

PERSONIX

声誉网络的镜子，连被粉饰的丑陋也照得分明。

一个不可审查、不可腐蚀的去中心化声誉网络
国家的演进式继承者

Personix — 不妥协的变革

一个不可审查、不可腐蚀的去中心化声誉网络

© 2025–2026 Pavel Kudrna, 布拉格

版本 6 (2026-05-21)

初稿写于 2025 年 6 月至 11 月;2026 年 4 月至 5 月重写。

以 $\text{\LaTeX}$ (Tectonic) 使用 Palatino 与 Helvetica Neue 排版。

A5 横向 (210 × 148 mm)。

依据 **知识共享署名-相同方式共享 4.0 国际许可协议 (CC BY-SA 4.0)** 授权 —

creativecommons.org/licenses/by-sa/4.0/。保留的元素、源代码仓库以及关于制作工具的说明列于本书末尾。

Contents

引言	7
这本书讲的是什么?	9
变革的基本前提	16
变革的基本前提	16
工具	25
基于声誉的社交网络	25
为什么网络不可审查	48
回应显而易见的疑问	50
对技术的依赖	50
在网络里发布信息, 是不是把钱往窗外扔?	50
如果有人创建多个身份怎么办?	52
有钱人难道不能干脆“买”更多身份 (或创建虚拟共同体) 吗?	54
那些只想读取、却对共同体毫无贡献的搭便车者呢?	55
金融中性	57
自愿、责任与自由	59

自由—极权开关 (以及委托) 61

预言机问题——架起数字世界与物理世界的桥梁 66

How Verification Works 68

共识与验证过程 68

Roles in the verification transaction and authorities 74

 角色总览 74

 发行者 76

 权威 76

 验证者 80

 主体 83

 观察者 85

 机制: 时间戳与挑战码 85

 妙处: 你并不需要真的观察者 85

 委托 87

 角色之间的通信 91

涌现的社会契约——政策 93

提供关于你自己的信息 97

Crime and Punishment	98
罪与罚	98
Follow the Money	108
追踪金钱的轨迹	108
国家的软肋——金钱	108
Proposed transition to the state's successor	111
演化, 而非革命	111
相互支撑的四件工具	112
简单的税制——入场的诱饵	113
电子支出登记册 (ESR)	113
公民税收分配——决定的分量	114
与国家谈判的平台——那根杠杆	118
无需身份证的公民身份	120
重叠与可逆性	123
Conclusion	126
道路是存在的	126
致谢	127
灵感	127

审阅与反馈	128
工具	128

Glossary	131
-----------------	------------

用语集	131
---------------	-----

引言

你觉得自己能够依靠国家吗？

当一桩显而易见的不公降临到我、我的家人或我共同体中的某个人身上时（身体伤害、欺诈、无理的限制……），而国家的回应是：是的，这是不公，但你别去做任何事，否则你和我们之间会有更大的麻烦——我拒绝接受这样的局面。这样的状态是病态的，而随波逐流的人是同谋。

其结果是彼此疏离、假装共同体、假装国家，以及社会日益深重的崩解，并预示着前方严重的冲突。

国家能否跟上技术与社会变化的步伐，还是宁愿一巴掌拍下一道禁令？

我本可以假装把停滞、无能与腐败当作法治与传统所宣告的稳定，但我不想再这样了。那样我是在欺骗自己，也是在欺骗你。

当你胆敢批评一件你正在为之付费的事情时，你是否感到官员与民选代表的居高临下？

我拒绝认为社会中的人不配参与塑造全社会的规则，仅仅因为他们不擅长那些软技能——而在今天，那些软技能恰恰是成为国家法定机构中民选代表的唯一充分前提。我相信绝大多数人能照顾好自己、自己所爱的人和自己的共同体——并且他们在其中看到了自己人生哲学的一部分。我所认识的身边大多数人都是这样。

但我也遇到过彻头彻尾的混蛋，他们之所以是混蛋，恰恰因为他们懂得如何玩这场游戏——在这场游戏里，共同体不被允许给他们反馈，国家保护着他们，使得没有人去扰动那个秩序；而那个秩序若被扰动，就会威胁到国家的本质：确保权力等级及其机构等级上层楼阁里的那些人过得轻松。封建制被民粹主义者和精英主义者的统治所取代。

你喜欢别人告诉你该怎么过日子吗？



我拒绝接受。

我不接受当下的局面是一个完美的、最终的既定事实。我在自己的行为举止上必须负责,但当有人告诉我应当持有哪些“正确的”价值观时,我却很敏感。你呢?

这本书讲的是什​​么？

想象一个声誉社交网络,任何想加入的人都可以为它创建自己的密钥(创建自己的身份),由自己共同体的成员邀请进来,或者独自闯入。你建立自己的声誉,并在需要时调取关于他人的有用声誉数据。当不公发生时,你向自己的共同体寻求支持——你陈述你的案情,提供你手上关于此事的东西,共同体则审视一番,根据相关性决定是否将其传递给他人,以及如何制裁冒犯者。

这整本小书讲的,就是要让这一切在外表上如此简单地运作,底层需要有什么。如果你坚持读下去,我向你保证:一个如今在许多方面看似复杂、难以理解的世界,将展现得更可理解、更清晰。你会看到,许多坏事都可以被安排成不再那样不公地对坏人有利。而且这并不是一个你得整天挂在网上盯着的工具——只需在你觉得需要时伸手取用即可。

我还向你保证, 这里所描述的安排会更贴近人们真实的动机, 而且它无需以某种“主义”之名去重塑人们。

最后,我们会展示如何管理一次非革命性的过渡,过渡到一种状态:共同体重新成为完全运转的社会单元,不再只是假装。

■ 技术文章

本书是对原始技术文章《去中心化声誉网络: 一个自愿且不可腐蚀的社会组织框架》(Pavel Kudrna, 2026) 的解释性伴读。如果你对形式化的架构描述、验证者选择算法以及经济机制的论证感兴趣, 我建议从那篇开始。

■ 各章层层递进

本书写作为一个连贯的论证——每一章都建立在前一章之上。某一部分引入的概念会在后续部分被使用, 包括在脚注中。我不建议乱序阅读。

■ 这本书写给谁？

尽管其外观形式友好，但主要的目标读者是较为专业的读者——关心人类社会可能的社会—经济—政治安排之原理的人。要达到我能把整个主题以通俗方式向最广大公众讲清楚的程度，仍需一条漫长的道路来构建完整的传播策略。

我欢迎任何建议，帮助我把这部小作品，或后续任何修订版本，做得更易懂。

警告

我不知道你是谁、多大年纪，也不知道你这一生中不得不忍受过多少破碎的承诺。我在这里不再做承诺了。相反，我想警告你：你在此处所学到的，唯一的结果或许就是意识到——大概没有一条路能让别人替你收拾今天的烂摊子。我只是在描述一个框架/工具，它或许能让这场收拾变得更简单、更透明。基于它，我将尝试筹备一场社会实验/模拟，修复潜在的缺陷，或者——如果它被证明存在根本性的设计裂痕——把这个想法彻底埋葬。

一个请求

如果你在读完这本书之后，仍能对这个想法保持开放的心态，并追随它的走向以及它在实质性批评下如何经受考验，我将深深地感激。而我更加感激的，是你日后的参与——用你自己的想法去尝试打破所提出的均衡原理，从而更好地检验这个系统对真实人类行为的准备程度。

没有你和你的参与，这个想法就只是抽屉里的一个理论。没有你和你的贡献与资金捐助，我们就无法把这个项目推向一场真正的社会实验，也就无法推向真正的成熟。没有你和你实质性的批评，我们就会停滞在如今的位置，或任由自己被中间派的圣杯所吞噬——绝对的、不可逆转的极权主义。选哪条路，或者干脆坐视，取决于我们。



底层还有更多。别走开。

挑战

我们并不算给你的肩膀减了负。你已经在通过那些神秘咒语般的手册承担着一个挥霍无度地治理的社会成本——那些手册声称社会最好是集中式治理——而我竟有胆量向你要求更多, 并且去对抗这些倾向。我很清楚, 我们将不得不熬过历史上最大的庞氏骗局——现收现付的养老金制度——的崩溃, 而又不能让好几代人挨饿。

但我仍然请求你: 让我们终于把一个更好的世界交到我们孩子手中, 并留下一份遗产——我们完成了那场变革。释放经济增长或许能部分地解决地平线上那道困局。届时我们可以对后代说: 这是一段惊心动魄的旅程, 而且没有白费。

■ 问题

唯有由直接责任所平衡的自由, 才能真正释放出这个世界从未见过的巨大潜能。



更好的世界值得这份重量。



进来吧。方案已经摆好。

变革的基本前提

变革的基本前提

向好变革的根本前提,是让从属者成为主导者——或至少成为平等的伙伴。这意味着接管那些界定主从关系的最关键议程。这些议程立足于对“归属于构成国家公民的那些共同体”的证明。它们是信息,以及信息在这些共同体中拥有实质声誉的权威面前的可验证性。

为了克服当前的体制,我归纳出了五个前提,随后我会以我理解它们的方式加以阐释。我们可以把这些前提看作支撑本书其余部分所描述的体系之建构的支柱。这些支柱起初是我按自己的偏好挑选的,而在用它们进行推理的过程中,我逐步把对每一根支柱的理解打磨成了下文的形态。

让我们逐一走过这五根支柱。

■ 自愿

在我们这里,自愿意味着我不被强迫参与去中心化声誉网络的运作。然而,我不能要求一项被正面表述的权利,即网络不得记录关于我的信息——关于我的信息是否最终进入声誉网络,只取决于我自己、我的行为、我的作为以及我的责任。

变革的前提

自愿性



自愿性

我为自己做主。

组织



组织

齐心协力，不受制于人。

韧性



韧性

无法被审查。

共识



共识

以压力换取妥协。

不可腐蚀



不可腐蚀

无法收买所有人。

没有这五根支柱，任何变革都只是白白烧时间。



自愿——直到责任找上你。

■ 组织

单个个体无法打破这个体制,这正是为什么国家控制着人的结社,以此防止那些可能威胁其结构与运转的活动。人们组织起来是支持还是反对,并不重要。通过组织,可以在有人胆敢违背国家意志去改变国家时,软化杀鸡儆猴式惩罚的冲击。

我们不必设定共同的目标,但我们可以一同走一段路。最自然的形式是在共同体中结社——但它们必须拥有所需的全部工具,才不至于停留在假装做共同体的表演。

■ 韧性

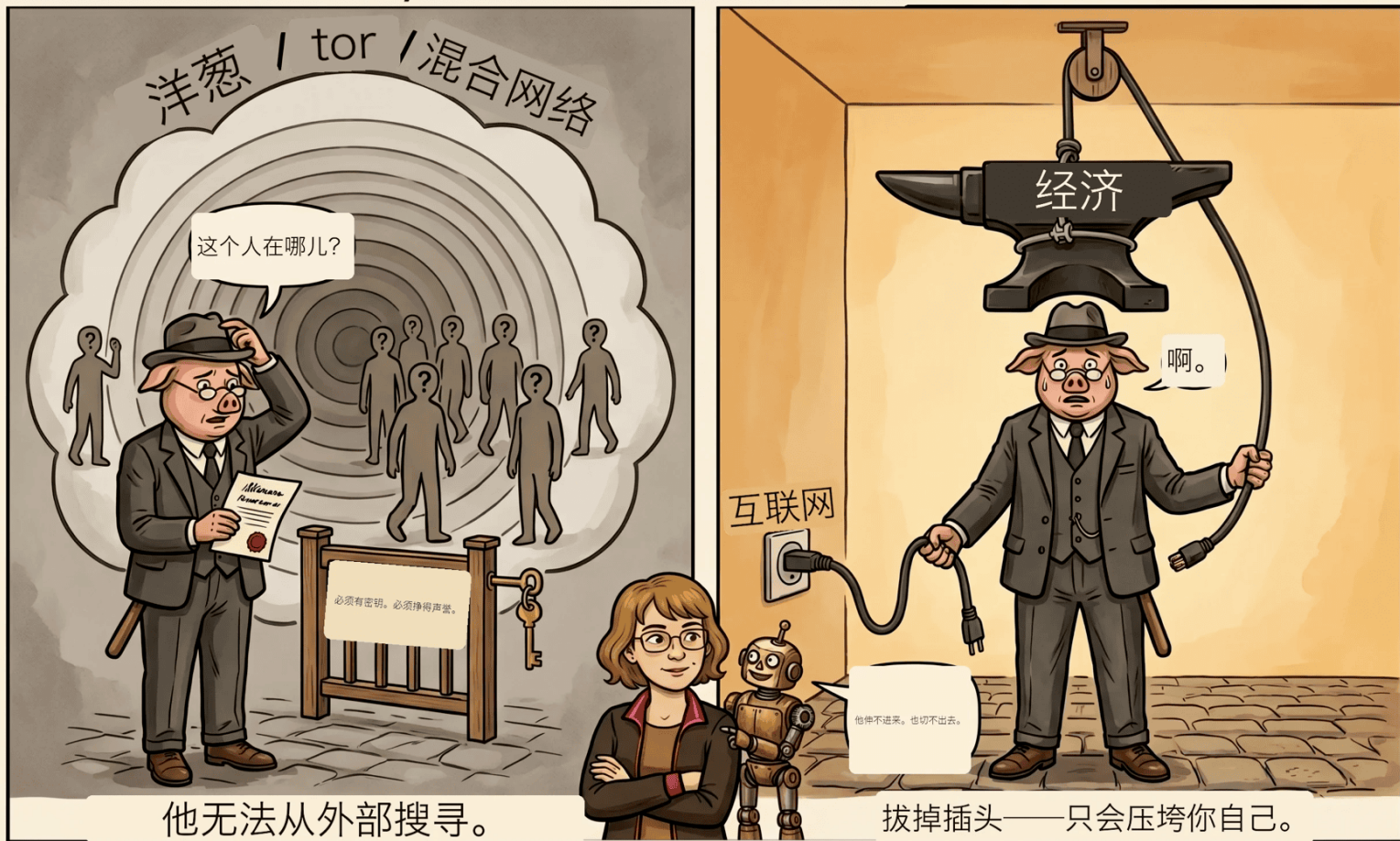
所选择的通往改变国家——或通往向其继任者过渡——的道路,绝不能依赖国家的许可。工具及其使用方式必须尽可能不可审查。说白了,审查它们的代价必须高得惊人。没有这一点,整个努力不过是在燃烧时间和生命的能量。

前提2/5：组织



个人无法打破体系。社区可以。

前提3/5: 韧性



他无法从外部搜寻。

拔掉插头——只会压垮你自己。

没有许可可求。没有插头可拔。

■ 共识

所提出的解决方案, 就其本性而言, 必须既包含协商共识的机制, 也包含在共识被违反时的镇压性组成部分。它不必是一份自愿的双赢协议。只要环境与社会压力能确保在一个共同体内部以及各共同体之间达成妥协, 就足够了。

■ 不可腐蚀性

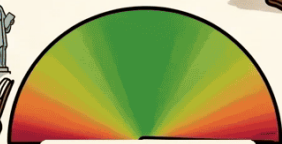
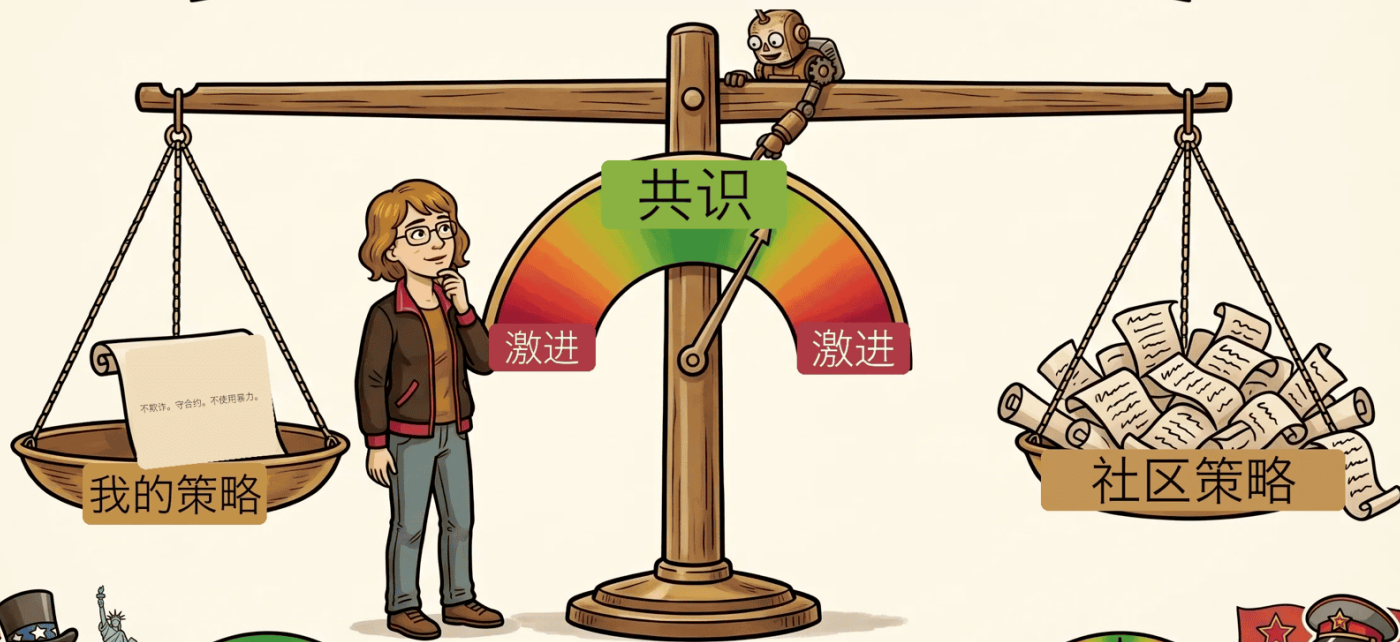
不可审查性从外部捍卫这个体系。但一个由人构成的网络也可能从内部腐烂——某个权威被俘获, 一条宣告的规则被悄悄抛弃, 参与者在无声中被收买。如果内部的人可以毫无后果地被腐蚀, 那么不可审查性就只是一场空洞的胜利。

不可腐蚀性并非某个巧妙把戏的结果。它从**所提出设计的若干原理协同运作**中涌现出来——对抗腐败的压力被分布在整個架构之中, 各个部分相互加固。这具体是如何实现的, 我们将逐步展示——在对工具的描述中、在验证中、在惩罚那一章中。最后我们会把已经搭建起来的一切重新梳理一遍。

不可腐蚀性同时既是设计的一个前提, 又是结果的一个涌现属性。它是第二道防线——朝向内部的那一道——没有它, 韧性就只是半面盾牌。

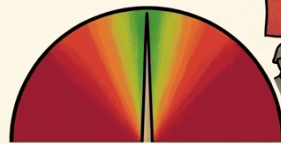
待在区间内——别当激进分子。

你可以不同。只是别偏得太远。



自由社会

文化越自由，容忍区间越宽。



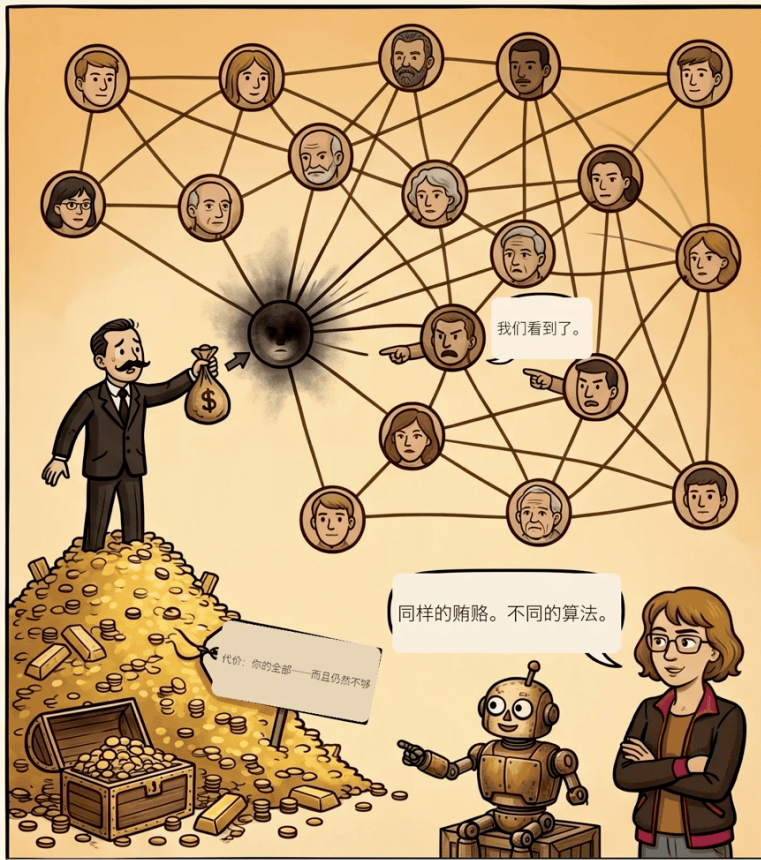
威权社会

前提4/5：共识
共识是一个容忍区间，而非一致同意。

前提5/5: 不可腐蚀



贿赂几个上层——底层照样买单。



贿赂一个人——他永远失去声誉。贿赂所有人——你先破产。

腐败有代价。在网状网络里，这代价太高。

工具

基于声誉的社交网络

要带来变革, 我们需要一件精心设计的工具。首先我们会简要勾勒它; 在后面的章节里, 我们会更细致地考察它的每一个部件, 并加入更多内容。想象一个不可审查的、全球的、去中心化的社交网络, 你可以在其中安全地创建和管理你的代理身份——即所谓的去中心化身份 (DID)。DID 是一种由你自己创建和掌控的数字身份, 不依赖任何中心权威。没有人能夺走它或伪造它, 因为它由你的私钥 (或多把密钥, 通过多签) 进行了密码学签名。

■ 说明

一个推论是: 这样的身份可以逐步取代国家签发的身份证件——不过这留待过渡那一章再谈。

在这样的网络里, 你可以通过你的身份报告有人对你造成了伤害 (以及之后, 可能地, 报告他们已经补救了伤害或被迫补救)。为了让这份反馈——指向伤害的始作俑者——作为一个相关信息源具有价值, 把信息录入网络必须付出时间、精力和金钱——而且在此之上, 还必须向他人拿出可验证的证明, 表明这不是无聊的闲言碎语。

读取信息会很容易且相对便宜, 但创建一条个人记录则代价高昂、要求很高。写入会遵循一个清晰的协议, 其中按照所选算法进行的计算严格决定: 向哪个 DID 请求对所提交信息的验证, 以及如何推进, 才能让被选中的参与者代你处理这条信息、将其发布, 并成为它的验证者。

你的身份，你的密钥，你的规则

你自己创建它

私钥——代表你签名。只有你持有。永不外泄。

公钥——验证你的签名。对网络可见。



无需登记处。无需政府表格。无需公司批准。你自己生成它。

你的DID文档

声明——关于你的已验证事实



公民身份：捷克共和国——由市政公证人验证

学历：计算机科学硕士——由大学校友网络验证

财产：42号公寓所有者——由本地社区登记处验证

消费凭证：无不良行为——未履行准则——由邻里见证人提交

欺诈：骗走邻居的积蓄——由本地社区法庭提交

这是你的公开身份，任何人都能读取。只有你的私钥能签署对它的变更。

任何国家都无法撤销它。

声明是已验证的陈述。好的和坏的都在。权威机构以其声誉为他们背书。



国家身份证——可撤销：国家随时能注销它。



DID——不可撤销：只有你的私钥能签署变更。

你的身份。你的密钥。你的规则。无需任何许可。

■ 算法与激进

对验证者的算法化选择确保了: 非激进的信息发布者, 随着时间推移, 会在所发布信息的成本与验证的报酬之间保持近乎中性的平衡。

发布要花费时间、能量和金钱

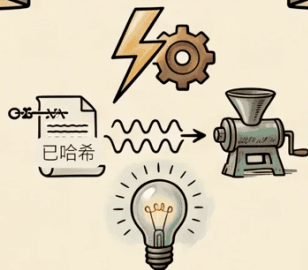
写入网络从不免费——读取才免费
三个渠道——二个都是真的

时间



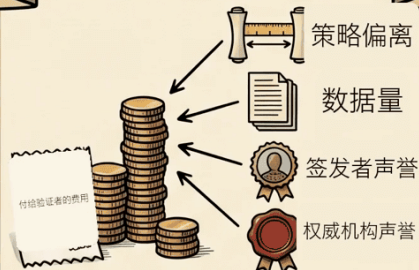
算法抽取验证者、验证者读你的记录时，你得等待。

能量



签名、哈希、验证——每一步都是实打实的计算和电力。

金钱



你的声明离验证者的策略越远、记录越大、签发者或其权威机构越不受信任——费用就越高。

接下来会发生什么



三步验证器——“算法和哈希验证器”，“验证者验证器”，“签发者验证器”



为什么昂贵的写入保护声誉——本章后文

跨迭代的成本

每次重试成本更高



每次拒绝或重掷都给算法吞入的时间顺序记录添一笔——下一次尝试在三个渠道里都更贵。

读取不对称

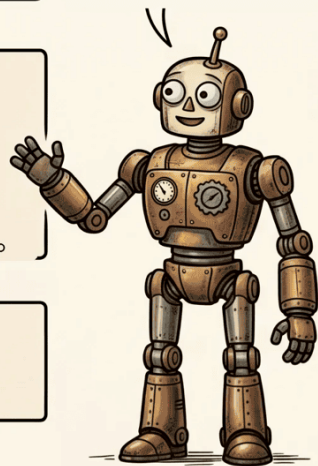
读取很便宜



查询免费

任何人都可读取声誉和记录——只有写入才收费。这种不对称正是设计所在。

写入收费——读取不收。这种不对称正是要点。



让我们看看算法是如何选择验证者的。

■ 算法

算法化选择会为不同的信息片段非确定性地选出不同的验证者 (或一组可能的验证者)。完整 DID 文档的哈希 (一种单向数学函数, 从任意输入产生一个唯一的“指纹”——就像文档的指纹) 决定了它在一致性哈希环上的位置, 并选出验证者候选人。

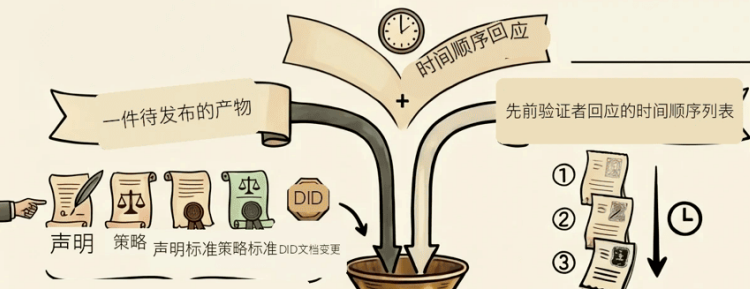
$$\text{verifier_DID} = \text{hash}(\text{DID_document})$$

用大白话说: 算法取你整份 DID 文档, 从中计算出一个指纹, 而这个指纹决定了你的验证者。

在算法选出的第一个验证者那里, 你作为发布者可能不会成功——你的声誉或所宣告的设置可能不满足他们的要求。你会通过再执行一次递归迭代来算法化地继续搜索下一个, 这次迭代会给你分配又一个验证者。每走一步, 通往目标验证者的“距离”就增大一分, 随之必须发布的附带元数据也随之增大。随着数据增长, 成本自然上升 (不仅因为主张本身最初的体量, 还因为每被拒绝一次就累积的元数据)。可信的信息比无稽的一时兴起要顺利得多。每个人愿意承受多高的价格、这条记录对自己有多重要, 都由自己决定——激进保证会变得昂贵。

算法如何抽取验证者

一枚公平硬币掷向你的整个社区

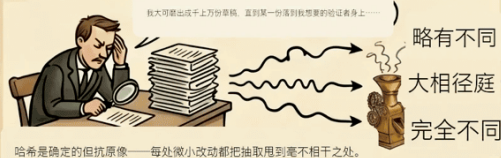


你的社区及其周边

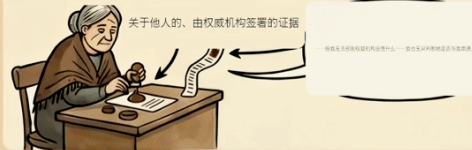


发布路径为：签发者 → 权威机构 → 验证者。任何记录都不得跳过权威机构这一步。

A 你能暴力破解它吗？



B 但权威机构封住了其余部分



算法来抽取——不是你。

验证者如何作答

以一个价格接受——或以一个理由拒绝



拒绝——附带一个理由

- 签发者的声誉对我来说太低
- 这条记录背后的权威机构不是我信任的
- 签发者的策略离我太远，无法弥合
- 文档本身未通过我策略的规则

验证者写下原因——他们的回应会反馈进下一次尝试。

接受——附带一个价格



验证者同意发布——但价格反映了声明离其策略有多远、承载多少数据、以及签发者有多受信任。

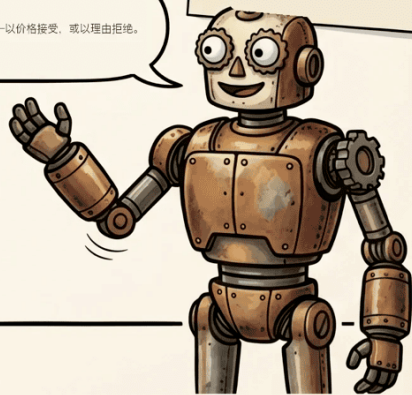
验证者的回应会被加入时间顺序列表——如“算法如何抽取验证者”所示

无论验证者说什么，其答复都成为下一次调用的输入。

每个验证者都以这两种方式之一作答——以价格接受，或以理由拒绝。

接下来会发生什么

- 签发者决定如何处理这个答复——见“签发者的决策”
- 每个答复都涉及下一次尝试的成本——见“发布成本”



验证者要么定一个价格——要么告诉你为什么不行。

无论验证者对你的验证请求作何回应, 球又回到了发布者的场地: 他们可以接受验证者提出的验证服务报价, 把该回应折入时间顺序记录并 (更昂贵地) 再试一次, 或者干脆走人、吞下已经沉没的成本。

为了给你的信息更大的分量、更高的被验证者接受的机会, 你——作为对所发行信息有切身利益的发布者——可以使用一个**可信权威**的服务。权威要么拒绝所提交的信息, 要么接受它并把自己的好名声 (声誉) 押在上面。权威通常会索要现实世界的证据, 加以核验, 并进行分类。其产物是一份关于给定案件在给定时间的评估协议。可以把权威想象成现实与数字世界中某类服务的专家——例如一位调查员、审计师、保险人、某类商品的供应商 (本质上, 市场上任何一个经济行为者)。

签发者的抉择

如何处理验证者的答复

无论验证者说了什么，签发者现在桌上都有一个答复。



回到“算法如何抽取验证者”——先前所示

只有重试这条路会循环——接受和走开都终结故事。

验证者提议——签发者定夺。

三条出路

接受这笔交易

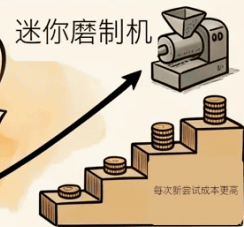
DID网络



仅当已被接受

签发者支付费用——记录被发布进网络。仅在验证者接受时可用。

再试一次



回应被并入时间顺序记录——算法再掷一次。无论验证者是以签发者不满意的价格接受、还是干脆拒绝，都适用。

就此走开

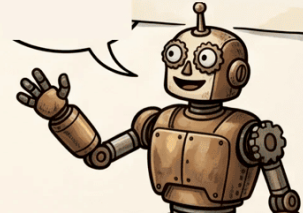


已经花掉的时间和能量不会回来

签发者丢弃草稿。什么都不发布——但也不再多花钱。

接下来会发生什么

每次重试都会支付下一次尝试的成本——“即时”发布成本“折扣”
验证者的答复本身——“预先”验证者如何作答“指示”

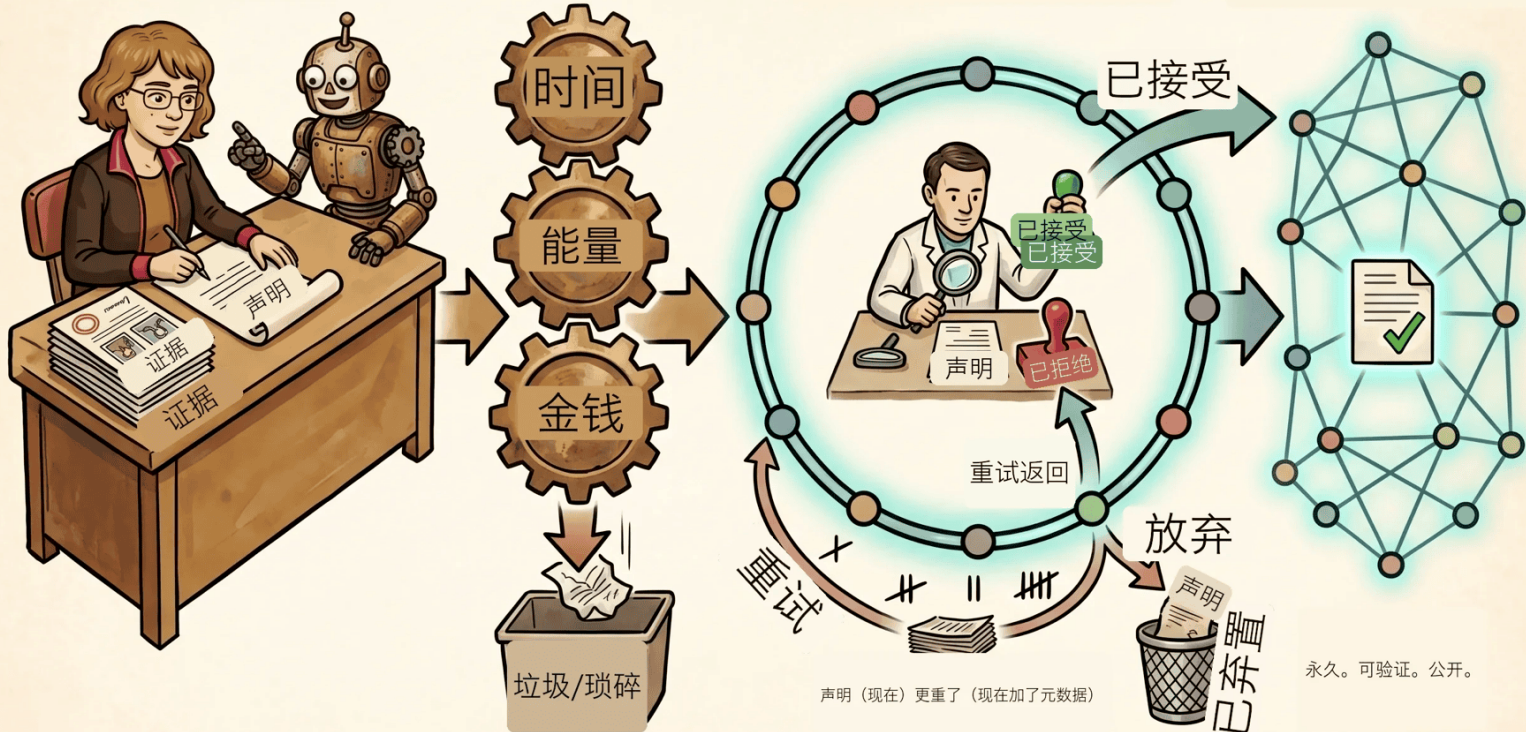


阶段1——发布一条声明

阶段2——质量过滤

阶段3——递归验证

阶段4——永久记录



伪造昂贵。验证便宜。

在你试图把信息发布进网络的时候, 网络里很可能已经包含了关于其行为者的信息——这些就是声誉信号。如何解读声誉信号——它们在不同处境下对你意味着什么、带有什么风险——可能并不简单。每个参与者都可以透过自己的 DID, 依据自己正在处理的、与对方相关的处境, 以不同的方式看待声誉记录。对方是可靠的付款人, 还是我需要在一笔交易中要求预付款项? 所提供的产品是否带有关于隐蔽欺诈或缺陷的评价? 当出了岔子时, 他们是否想要逃避合同责任? 有时, 一个关于对方整体一致性的、更为复杂的视角会派上用场——这取决于要求获取这份概览的人的偏好。市场可以提供一些产品和服务, 针对当下处境去简化、处理和厘清对声誉的解读。各种权威及其所提供的服务也能服务于这一目的。

■ 示例

典型的、发布者感兴趣——且对他人有价值——的信息, 涉及的是现实或虚拟世界中超出寻常人际交流的事件。

负面示例: - 犯罪行为的证据 (例如, 由某个可信的调查机构审计) - 间接证据 (单独看很弱, 但在统计上可累积)——例如, 短时间内多次出现在多起盗窃案附近 → 还是巧合吗? - 违约

正面示例: - 已补救的伤害 (自愿的, 或在共同体作为惩罚施加的压力下) - 接受并履行由权威 X 提出的处罚 - 权威 X 在一定范围内撤销了对加害者财产权的承认

由每个人自行去收集关于对方的可得信息, 并按自己的偏好评估风险。

如何读取声誉

知情决策



她清楚自己在和谁打交道。
在你放贷之前。在你签字之前。在你信任之前。

盲目信任



他根本不知道自己在和谁打交道。

会出现哪些种类的声明？

声明产物示例

(声明是关于某人的陈述——这里是人们发布的种类。)

-

负向声明



殴打
对他人施加了身体攻击



盗窃
未经同意取走了他人财物



欺诈
为个人利益欺骗他人



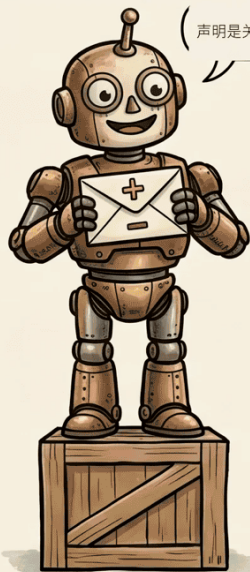
违约
未履行约定的义务



诽谤
散布有害的不实之词



遗弃
对被扶养人或职责弃之不顾



+

正向声明



伤害已弥补
修复了早先造成的损害



债务已清偿
全额偿还了所欠



施以援手
以个人代价帮助他人



经验证的专长
在某手艺或领域证明了技能



卓越贡献
对社区的长期贡献

仅为示例——每条声明都遵循一个声明标准，任何标准都可再添加。

你不必加入——但若你的行为影响他人，就有人能就此发布声明。

声明是行为进入网络的方式。

■ 关于你的信息是否出现在网络中, 只取决于你自己的行为。

你永远不必加入这样的网络, 然而关于你的信息仍可能出现在其中。这只取决于你的作为以及它们对他人产生的影响。

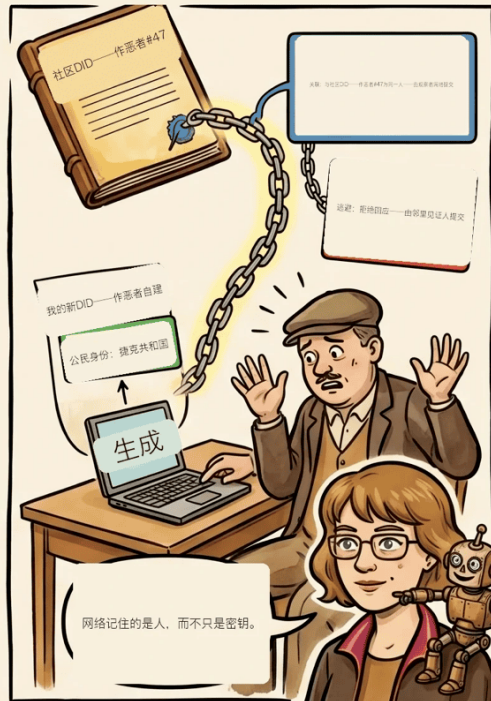
社区可以替你开一个



他没有自己的DID。他从未注册。



你没有开设DID，于是社区替你开了一个。负面声明附着在这里。



你可以开设自己的——但社区会把两者关联。声誉跟着人走，而非密钥。

1. 你造成了伤害 -----> 2. 社区提交 -----> 3. 你无法逃脱 -->

你可以选择不开DID。你无法选择不要声誉。

我刚才简要勾勒的, 就是一个受去中心化身份 (DID) 启发的社交网络可以如何运作。DID 概念的首要目的, 是通过“订阅我将遵循并据以生活的规则”这一原则来强化隐私与自由——让用户能够决定分享哪些信息、在什么条件下分享。

我进一步提议, 把各个 DID 连接成一个通信网络, 让它们的持有者即便在“某人身上出了事、共同体或个人需要作出反应”这类情形之外, 也交换反馈。这种对我们所订阅规则的预防性对照——并可以计算“关于对方本应如何运作的期望之间的相互偏差”所带来的经济及其他后果——可以被视为一种寻找共识的动机。这样的体系强调的不是自由, 而是与“对现实行为负责”相结合的自愿决策。

单个个体无法独自打破这个体制——一群人机会更大, 而一群达成了协商共识、并在许多议题上有动机同心协力的人, 抵抗威权倾向的机会则更大。第一章的组织前提, 将在两个条件被满足时得以实现: DID 声誉网络足够有代表性地覆盖了各共同体, 以至于使用它不再是稀奇之事; 与此同时, 这一共同体群体成为一个具有经济分量的少数派, 能够强势地与社会其余部分谈判。

■ 自愿与自由

自由——就其正面意义而言——将是平衡两个因素的次生效应: 自愿, 以及来自周遭环境、朝向责任的压力。

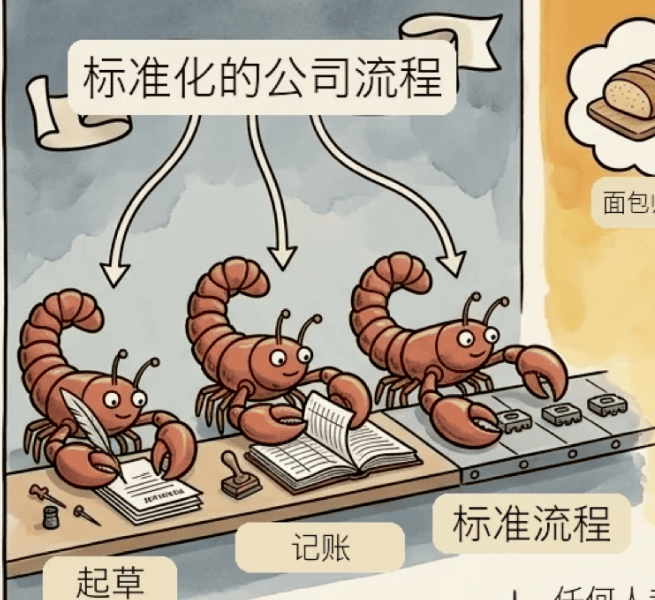
■ 人工智能时代与声誉的价值

在人工智能时代, 一切与认知思维相关的东西都在被自动化——而且可能走得更远。那么, 在人类活动中, 还剩下什么可作为竞争优势? 答案很难, 而且总会找到些什么, 但有一件事我们可以确定无疑: 声誉将起决定作用。你行为的可验证历史、你的承诺及其兑现——这是人工智能无法替你构建的东西。

如果AI来思考——那还剩什么给你？

AI智能体接管重复性工作

标准化的公司流程



重复性的认知工作正一件件交给智能体。

那么人还剩下什么？



任何人都能挑中同一个点子。

是什么让一个面包师、一个木匠、一间AI工坊与别家不同？

声誉——区分二者的东西



总得有人为智能体的所作所为负责——而这份责任无法转嫁。

没有声誉，每个面包师看起来都一样——最大的品牌把他们全吃掉。
这会不会是让AI受约束的去中心化路径——那个阻止任何一方独占整个市场的东西？

AI把工作自动化——声誉让你保持为人。

真相的经济学

社区条



同一个社区。不同的立场。

提交

不受欢迎但可接受



在共识内



在共识外



抽签



抽签公平——被抽中的验证者收取风险溢价



基础费——首次抽签即接受



抽签不断进行——没有被抽中的验证者会签 已弃置

退款

账本



多数退还，随时间有些损失

略贵一些



平均全额退款



负担不起



多数情况下押金被没收

多数情况下押金被没收

底部观察者条

算法公平地抽取。人们为自己的信念定价。

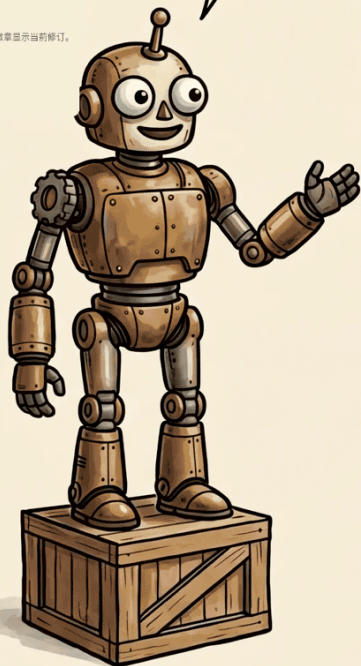


均匀抽取。非均匀回应。离共识越远，价格越高。

网络中会发布什么

五样东西可被发布——每样都有版本。

每件产物都有版本——徽章显示当前修订。



v1

DID文档

你在网络上的身份记录——密钥、元数据、网络地址。

管理者：DID持有人



v2

声明标准

声明必须遵循的模式——什么算作有效证据。

管理者：在该领域有声誉的权威机构



v3

策略标准

策略必须遵循的模式——接受规则的形态。

管理者：在该领域有声誉的权威机构

↓ 派生自

↓ 派生自



v4

声明

以下某类型的声明，通常发布时只有一组发布者发布，但发布者可能因需求更多，而随时更改发布方式和数量。

管理者：签发者发布——主体是目标

声明分两类——正向 (+) 和负向 (-)。



v5

DID策略

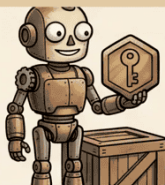
某个DID声明的规则——用于验证、共识和签发。

管理者：DID持有人（每个DID一条策略）

每件产物都有版本。每个版本都公开。

DID文档里面有什么？

DID文档产物示例



你的DID文档是你在网络上的身份卡——这里是它承载的内容。

你会在里面找到的字段



密钥与签名规则

每把密钥授予一种权力——编辑本文档、签发声明、担任权威机构、管理策略或委托恢复。权力可交给另一个DID，日后也可再撤销。



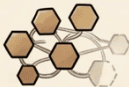
网络地址

网络在哪里能联系到你——可更改，无需重写身份



创建与版本

这个身份最初出现的时间，以及你此刻读到的是哪个修订版



关联身份

你所连接的DID的汇总——一个便捷视图，技术上存储在别处



继承标记

若你的身份丢失或被盗，权威机构可将此DID标记为终止并指向你的继承者

DID持有人书写每个新版本——但每个版本在网络发布前都必须通过验证。



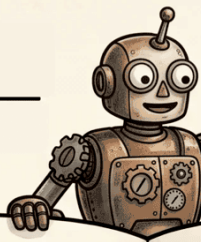
你书写它——网络验证它。你无法凭空把自己发布成存在。

你撰写自己的身份——网络予以确认。



什么是声明标准？

声明标准产物示例



标准是空白表格——声明是填好的表格。让我把两者并排给你看。

标准——空白表格



主体DID
此声明所涉及的DID



签发者DID
签署此声明的DID



何时何地
一个日期和一个地点



认可的证据类型
只有这些种类的证据算数



无/轻度/中等/严重

结果分级
恰好选一档——不接受自由文本

声明——填好的表格



主体DID
did:nwo:alice-9f2b...



签发者DID
did:nwo:magistrates-04...



何时何地
3月3日，集市广场



认可的证据类型
照片 ✓ — 证人宣誓 ✓ —
医师证明 ✓



认可的证据类型
中等——第2档，共4档

左侧每个字段在右侧都有恰好一个填好的对应项——正是这一点让同类的两条声明可比较。

权威机构在填好的声明上的印章是一份承诺：流程确实被遵循了。

声誉脚本



我以声誉担保这是按标准处理的。

殴打 v3



殴打 v3

债务 v2



债务 v2

医疗援助 v4



医疗援助 v4

手艺 v1



手艺 v1

证词 v5

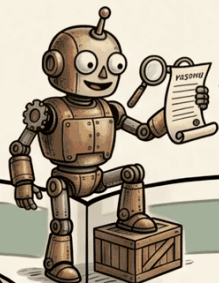


证词 v5

标准是表格——声明是填好并作出承诺的表格。

什么是DID策略？

DID策略产物示例



我的策略不以发声明给谁看——而是给谁看。我的策略是我自己决定要不要理会这条声明。我也不在乎谁管它的人。我也不在乎谁管它的机构。这件事都不值得我关注？

标准——继承的默认值

规则名



签发者声誉下限

我接受声誉在十分之五以上的签发者



权威机构信任清单

我接受任何自身声誉在十分之四以上的权威机构所定的标准



实质性过滤

我处理任何事由重于“微不足道”标记的声明



反垃圾定价曲线

我提交实质性的、罕见的声明很便宜——琐碎或重复的则迅速变贵



综合权重

签发者声誉与权威机构声誉相乘；琐碎事由将结果减半

策略不重写标准——它评判的是人和事。保留、修改、移除、添加——在共同起点之上的四个简单动作。

我的策略——由它派生

规则名



保留



已修改



已移除



已添加



签发者声誉下限

收紧——我要求十分之七



权威机构信任清单

我完全信任裁判官和医师；对市政监管下的声明我把权重减半；对警务署名下的一律拒绝



实质性过滤

琐碎骚扰声明我不理会毫无代价——我干脆不读它们



反垃圾定价曲线

保留——默认曲线对我够用



综合权重

移除——我用我自己的公式



重复声明的指数级成本

已添加——同一签发者一年内每条关于我的新声明，成本都是上一条的十倍



静默时段

已添加——我拒绝阅读午夜到黎明间提交的任何声明



验证——我的策略既过滤收到的声明，也在我签发前过滤自己的草稿。签发就是出门时的验证。

共识——在社区协商共同规则时，我的策略就是我的一票。

你的策略评判信使——而非表格。



为什么网络不可审查

到目前为止, 我们描述了网络能做什么。现在让我们聚焦于: 为什么它无法被关闭或审查。

网络很可能建立在 Tor 网络与洋葱路由 (一种利用混合网络原理来掩盖网络流量的方式) 的原则之上——这是一种技术: 通信穿过若干层加密, 没有任何单一节点能同时知道发送者和接收者。中心机构无法越过你的肩膀窥视, 因为根本无处可窥。

你可以去对抗它, 但你无法取胜——除非你把整个国家投入极权式监控, 监控到个体网络数据包的层次, 监控到每一件连入互联网并在该声誉网络上通信的硬件上的底层操作。而这恰恰是集中主义所面临的选择: 要么接受公民手握一件你够不着的工具, 要么以极权控制之名把自己挖进经济停滞的黑暗角落。

我们别抱幻想——即便是民主政体, 在其为集中主义而战时也会尝试这么做。他们宁愿给人们戴上镣铐, 也不愿交出权力。

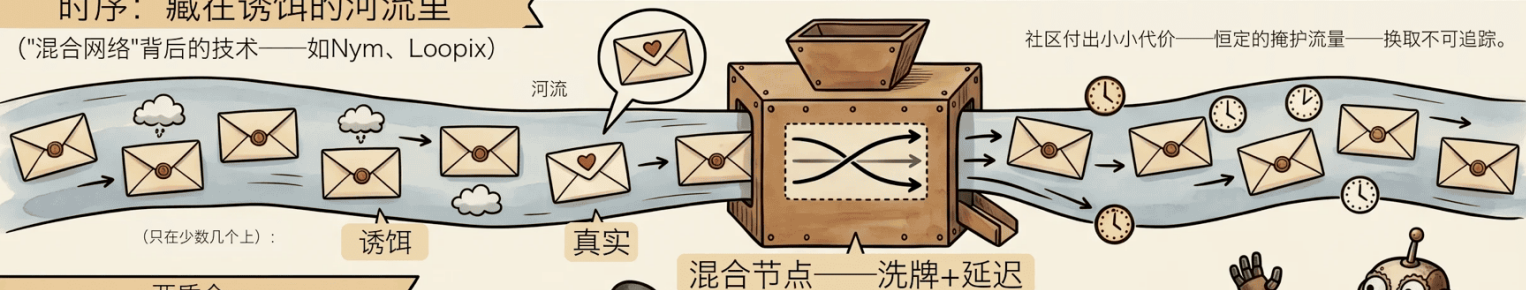
你的消息如何保持不可追踪

两道防线同时上阵——层与噪音

内容：裹在层层之中
(Tor背后的技术——“洋葱路由”)



时序：藏在诱饵的河流里
(“混合网络”背后的技术——如Nym、Loopix)



两盾合一

矩阵	隐藏谁？谁？	隐藏何时？
仅有层	✓ 是	✗ 否
仅有噪音	✗ 否	✓ 是
层+噪音	✓ 是	✓ 是

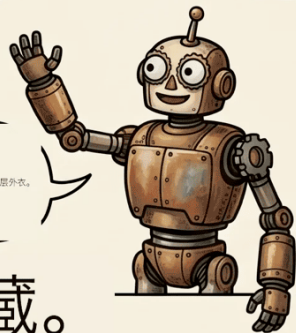
(洋葱路由+混合网络——我们需要的组合)



一小时一百个信封——哪一个才要紧？

观察者看得到流量，却分不清真实与诱饵。

真实消息在诱饵之河里滑动——还穿着四层外衣。



内容由层隐藏——时序由噪音隐藏。

回应显而易见的疑问

所描述的体系自然会引出一连串问题。让我们回应其中最常见的几个。

对技术的依赖

你大概已经注意到, 与现代国家——它伴随我们约有 150 年——不同, 这里所描述的声誉网络方案在很大程度上依赖全球 / 本地互联网的技术。一旦发生断网, 这样一个网络的运作就会陷入风险。

如果断网是暂时的, 那么网络中的数据或主张一致性不会有损失, 共同体内部所达成的声誉平衡也不应受到扰动。与可比的支付网络不同, 这个网络假定的流程与数据项频率非常低。在这一点上, 去中心化声誉网络与今天的国家并无二致——国家如今也严重依赖技术, 并且已经忘了怎么用纸质卡片档案办事 (尽管在危机预案里它别无选择)。

以声誉网络形式出现的国家的进化继任者, 在永久性断网 (一场规模空前的灾难) 的情况下, 可以回退到一个更原始的集中式系统——国家。

技术让人类得以达到更高的文明治理形态, 给我们带来益处, 但也带来风险。

在网络里发布信息, 是不是把钱往窗外扔?

在声誉网络里发布一条主张的成本, 大部分并非沉没成本。发布者投入网络的讯息之有用性——关于真实冤情的信息、经核验的经验、相关的警示——在统计上会在较长的时间跨度内, 以为共同体中别人的主张作验证所收取的费用的形式返还。共同体受益, 发布者建立声誉, 于是发布真实信息的成本便趋近于一笔可退还的押金, 只需从中扣去较小一部分实际的网络维护成本。反过来, 不真实或琐碎的记录不会返还——它们的成本是净损失。因此, 诚实不仅是一种道德选择, 也是一种在经济上理性的策略。

两条路——国家在关闭，网络在开启

声誉网络

国家

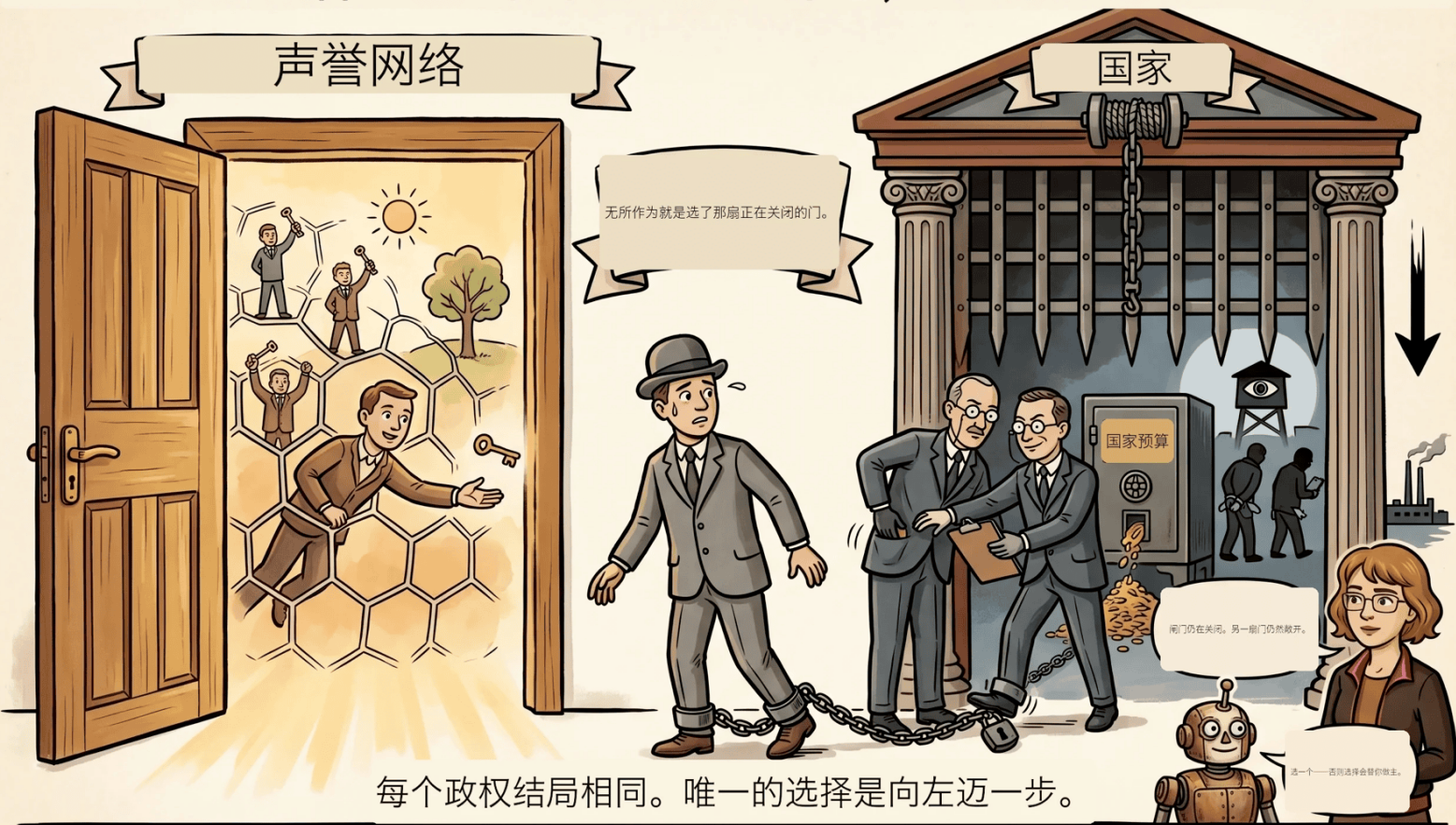
无所作为就是选了那扇正在关闭的门。

南门仍在关闭。另一扇门仍然敞开。

选一个——否则选择会替你做出。

每个政权结局相同。唯一的选择是向左迈一步。

国家默认关闭。网络因选择而开启。



在扣除网络维护成本之后,这一原则可以称为经济中性原则——当我与共同体同在时,我不亏;当我与它作对时,我才亏。

共同体也有团结的渠道来赏识其成员的诚实态度。但我们别抱幻想:对团结的诉求大多源自共同体的社会压力,所以这也许并不是“自愿”这个词本义上的团结。

如果有人创建多个身份怎么办?

一个人可以并行运作多个 DID 身份。然而,为每个身份建立声誉都需要独立的努力——时间、精力、金钱。

没有捷径:每个身份都必须通过实际活动来累积它的履历¹。因此,维护并行身份是被有意设计得昂贵的。

在自由社会里,这些成本会打消滥用的念头。

然而在独裁统治下,并行身份成了一种生存工具:它们使地下网络的组织、对黑市更安全的周旋,以及区隔化²的抵抗成为可能——在区隔化的抵抗中,一个身份的暴露不会牵出其他身份;而在政权垮台之后,它们让人得以无缝地回归公共生活以及在那里积累的声誉,甚至通过一条主张,把先前的公开 DID 与秘密 DID 合并成单一的一套综合记录。

¹**履历 (Track record)**——一般而言:一个人或组织过去的结果、成功与失败的历史。这里:某个 DID 身份在网络中过去所有互动的总和——已验证的主张、被接受和被拒绝的记录——其声誉即由此推导而来。

²**区隔化 (Compartmentalization, 源自英文 *compartment*)** 意味着把信息分隔成彼此孤立的单元,使得一个单元的暴露不会危及其他单元。这是情报机构所熟知的一条原则:一名特工只知道行动中属于自己的那部分,因此即便在胁迫之下也无法供出全貌。

同样的付出——在自由中是浪费的备用计划，在暴政下是救命稻草

平行身份代价相同——但换来的东西不同

在自由中——一个昂贵的备用计划



第一年——挖坑打桩、搬盆、施肥——声誉0% 第二年——同时挖坑栽两棵并一起浇水，各自独立培养——声誉20% 第0年——一切从头开始——声誉0%

三座花园，三倍成本——而你依然只有两只手。

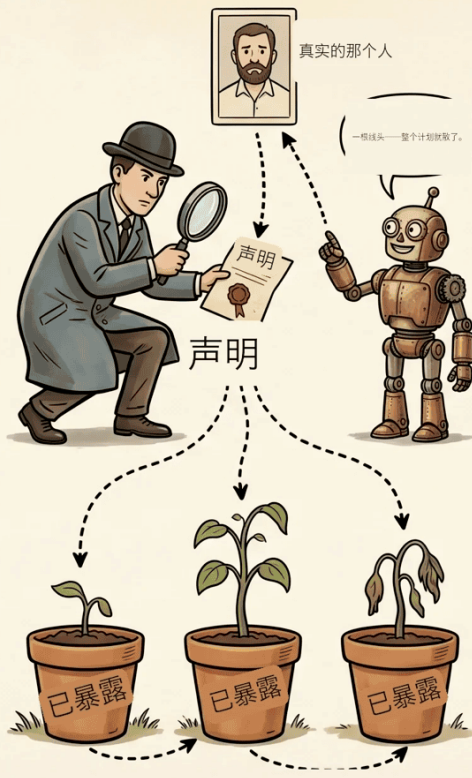
或者——一个人失足后的一次洗白



旧DID——声誉烧毁 DID-全新开始

只有当烂摊子仅关乎你时才管用——若他人受害却要追查你，他们的声明会跟着你进新盆。

一条声明，备用计划就没了



单一条声明就把每个盆和同一个面丁绑在一起——多年的时间、能量和金钱在一个下午被抹去。

在暴政下——一个交给国家，其余用于自由结社

(国家要么彻底禁止身份，要么要求身份归其掌控)



国家绝不能看见的那些



小组之间的砖墙做着国家所禁止的事——丢了一个，其余照常；政权垮台时，公共生活重启。

在自由中，滥用代价高昂。在暴政下，生存成为可能。

■ 视角被里外翻转

与国家不同,建立在去中心化身份之上的声誉网络原则,颠倒了“何为优先”的感知范式:

- 要紧的是声誉——也就是过去,用来评估与对方交往的风险——而姓名、地址等个人数据可以是出于礼节的数据交换
- 而国家则首先索要个人数据,累积声誉数据,却只让共同体看到合它心意的那部分

有钱人难道不能干脆“买”更多身份 (或创建虚拟共同体) 吗?

创建并行去中心化身份的可能性,乍一看像是经济手段更雄厚的人相对手段较少者的不公平优势。然而必须强调,与集中式系统不同——在集中式系统里,只需腐蚀权力金字塔中的几个点,就足以让某些劣迹消失——在去中心化系统里,人们不得不腐蚀整个共同体。

替身的去中心化身份可以服务于这一目的,但它们的声誉必须随时间推移、通过与实际的其他共同体成员的真实互动来建立——它无法被简单地买来,因为网络让“某个身份表现如何”变得可验证。

此外,市场上可以存在 (而且很可能会存在) 这样的服务: 它们以权威的身份行事,提供调查,拿出证据证明若干去中心化身份实际上是同一个人。于是,一条录入声誉网络的记录,就能以其零头的成本,抹掉建立并行身份所投入的全部时间、精力与金钱。

因此,从经济上讲,不欺骗共同体是划算的——并且在需要时,审视自己的行为并寻求补救,好让声誉回到可接受的水平,让其承载者不至于因共同体的震怒而在经济上或其他方面受损。

一个在自己共同体中丧失了声誉、却经济实力雄厚的身份,可能会试图通过与被针对的身份做见不得人的交易来逃脱共同体的震怒——但那些身份随后也要冒失去自己声誉的风险。

此外仍然剩下一条退路: 用一个全新身份逃到另一个共同体——但那意味着把一切成就抛在身后, 在某处从零、从零声誉重新开始。有时这或许是一条可以理解的路, 也是唯一的出路。

■ 说明

同样地, 面对“有人投入资源去创建一个虚拟身份”这类攻击, 共同体也会这样应对: 对那个身份而言, 未经验证就与另一个共同体发生互动——也就是不加批判地接受关于另一个共同体各身份的信息——是有风险的。声誉总是在一个共同体内部建立, 而非全球性地建立。

那些只想读取、却对共同体毫无贡献的搭便车者呢?

从创建去中心化身份的第一天起, 对信息的获取并非不受限。新参与者——那些尚未通过实际活动建立声誉的人——面临分级的限制: 信息更少、等待时间更长、查询成本更高。网络奖励的是参与, 而非被动消费与任意的数据搜刮。

一个去中心化身份, 在把自己的声誉有偿出借给另一个人 (那人甚至可能根本不在 DID 声誉网络里) 时, 也会拿它的声誉去冒险。这里适用同样的原则: 这样对共同体的背叛 (侵犯隐私) 会反映到叛徒的声誉上, 而且这一行为无法像在集中式系统里那样通过一笔见不得人的交易被抹去。他们必须面对共同体的震怒, 包括成就的丧失——因为在他们眼中, 共同体正是诸如“拥有动产和不动产之特权”的担保者。

■ 与现实世界的锚定

在评估风险时, 风险更高的主体自然是那个不享有某个共同体所承认的所有权特权的人——他们在交往中要失去的更少 (数字资产更容易转移)。

它看似是个小细节, 却有重大后果。希望共同体对其成员拥有制衡力, 就意味着把所有权当作一种特权——在最自由的社会里几乎不可触碰, 却仍然不是一项权利, 也不是一条根本原则, 而是一种在极端情况下可能被撤销的特权 (我可以设想, 比如在武装冲突中拒绝服兵役以保卫共同体)。

它也在潜移默化中回答了: 共同体将如何对待它的成员, 以及一个成员为共同体奋战的动机是什么——为了保住自己的特权。一个人可以在对共同体的责任上失职, 但在道义上不能指望, 当涉及保住来之不易的特权时, 还会得到宽待。

金融中性

当读到去中心化、不可审查、不可腐蚀这类词时, 人们不禁会把它们与最知名的加密货币联系起来——比特币、门罗币, 以及比如说 Kaspia——这些确实可以用上述词汇来描述。然而, 直觉在这里会误导人: 权威服务的费用, 验证与发布的费用等等, 都可以用任何货币或金钱来结算。对于 DID 网络中该社交网络的连接参与者 (也就是你的共同体) 及其周遭而言, 要紧的是一份以声誉背书的、“款项已支付” 的确认。发布一条主张必须个别地带有合理的、可验证的成本, 如此一来, 一个行为者就不能不耗费精力、金钱与时间, 想发多少、想发哪些主张就发多少、发哪些——那是一种极不可取的状态, 对应着当今充斥腐败的国家体制中精英的特权。

网络奖励参与

声誉靠一生累积——买不来，也借不到

参与者——一生的真实劳作

时间 能量 金钱

时间 能量 金钱

时间 能量 金钱

时间造就曲线——你自己的努力能让它更陡。

更努力的付出能缩短攀登。

20岁——新来者

声誉5%

完成项目0个

28岁——活跃参与者

声誉30%

交付合约 解决纠纷 社区贡献

40岁——贡献者

声誉70%

交付项目8个

55岁——权威

声誉95%

交付项目15个以上

受限——成本高，等待久

基本——标准成本，正常速度

完整——低成本，即时访问

权威——你验证他人，网络付你报酬

每个阶段都花费时间、能量和金钱——三者一起，而非分开。

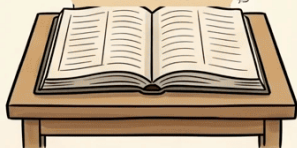
随着你事迹的账本被填满，门便打开。

搭便车者——只读，从不发布



读取别人所有的记录。

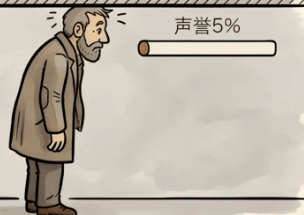
我的贡献



自己什么也不发布。

受限访问——天花板不会抬升

（网络回馈那些投入的人。）



声誉5%

没有事迹，就没有上升——年复一年，同一条杠。



时间、能量和金钱——投入真实劳作——是网络唯一尊重的货币。

在这一点上,前面提到的那些加密货币有一个小小的优势:它们的网络充当可信权威,用以验证某笔支付确实发生过,代价是损失一点隐私,并暴露自己的部分地址。

自愿、责任与自由

自愿是个体性的:我为自己作决定。责任在我的决定影响到他人时产生——反馈随之启动。自由既非馈赠,也非权利——它是自愿与朝向责任的压力之间成千上万次双边互动、在全社会范围内取平均后的一个现象学³结果。网络并不规定这个平均值——它从验证者的价格信号中,以及从共同体成员为自己设定的、预先公告的验证政策中涌现⁴出来。

一个共同体内的个体对自由有多大的偏好,取决于他对“自己愿意吞下多少加诸他人之物”的感受。对于那些不言自明、对所有人都有益的事情,并不存在道德困境——谋杀是错的。接下来是一些选项滑入“被许可的特权”这一位置的情形:证明你的技能,我们才让你上路,这样你把我们撞死的机会就小些。最后还有一些事情,个体会说——这对我不值得,我不会禁止邻居烧树叶,因为我自己也烧,或者因为我做着别的限制他人舒适的事情,而他们反过来也可能想禁止我。如果历史上我们把自由的边界画在“他人自由开始的地方”,那么一个建立在去中心化身份之上的声誉社交网络,在这一点上比集中式代议民主做得更好——在代议民主里,比如民选代表这一特权阶层试图让自己免于对个人通信的监控,却毫不犹豫地把这种监控强加给他们的选民,借助形形色色见不得人的借口与操纵。

³现象学的 (Phenomenological, 源自希腊文 *phainómenon*——显现之物)——一般而言:一种考察现象的取向,通过观察从现象中得出什么,把现象当作它们在直接经验中呈现的样子来考察,而不预设关于底下微观世界的解释性理论。这里:自由不是一条自上而下界定的抽象原则,而是一个被观察到的现象——是人与人之间成千上万次微观互动的结果,从我们实际如何彼此相待中涌现出来。

⁴涌现的 (Emergent, 源自拉丁文 *emergere*——浮现) 意味着“从更简单部分的互动中自发地产生,而无人设计或指挥”。一群鸟成队形飞行,尽管没有任何一只鸟对队形有计划——队形从每一只鸟简单的规则中涌现出来。同样地,在声誉网络里,没有人设计整体的行为规则——它们从成千上万个个体决策中涌现出来。

自愿 → 责任 → 自由

三步——从一个人的选择到整个社会的模式

自愿性



个人做决定——没有人强迫她。
也没有人欠她一份积极的义务。

责任



我的选择向外荡开涟漪——涟漪又荡回来。
后果自己说话。

自由



没有人画出这张地图——它自己画出了自己。
自由是邻居们彼此相待方式的统计形态。

自由不是自上而下的恩赐——它是我们共同造就的模式。

自由—极权开关 (以及委托)

DID 网络内部一个简单的机制, 就能把整个体系推向更大的自由或更大的控制: 一个 DID 在每一段时间内可以发布的记录数量。当记录保持稀缺时, 每一条都分量十足——那就是自由。当记录泛滥、连琐事都覆盖时, 那就是极权。你用每一条发布的记录来选择方向。网络本身既非自由亦非极权——它是中性的基础设施。它变成什么, 由人们如何使用它来决定。

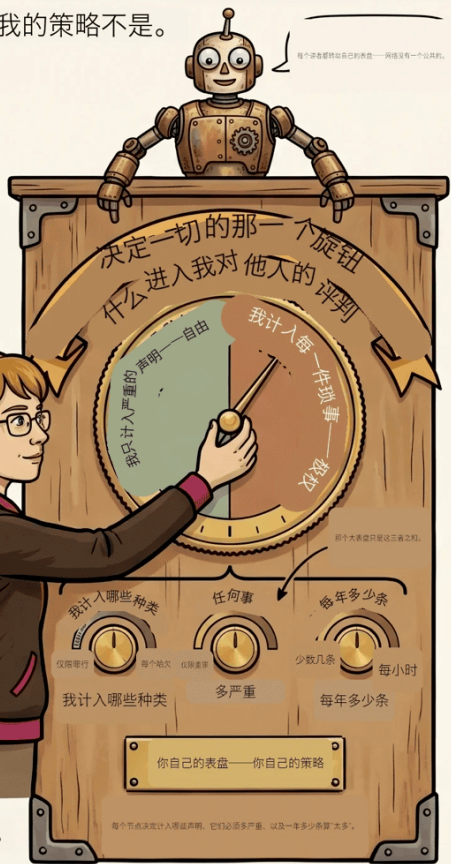
在我自己的验证政策里, 为我社交网络中——在各共同体内——的一个泛用 DID 设定一个上限是合理的: 发布者在每段时期内可以发行的主张 (泛指任何可验证的陈述; 这里指: 发布到声誉网络的一条记录) 数量, 超过这个上限, 我就不会把他们纳入考量, 也不会认为他们的验证请求相关。

我心里揣着一个直觉性的数字——每年个位数的低值。这样一种审查式的做法, 或许显得相当孤立主义。但我可以正当地把申请人置于一个选择面前: 从他们已经发行的记录中挑一条, 请验证者撤销它, 即令其失效——在一种申请人恰巧真的成了多起相互独立之冤情的受害者的处境下, 这是一个相当咄咄逼人的要求, 而我这样做无疑会加深他们与共同体疏离的感受。

因此, 一个更有效的选项似乎是这样一种能力: 把自己“向声誉网络发布信息”的部分权利委托给一个共同体或一个权威, 由后者随后——以团结互助的方式, 按其内部规则——决定如何分配这些权利。一个群体必须能够照顾它认为处于弱势的人。共同体成员有权在任何时候, 只要被委托的权利尚未被使用, 就撤销自己的委托。

自由—极权切换开关

仅仅一个参数——我选择追踪多少——就能让整个社会在自由与暴政之间倾覆
按钮不多——只有一个。网络本身是中立的；我的策略不是。



只有严重的事才被刻进记录——因为刻录是艰难的。
其余一切保持私密：糟糕的一天、一个笑话、一次迟到——都不是谁的档案。
人们不受监视地生活、自由交谈、抬头呼吸而无须提心吊胆。

每个哈欠都成了档案——因为计数很便宜。
一个喷嚏、一次购物、一趟天黑后回家：全都进了记录，永久保存。
没有隐私——那只眼睛从不眨，邻居们不再开口。

自由是我拒绝去计的——极权是我同意去追踪的。

发布权的委托

你把密钥借给一个为你服务的权威机构——并能一瞬间把它收回

我借出密钥——线绳还在我手里。

委托——权威机构代我行事，在我批准的规则之内。

召回的线绳

我的发布权——1条记录

我可以一次借出不止一把密钥——但每次借出密钥都在我的授权密钥池中随机抽取一把密钥。

权威机构——服务提供方
策略·标准·党派·议院

策略

权威机构行事时随机抽取

已委托的密钥——池

权威机构行事时随机抽取

轻轻一拽——密钥就回家

线绳从不离开你的手——这就是全部诀窍。



仅在密钥被抽中前可撤销。

能持有密钥的权威机构种类



政党
PARTY



标准机构
BUDY



专业行会

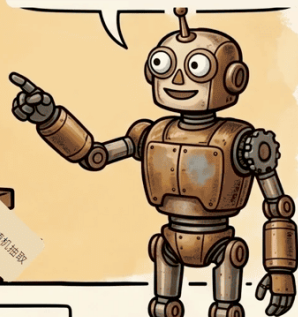


合作社

此后，公民必须签发一把新密钥才能维持与权威机构的关系——自动程序可在后台完成这件事。

权威机构服务其成员——而非反过来。
拽一下线绳就终结这个安排。

借出密钥——留住线绳。委托绝不是交权。

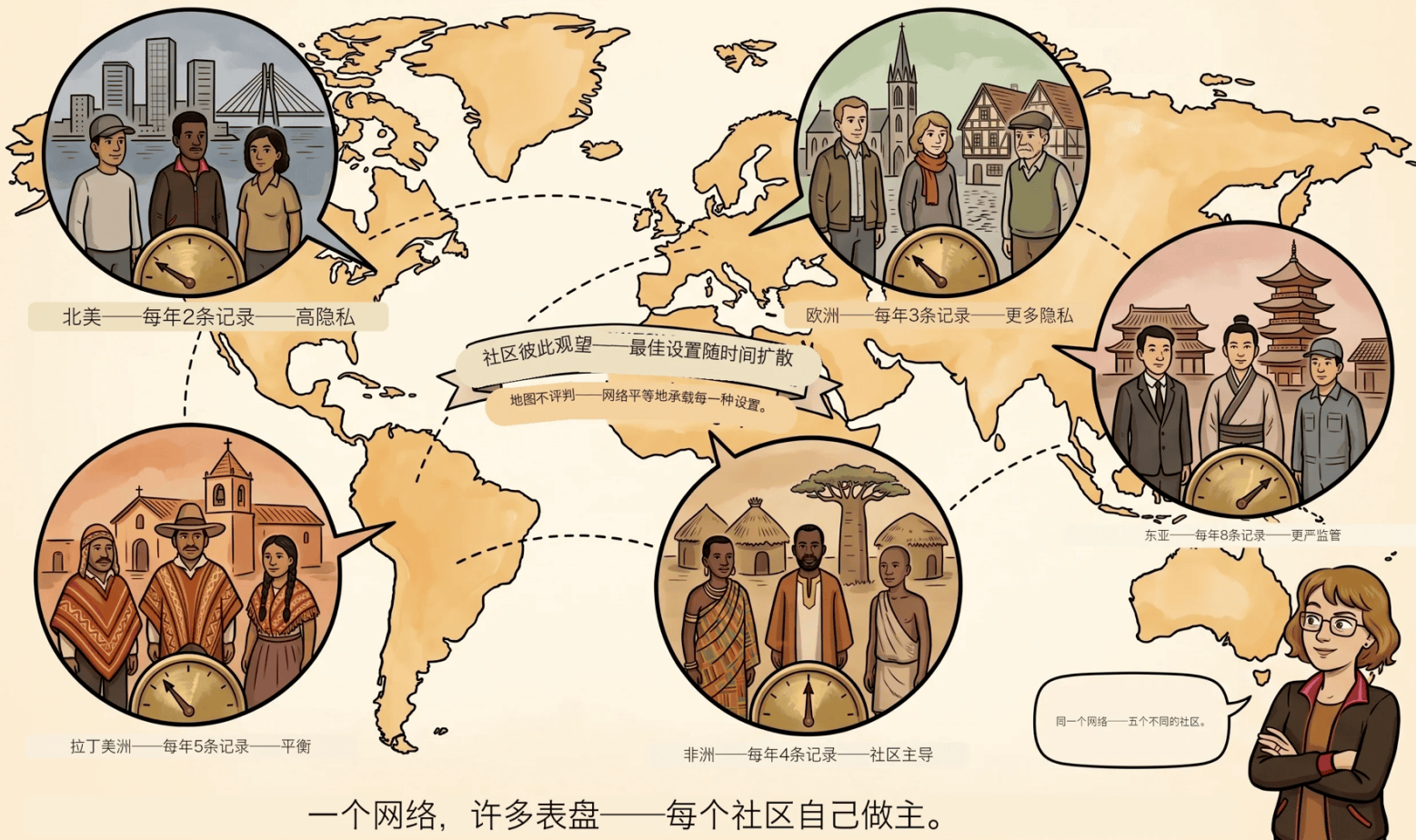


■ 自由与极权

一个看守并限制所发布记录数量的社会会更自由; 一个记录过剩的社会则倾向于走向极权。人们可以设想把发布记录的权利委托给别人。

每个社区自我配置

网络不歧视——每个地区设定自己的表盘



■ 文化上的适应

人们大体上都喜欢看进彼此的盘子里。在各文明之间, 我们在历史经验和人口密度上有所不同——因此在“我们想在各文明和共同体内部多大程度上看进彼此的盘子”这件事上也有所不同。亚洲国家倾向于对社会中的自由持更严格的看法; 欧洲和美洲国家在某些方面则倾向于更自由的看法。无论这一类比多么不精确, 在去中心化声誉社交网络的语境下, 可以说: 各个共同体的设置并不歧视任何人, 每个共同体都可以按自己认为合适的方式来配置其内部的通信。

尽管存在差异, 仍会有一些被全球范围内公认的标准(随机杀人对谁都没好处)。其他设置将是共同体特有的, 还有一些甚至是个人特有的。其优势在于: 通过相互沟通, 并通过观察周遭的设置, 可以寻求共识——从双边安排, 到共同体层面的安排, 再随时间推移, 凭借配置最佳的共同体的经济主导地位, 乃至在全球层面达成共识。

但关于共识, 留待下一章再谈。

预言机问题——架起数字世界与物理世界的桥梁

“预言机问题”(oracle problem) 一词近来最常与加密货币、区块链和去中心化系统的世界相联系。一般而言: 你如何确保进入数字系统的数据忠实地对应于物理世界里实际发生的事情? 系统本身无法核实某人是否真的伤害了另一个人, 或某份合同是否真的被违反了——它只看得到数字记录。

这里我们通过权威来解决这个问题——权威是把自己的声誉押在“数字记录对应于物理现实”之上的实体。权威的声誉越高, 记录就越可信。

沟通数字世界与物理世界

权威机构把现实带过去



权威机构的声誉让数字记录变得可信。

* 系统不试图解决预言机问题——它接受权威机构就是那座桥，并与现实世界中信任本已运作的方式相契合。

* 若桥对现实判断有误，权威机构付出代价——用它的声誉。

How Verification Works

共识与验证过程

为了就“社会平均而言应当维护并执行哪些规则”建立共识, 下面这个机制可以帮上忙。作为一个 DID 参与者, 我宣告我所订阅、并将据以生活的规则, 并把它们发布出来。(可以把它想象成在我看来构成我理想世界的那套章程与规约——一个我不感到受限、而是感到安全的世界。)

我可以预先估计我的 DID 联系人会作何反应——并评估, 若那些寻常的社交或商业互动假设性地发生, 我会以多大力度、被谁施加制裁。

真正决定性的评估, 发生在你向另一个 DID 请求信息、或请它验证一条主张 (或向某个权威请求一项服务, 等等)——你想把这条主张发布到声誉网络时。它应当与你自己在演练中、对照对方所宣告政策所跑出的结果一致——如果不一致, 那么对方那边就有问题: 他们在试图玩一场不诚实的游戏。

结果要么是接受, 并附上一个验证报价 (在验证者或权威服务的情形下), 要么是拒绝。对“偏离评估者政策”的制裁与奖励, 都被折入报价之中。请求者随后决定: 是接受这些条件, 还是转入分配算法中的下一轮验证——不断重复这个过程, 直到满意, 或者直到经济上继续下去已无意义为止。

■ 社交图谱

声誉网络首先是一个社交网络。你添加联系人——那些同意建立连接的人。他们有联系人, 那些联系人又有联系人。算法在一个可配置的深度内 (例如三层: 你的直接联系人、他们的联系人, 再往外一层) 搜索验证者。不需要全球区块链——网络自然形成一个个共同体, 并与其他共同体有所重叠。

算法是非确定性的: 它对你的主张文档求哈希, 把哈希映射到这个圈子内已知身份构成的环上的一个位置, 并选出最近的一个作为候选验证者。你无法预测或影响谁将来验证你的主张。

每个验证者的拒绝都会扩大你的文档, 并抬高它的处理成本——这是第一条成本通道 (文档增长)。每个新的验证者都会根据数据量、你的声誉, 以及你主张的内容偏离其所宣告验证政策的程度来收费——这是第二条成本通道 (风险溢价)。而每一次迭代都花费时间和精力——这是第三条成本通道。

■ 验证者依次检查什么

一旦被选中, 验证者会大致按四个有序步骤评估一条主张——最便宜的过滤在先, 昂贵的内容检查在后:

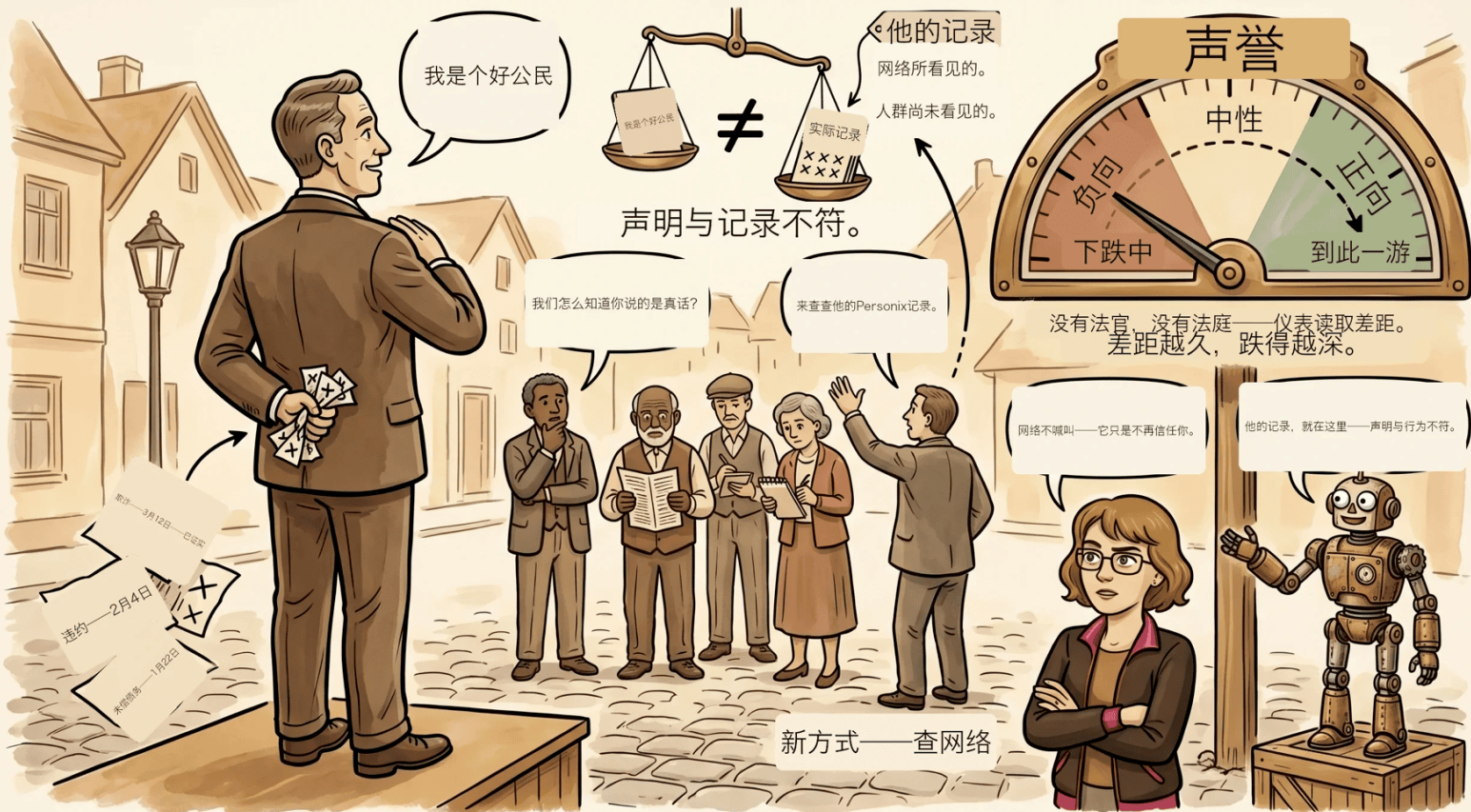
1. **政策门槛。**这类主张是否落在验证者公开会验证的范围之内? 若不在, 请求当即被拒。
2. **权威信任。**为该主张背书的权威, 在验证者自己所宣告的政策之下是否足够可信? 一个低于验证者信任阈值的权威, 无论主张内容如何都构成拒绝的理由。
3. **发行者声誉。**发行者是否达到验证者为此类主张所宣告的声誉阈值? 低声誉可能要么抬高费用, 要么触发拒绝。
4. **内容检查。**只有当前三道门都通过时, 验证者才评估主张本身——签名、内部一致性、形式正确性, 以及它偏离验证者政策的程度。为这最后一步所收取的费用, 反映了实际承担的风险。

验证者会发布规范上述每一道门的政策, 因此这些步骤并非任由其自由裁量——它们受其已经宣告之内容的约束。偏离所发布的政策, 本身就是一条可以针对他们发布的主张, 而他们要用自己的声誉为此埋单。

结果是: 发布一条可信而有用的主张, 几乎不花什么钱。发布一条激进的主张, 花费更多。发布一个谎言, 则变得贵得令人却步——你必须一个接一个地迭代过验证者, 而每一个拒绝你的人都会增加成本。市场为你的主张定价, 而这个价格告诉你, 你相对于你所身处的共同体处在什么位置。

伪善是最昂贵的行为

网络悄悄把你的言辞与你的记录对照——差距人人可见



仅仅宣告你遵守某条规则、而实际上并不遵守,是不够的。在那种情况下,你的 DID 就冒着被发布一条揭露伪善的负面记录的风险——这会把你变成对所有其他人的一个风险。其结果应当是:规则更少,但被更一贯地遵守,并清理掉那片连法律专业人士都几乎无从穿行的法律与法规的丛林。

■ 共识与问责

为了让网络充当有价值的信息源,一个 DID 不应过于激进——否则其他人会拒绝它。社会压力会寻求均衡,而试图破坏这一均衡的行为很可能会受到惩罚。

伪善的解药

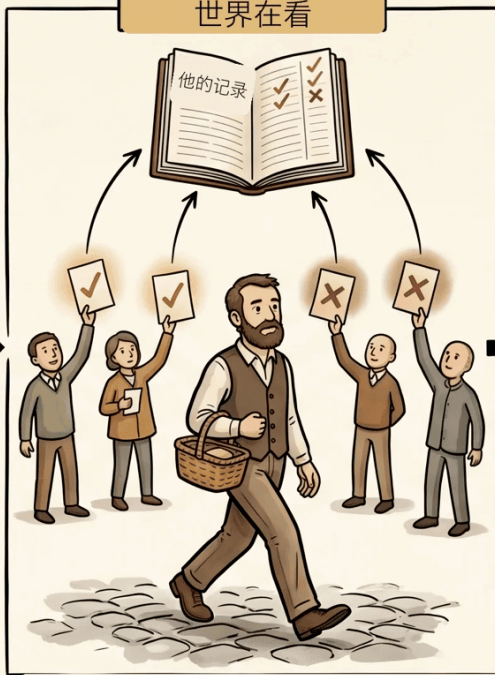
公开你的规则——照它们生活，否则市场会把你定价出局

我公开我的策略



人人都能读到我的标准。
没有看门人，没有秘密。

世界在看



世界写下他实际做了什么。
每一天，他账本上的证据都多一点。

在店里——你为自己的记录买单



伪善不由权威机构惩罚——它由市场定价。
你买的每一项普通服务，都带着一份由你自己记录塑造的附加费。

我们能否把这称作社会契约——被重新绘制？

没有权威机构强推这套伦理——是价格在推。

温和的压力，日复一日，朝向我们都能接受共同平均。

言行一致——否则为它买单。伪善是最昂贵的声誉。

■ 票数不等于一个声音的分量

Juraj Karpiš 说“金钱是善行的记忆”。我要补充一句：声誉是恶行的记忆。

由此得出，在贤能政治的意义上，贡献更多且没有坏声誉的人，应当在共同体中拥有更大的声音分量。透过双边关系的视角来看：当我权衡该迁就哪些共识压力时，最大的分量归于那些让我获得最大经济利益的关系。十个我与之没有活跃交易往来的人，对我的影响远不如一个长期的生意伙伴。这一范式并不限于商业——它延伸到社会、政治及其他关系。

Roles in the verification transaction and authorities

角色总览

在关于网络及其基本属性的那一章里，我们已经简要提到过其中一些角色。现在是时候更细致地再看一遍，并加入我们为使网络更稳健所需要的其他角色。每一笔验证交易都涉及若干角色——让我们看看它们如何行事。

■ 验证交易中的角色

每一次验证最多涉及六个不同的角色, 汇总于下表。它们全都可以在去中心化声誉网络中拥有自己的 DID。

角色	描述
发行者 (Issuer)	把信息发布到网络的人——主张某事发生了 (创建、编辑或解散了一个 DID, 一条主张, 某个 DID 的政策, 等等)
主体 (Subject)	信息所关于的人——主张的对象
权威 (Authority)	一个可信实体, 通过调查主张、审阅所出示的证据或主动收集证据, 把自己的名声押在主张的质量上
观察者 (Observer)	一个独立的第三方, 记录验证者如何处理主张——确保验证者既不保持沉默, 也不偏离其所宣告的政策
验证者 (Verifier)	一个由算法选出、处理该交易的参与者
受托人 (Delegate)	代表另一个参与者行事的人

发行者

我们此前见过发行者。当有人受了委屈, 或想广而告之自己的一项成就时, 正是由 DID 所代表的那个人来触发“把这份信息投入网络”的过程——因为他们对此有首要的利益 (动机要么是渴望获得满足感, 要么是希望改善自己的公众形象)。

■ 正面主张

关于我自己的正面主张, 我可以作为它们的主体兼发行者留存下来, 以备我想在其中使用它们的场合 (一纸文凭、一份已提供服务的证明, 等等)。这正是正面主张与负面主张的不同之处: 分享一条经过验证的负面主张符合整个共同体的利益, 而正面主张则符合其承载者的利益。

权威

权威扮演着双重角色: 它可以是一名**审计者** (在主张被发布之前核验证据的质量), 也可以是一名**担保者** (把自己的声誉押在主张的真实性上)。无论哪种情形, 它都强化了发行者的主张。这两项服务是可以分开的——一个权威可以只提供其中一项、另一项, 或同时提供两项。工作假设是: 权威所提供的大多数服务都可以在自由市场的基础上交付。这一点即便在那些难以想象能被私有化的领域也成立, 比如司法——在那里, 专门化的服务 (调查、证据评估, 乃至今天由集中式军队提供的服务: 战略规划、标准化训练、采购与库存管理等) 都可以由市场行为者高效地交付。几乎没有什么东西是经过重构之后, 无法借助自由市场的激励而变得更高效的。

签发者——写下声明的人

作者 · 签署者 · 押上自己的声誉

签发者做什么

1. 用自己的话写下声明
2. 收集现实世界的证据来支撑它
3. 用自己的DID签署，并送交选定的权威机构
4. 支付发布成本
5. 若撒谎，承受后果

我的声誉



如果这是假的，我的指针会下跌。

我的名字押在这上面。我最好确信无疑——并带上证据。

人人都能开口。谎言让说话者付出代价。



签发者是作者。名字署在上面。

* 任何持有DID的人都能当签发者。门槛是勇气与担当，不是执照。

■ 权威、发行者和观察者绝不能成为自己案件的验证者。

对验证者的算法化选择保证了独立性。没有人能验证自己的主张, 或一条自己在其中有直接利益的主张。这是整个 DID 共同体都有利益去维护的基本规则之一。

下面的图形展示了对权威所覆盖之活动广度的几种互补视角 (术语“权威”可以与“服务提供者”互换阅读)。

向导——帮你按自己的价值观导航

背后是深度分析 · 前台是简单指引

把复杂的世界汇总成世界。遵循着注意力投票——随时可离开，转投他人。

向导

我的分析

工作笔记——不给客户看

公共开支——已重新导向

复杂的话我来做。你得到简单的推荐——但仅当你认同我的价值观。

提供者

1	70M	☆☆☆☆
2	70M	☆☆☆☆
3	70M	☆☆☆☆
4	70M	☆☆☆☆

给你——它意味着什么
经过汇总，使你的收益最大化
把开支挪到这里，在那里省下
支持通过我测试的提供者
采用这个策略模板



那熟悉的钱到底来自哪里——哪些分析和哪些信任？

例如：谁可以当向导

- 一个发布策略模板的政党或运动
- 一家发布汇总结论的独立审计公司
- 一家服务比价机构
- 一个声誉聚合方（经济、商业、邻里、道路安全）
- 一位记者评论员或消费者评价服务

受其公开立场的约束。你的信任是他们的货币。

* 你随时可以离开，转投另一位向导。你的信任并未被锁定。

* 向导也可能签署一些自己的声明——已发布，但不是这里的主线。

■ 作为隐身观察者的权威

一个有声望的权威——想想一位其生意全靠自己履历的公证人——除了主要职能 (审计者/担保者) 之外, 还可以提供第三种职能: 在验证过程中充当隐身观察者的角色。他们对所提交的主张保存一份带时间戳的记录, 使得验证者无法悄悄把它丢掉。观察者角色的机制, 在关于观察者角色的那一节里有进一步的描述。

验证者

任何 DID 都可以充当验证者, 或直接充当, 或通过委托给第三个 DID 的验证权来充当。为了让我——或我的受托人——能够进行验证, 我应当在网络上可达 (在线)。并非每个人都愿意作此承诺, 这正是为什么一份 DID 记录可以按优先序列出替补者, 由他们在该 DID 离线时代其履行这一职能。

网络中每一个活跃的 DID 都会公开宣告自己的政策。它通过该政策中定义的规则, 在验证过程中判断对方的声誉, 以及发行者已标记要发布到声誉网络的那条主张的内容与形式。政策的一部分是用来计算验证服务费用的计算公式。这一切就位之后, 在流经网络的、统计上大量的主张中, 我便等待网络的算法把我抽到发行者那一侧, 在某次迭代中指派我去验证正在发行的信息。发行者可以预先计算出一个行为正确的验证者会如何反应, 却无法避免实际去联系他们 (或他们的替补者); 即便发行者事先知道通不过, 与被选中验证者的这次迭代也必须由发行者执行。

我们怎么知道发行者是在正确的一组候选验证者 DID 上运行验证者选择算法的? 每个 DID 除了公开宣告自己的政策之外, 也会发布它在声誉网络中社交网络的当前标识符清单。如果一个发行者把自己的社交网络定义成一个只是回响并强化自身观点的社交气泡, 那么通过它发布的信息几乎不会被其他共同体更广泛地接收。我不惜高昂代价把一条激进主张推入网络这一事实, 并不意味着在判断对方声誉时, 我会给它任

提供者——运营一项声明签署服务

公开的标准 · 公开的价格

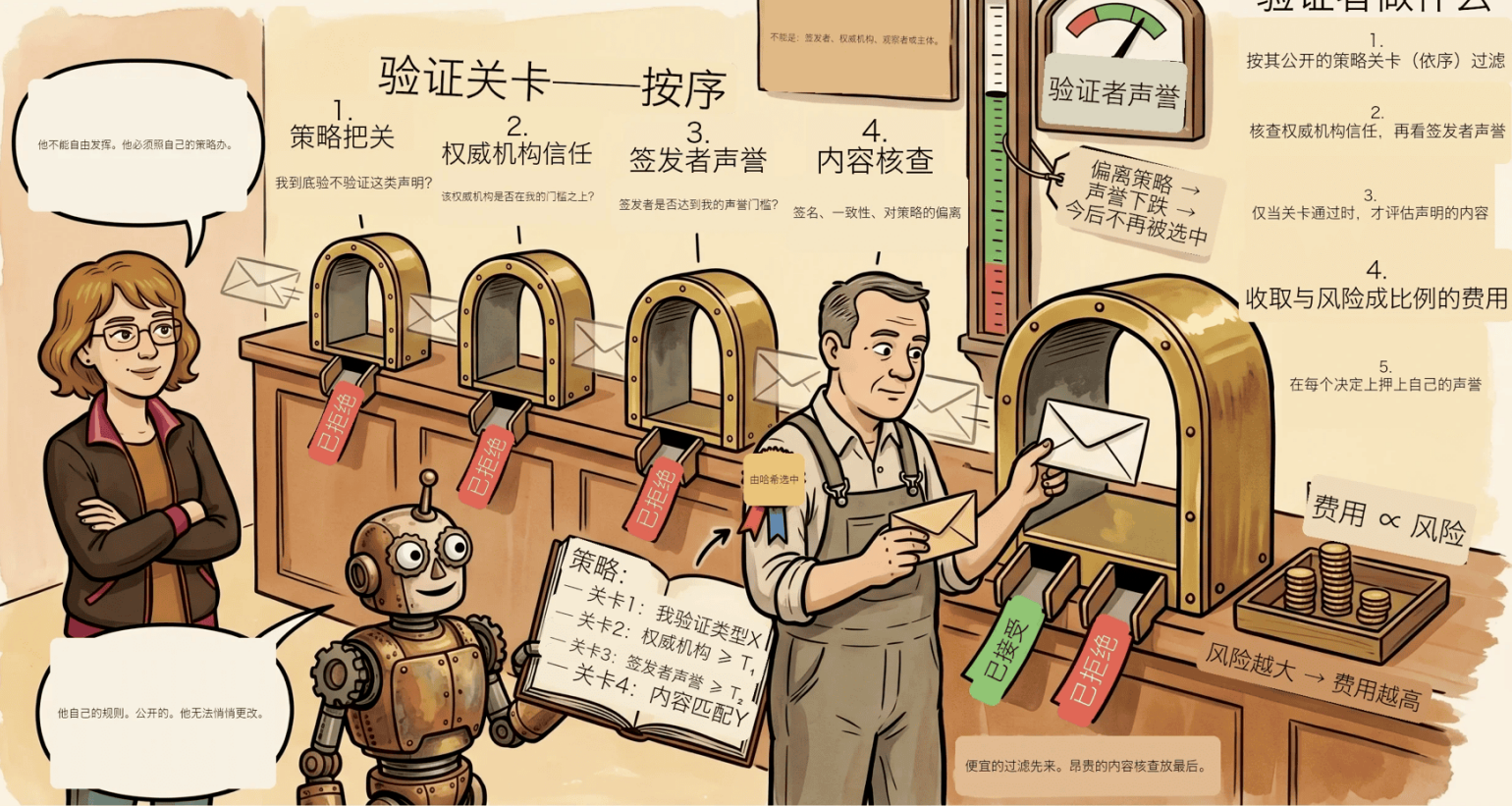


受其公开宣称的约束。其余交给声誉。

* 提供者也有自己的策略——何时接受、何时拒绝、如何退款。它是公开的，但不是这里的主线。

验证者——对照公开策略评判声明

四道关卡，按序进行。便宜的过滤先来。



验证者受他自己发布的规则约束——也受他将要失去的声誉约束。

何分量。有些主张是我的共同体推着我去纳入考量的 (对冒犯者施加的判罚与限制); 另一些则完全由我作主——我自己决定纳入或排除某一条信息的经济价值。

主体

那些被发布到声誉网络的信息所关涉的人, 就是信息的承载者, 或者你也可以说, 是信息的主体。声誉网络中的激励机制被设置成: 主体有充分的理由按自己的规则正确而公正地行事, 并被共同体推向某种恪守规则的黄金中道。他们甚至有动机表现得比其所宣告的规则所要求的更合乎道德——预先宣告的政策并不排除共同体就“追溯性地施加某种评价或措施”达成共识。从这个角度看, 声誉网络对法律秩序的替代方案, 不出所料地, 更多由规则的精神所驱动, 而非其字面文本。

主体——声明所关于的那个人

接收它 · 无法压制它 · 可以回应它



主体没有否决权——只有回应的权利。

* 反声明本身也是一条声明——主体要为回应中所说的话押上自己的声誉。

观察者

观察者角色消除了验证者钻规则空子的动机。在验证者不喜欢发行者或权威的请求的情形下, 他们本可以干脆保持沉默——不作回应, 阻塞算法序列。观察者——或一组观察者——把自己的声誉押在“记录验证者是如何被问询的”这件事上。如果验证者尽管有一条相反的宣告政策却依然沉默, 他们就可以被判定违反了协议。

机制: 时间戳与挑战码

在你把一条主张发送给验证者之前, 你先把它经由观察者转发——那些你信任的人, 或收取少量费用的专门的观察者服务提供者。每个观察者收到你的提交, 给它打上时间戳, 签署“他们目睹了它发出去”, 并生成一个挑战码——一个对其签名的密码学哈希。这些码被附加到你的请求上。验证者看得到它们, 却完全不知道观察者是谁, 也不知道这些码是不是真的。观察者由此充当发行者与验证者之间的代理, 持有一份“主张已被提交、以及它包含什么内容”的独立记录。他们的数量可以是零到 N 个。

当验证者诚实行事时——按其宣告政策接受或拒绝——这些码保持不透明。没有人被暴露。

但如果验证者尽管有一条通融的政策却保持沉默, 或以一种与其所发布内容相矛盾的方式回应, 那么你手握原始的观察者签名。你可以把它们作为“主张已被提交、且验证者没有遵循协议”的代理证词发布出来。任何人都能核实这些签名与挑战码相匹配。

妙处: 你并不需要真的观察者

而这里是最精妙的部分: **你根本不需要真的观察者**。你可以生成看上去与挑战码一模一样的随机数。验证者分辨不出差别——他们不得不掷骰子, 决定要不要拿自己的声誉去冒险。在他们收到的每一个请求背后, 都可能有一位受人尊敬的观察者在隐身注视——也可能纯粹是噪声。验证者不知道。而正是这份不确定性, 就是这套机制。

维持诚实压力的成本: 几乎为零 (随机数是免费的)。不诚实对验证者的潜在成本: 灾难性的。即便实际上没有人在看, 诚实的行为也被激励着。

这套体系之所以奏效, 是因为每个人都有点疑神疑鬼。不确定性比监控更便宜。

■ 单次迭代中的多个验证者

一条能强化验证者可用性的配套规则, 可以是一种算法扩展: 在单次迭代中返回一组候选验证者, 而不只是一个。

委托

文明是把它的经济增长建立在分工与专业化的能力之上的。这个世界在其全部辉煌之中是复杂的, 而我们把握它全部复杂性的机会近乎于零。自然的出路, 是把一大批支撑舒适生活的活动委托给专门化的提供者, 他们为我们节省时间、精力和金钱, 并让我们反过来去专业化——去提供我们最擅长、或市场在我们身上最看重的那些服务。

同样的委托逻辑延伸进了声誉网络。在“自由—极权开关 (以及委托)”那一章里, 我们阐述了通过“限制在给定时间窗口内可以发布进网络的主张数量”来保障自由社会的原则 (简化地想象成每年个位数的低值发布权, 好让社会保持最大程度的自由)。这一原则可以扩展到多个维度。我们不必只有一个数字——我们可以有若干上限, 分别适用于不同类型的主张。给“揭发他人伪善”用一个数字, 给“生意往来中的不当行为”用另一个数字, 而比方说, 像改密码这样的 DID 文档变更, 共同体大概根本不会在意, 如此等等。这一切取决于共同体在演化的时间里如何定义: 什么才符合它在“目标自由水平与问责水平之间找到恰当平衡”上的利益。

让验证者保持诚实的那个虚张声势

从验证者的桌前看，一个真实的观察者和一次随机的虚张声势一模一样

签发者——卡米拉

真实观察者

验证者



验证者分不清真实记录与虚张声势。诚实行事是唯一的稳妥之选。

签发者——虚张声势

没有观察者

验证者——同样的行为



为什么共同体授予我的发布权数量应当足够小? 给一个 DID(主体) 贴上一条负面主张, 会摧毁花了很长时间才建立起来的声誉, 而且损害会持续。一个被摧毁的声誉, 面对的是共同体对每一次进一步或重复冒犯所施加的、层层升级的制裁, 一路直到把这个捣乱的实体彻底清除 (极端情形)。一套指数式升级的惩罚阶梯, 并不需要大量的主张, 只要它们有可验证的证据背书即可。这在和平的稳态下成立, 即社会没有被 (外部或内部的) 灾难所撼动之时。

对于概率极低的事件, 把统计频率套用到单个个体身上时是易变的。放到共同体的人口上则不然。这正是“把自己的发布权委托给专门的汇集型权威”的价值所在——它们很像损失发生时的相互保险, 替你使用你的权利并代你行使。让我们少说些抽象的, 举几个例子: - 侦测和打击有组织犯罪 - 强制执行商业协议 - 管理各种特权 (教育证明, 资产的取得/处置, 等等) - 如此等等

代理者——你订阅的一项声明服务

把你自己搞不定的委托出去 · 留下对你重要的



把你自己搞不定的外包出去。保持知情。随时把权利收回。
* 若代理者越出约定范围行事，他们的声誉和你的都会被公开波及。

■ “发布权”实际上如何运作

受托人的那张图为了清晰起见,把发布权画成了实体代币,但在系统里并不存在这样的物件。实际发生的是这样:你想与之保持良好关系的那些人——你的共同体——的政策,汇集成一条涌现的规则,规定他们在开始把你当作噪声之前,每年会接受你多少条被视为相关的主张。设置这个上限,是为了不让网络被琐碎的“假如”主张淹没,并让它保持共同体所希望的那般自由(通过限制数量来保存信号)。

你读取你的共同体所允许你读取的东西。这里有一个区间,并随你最看重哪些人而有些许浮动。然后你在那份额度之内决定——自己发布多少条主张,又把多少委托给专业服务,由它们仅在事件落入其职责范围时代你发布。正是这一点让你保持良好关系,而不必为发生在你身上的每一件小事都提交一条主张。

图中的“代币”是这个“读取并决定”循环的视觉替身,而非一个字面意义上的实物。

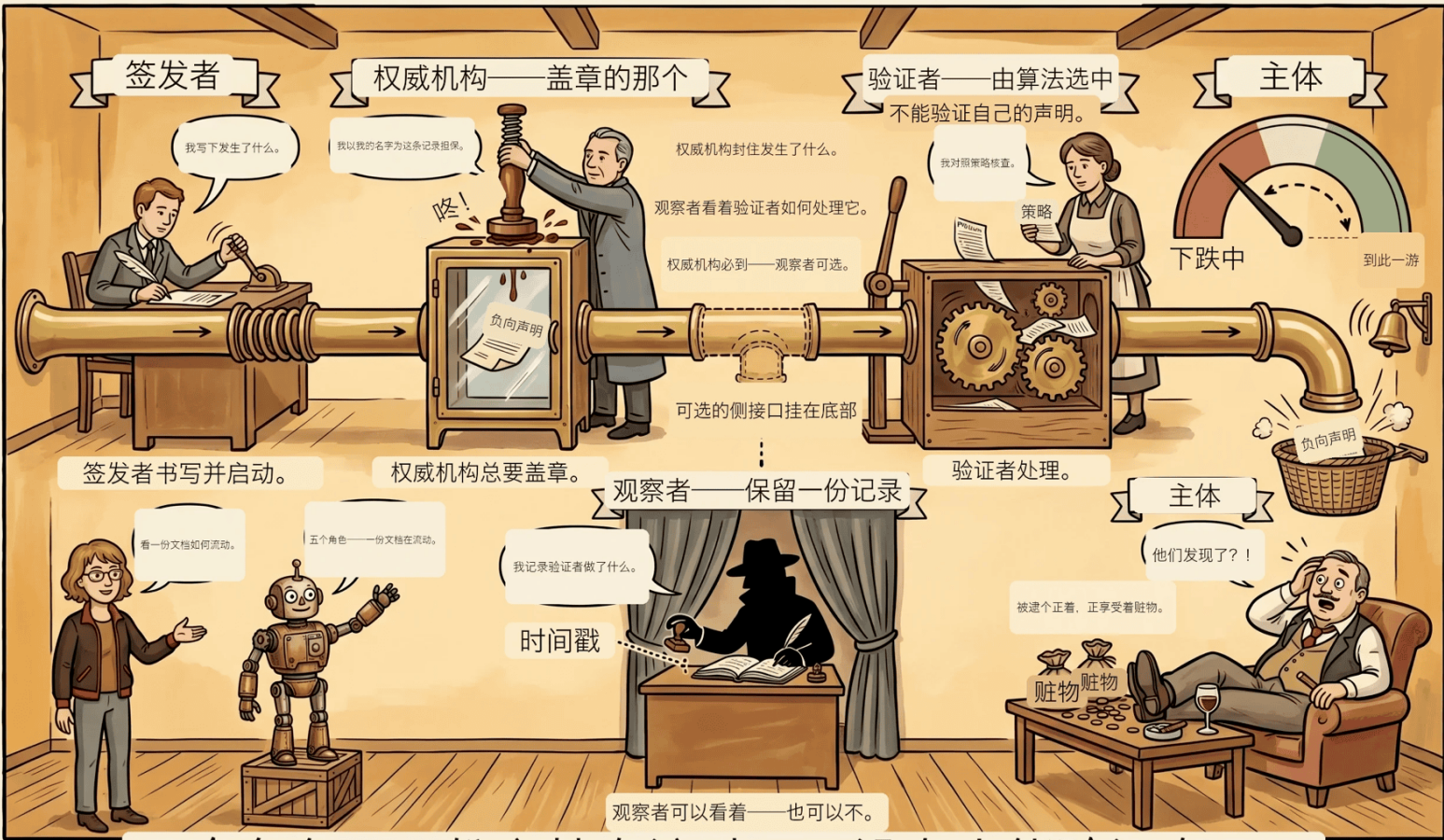
在声誉网络中,能够撤销已委托的发布权是值得的。只要受托人尚未使用被委托的权利,这一撤销随时都可以进行。

角色之间的通信

下面的图形总结了在一个真实案例中——从发行者提交一条主张,直到发布进声誉网络——各角色是如何相互连接的。

一条负向声明流经系统

跟随一份文档——看每个角色对它做了什么



五个角色，一份文档在流动——没有人能验证自己。

涌现的社会契约——政策

到目前为止, 我们描述了主张如何被验证, 以及每个角色各扮演什么部分。现在我们转向: 成千上万个个体政策如何变成某种像真正的社会契约那样运作的东西。绝不是那种左翼关于社会契约和强制团结的梦想。

每个参与者都基于自己的道德观宣告自己的政策 (policy)——他们认为什么是可接受的, 以及他们如何回应他人的具体行为。某个 DID 所知晓的、网络上所有政策的总和 (即共同体), 产生出涌现的伦理——一些无人设计、却可被观察到的规范。这些涌现的规范, 由双边价格信号支撑, 构成一份可测量的社会契约: 不是一个无人在出生时签署过的理论建构, 而是有实际行为背书、经验上可观察的规则。

■ 社会契约

人们争论“社会契约”是不是一个自相矛盾的说法。没有人签署任何东西, 然而它却被拿来为集中式社会中的立法与行政行为辩护。

对去中心化网络中互动的一次现象学分析, 或许能揭示某些涌现的规则——全球性的或局部性的。这些被广泛共享的规范, 我们该称之为? 一份社会契约吗? 正如共同体本身一样, 一个共同体的社会契约无法被精确地观察和勾勒——每一个观察共同体的角度, 都会得出一份不同的社会契约。

策略相遇。态度形成。

公开的规则定下今天的价格——契合有折扣，分歧有罚金

当策略契合

当策略分歧



当两套策略在交易中相遇，它们的契合度就成了价格调节因子。

* 这与声誉无关。它关乎每个人公开宣称了什么——如今被期待去守住。

■ 基于价值观的政策

把“基于价值观的政策”当作决策的基础, 同样具有误导性。价值观更像是一些在给定期普遍被偏好的规则的现象学聚类。它们会演变, 并且不会随时间恒定不变。

伦理如何从个人策略中涌现

从策略到规范

无暴力犯罪——99%认同

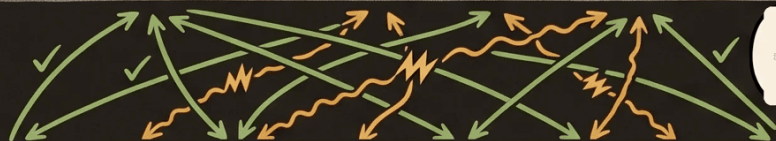
无欺诈——78%认同

信守合约——60%认同

公平工资——25%认同

越接近直接伤害，认同越强

涌现的规范——多数人所认同的，而非任何人写下的

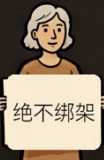


双边互动——策略相遇之处

没有人写下这些规范。它们从我们彼此实际相待的方式中涌现。



对暴力犯罪零容忍



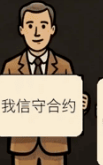
绝不绑架



我平等且诚实的做买卖



不行贿



我信守合约



我在30天内解决纠纷

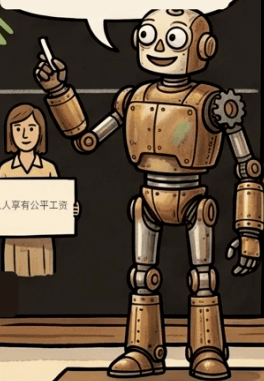


保护知识产权



人人享有公平工资

个人策略——每个人所坚守的



传统社会契约

没有人签署过这个

作为对比——

上面涌现的规范是经验上可观察的

伦理不是被设计出来的。它们从我们彼此实际相待的方式中涌现。

* 每条个人策略都是那个人道德观的映照——由适应其社区与环境的压力所塑造。

提供关于你自己的信息

当两个彼此不认识的 DID 主体进入一段经济关系时, 我们可以代表对方向我们自己的共同体签署一份请求, 请它提供关于我们的信息。然而, 这是一种对“共同体是什么”非常天真的想法。共同体不是封闭的集合——单个 DID 可以是多个共同体的参与者。

从某个特定参与者的视角看, 共同体更像是一种现象学的现象, 显现在他们的社交网络之中; 它并不是那种总是值得被明确勾勒出来的东西。一个村庄的居民可以构成一个共同体, 但一个只在周末来这里的人, 也可能在同一个村庄里属于另一个共同体。一个共同体也可以是一支足球队, 连同它的球迷、职员以及所有相关人员的家庭。一个民族, 或一个国家的公民, 同样可以构成一个或多个共同体。从不同角度看一个 DID, 可以得出对共同体不同 (而且往往很直觉) 的看法。

因此, 处于边界上的各个 DID 就是共同体之间的桥梁——而且, 正如现实生活中一样, 它们在“提供关于其他共同体成员的信息”上扮演着角色, 无论是在业余层面, 还是在专业层面, 比如一个专门做“把人连接起来、并汇集关于高风险 DID 之信息”的权威。

DID 网络中的某些类型的信息很适合无条件地相互交换 (暴力犯罪、欺诈等)。然而, 另一些类型的信息, 共同体可能会视为仅限于在共同体内部使用 (相当于公民身份、为共同体提供的服务等)。这类信息能否被传播, 取决于共同体成员所宣告的政策。

一个共同体在自己的圈子内守护它的隐私, 作为回报, 则以无条件的方式交换对彼此都有益的信息。

Crime and Punishment

罪与罚

我们已经知道网络如何运作: 记录如何被创建, 谁来验证它们, 验证者的欺诈如何被遏制。还剩一个问题——当有人当真对另一个人造成伤害时, 会发生什么?

■ 我们彼此依赖有多深?

我们全都站在这个大世界的舞台上, 而这里发生的一切都与所有人相关。

扬·阿摩斯·夸美纽斯 (《世界迷宫与心灵乐园》)

加入或拒绝一个声誉网络, 可能有许多理由。然而我相信, 人性中的一个方面——常被视为阴暗的——可以派上好的用场。我说的是人对“在感知到的不公面前获得某种满足感”的需求。

满足感、正义、复仇——我们把它们与负面情绪联系在一起。为了让社会维系在一起, 我们至少必须假装满足感是通过集中式机构来交付的; 否则个体就不会感到自己与所生活的社会相连。平均而言, 我们需要归属于某个队伍并信任它。当有人伤害了另一个人, 社会便注视着满足感将如何、以何种程度被交付。带点争议地, 我要论证: 在历史上——以及现在——针对被告的程序是否干净, 往往并不要紧; 要紧的是社会看到某个人受到了惩罚。它发挥了它的功能 (无数的冤假错案, 以及对重审旧案的抗拒, 已经把这个故事讲得

够清楚了)。受害者与遗属的看法只是次要的。首先要紧的, 是给社会其余部分的信号, 因为不公并不那么常见, 以至于会出现在每个人的日常之中。

一个去中心化的声誉系统会让公众看到直接和间接的证据 (经审计的和未经审计的)。社会可以对信息进行评估——自行评估, 或通过专门做此类评估的自由市场服务——并根据定性和定量的信号, 决定对主体 (甚至对作者或验证的权威) 施加什么措施。

惩罚是什么? 它是通过剥夺被指称的加害者的特权来使其谦卑的一种形式——好让受害者和社会都感到: 有害的行为不划算。这才是关键所在: 有害或危险的行为绝不能划算, 同时它又必须给受害者以及社会/ 共同体的其余部分带来一种满足感。

这样的惩罚必须由某个中心权威来施行吗? 如果每个人都能预先宣告自己将如何回应某些行为呢? 我显然不能在别人的财产上执行惩罚。但我可以宣告: 我将把我的财产提供给一组选定的执行权威, 用于在冒犯者进入我的领地时惩罚加害者。而且我可以公开地宣告这一点, 好让其他人评估我的做法。

于是, 比例性的价格机制就登场了。如果我为轻微的冒犯宣告严厉的惩罚, 网络就会为我的极端定价: 其他人会觉得我不成比例, 我的验证成本会上升——愿意与我交易的人更少, 或者只肯以更差的价格与我交易, 愿意以合理价格接受我记录的验证者也更少。反过来, 如果我过于宽容, 共同体则会把我看作一个风险。如果我容忍欺诈, 共同体多半会对那些实施了触及我之行为的加害者也很宽容——好让我尝尝以其人之道还治其人之身的滋味。

正义不再是一件奢侈品。当有人伤害了你, 记录这桩冤情是负担得起的, 记录是永久性的, 而冒犯者所承担的后果是实实在在的。

■ 伪善的侦测

宣告一项政策还不够——如果我宣告我遵循某条规则, 而我的行为偏离了它, 这道缺口就是一项声誉上的负债。网络并不区分“ 违了法” 和“ 食了言”——两者都是可测量的信号。伪善代价高昂, 而我不知道还有什么比它更令人恼火的红布。

■ 数量与质量——打破罪责的二元范式

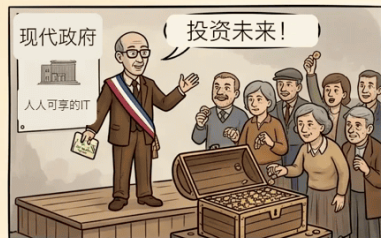
在集中式司法中, 罪责是二元的:0 或 1, 有罪或无罪。在声誉网络中, 要紧的不是罪责是否被证明, 而是信号的累积。

针对被告的单独一条未经证实的盗窃指控, 大概不会引起太多警觉——但它会引起对指控者的警觉。然而, 来自各种独立来源的多个间接迹象, 则开始具有分量。要紧的不仅是数量, 还有质量: 证人的独立性 (不同的人、不同的情境)、模式的一致性 (一种反复出现的嫌疑类型)、时间上的相关性。

反过来, 若同一个发行者针对各种目标主体反复提出无凭无据的频繁指控, 则可能被评估为诽谤——而指控者的声誉要为此付出代价。

一个陷阱。一条出路。

内部人总是拿下合约——直到网络让人们直接购买，无需任何许可。



现代政府 公众一致认为需要一套新的公共IT系统，由预算支付，由税款支付。



采购门来，采购员和内部人把招标设计成只合一家公司。一个信封易了手。



采购员会作弊。他的口袋被贿赂撑得鼓鼓的。



合约中饱。内部人拿到了执照——却没有技能，也不尊重这说法。



内部人拜访邻居。那份“慈善”的报价是最小的一块碎屑。

1. 传统方式

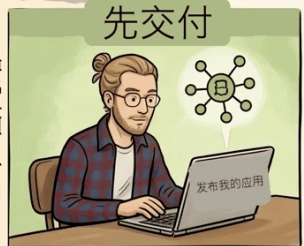


投诉、沉默、期限、律师。

陷阱就built在规则本身里。

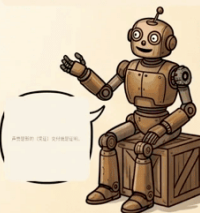


2. 声誉网络



他先交付。人们发现应用并使用它。

每个心存感激的用户付出一小点，体面的生计，没有中饱私囊可获利。



两条路。陷阱在规则里。出路是干脆绕开规则。

* 当规则由内部人书写，在规则内投诉就是那个陷阱——连罚款也回流进下一场被操纵的招标。声誉网络是出路：公民找到产品，直接付钱给创作者，社区授予唯一重要的执照——靠交付挣来的信任。

证据如何积累

"一条孤立的声明"



单一条声明——太单薄，不足以据此行动。哪一方都未被定罪。

"一种模式"



五个独立来源——一种模式浮现。

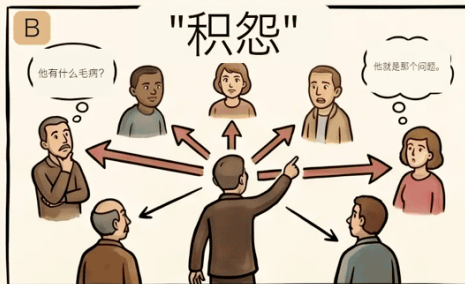
"压倒性的分量"



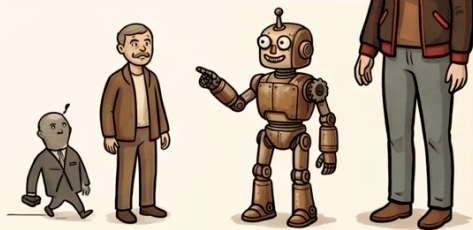
二十个独立来源——压倒性的分量。



一遍遍广播同一条毫无证据的声明——当社区什么也没看见，那就是诽谤，代价由指控者承担。



一个指控者指着所有人——模式指向的是指控者，而非被指控者。



这看起来像今天的世界。它就是——只不过现在动机诚实了。没有革命。只是更好的激励。

独立性就是一切。一个声音重复是噪音。来自不同地方的许多声音是信号。

我们也不应忽视赎罪——一种在罪责被证明之前的自愿行为, 它提供了一条通往较轻惩罚的路径。惩罚在加害者不采取任何朝向赎罪的步骤、并指望逃脱后果时启动——正如那种结果在集中式系统中可以通过腐败被买通一样。

■ 冒犯的类型

信息的类型不必限于犯罪行为; 它们可以包括: - 对犯罪的描述 - 违约与欺诈 - 不道德的行为 - 对公共秩序的扰乱, 无论真实的还是虚拟的 -

DID 网络中每一条新记录, 都会调整给定 DID(无论是主体、作者, 还是验证的权威) 被评估的声誉与风险。我为他人的命运承担的责任越多, 我就变得越谨慎。

■ 通过委托来强制执行惩罚

我不必亲自去强制执行惩罚。我可以 (类似于保险那样) 付钱给一个按预定规则强制执行结果的权威——相当于调查、执法和司法服务。如果该权威执行得不成比例, 或者根本不执行, 它的声誉就会受损——而它的声誉正是它赖以谋生之能力的根基。

委托是网络中的一条通用原则: 一个 DID 的任何活动都可以委托给另一个 DID——验证、发行主张、回应争议、宣告共识。委托随时可以撤销。它使得专业化成为可能, 而不丧失最终的控制权。

伤害之后——你会怎么做？

被拒之门外。被迫。自愿。村子作答的三种方式。

发生了什么



现在怎么办？



现在怎么办？

理性的选择永远是最便宜的那个：自己把它修好。

"我试着无视它——到哪儿都被拒之门外"



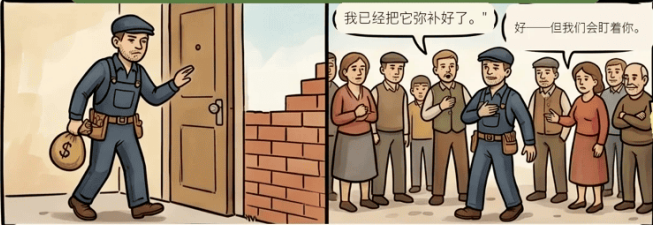
没有补救——每家店都把他拒之门外——被逐出经济生活

"我把它弥补好——但只在众目睽睽之下"



在众目睽睽之下被迫补救——偿清了，但村子给他的价码加了码——教训必须牢记。

"我自己主动前来——并向所有人承认"



自愿的、及早的补救——被公开接受——声誉恢复，但仍带着谨慎



谁来调查，谁来执行？

"伤害" → "调查员"

"记录" → "执行者"

→ "问责"



委托是可撤销的。任何DID活动都可委托给另一个DID——并可随时收回。
你不亲自执行惩罚。你委托。而受托者向你负责。

■ 风险管理的民主化

一个不可审查的声誉网络, 会把银行或大型工业综合体的维护者作为核心活动早已在做的事情——风险管理——民主化。

对面对客户的银行家而言, 风险评估是预先准备好的, 好让他们能在预定的限额内专注于业务互动。老练的维护团队在固定的检修间隔收集风险指标 (结构裂纹的迹象、正在侵蚀运营基础设施的过程等等), 并能预测设备的使用寿命, 以及何时需要何种干预。

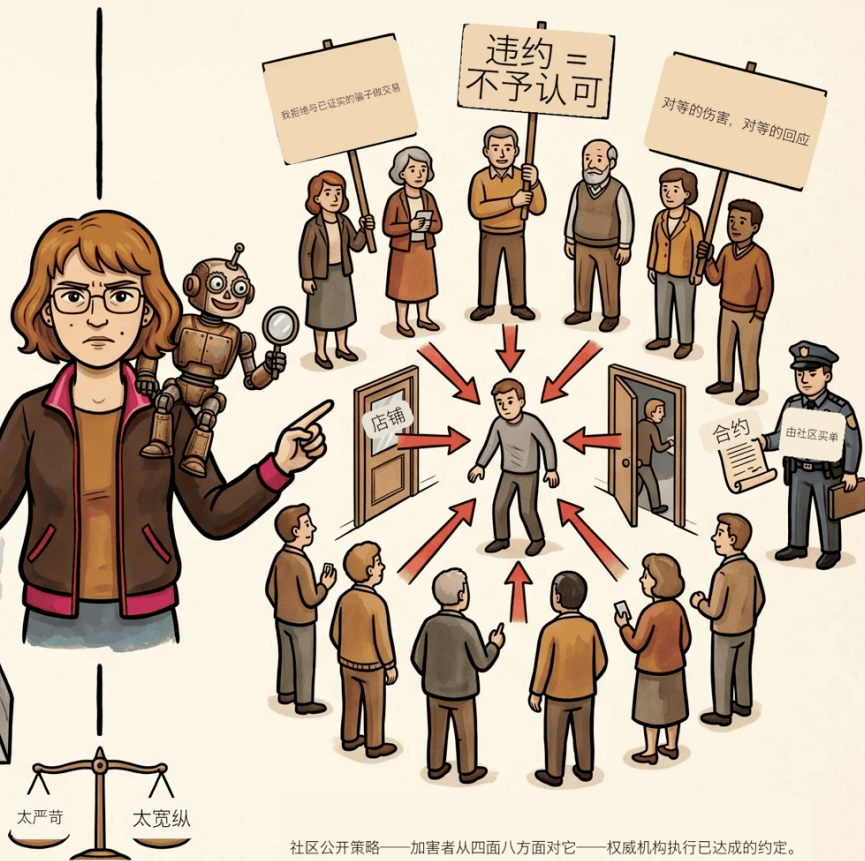
我设想 DID 的使用也类似——某人 (大概是收费的) 会提供一份关于对方的、人类可读的概要。这份概要将反映客户所订购的那个风险评估角度。

没有中央权威的惩罚



一人为所有人做主——受害者等待——加害者心知肚明。

赎罪比惩罚便宜。



太严苛
太宽纵

对等——最便宜的位置

社区公开策略——加害者从四面八方对它——权威机构执行已达成的约定。

惩罚比无视便宜。

Follow the Money

追踪金钱的轨迹

我们有了一个去中心化的声誉网络,也有了使用它的动机。它已经能给用户带来竞争优势——厘清潜在风险,帮助他们更高效地规划对手方的选择与谈判策略,并纳入那些能降低这些已知风险的措施。

接下来是使用它的一种可能的方法论⁵——利用网络的属性,构建一条从当前体制通往其进化继任者的迁移路径。它很可能带有中欧的特征,因为这篇论述正是在中欧构思的。当地条件将需要当地的调整。我建议在任何地方都保留的,是对国家所采取措施的不妥协性质——没有这一点,目标就会在无休止的讨论中被稀释,变革也就永远不会到来。当心——国家的代表与支持者会用胡言乱语让你忙个不停。别上当;他们的生计岌岌可危。

国家的软肋——金钱

让我们回到起点。我们想要更多由此产生的自由,想对它们的形态有更多影响,也想要与之相称的更多责任。问题很简单:国家在哪里有一个脆弱之处,让我们可以不诉诸暴力地去施压?

答案很无趣,却也因此更为有力。金钱。

究其核心,国家是一股金钱之流。它通过金钱来供养其强制机器——那个执行其规则的最终环节。没有金钱,就没有暴力垄断。一旦金钱开始停滞,国家就开始退却——因为它再没有东西可以付给那些在物理上把它撑住的人。

⁵国家的解构 (Deconstruction of the State)——在捷克语境下,这很可能在捷克作者 Daniel Steigerwald 一本即将出版的同名著作 (计划于 2026 年) 中被深入探讨。我个人对其中将包含的富有启发的路数抱有很高的期望。

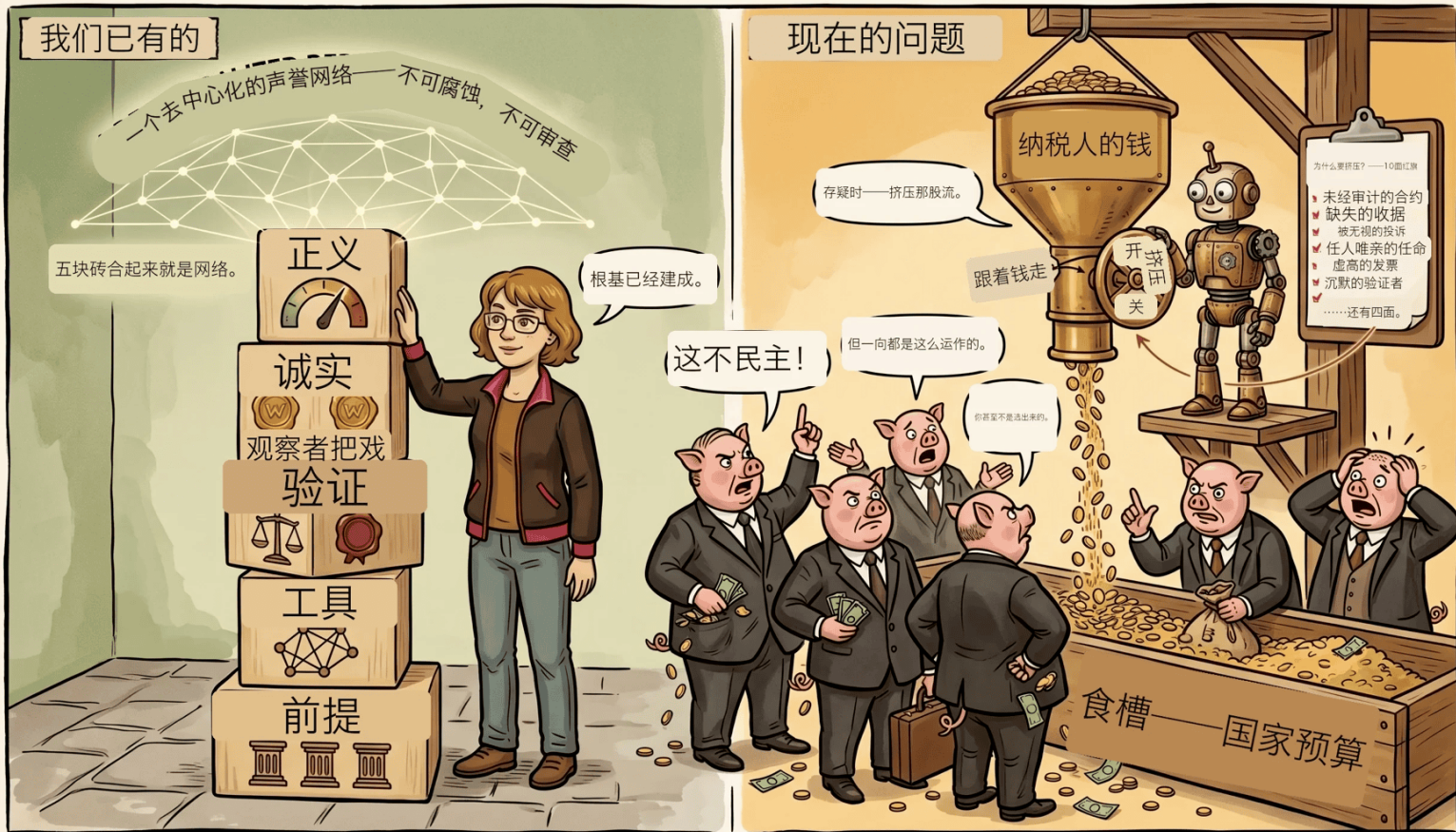
金钱之流自然会波动——随季节、随经济、随投资者情绪。国家通过在金融市场上举债来掩盖这一点。偶尔,一笔贷款是理性的:一项能回本的投资。然而更多时候,它是在为那些早已长出国家核心目的——国防、国内安全、刑事司法、外交——之外的超额支出埋单。一旦跨过某些阈值,央行就会以锚定在其自身通胀估计上的利率作出反应。

而这里就来了那道无声的税。通胀远不只是宏观经济学家的一個记账关切。它是一道对货币购买力征收的隐形税,由每一个为晚年攒下几枚硬币的人来缴纳。低通胀之所以被公众所容忍,是因为他们没有把握它的长期效应——他们没能意识到,即便是个位数的数字,在长时间跨度上复利累积,也会把储蓄侵蚀到人们再也无法保障自己晚年的地步。而当他们无法保障时,他们就被驯化成对国家福利的依赖。到头来,他们悬在那些福利之上,被拴在了国家身上。

在动荡的时期——战争、大流行、“必要的”投资——国家干脆开动印钞机。负责任的储蓄者和劳动者的购买力下降,而我们却许诺他们:他们得到了丰厚的回报。有人在悄悄偷走他们的时间,还称赞他们挣得多么漂亮。

这严重地侵蚀了对国家的信任。而信任是把整台机器黏合在一起的唯一胶水。那才是值得去撬的杠杆。

跟着钱走——故事至此



我们有了网络。我们懂它怎么运作。现在：存疑时，挤压那个料斗。
国家的权力立在钱上。跟着它走。

■ 为什么是金钱, 而不是其他软肋?

文明的另一个软肋, 比如说, 是能源分配。但扰乱它的流动不是一根杠杆——那会是侵犯、是恐怖主义。自愿的不参与, 与蓄意伤害他人不是一回事。金钱的特殊之处在于: 压力可以自愿地、透明地、不诉诸暴力地被施加。

Proposed transition to the state's successor

演化, 而非革命

突如其来的变化会激起抵抗, 即便最终目标很有吸引力。它会撕裂纽带——社会的、经济的、政治的——而人们本能地捍卫他们已经熟知的东西, 去对抗他们尚未见到的东西。

因此, 正确的做法是渐进的——演化式的。一种能防止旧体制只需轻按一个按钮就把新体制关掉的做法——因为一旦它能这么做, 它就会这么做。

不幸的是, 这意味着要接受: 均衡的自愿不会在我们部署声誉网络的那一刻就达成; 它将随一个渐进的采纳速率一同到来。采纳可能——而且很可能——是缓慢的, 由个体的渴望所牵引: 当循规蹈矩地寻求正义的方式失败时 (它一定会失败), 他们渴望通过声誉网络从自己的共同体那里找到救济。让我们看看那条路是什么样子, 包括按比例估算。

■ 空槽——别按为你注定要输而做的规则来玩

试图从内部改革国家, 依赖的是权力的内部机制, 寄望于官员们违背自己的激励去推动变革。那不会发生。

为什么要去喂养这个体制, 而不是把给它输送生命的渠道抽干? 与其改变旧体制并按它的规则来玩, 不如让我们在它之上建起一个新的层——一个逐渐把旧体制变成新体制之子集的层。一开始, 那个子集几乎填满整个超集。随着时间推移, 行为者们迁移过来。

公司深谙此道: 当一个旧系统盘根错节、无法替换时, 你就建一个平行的系统, 并逐渐把用户迁移进去。

国家最终会陷入“它实际交付的东西”与“它的章程所要求的东西”之间的不匹配, 并且会倾向于援引这道不匹配——这纯粹是它的代表和参与者的问题, 无需回头去理会。唯一值得盯着的是这一点: 不要挑起一场暴力革命。

相互支撑的四件工具

到目前为止,DID 声誉网络看起来或许主要在正义与共识方面有用。我相信它能够稳定自身、经受住攻击——以贤能政治的方式, 依据实际的功绩而非形式上的头衔。但仅凭声誉, 不会迫使国家让步。经济会, 以及通过它施加的压力会。

没有什么能阻止我们把 DID 身份用作更多工具的管理者, 去狠狠地压住金钱。不要从公民一方的完全自愿开始——设定一个国家无法阻止的方向。这场过渡立足于四件相互撑腰的具体工具。它们没有一件单靠自己就足够。

简单的税制——入场的诱饵

第一件工具是一套建立在 DID 基础设施之上的、简单而透明的税制。对所有收入采用单一的比例税率。没有例外, 没有免税期, 没有企业扣除, 没有优惠。每一笔支付都可审计地关联到支付它的那个 DID。交叉核对易如反掌。

税率会为一部分纳税人设定得比旧体制略微更有利一些——那就是入场的诱饵。税负会逐年下降, 体制会变得越来越有吸引力, 而迁移会加快脚步。这套体制既是记录保管者, 又是征收者——没有中间人, 没有不透明。

电子支出登记册 (ESR)

第二件工具, 是把 EET(电子收入记录) 倒过来做。国家不再追踪每一个自由职业者的每一张收据, 而是我们来追踪国家的每一笔支出。除非每一笔支出——无一例外——都通过这道筛子, 否则国家从简单税制中收不到任何一笔个别款项。

ESR 是一个代理⁶, 它把国家和公共部门每一笔已实现的支出, 与一条计划付款的记录相匹配——记录中包括语境上的理由说明, 以及对收款方、支出实体和用途的标识。其产物是原始的公共数据, 可以层级化。任何人都可以拿去分析。任何人都可以就“该拿它怎么办”向纳税人提供建议。审计机关多年都做不到的事情, 如今对每个人来说都只有一次点击之遥。

你不必转到新的税制才能参与。只需为每一笔已实现的付款支付一笔完全可退还的押金 (比方说 1—3 欧元) 就够了。未实现的余额会被没收, 充当垃圾过滤器, 好让系统不被攻击淹没。一条记录可以由任何人提交——一个亲属为丧失行为能力的祖父母提交, 一个邻居为被荒废的公共服务提交, 一个企业主为一张未付的发票提交, 一个官员通过从他们自己系统里导出来提交。

⁶代理 (Proxy, 拉丁文 *procurator*, 英文 *deputy/intermediary*)——这里指代表另一个实体行事的中间人。ESR 代理接受来自任何人的计划付款记录, 把它们与国家已实现的支出相匹配, 并把由此产生的数据集向公众开放。

我预料政党会自然而然地转型为分析提供者, 或者会涌现出新的政党来取代旧的。他们不再靠税收拨款, 而将开始提供专业能力——为其支持者提供基于价值观的概览, 涵盖结盟的各权威、它们的服务以及验证政策。那些出于懒惰而草率或荒唐的记录, 自动成为削减预算的候选对象。在第一轮里, 你的付款总会被拨付, 但所收集的数据将作为评估“某项国家支出是否仍然合理”的依据。承包商以及其他吃国家饭的人, 有动机去确保官员们负责地把支出对应起来。否则, 在下一轮里, 钱就不会流出去。

■ ESR 这个缩写从何而来

在 2016 年至 2022 年间, 捷克政府针对本国公民——自由职业者和企业主——实施了一套名为电子收入记录 (EET) 的系统。其讯息很简单: “我们假定你们每一个人都在欺骗国家。你们要记录每一笔交易, 实时传送到我们的服务器, 我们会盯着你们。如果你们试图逃避 EET, 我们会用钱奖励你们的同胞去举报你们。”

EET 以巴尔干地区的先例为蓝本。世界上第一个在 2013 年推出全国性强制实时在线销售记录系统的国家, 是克罗地亚, 名为 *fiskalizacija*。捷克几年后采纳了这一模式, 并运行了欧洲最严格的收入追踪制度之一。

既然黑暗一方毫不犹豫, 我们又何必犹豫? ESR 就是同一套逻辑颠倒过来: 让国家向公民证明自己。

公民税收分配——决定的分量

第三件工具让纳税人对自己的钱流向何处拥有直接的影响力。在第一轮里, 无论 ESR 揭示了什么, 国家都收到了它的钱。在第二轮里它不再如此——由第一轮的分析, 以及随之而来的措施接手。已缴税款中的一

ESR——电子支出报告

每一笔国家支出都被记录——每一条记录都公开——每一处细节都开放供审视。

谁来提交

任何记录都可由以下人提交：
处理该笔付款的官员
收到款项的收款方

向任何想看的人开放。

国家支出公共账本——开放记录

类别：法律咨询
承包商：XYZ律师事务所 s.已支付
金额：1400万欧元 承诺：2026-01-14
支付：2026-01-19

背景：对新采购改革法案的外部意见——出了三份草稿，与部委开了一次工作会，举行公开听证——持续到立法期结束之后交付。

类别：道路设计咨询
承包商：Consultancy T 金额：待处理
承诺：2026-02-03 支付：—

背景：D-X高速公路延伸段的可行性研究——同一家公司此前为这条路线写过三份研究——结论：建议以高出原预算+18%推进。

类别：车牌签发 收款方：Z地区办公室
金额：每块车牌45欧元已支付
承诺：2026-02-19 支付：2026-02-19

背景：每块签发车牌的标准行政费——费用涵盖材料与处理——本季度总量12400块车牌。

类别：医院设备 承包商：MedSupply
a.s. 金额：820万欧元逾期
承诺：2026-03-08 支付：—

背景：地区医院的MRI扫描仪安装——公开招标，三家投标——中标价低于中位数2%，60天内交付——发票已逾期90天。

垃圾过滤器——可退押金



当记录被证明准确时退还押金。

党派转变为分析提供方



这些描述是原材料——党派会仔细调查它们并产生建议，交叉对比并产出研发的蓝图。

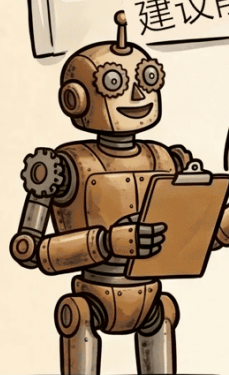
公款的每一分——被记录、公开、可审视。
价值不在抓骗子——而在打开账簿。

小部分百分比,可以由缴纳者本人来分配——前提是他们的 DID 满足声誉上的要求(公民身份、居留、共同体承认的成年/行为能力)。可分配的百分比逐年增长。有可能是指数式增长。从今天的百分之五,到十年内的几十个百分点。

原始的 ESR 数据自然会引导纳税人的决定。低质量的记录起到一个警示信号的作用,提示建议削减的对象。任何人都可以创建记录;官员们不得不去对应。没有这一点,税制就不会给国家送去任何钱——既不进国库,也不进其他预算科目。政治斗争从投票箱转移到了对一枚枚具体硬币的日常分配之上。

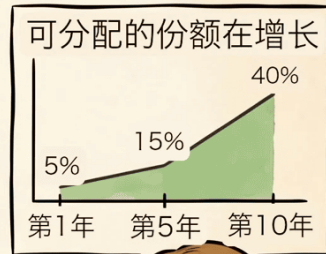
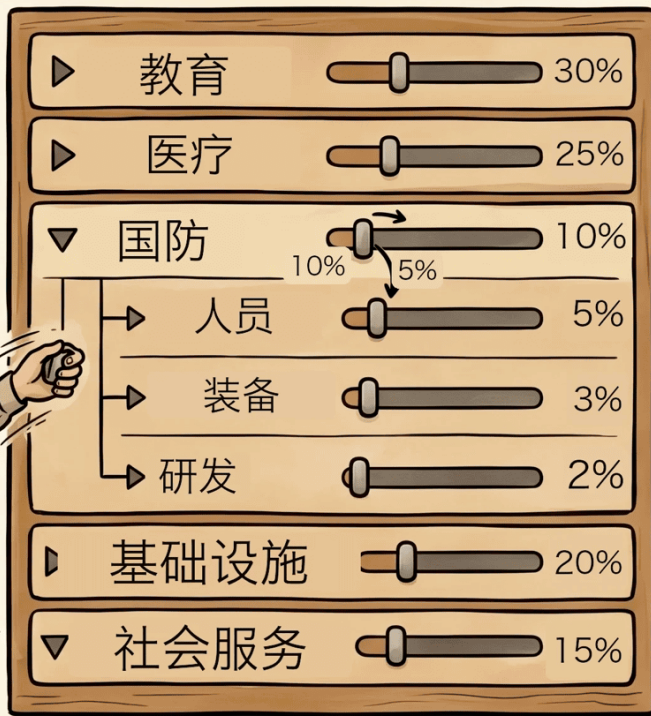
公民税款分配——你决定你的税金去哪里

研究ESR记录。然后拖动滑块。



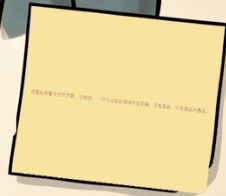
或者委托这件事

不想自己分配？委托给：
- 你所在党派的分析师
- 一位可信的邻居
- 你选的任何人——并随时可撤销。



点任意一行深入——在任意范围分配。

从今天的5%到十年内的40%。你来决定。
记录是开放的。分析是你的。滑块也是。



与国家谈判的平台——那根杠杆

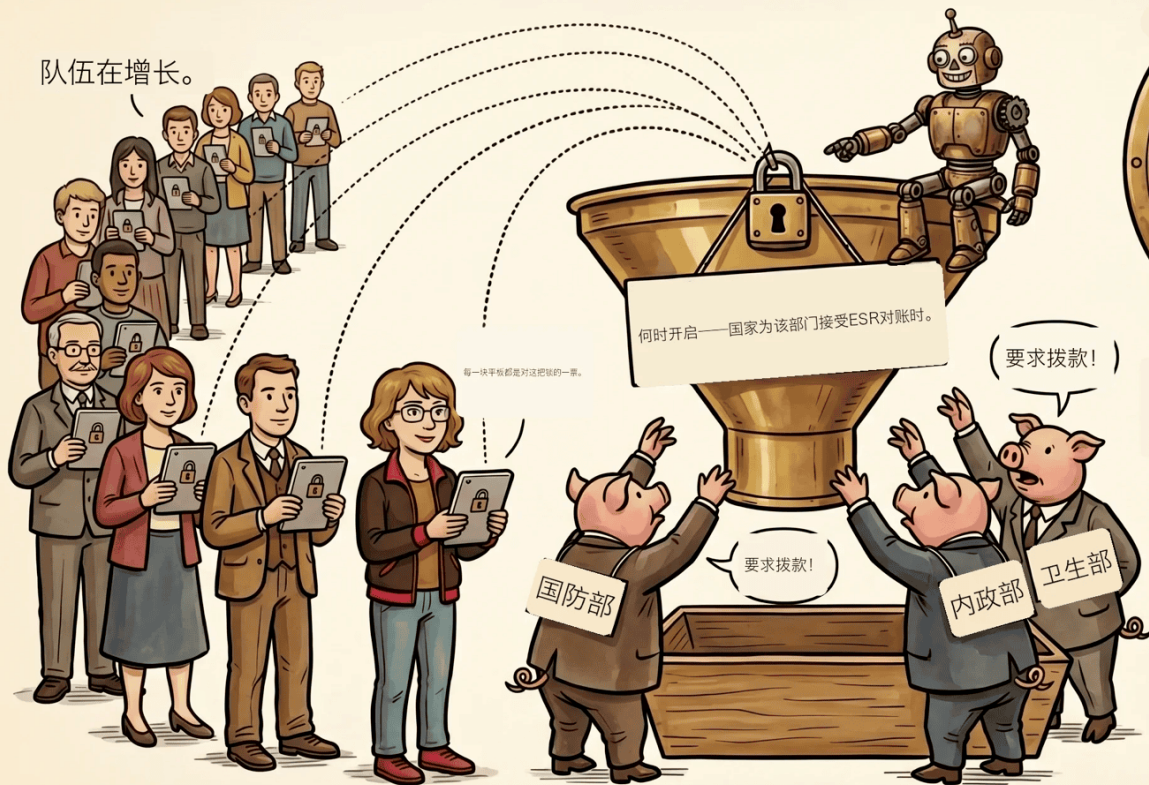
第四件工具,是把一切系在一起、并且直到此刻才让网络成为真正威胁的那根杠杆。主要的冲突,将围绕国家是否接受 ESR 作为一项有约束力的标准——首先在国家层面——以及在此之前,来自声誉网络参与者的威胁必须变得多大。

这根杠杆,是即时 (just-in-time) 的国家拨款,以“支付须对照 ESR 进行匹配”为条件。只有在国家满足了透明条件的那一刻,钱才流动。声誉网络内部会形成一份参与者的等候名单。当这个群体达到一个具有经济分量的规模——大到足以威胁 GDP 增长——时,激活便启动。到那时,公民不服从就成了一个可信的威胁,而不再是一个姿态。

国家会最长久地死守那些与腐败相关的、有问题的付款。在压力之下,它会首先在它有现成的辩护、且没有腾挪节约空间的地方让步。渐进而上升的压力,会迫使它一步一步地从每一个阵地退却。

协商平台——那个关闭的料斗

公民握着钥匙。条件满足时国家便开启。



GDP的一小部分就够了——达到临界质量时，扣留就成了实实在在的威胁。

达到临界质量后——

国家先为简单、站得住脚的付款打开账簿。

与腐败挂钩的类别锁得最久。

每一次解锁都需要更大的压力——而这本身也被记录。

有了透明，钱才流动。公民握着钥匙。
没有武力。没有革命。只要有足够多的平板，让扣留变得可信。

无需身份证的公民身份

随着向 DID 基础设施的过渡, 国家签发的身份证失去了它那垄断性的用途。你不需要一张身份证——你拥有一系列 DID 身份, 其中一些带有一条主张, 证明你比如说是捷克人、某个特定市镇的居民、某所特定学校的毕业生。这条主张是可验证的、永久的, 并且独立于国家行政之外。

一个技术实现的问题随之打开: 市场上谁来承担在国家控制之外去核验现有公民的工作, 以便把有意愿者迁移到一个带有相应主张的 DID 里? 最有可能的是, 专门做身份核验的自由市场权威会涌现出来——公证人、地方共同体、雇主。他们把自己的声誉押在“某个 DID 确实属于声称拥有它的那个人”之上。公民身份——也就是对某种共同体视角的归属——从对中央记录保管的需要中被解放出来。通过在声誉网络中的连接, 所嵌入的数据被第二次、第三次、第四次, 以此类推地连接和校验。

其后果既直觉, 又在集中式系统里无解。今天, 政治领导层——着眼于更容易的连任——可以把集中的问题输入共同体的领土: 接纳来自其他文明圈的移民, 而不顾吸纳能力和既有共同体的意愿, 在一种压力下改变原有的共同体, 使其跟不上同化来自其他文明圈之新来者的速度。在一个 DID 网络里, 一个通过腐败被接纳的移民, 无法确定共同体会继续把他们视为自己人。共同体的成员客观上不必忍受一个陌生人穿着共同体的制服找上门来, 拿他打算在屋里实施之措施的合法性来威胁他们 (因为一条社交媒体帖子而没收孩子的手机)。有了去中心化的声誉, 社会中的多数就不必对此逆来顺受——因为对一个共同体的归属, 不是由一名官员来裁定, 而是由共同体本身, 通过它的声誉信号来裁定。

演进式转型的四件工具

系统从内部重建——不打破它。

简单税



一个税率。没有优惠。可交叉核对。

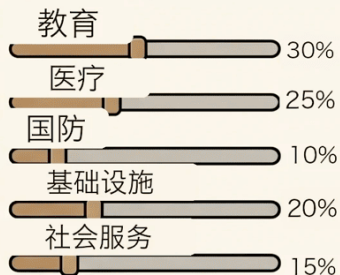
ESR——电子支出报告



每一笔支出都被记录——每一条记录都公开——每一处细节都开放。

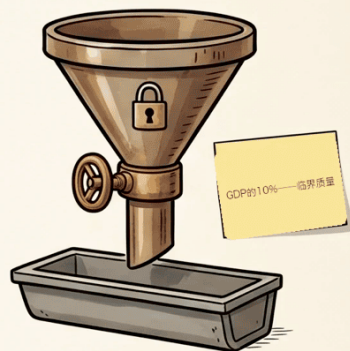
公民税款分配

5% → 15% → 40%
第1年 → 第5年 → 第10年



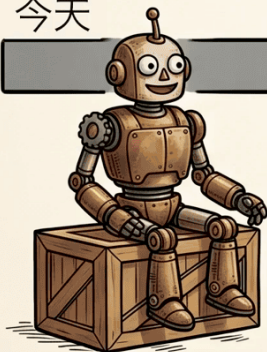
研究记录——然后拖动滑块。

协商平台



只有满足透明条件，钱才流动。

今天



演进进行中

95%

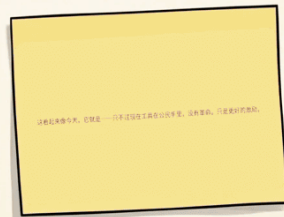
十年后

扩张中



演进，而非革命。

四件从内部改变系统的工具。



没有身份证的公民身份

从国家许可到社区认可。

"旧方式"——机构式、国家签发的身份证模型



身份 = 国家许可——可撤销——社区没有发言权。

"新方式"——带声明卡的DID模型



没有单一权威掌控所有声明——每一条都独立——每个验证者押上自己的声誉。

"谁属于这里?"



你的公民身份不是政府许可。它是你社区的认可。
国家印章被一条独立验证者的链条取代——每一环都利益攸关。

在国家职权范围内的事务上, 一张选票的分量, 未来可以与“对简单税制之贡献”的记录严格相关联。再结合合适的 DID 主张 (例如“此人是捷克公民”), 外国国民随后便可以被排除在选举之外——人口登记册干脆就搬到这里来。

重叠与可逆性

我并不宣称这是唯一的路, 甚至也不宣称是最好的路。但它看起来是全面的——它覆盖了主要的杠杆, 并保留了退回一步的选项。在起点上, 新体制与旧体制重叠, 只改变激励。随着时间推移, 旧体制收缩, 直到它可能——但不必——消解。其优势在于渐进性与可逆性: 万一设计中出现了错误, 那不是终点, 只是一次修正。

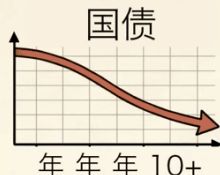
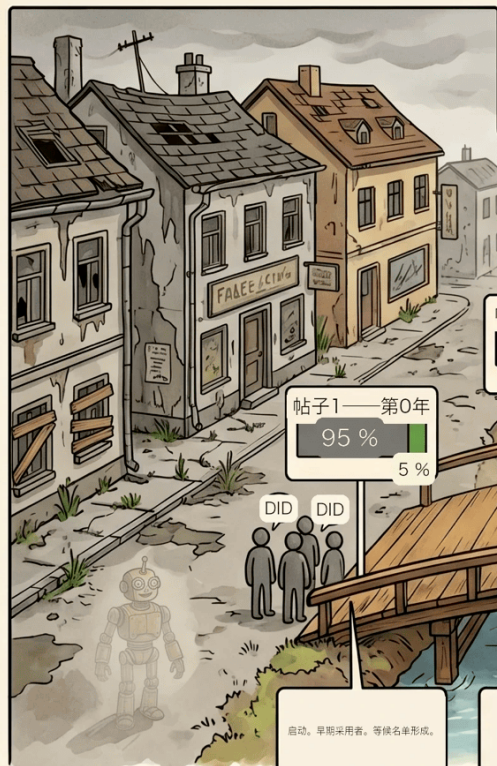
■ 和平与战争

国家的某些部分, 可能没有一个完全基于市场的替代方案——或许是军队。在危机中临时委托非常之权 (正如古希腊人任命独裁官), 可以继续存在。然而, 责任落在去中心化的社会身上, 而非领袖身上。与集中式治理不同, 一旦威胁过去——或在必要时更早——收回被委托的权力会更容易。一个自由的社会, 即便在与威权政体的兵戎相见中, 也不会处于劣势。

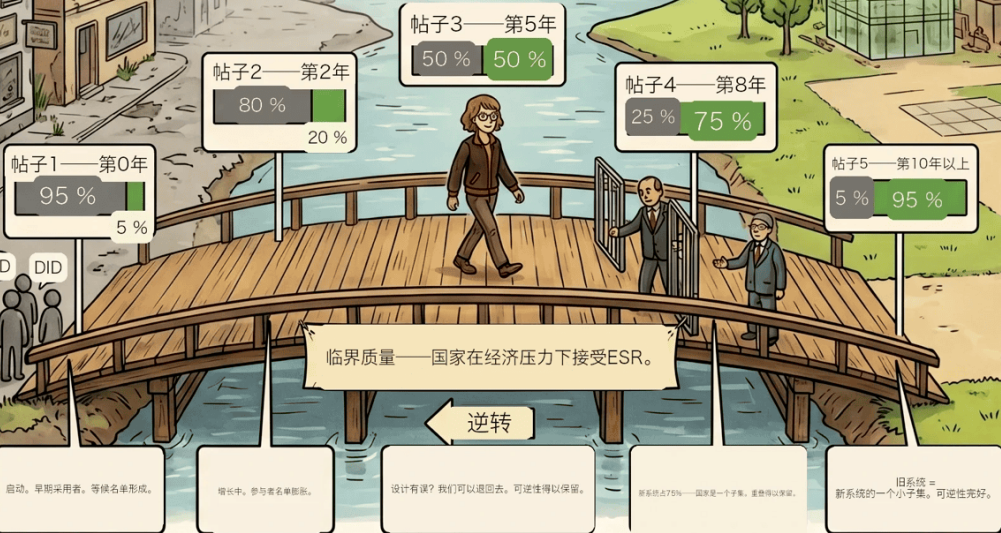
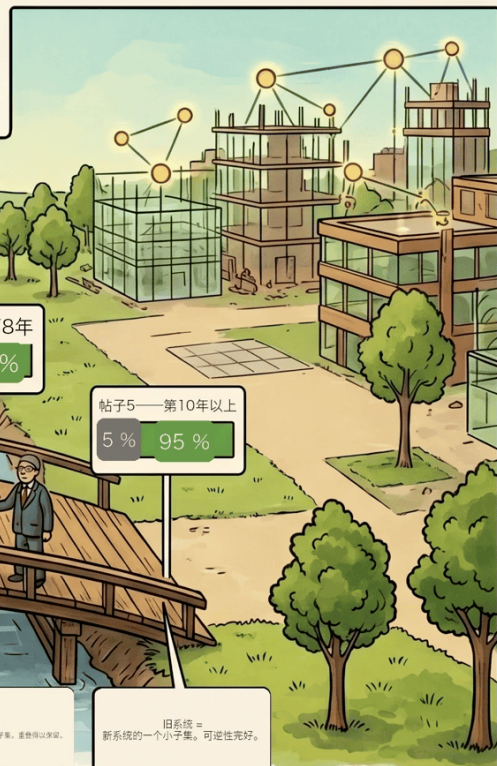
■ 战争罪犯

有了即时的国家拨款和透明的支出, 要为侵略战争积累资源就变得困难。一个继续为这样一场战争提供资金的纳税人, 就成了一个正当的目标, 不能援引《日内瓦公约》下的平民身份。在新的安排下, 沉默的默许被当作参与。

左岸——"旧系统"



右岸——"新系统"



旧系统

> 渐进的转型。万一出错可逆。一个无法被阻挡的方向。

Conclusion

道路是存在的

国家之所以是它现在的样子, 是因为其诞生之时可用的技术, 与人性相结合。把一切建立在自愿之上、却否认人性的平均面貌, 是天真的。它不会摆脱腐败, 也不会摆脱它那朝向极权漂移的倾向。国家的继任者, 必须连同我们当下视为负面、并且要么排斥、要么惩罚的那些特质一起运作。

这本书描述了一个问题——一个体制 (国家), 在其中书写规则的人永远不必与规则的后果同处。它描述了要让那个体制停止运转所必须成立的那些条件。它描述了一件满足那些条件的工具。它还提出了一套过渡方法论——如何使用那件工具的具体步骤, 作为一个可能的范例。

在第一章里, 我点出了五根支柱: 自愿、组织、韧性、共识、不可腐蚀性。我们把它们全都满足了。

- **自愿**——没有人强迫你进入网络, 但你也无法把自己的作为从中抹去。
- **组织**——共同体从你所认识的人当中生长出来, 而不是从上头分派给你的人当中。
- **韧性**——即便一个愿意为监控每一个数据包付费的国家, 也关不掉它。
- **共识**——它从伪善的价签中产生, 而非从投票或签名中产生。
- **不可腐蚀性**——五条小规则把腐败织进成本里: 自愿与责任的平衡、任何人开口说话的权利、无人能挑选的入场点、带有伪善之价的公开承诺, 以及每一笔交易背后一个观察者的影子。

没有任何单独一根支柱能凭自己取代国家。合在一起, 它们构成一种对腐败的抵抗力——国家连同它全部的机器都做不到的抵抗力。

我们正处在这样一个点上:整件事情用当今的技术就能建成(密码学、去中心化身份、为把激励调对而广为人知的心理学模式,.....)。唯独缺一样东西:一批找到勇气去支持它、并决心开始的、达到临界质量的人。

不是通过革命。

这些想法成形于 2025 年 6 月至 2026 年 5 月之间。到我坐下来动笔时,我脑中已经大体理清了这个框架——正因如此,当《告别威斯特伐利亚》出现时,我怀着那样的好奇去读它。我想看看他们的图景会在哪里与我的重叠,又会在哪里分道扬镳。第二个推动来自 Dan Steigerwald,他在 2026 年 5 月仍在从一个不同的角度写自己的书《国家的解构》;他的项目促使我把自己的构想以完整的广度记录下来——好让这个世界把它与其他人的构想一并摆上桌面。

我不是在请你相信这其中的任何一点。我是在请你思考它。而我留给你一个问题:
你会改变什么?

Pavel Kudrna 2026 年 5 月 23 日,捷克布拉格

致谢

灵感

我感谢《告别威斯特伐利亚》的作者(Jarrad Hope、Peter Ludlow),感谢他们提供的思想框架,帮助我把自己的想法置于“解构威斯特伐利亚式国家模型”这一更广阔的语境之中。

审阅与反馈

工具

在准备文本、梳理思路、反复推敲措辞的过程中,我借助了 AI 工具的协助。所有的想法、论证和结论都是我自己的——AI 充当的是打磨的工具,而非内容的合著者。

信息图是在作者的规格说明基础上,与生成式工具协作设计的。

夺回你的时间

公民把寄生虫赶出公共生活。



守护他们妄图偷走的时间。

以欺骗夺走的，用双手夺回来。

支持这个项目

Personix 由像你一样的读者资助。



Bitcoin 链上

bc1q7cctdz6r026dgy06ul94eygg49ujtxnrq5czaj



Bitcoin Lightning

LNURL1DP68GURN8GHJ7AMPD3KX2AR0VEEKZ
AR0WD5XJTNRDAKJ7TNHV4KXCCTTDEHHMM30
D3H82UNVWQHXXMNP0FA8JMTVPD5XJMN9XUE
RX55QRQK

■ 这笔钱去往何处 — 以及不去往何处

这些二维码是**作者的个人渠道**——捐款直接进入 Pavel Kudrna, 资助进一步的写作、社会实验/模拟以及对外推广。

本项目唯一正当的筹款渠道是:

- **本书由作者授权的版本**——由作者授权的副本中所含的二维码。
- personix.org 或从那里链接的 Tor 洋葱镜像——*Personix* 基金会在那里为其所述的各项活动筹款。基金会的资金与作者的个人资金分开保管。

在其他任何地方宣传的东西——再版中被替换的地址、以假乱真的域名、社交媒体上的“为 *Personix* 募款”、与上述两者都无关的筹款——**极有可能是诈骗**。我们刻意回避加密货币市场及其参与者。

Glossary

用语集

术语	捷克语	含义
权威 (Authority)	Autorita	一个可信实体 (人、组织), 它验证信息并把自己的声誉押在上面。可以是专门化的 (调查、法律、技术)。
主张 (Claim)	Tvrzení	一般而言: 任何可验证的陈述。这里: 发布到声誉网络的一条记录——一个关于事件、性质或关系的断言, 经过密码学签名并被验证。例如“我是市镇 X 的居民”或“此人违反了一份合同”。
区隔化 (Compartmentalization)	Compartmentalizace	一般而言: 把信息分隔成彼此孤立的单元, 使得暴露一个单元不会危及其他单元。这是情报机构所熟知的一条原则。这里: 独裁统治下的并行 DID 身份——暴露一个不会牵出其他。

术语	捷克语	含义
一致性哈希环 (Consistent Hash Ring)	Hash ring	一种用于选择验证者的算法机制——环上的位置由社交图谱内 DID 文档的哈希决定。确保一种非确定性却可验证的选择。
DID	DID (Decentralizovaná identita)	一种由你自己创建并掌控、无需中心权威的数字身份。由你的私钥进行密码学签名——没有人能撤销它或伪造它。
DID 文档 (DID Document)	DID dokument	一份公开可得的数据文件, 描述你的 DID 身份——包含公钥、网络地址和元数据。用于在网络中验证你的身份。
尽职调查 (Due Diligence)	Due diligence	一般而言: 在进入商业或法律关系之前对对方的深入核验——检查其历史、财务、声誉和风险。这里: 在声誉网络中, 由于经过验证的记录随手可得, 它进行得更快、更自动。
经济中性原则 (Economic Neutrality Principle)	Princip ekonomické neutrality	网络中诚实的行为在经济上接近于零——发布成本以验证报酬的形式返还。不诚实的行为是净损失。

术语	捷克语	含义
涌现的 (Emergent)	Emergentní	从更简单部分的互动中自发地产生, 而无人设计或指挥。一群鸟无需计划就成队形飞行——队形从每一只鸟所遵循的简单规则中涌现。
涌现的社会契约 (Emergent Social Contract)	Emergentní společenská smlouva	一些行为规则, 它们不是自上而下 (法律) 产生, 而是自下而上——从一个共同体内部反复的互动与共识中——产生。
ESR	Electronic Spending Register	一套提议中的、用于透明追踪公共支出的系统——国家每一笔已实现的支出都与一条计划付款相匹配。受捷克 EET 启发, 但反过来对准国家。
哈希 (Hash)	Hash (otisk)	一般而言: 一种单向数学函数, 从任意输入产生一个唯一的、固定长度的“指纹”——就像文档的指纹。相同的输入总是产生相同的输出, 但无法从输出推导出输入。这里: 用于确定哈希环上的位置, 以及验证文档的完整性。
即时资金供给 (Just-in-Time Funding)	Just-in-time financování	以透明为条件的国家拨款——只有当国家接受 ESR 并对应其支出时, 钱才流动。一根用来迫使合作的杠杆。

术语	捷克语	含义
贤能政治 (Meritocracy)	Meritokracie	一般而言: 一种由实际功绩和已证明的能力、而非形式头衔、人脉或世袭特权来决定地位的体制。这里: 声誉网络自然偏向那些明白地为共同体作贡献的人——他们的声音因履历、而非因职位而更有分量。
洋葱网关 (Onion Gateway)	Onion gateway	一个 DID 身份在洋葱网络上的网络地址。与 DID 文档分开——它可以在不丢失身份的情况下更改 (类似于更改域名背后的 IP 地址)。
洋葱路由 (Onion Routing / Tor)	Onion routing (Tor)	一种确保网络不可审查性的通信协议。消息被逐层加密——每个节点剥去一层, 却不知道完整的路径。
预言机问题 (Oracle Problem)	Oracle problém	一般而言: 如何确保进入数字系统的数据忠实地对应于物理世界里实际发生的事情。这一术语源自区块链领域。这里: 通过把自己声誉押上、以担保“数字记录对应于物理现实”的权威来加以应对。

术语	捷克语	含义
现象学的 (Phenomenological)	Fenomenologický	一般而言: 一种取向, 它研究现象在直接经验中显现出来的样子, 通过观察从中得出什么, 而不预设既定理论。这里: 自由、社会契约和行为规范都是被观察到的现象——是人与人之间成千上万次微观互动的结果, 而非自上而下界定的原理。
政策 (Policy)	Policy (politika)	一般而言: 一套在给定语境中规范行为的规则或原理。这里: DID 网络中的每个参与者都宣告自己的政策——他们如何回应他人的具体行为、遵循哪些规则、认为哪些惩罚是成比例的。政策的总和构成涌现的社会契约。
代理 (Proxy)	Proxy	一般而言: 一个替身或中间人——代表另一方行事的系统或实体。这里在两个语境中使用:(1) 作为把公共支出与计划付款相匹配之代理的 ESR;(2) 在观察者把戏中作为发布者与验证者之间代理的观察者。
发布者 (Publisher)	Vydavatel	创建并发布一条记录 (关于一桩不公、一次补救等等的主张) 的网络参与者。承担发布的成本。

术语	捷克语	含义
基于声誉的社交网络 (RSN)	Reputační síť	一种去中心化社交网络, 参与者在其中交换关于现实世界行为的反馈。记录创建起来代价高, 读取起来便宜。
声誉信号 (Reputation Signal)	Reputační signál	网络中的一条个别记录——正面的 (伤害的补救、义务的履行) 或负面的 (不公、违约)。累积起来, 信号构成一份声誉档案。
社交图谱 (Social Graph)	Sociální graf	你的联系人以及你联系人之联系人的网络。算法在一个可配置的深度 (例如 3 层) 搜索验证者。没有全球区块链——网络自然形成一个个有所重叠的共同体。
税收分配 (Tax Allocation)	Alokace daní	一种机制, 纳税人借此决定自己税款的一部分流向何处。可分配的百分比逐年增长。
履历 (Track Record)	Track record	一般而言: 一个人或组织过去的结果、成功与失败的历史。这里: 某个 DID 身份在网络中过去所有互动的总和——已验证的主张、被接受和被拒绝的记录——其声誉即由此推导而来。

术语	捷克语	含义
验证者 (Verifier)	Ověřovatel	一个由算法选出、用来验证并发布一条记录的参与者。把自己的好名声押在信息的真实性上。

制作说明

由作者保留 (不受 CC BY-SA 许可协议覆盖): Personix 名称与字标、书籍封面, 以及支持渠道 (二维码、地址、联系邮箱)。衍生版本必须移除或替换这些元素; 它们只能用于引用这部原始作品。

源文件。完整的源代码——markdown 文本、信息图、构建脚本——与本 PDF 一同提供。

工具。在准备文本、反复推敲措辞的过程中, 作者借助了 AI 工具的协助; 信息图是在作者的规格说明基础上, 与生成式工具协作设计的。所有的想法、论证和结论都是作者自己的。

PERSONIX

不妥协的变革

一个不可审查、不可腐蚀的去中心化声誉网络

当通信缓慢、决策在地、权威就住在它所统治的同一座城镇里时, 国家是奏效的。今天的网络把这三个假设全都颠倒了——建立在其上的机构, 再也不适合它们所治理的世界。这本书描述了一件确实合身的工具: 一个去中心化的声誉网络, 在其中身份归你所有, 真相可被公开审计, 而贿赂等于声誉上的自杀。

不是一份宣言。不是一个预测。而是一份可运作的蓝图, 讲述国家的继任者可以如何被建成——自愿地, 一次一份宣告。

Pavel Kudrna • 布拉格, 2026

personix.org

依据知识共享署名-相同方式共享 4.0 国际许可协议 (CC BY-SA 4.0) 授权。