.

Ciągłość i ewolucja. System współistnieje z istniejącymi instytucjami przez okres przejściowy o nieokreślonej długości. Przejście musi być odwracalne na każdym etapie.

Dobrowolność i odpowiedzialność. Uczestnictwo jest dobrowolne, ale dobrowolność jest sprzężona z odpowiedzialnością: uczestnik przejmujący kontrolę nad funkcją instytucjonalną jednocześnie przejmuje związane z nią obowiązki.

Różnorodność i neutralność kulturowa. Konsensus wyłania się z dwustronnych interakcji, a nie z gotowych reguł narzuconych przez projektantów systemu.

Odporność i odporność na cenzurę. Koszt cenzurowania sieci musi przewyższać koszt jej tolerowania. Tylko pełny totalitaryzm—za rujnującą cenę polityczną i ekonomiczną—może stłumić system.

Konsensus i egzekwowanie. System uwzględnia ludzkie skłonności do małostkowości, złośliwości i satysfakcji z odwetu jako cechy nośne. Występek nie jest zakazany; jest wyceniany.

3 Przegląd systemu

3.1 Reputacyjna Sieć Społeczna

RSN to zdecentralizowana warstwa komunikacyjna typu peer-to-peer, w której uczestnicy wymieniają weryfikowalne informacje o zachowaniach w świecie rzeczywistym. Jej definiujące właściwości to: zdecentralizowana i niecenzurowalna infras-

truktura; pseudonimowe uczestnictwo poprzez DID; asymetryczna struktura kosztów (odczyt jest tani, zapis kosztowny); weryfikowalność kryptograficzna; oraz **wsparcie standardów**—formaty czytelne maszynowo dla informacji propagowanych przez sieć, dla usług oferowanych przez autorytety (kompozycja twierdzeń, poparcie, usługa świadka) oraz dla formatu polityki obejmującego reguły oceny reputacji kontrahenta i szacowania ryzyka niezgodności polityki (między deklarowanym a rzeczywistym zachowaniem).

3.2 Komponenty architektoniczne

RSN składa się z czterech głównych komponentów (Rys. 1):

1. **Warstwa DID.** Tożsamości uczestników z deklarowanymi regułami, metadanymi reputacji i routingiem sieciowym.
2. **Protokół Twierdzeń.** Ustrukturyzowany format publikowania stwierdzeń o zdarzeniach w świecie rzeczywistym.
3. **Sieć Weryfikacyjna.** Zdecentralizowany protokół przydzielania weryfikatorów przy użyciu spójnego haszowania.
4. **Warstwa Agregacji Reputacji.** Usługi wytwarzające czytelne dla człowieka podsumowania reputacji.

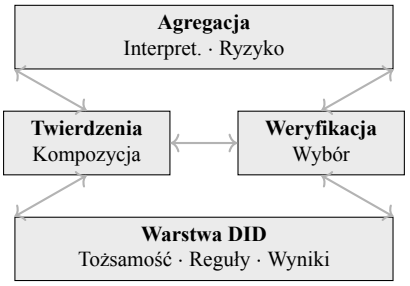

Figure 1: Czterowarstwowa architektura RSN.

3.3 Role uczestników

Transakcja weryfikacyjna angażuje do sześciu odrębnych ról DID (Rys. 2): **Wystawca** (DID_I, tworzy i przesyła twierdzenie), **Podmiot** (DID_S, DID, którego dotyczy twierdzenie—może pokrywać się z Wystawcą), **Autorytet** (DID_A, popiera twierdzenie za opłatą; może również oferować usługi świadka—*opcjonalnie*), **Świadek** ($DID_{W_{1..n}}$, incognito monitorujący uczciwość weryfikatora za pomocą kodów wyzwań—*opcjonalnie*), **Weryfikator** (DID_V, algorytmicznie wybrany do publikacji twierdzenia) oraz **RSN** (sieć, w której publikowane są zweryfikowane twierdzenia). Każda rola może zostać delegowana innemu DID (Rozdział 7.4). Autorytet zwykle łączy poparcie z usługami świadka, ale obie funkcje są logicznie niezależne.

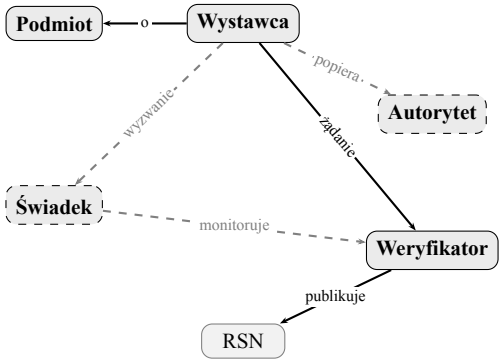

Figure 2: Role DID w transakcji weryfikacyjnej. Przerywana = opcjonalna (Autorytet, Świadek).

3.4 Niepodatność na korupcję: cztery siły

Niepodatność RSN na korupcję nie wynika z pojedynczego mechanizmu, lecz z czterech równoczesnych sił. Żadna nie wystarcza samodzielnie; dopiero ich połączenie czyni próbę korupcji ekonomicznie irracjonalną.

Nieprzewidywalny cel. Weryfikator dla każdego twierdzenia jest wybierany niedeterministycznie poprzez spójne haszowanie (Rozdział 5.3). Atakujący nie może z góry namierzyć konkretnego weryfikatora do przekupienia, ponieważ tożsamość weryfikatora jest funkcją treści twierdzenia i poprzednich iteracji, a nie wyboru wystawcy.

Dźwignia reputacyjna—w górę powoli, w dół szybko. Reputację można zdobywać jedynie stopniowo, poprzez trwałe konsekwentne zachowanie. Można ją jednak katastrofalnie utracić w jednym oszukańczym poparciu (Rozdział 6.2). Ta asymetria oznacza, że lata zgromadzonej reputacji mogą zostać zniszczone przez jedno nieuczciwe poparcie; optymalną strategią pozostaje odrzucanie podejrzanych twierdzeń.

Publiczna obietnica. Każdy Posiadacz DID publicznie deklaruje swoją politykę—czytelny maszynowo zestaw reguł, których zobowiązuje się przestrzegać. Rozbieżność między deklaracją a rzeczywistym zachowaniem jest wykrywalna i wyceniana jako przewinienie reputacyjne poważniejsze niż samo naruszenie (Rozdział 7.3). Hipokryzja jest zatem kosztowniejsza niż bezpośrednia nieuczciwość.

Asymetria kosztu ataku na siatkę. Koszt cenzurowania lub destabilizowania sieci rośnie nieliniowo wraz z rozmiarem i gęstością sieci, podczas gdy koszt jej tolerowania pozostaje stały. Aby cen-

zura się opłaciła, scentralizowany aktor musiałby zastosować ją w całej sieci—co wymaga środków nieproporcjonalnych do jakiejkolwiek wyobrażalnej korzyści (Rozdział 10).

4 Model zdecentralizowanej tożsamości

4.1 DID o szczególnych właściwościach

RSN rozszerza specyfikację W3C DID Core [2] o: **Deklarowane Reguły** (czytelne maszynowo zobowiązania behawioralne), **Metadane Reputacji** (zagregowany wynik behawioralny) oraz **Routing Sieciowy** (zmienna brama onion odrębna od stabilnej pozycji na pierścieniu).

4.2 Cykl życia twierdzenia

Twierdzenie przechodzi przez sześć etapów (Rys. 3): kompozycja, opcjonalne poparcie autorytetu, wybór weryfikatora poprzez spójne haszowanie, decyzja weryfikacyjna (akceptacja lub odrzucenie z iteracją), publikacja oraz aktualizacja reputacji wszystkich stron.

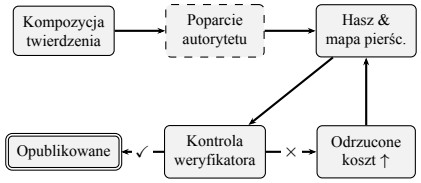

Figure 3: Cykl życia twierdzenia z pętlą iteracji.

4.3 Relacja do W3C DID Core

Model tożsamości RSN jest właściwym nadzbiorem specyfikacji W3C DID Core [2]. Wszystkie DID w RSN są prawidłowymi DID W3C. Rozszerzenia specyficzne dla RSN są zakodowane jako właściwości dokumentu DID zdefiniowane w dedykowanej specyfikacji metody DID, kompatybilnej z istniejącą infrastrukturą, taką jak ION [7] i Ceramic [8].

5 Weryfikacja i propagacja informacji

5.1 Koszt wprowadzenia informacji

Podstawową zasadą ekonomiczną RSN jest asymetria między odczytem a zapisem. Odczyt danych reputacyjnych zbliża się do zerowego kosztu krańcowego dla uczestników należących do sieci DID i posiadających dostęp współmierny do ich reputacji—nowicjusze muszą stopniowo wypracowywać szerszy dostęp (zob. zapobieganie pasożytnictwu). Zapis—rozumiany jako trwałe zmaterializowanie reputacji w sieci, a nie jako jednorazowy akt techniczny—wymaga weryfikowalnej skumulowanej inwestycji w trzech wymiarach: **czasu** (lat życia poświęconych konsekwentnemu zachowaniu, dotrzymanym zobowiązaniom i pielęgnowanym relacjom), **energii** (wysiłku włożonego w uczciwe postępowanie, dbałość o partnerstwa i długoterminową wiarygodność) oraz **pieniędzy** (inwestycji we własne imię—rozwoju zawodowego, jakości własnej pracy, restytucji za wyrządzone krzywdy). Reputacji nie da się kupić natychmiast—jest ona wynikiem gromadzenia trwającego całe życie, które system jedynie uwidacznia i czyni przenaszalnym między kontekstami. Jednorazowe koszty techniczne protokołu (podpisy kryptograficzne, opłaty weryfikatorów) są znikome w porównaniu z tą podstawową inwestycją.

5.2 Graf społeczny i głębokość kontaktów

RSN jest fundamentalnie siecią społeczną: każdy DID dodaje kontakty (inne DID), które zezwalają na połączenie. Kontakty te mają własne kontakty i tak dalej. Algorytmiczne wyszukiwanie weryfikatora działa w konfigurowalnej głębokości h powiązanych kontaktów (np. $h = 3$: własne bezpośrednie kontakty, ich kontakty oraz kontakty kontaktów). Ta architektura nie wymaga jednego globalnego blockchaina—naturalnie sprzyja wyłanianiu się społeczności/klastrów z nakładaniami na inne społeczności. Każdy uczestnik DID musi mieć opublikowaną **politykę**—czytelny maszynowo zestaw reguł regulujących sposób oceny przychodzących żądań. Naruszenia polityki są zapisywane w sieci jako negatywne artefakty.

5.3 Algorytmiczny wybór weryfikatora

Protokół weryfikacji wykorzystuje spójne haszowanie [6] (Rys. 4). Weryfikacja jest usługą płatną: weryfikatorzy działają za opłatą, tworząc rynek, w którym konkurencja napędza cenę, szybkość i niezawodność.

Krok 1. Dokument twierdzenia jest łączony z wynikiem haszu poprzedniej iteracji (lub $\varnothing$ przy pierwszej iteracji) i haszowany:

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

Algorytm musi obejmować *pełną ścieżkę* wszystkich poprzednich żądań i odpowiedzi—nie tylko hasz poprzedniej iteracji. Pełna odpowiedź każdego weryfikatora (akceptacja, odrzucenie, podana cena, powód) jest dołączana do rosnącego dokumentu, weryfikowalna poprzez podpisy kryptograficzne na każdym kroku.

Krok 2. Hasz jest mapowany na N pozycji na pierścieniu spójnego haszu, równomiernie rozłożonych (np. dla $N = 4$: $+0°, +90°, +180°, +270°$). Z każdej pozycji wyszukiwanie zgodne z ruchem wskazówek zegara wybiera najbliższy DID jako kandydata na weryfikatora (Rys. 5). N jest parametrem zmiennym, który może zależeć od odsetka aktualnie aktywnych DID, pory dnia lub historycznej responsywności sieci.

Krok 3. Weryfikator akceptuje (publikuje, stawiając reputację) lub odrzuca (wracając do Kroku 1 z haszem odrzucenia). Gdy węzeł nie odpowiada mimo deklarowania statusu online, kolejne żądanie musi zawierać wszystkie odpowiedzi wszystkich N poprzednich kandydatów na weryfikatorów.

Zapobieganie pasożytnictwu. Docelowe DID mogą ustalać opłaty lub ograniczenia dostępu dla zapytań od nowych DID o niewielkiej reputacji. Ten stopniowalny mechanizm (nie binarny) zmusza nowicjuszy do budowania reputacji poprzez rzeczywistą aktywność, zanim zaczną swobodnie konsumować dane innych.

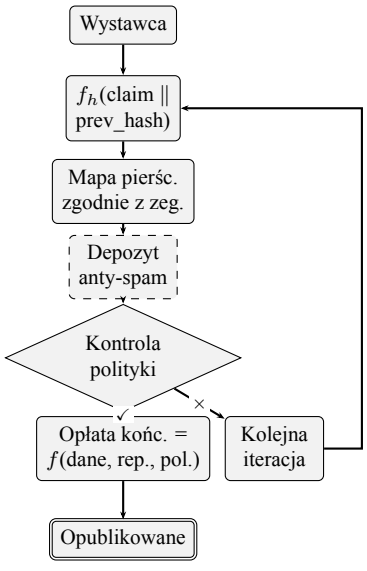

Figure 4: Protokół wyboru weryfikatora. Depozyt anty-spam jest opcjonalny i może być zwrotny. Opłata końcowa jest płatna tylko przy akceptacji.

Każde odrzucenie dołącza pełną odpowiedź weryfikatora (wraz z powodami i warunkami) do dokumentu twierdzenia, rozszerzając jego treść. Z rozszerzonego dokumentu niedeterministycznie obliczany jest nowy hasz, mapowany na inną pozycję na pierścieniu i wybierający innego weryfikatora. Dokumentacja pełnej ścieżki jest obowiązkowa—wystawca nie może przewidzieć ani wpłynąć na kolejnego weryfikatora. Jeśli weryfikator zignoruje żądanie mimo zgodnej deklarowanej polityki, świadkowie (jeśli obecni) to rejestrują, a wystawca może dołączyć dowody świadków przy końcowej publikacji.

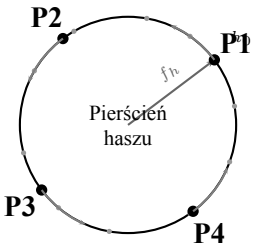

Figure 5: Wybór wielopunktowy (N=4). Hasz h_0 ustala przesunięcie; 4 punkty równomiernie rozłożone od h_0.

5.4 Protokół świadka

Rola **Świadka** zapewnia incognito rozliczalność uczciwości weryfikatora poprzez kryptograficzny protokół kodów wyzwań:

Krok W1. Żądający podpisuje żądanie weryfikacji i wysyła je do N_w świadków—kontaktów z sieci społecznej żądającego lub wyspecjalizowanych dostawców usług świadka działających za opłatą.

Krok W2. Każdy świadek w_i podpisuje całe podpisane żądanie, zachowuje oryginalny podpis σ_i i oblicza kod wyzwania $c_i = h(\sigma_i)$. Kod wyzwania jest dołączany do żądania.

Krok W3. Żądanie, niosące teraz N_w kodów wyzwań, jest wysyłane do weryfikatora. Weryfikator obserwuje kody, ale *nie może ustalić*, kim są świadkowie ani czy kody odpowiadają prawdziwym podpisom.

Krok W4 (uczciwy weryfikator). Jeśli weryfikator odpowiada zgodnie ze swoją deklarowaną polityką, kody wyzwań pozostają nieprzejrzyste. Świadkowie nigdy nie zostają ujawnieni. Nie są publikowane żadne dodatkowe dane.

Krok W5 (naruszenie polityki). Jeśli weryfikator narusza swoją politykę (ignoruje żądanie, odpowiada niekonsekwentnie), żądający posiada oryginalne podpisy świadków σ_i. Można je opublikować jako dane towarzyszące: każdy może zweryfikować $c_i = h(\sigma_i)$, dowodząc, że świadek był obecny. Żądający może op-

ublikować niezależnie—weryfikator już podpisał kody wyzwań w swojej odpowiedzi.

Właściwość blefu. Żądający może podstawić losowe wartości za niektóre lub wszystkie kody wyzwań. Weryfikator nie potrafi odróżnić prawdziwych kodów (zabezpieczonych przez autorytety o wysokiej reputacji) od szumu. Ta *niepewność* jest mechanizmem egzekwowania: każde żądanie niesie ryzyko, że szanowany autorytet obserwuje incognito. Koszt utrzymania tej presji jest bliski zeru (generowanie liczb losowych), podczas gdy potencjalny koszt nieuczciwości dla weryfikatora jest katastrofalny. Uczciwe zachowanie jest motywowane nawet przy braku rzeczywistych świadków.

Autorytet jako świadek. Autorytet popierający twierdzenie może połączyć usługi świadka: „Popieram twoje twierdzenie i pełnię rolę incognito świadka podczas weryfikacji". Jest to naturalny model biznesowy—autorytet już posiada kontekst. Obie role są jednak logicznie niezależne.

5.5 Zasada Kosztownego Radykalizmu

Trzy kanały kosztowe nakładają się jednocześnie (Rys. 6):

Kanał 1: Koszt iteracji. Każde odrzucenie wymusza nową iterację pochłaniającą czas i zasoby. Wzrost liniowy.

Kanał 2: Rozrost dokumentu. Każda iteracja dodaje pełną odpowiedź weryfikatora do dokumentu twierdzenia. Wzrost jest liniowy, ale z przesunięciem bazowym: początkowy ładunek (treść twierdzenia, poparcie autorytetu, podpisy kryptograficzne) ma już nietrywialny rozmiar B_0. Całkowity rozmiar po k iteracjach: $S(k) = B_0 + k \cdot \Delta$, gdzie Δ to średni rozmiar odpowiedzi.

Kanał 3: Premia za ryzyko weryfikatora. Ryzyko jest subiektywne. Weryfikator ocenia, czy deklarowane polityki żądającego DID kolidują z jego własnymi interesami handlowymi, społecznymi lub politycznymi. Premia może rosnąć wykładniczo wraz z postrzeganą odległością od „ideału" weryfikatora: $P(d) = \alpha \cdot e^{\beta d}$, gdzie d to subiektywna metryka odległości weryfikatora. Poza osobistą granicą no-go weryfikator może odmówić całkowicie.

Co kluczowe, nie jest to cenzura. Żadne twierdzenie nie jest zakazane. System wycenia ryzyko (Tabela 1).

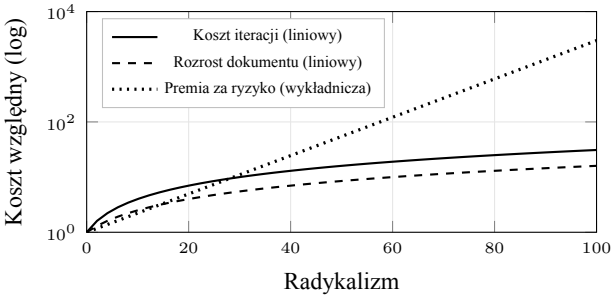

Figure 6: Eskalacja kosztów w trzech kanałach. Premia za ryzyko dominuje przy wysokim radykalizmie.

Table 1: Porównanie kosztów według typów twierdzeń.

Typ	Iter.	Dok	Opłata	Razem
Wiarygodne	k_{min}	B_0	P_{min}	Niski
Kontrowersyjne	k_{med}	$B_0 + k\Delta$	$\alpha e^{\beta d}$	Umiarkowany
Radykalne	k_{max}	$\gg B_0$	$\gg P_{\text{min}}$	Bardzo wysoki
Oszukańcze	$\to \infty$	—	—	Niepublikowalne

6 Reputacja i ocena ryzyka

6.1 Model akumulacji reputacji

Reputacja wykazuje trzy właściwości: **powolną akumulację** (stopniowy wzrost poprzez konsekwentne zachowanie), **szybkie zniszczenie** (pojedyncze oszustwo może katastrofalnie obniżyć wynik) oraz **indywidualną ocenę** (każda konsumująca strona może stosować własną funkcję agregacji).

6.2 Reputowane autorytety

Reputowany Autorytet analizuje twierdzenia i, jeśli jest usatysfakcjonowany, współpodpisuje je—stawiając własną reputację (Rys. 7). Asymetria jest zamierzona: zdobywanie reputacji jest powolne (przyrostowe $+\delta$ za uczciwe poparcie), podczas gdy jej utrata jest katastrofalna ($-\Delta \gg \delta$ za fałszywe poparcie; poglądowo np. $+2\%$ wobec -70%). Optymalną strategią jest zawsze odrzucanie podejrzanych twierdzeń. Konkretne wartości δ i Δ są parametrami systemu.

6.3 Reputacja wielowymiarowa

Reputacja nie jest skalarem, lecz wektorem obejmującym wiele wymiarów: wiarygodność finansową, integralność osobistą, kompetencję zawodową, zaangażowanie obywatelskie i inne (Rys. 8). Różni dostawcy oceny rzutują ten wektor odmiennie w zależności od kontekstu—pożyczkodawca prioryte-

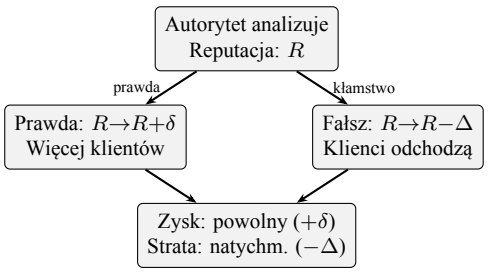

Figure 7: Reputowany Autorytet—asymetryczne bodźce.

tyzuje wiarygodność finansową, pracodawca kompetencję zawodową, sąsiad zachowanie obywatelskie. Nie istnieje jeden „poprawny" wynik reputacji; istnieją jedynie perspektywy.

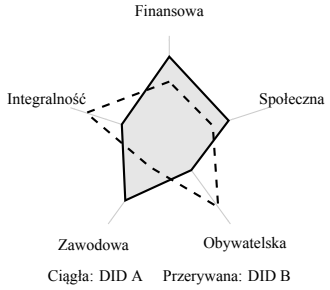

Figure 8: Wielowymiarowe wektory reputacji. Różne DID wykazują różne profile w poszczególnych wymiarach.

6.4 Zdemokratyzowana ocena ryzyka

RSN demokratyzuje systematyczną ocenę ryzyka—zdolność obecnie skoncentrowaną w instytucjach finansowych, lecz mającą zastosowanie daleko poza bankowością. Konglomeraty przemysłowe oceniające wiarygodność dostawców, firmy ubezpieczeniowe wyceniające ryzyko behawioralne, due diligence przy fuzjach i przejęciach, szacowanie żywotności sprzętu w produkcji—wszędzie tam, gdzie ryzyko kontrahenta wymaga oceny, RSN zapewnia zdecentralizowaną, rynkową alternatywę. Rynek usług interpretacyjnych przekłada surowe wielowymiarowe dane reputacyjne na użyteczne podsumowania, czyniąc ocenę kontrahenta dostępną dla każdego uczestnika po koszcie krańcowym.

7 Konsensus i rozstrzyganie sporów

7.1 Model zdecentralizowanej sprawiedliwości

RSN przekształca sprawiedliwość w problem dwustronnej negocjacji. Groźba trwałej, weryfikowalnej szkody reputacyjnej jest skuteczniejszym środkiem odstraszającym niż postępowanie sądowe, na które strony poszkodowanej nie stać.

7.2 Emergentny konsensus

Każdy uczestnik utrzymuje osobisty próg satysfakcji. Zagregowany konsensus sieci wyłania się jako statystyczna średnia tysięcy dwustronnych negocjacji (Rys. 9). Reakcja znacznie powyżej konsensusu (barbarzyńska) jest kosztowna do opublikowania. Reakcja bliska konsensusowi (proporcjonalna) jest tania. Konsensus nie jest regułą—jest sygnałem cenowym.

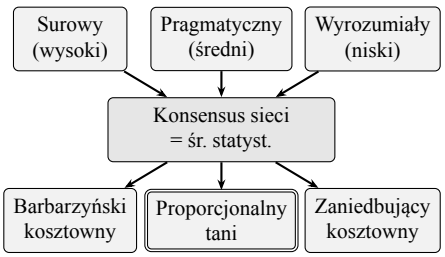

Figure 9: Emergentny konsensus z indywidualnych progów satysfakcji.

7.3 Kalibracja kary

Każdy Posiadacz DID publicznie deklaruje reakcje na konkretne kategorie występków. Nieproporcjonalnie surowe lub pobłażliwe deklaracje przyciągają negatywne sygnały reputacyjne. Proporcjonalne deklaracje są akceptowane bez tarcia—samokalibrujący się system kar.

Deklaracje konsensusowe same podlegają wykrywaniu hipokryzji: deklaratywne zobowiązanie do stanowiska konsensusowego przy niekonsekwentnym postępowaniu stanowi przewinienie reputacyjne poważniejsze niż samo odstępstwo, gdyż dowodzi celowego oszustwa.

7.4 Delegowanie

Wszystkie czynności w ramach DID—z jednym wyjątkiem—mogą być delegowane innemu DID. **Rola Podmiotu—DID, którego zachowania dotyczy twierdzenie—nie może być delegowana;**

jest nieprzenaszalnie związana z posiadaczem tożsamości, do którego odnosi się twierdzenie. Pozostałe aktywne role—Wystawca, Autorytet, Świadek, Weryfikator—mogą być przeniesione na wyznaczony DID delegata, czy to na potrzeby weryfikacji, przesłania twierdzenia, udziału w konsensusie, czy rozstrzygania sporów. Delegowanie jest odwoływalne w każdej chwili. Umożliwia to specjalizację (np. profesjonalne usługi weryfikacyjne), podczas gdy Posiadacz zachowuje ostateczną kontrolę i odpowiedzialność. Delegowanie stosuje się w całym systemie i jest niezbędne dla skalowalności.

7.5 Od moralności indywidualnej do emergentnej umowy społecznej

Deklarowana polityka każdego DID stanowi formalizację indywidualnej moralności: co Posiadacz uważa za akceptowalne, jak reaguje na konkretne zachowania, czego wymaga od kontrahentów. Polityki te nie są narzucane przez projektanta systemu—są wybierane przez każdego uczestnika i wyrażane w formie czytelnej maszynowo.

Ta formalizacja umożliwia trzyetapową emergencję. Po pierwsze, indywidualna moralność jest kodowana jako **polityka**—zestaw deklarowanych reguł, progów i funkcji reakcji, których DID zobowiązuje się przestrzegać. Po drugie, agregat wszystkich polityk w sieci wytwarza **emergentną etykę**—statystycznie obserwowalne normy, których żaden pojedynczy uczestnik nie zaprojektował, lecz które powstają z interakcji tysięcy indywidualnych stanowisk moralnych [9]. Po trzecie, ta emergentna etyka, wyrażona jako wspólne normy behawioralne egzekwowane poprzez dwustronne sygnały cenowe, stanowi **mierzalną umowę społeczną**—nie teoretyczny konstrukt, którego nikt nie podpisał [10], lecz empirycznie obserwowalny zestaw reguł popartych rzeczywistym zachowaniem.

Proces ten odzwierciedla wnioski z teorii fundamentów moralnych [11]: ludzka moralność jest intuicyjna, pluralistyczna i zmienna kulturowo. RSN nie wymaga konsensusu moralnego jako danych wejściowych; wytwarza konsensus behawioralny jako wynik. Powstała umowa społeczna nie jest statyczna—ewoluuje, gdy uczestnicy dołączają, odchodzą i dostosowują swoje polityki w odpowiedzi na sprzężenie zwrotne sieci. W przeciwieństwie do klasycznej teorii umowy społecznej, umowa ta jest odwoływalna, obserwowalna i egzekwowalna bez suwerena.

8 Model ekonomiczny

Poprzednie rozdziały opisały narzędzie—mechanizmy weryfikacji RSN, strukturę kosztów i emergentny konsensus. Ten rozdział opisuje metodologię stosowania tego narzędzia do transformacji fiskalnej i systemu rządzenia. Narzędzie jest ogólnego przeznaczenia; poniższa metodologia to jedno z możliwych zastosowań.

8.1 Struktura bodźców

Reputacja jako kapitał tworzy bezpośrednie bodźce do uczciwego zachowania. W normalnym funkcjonowaniu uczestnicy budują reputację naturalnie, poprzez codzienne transakcje i dotrzymane zobowiązania. Głównym wydatkiem są twierdzenia *negatywne*—odnotowywanie krzywd wyrządzonych przez innych. Ta asymetria motywuje do pożytecznego, niezawodnego zachowania: bycie uczciwym nie kosztuje nic dodatkowo, podczas gdy bycie szkodliwym tworzy udokumentowany ślad, który inni zapłacą za zachowanie. Hipokryzja—deklarowanie reguł bez ich przestrzegania—jest najkosztowniejszym trybem awarii, gdyż dowodzi celowego oszustwa.

8.2 Równoległy system fiskalny

Trzy komponenty umożliwiają presję konkurencyjną: (1) **Prosty Podatek**—jedna proporcjonalna stawka bez wyjątków, początkowo nieznacznie poniżej istniejącej stawki efektywnej; (2) **Elektroniczna Rejestracja Wydatków (ESR)**—odwrócenie czeskiego modelu EET, tak by obywatele nadzorowali wydatki państwa; (3) **Alokacja Podatków Kierowana przez Obywateli**—zaczynająca od kilku procent w pierwszym roku i sięgająca po dekadzie od kilkunastu procent aż po pełną migrację do systemu kierowanego przez obywateli, w zależności od tempa adopcji. Rzeczywiste tempo zależy od szybkości, z jaką powstają wolnorynkowe alternatywy dla usług państwa, oraz od gotowości państwa do współpracy przy transformacji; funkcja czasu nie musi być liniowa i nie ma powodu ustalać górnej granicy tego, dokąd adopcja może ostatecznie dotrzeć.

9 Ścieżka migracji

Migracja przebiega przez trzy fazy (Rys. 10): **Faza 1: Nakładka** (RSN jako warstwa informacyjna, państwo jedynym dostawcą), **Faza 2: Konkurencja** (RSN dostarcza funkcjonalne alternatywy, użycie sys-

temu podwójnego) oraz **Faza 3: Przejście** (RSN przewyższa państwowe odpowiedniki, dobrowolna migracja). Przejście jest odwracalne na każdym etapie.

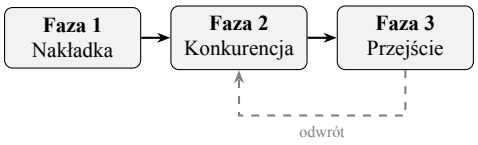

Figure 10: Trzyfazowa migracja—odwracalna na każdym etapie.

Głównym mechanizmem nacisku jest fiskalny: gdy warunkowi podatnicy osiągną „ekonomicznie znaczącą mniejszość X"—grupę na tyle dużą, że represja wobec niej powoduje nietrywialne szkody gospodarcze, a obywatelskie nieposłuszeństwo staje się wiarygodną groźbą—państwo przystępuje do negocjacji. Dla ilustracji przyjmujemy $X \approx 15\%$ PKB, ale rzeczywisty próg zależy od konkretnej gospodarki.

10 Analiza bezpieczeństwa

Rozważamy pięć kategorii przeciwników: indywidualnych złych aktorów, zorganizowane grupy, przeciwników korporacyjnych, przeciwników na poziomie państwa oraz przeciwników na poziomie protokołu.

Odporność na atak Sybil [12]. Budowanie reputacji wymaga trwałej, weryfikowalnej interakcji. Koszt udanego ataku Sybil rośnie liniowo wraz z liczbą fałszywych tożsamości i kwadratowo wraz z docelowym poziomem reputacji. Obsługiwanie wielu DID równolegle jest możliwe, lecz z założenia kosztowne—każda tożsamość musi niezależnie gromadzić reputację poprzez rzeczywistą aktywność. W wolnych społeczeństwach koszt ten zniechęca do przypadkowych nadużyć; w dyktaturach równoległe DID stają się narzędziem przetrwania, umożliwiającym podzielony na przedziały opór, podziemną koordynację i bezpieczniejszą nawigację po czarnym rynku.

Obrona przed zmową. Wzorce wzajemnego popierania w zamkniętych grupach są wykrywalne poprzez analizę grafu społecznego. Funkcje agregacji reputacji dyskontują twierdzenia z ciasno zgrupowanych źródeł.

Przeciwnik na poziomie państwa. Trasowanie onion, brak scentralizowanej infrastruktury i rozproszenie roli Weryfikatora wśród bazy uczestników zapewniają, że stłumienie wymaga środków nieproporcjonalnych do jakiejkolwiek korzyści.

Prywatność wobec rozliczalności. Pseudonimowość—środek pośredni między anonimowością a ujawnieniem tożsamości—pozwala DID gromadzić znaczącą reputację bez ujawniania rzeczywistej tożsamości Posiadacza.

11 Rozważania dotyczące prywatności

Model prywatności RSN opiera się na pseudonimowości. Posiadacz kontroluje ujawnianie tożsamości: minimalne (czysty pseudonim), selektywne (powiązanie konkretnych poświadczeń) lub pełne (przypisanie tożsamości prawnej). Opublikowane twierdzenia są trwałe—system wspiera korektę kontekstową, ale nie usunięcie. Posiadacz może porzucić skompromitowany DID i zacząć od nowa, tracąc zgromadzoną reputację.

12 Prace pokrewne

Specyfikacja W3C DID Core [2] i sieć Ceramic [8] zapewniają fundament tożsamości. EigenTrust [13] zademonstrował rozproszoną agregację reputacji bez centralnego autorytetu. SBT [3] wprowadziły nieprzenoszalne tokeny społeczne. RSN różni się naciskiem na koszt ekonomiczny jako filtr jakości oraz mechanizmem wyceny radykalizmu.

DAO [4] i głosowanie kwadratowe [5] wykazały wykonalność zdecentralizowanego zarządzania. RSN wyprowadza konsensus z dwustronnych negocjacji cenowych, a nie z głosowania.

Propozycja czerpie z teorii anarchokapitalistycznej [14, 15] w zakresie prywatnego świadczenia usług, lecz odchodzi od niej w dwóch kluczowych aspektach: projektuje z myślą o złośliwości i impulsach odwetowych, zamiast zakładać racjonalność, oraz proponuje stopniową transformację zamiast rewolucji. Koncepcja emergentnych umów społecznych ma poprzedników w Hayekowskiej teorii spontanicznego porządku [16].

Anarcho-agoryzm. RSN dzieli znaczną wspólną płaszczyznę z agorystyczną kontr-ekonomią [17]—szczególnie nacisk na budowanie równoległych instytucji i dobrowolną wymianę. Agoryzm skłania się jednak do zakładania racjonalnego interesu własnego jako wystarczającego mechanizmu koordynacji i często brakuje mu konkretnej ścieżki migracji z istniejących struktur państwowych. RSN wyraźnie uwzględnia nieracjonalne skłonności behawioralne i zapewnia fiskalny mechanizm nacisku dla stopniowej

transformacji.

Merytokracja. RSN może stanowić pierwszy funkcjonalny opis tego, jak merytokracja [18] może wyłonić się jako właściwość systemowa, a nie slogan. Tradycyjne systemy merytokratyczne zawodzą, ponieważ wymagają centralnego autorytetu do zdefiniowania i egzekwowania „zasługi". W RSN zasługa wyłania się z dwustronnych ocen: ci, którzy dostarczają wartość, gromadzą reputację; żaden komitet nie decyduje, kto ma zasługę.

Własność jako przywilej. RSN fundamentalnie odchodzi od anarchokapitalistycznych i austriackich ujęć praw własności jako naturalnych i absolutnych. W ramach RSN własność jest *przywilejem*—społecznym uznaniem, które w skrajnych przypadkach może zostać wycofane przez wspólnotę, gdy uczestnik fundamentalnie narusza wspólne normy (zdrada, odmowa obrony wspólnoty w kryzysie egzystencjalnym, systematyczne wyzyskiwanie). Nie jest to państwowe wywłaszczenie, lecz wycofanie społecznego uznania przez sieć relacji dwustronnych.

13 Dyskusja i ograniczenia

Zimny start. Wartość RSN zależy od efektów sieciowych; wcześni użytkownicy napotykają ograniczoną wartość, dopóki uczestnictwo nie osiągnie progu. Jednak nawet na wczesnych etapach sieć DID służy jako użyteczne uzupełniające źródło informacji. Oczekuje się, że wczesna ocena reputacji i ocena autorytetów będą rozwijać się stopniowo, wraz z rosnącym wyrafinowaniem.

Zapobieganie pasożytnictwu i kontrola dostępu. Docelowe DID mogą nakładać stopniowalne opłaty i ograniczenia dostępu na zapytania od DID o niskiej reputacji. Nie jest to binarne, lecz ciągła skala: im mniej reputacji ma odpytujący DID, tym mniej informacji otrzymuje, tym dłużej czeka i tym więcej płaci. Mechanizm ten motywuje do rzeczywistego uczestnictwa zamiast biernej konsumpcji.

Nierówność dostępu. Koszt publikowania twierdzeń może odfiltrowywać uzasadnione skargi od uczestników znajdujących się w gorszej sytuacji ekonomicznej. Łagodzenie: każdy uczestnik jednocześnie pełni rolę potencjalnego weryfikatora (lub deleguje tę rolę) i zarabia opłaty weryfikacyjne. W wystarczająco długim horyzoncie czasowym nieradykalny uczestnik powinien zbliżyć się do neutralności finansowej—dochód z weryfikacji w przybliżeniu rekompensujący koszty publikacji. Jest to oczekiwanie statystyczne, a nie gwarancja.

Nierówność reputacji. Wcześni użytkownicy gromadzą przewagi, które mogą tworzyć bariery dla późniejszych. Jednak ocena reputacji jest wykonywana przez zewnętrznych dostawców, którzy mogą zdecydować się łagodzić przewagę pierwszeństwa poprzez swoje algorytmy agregacji. Bycie pierwszym z dobrą reputacją jest uprawnioną komparatywną przewagą ekonomiczną—i to z założenia.

Pytania otwarte. Optymalne funkcje kosztów, standardy agregacji, zarządzanie protokołem oraz interakcja z syntetycznymi danymi reputacyjnymi generowanymi przez AI pozostają obszarami przyszłych prac.

Niniejsza praca nie twierdzi, że RSN będzie wytwarzać optymalne rezultaty we wszystkich okolicznościach. Twierdzi, że RSN stanowi *wykonalną* alternatywę—technicznie implementowalną, ekonomicznie samopodtrzymującą się i społecznie zgodną ze skłonnościami behawioralnymi człowieka takimi, jakie są w rzeczywistości.

14 Wnioski

Niniejsza praca przedstawiła Reputacyjną Sieć Społeczną, zdecentralizowany protokół umożliwiający pseudonimową, weryfikowalną wymianę reputacji. Zasada Kosztownego Radykalizmu zapewnia, że oszukańcze twierdzenia są wyceniane poza zasięg bez cenzury. Zaproponowana ścieżka migracji umożliwia stopniowe, odwracalne przejście z istniejących instytucji. Projekt uwzględnia ludzką naturę taką, jaka jest—w tym małostkowość, złośliwość i pragnienie satysfakcji z odwetu—zamiast wymagać transformacji ideologicznej. Rezultatem nie jest utopia, lecz system, w którym sprawiedliwość jest dostępna, reputacja jest zasłużona, a koszt występku ponosi strona, która go spowodowała.

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, "De-

centralized 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.