

The Unbreakable Promise

How institutions made of code could help save humanity

Şen Ferhat

July 2026, working draft v0.1

*Humanity paid dearly to get this far.
Seal what follows; let this be a beginning.
This seal was paid for dearly, and is being paid for still.*

*İnsanlık buraya gelene kadar çok bedeller ödedi.
Bundan gerisini mühürleyin; bu bir başlangıç olsun.
Bu mühüre çok bedel ödendi, ödeniyor.*

This book is part of the STOA project.
Companion academic text: “The Budget That Cannot Be Overspent:
From Governance Notaries to Governance Mechanisms” (docs/paper).

Draft version, please consult the author before quoting.

Contents

Preface	v
1 The Age of Broken Promises	1
2 Institutions: Humanity's Operating System	2
3 The Notary and the Mechanism	3
4 How a Promise Becomes Unbreakable	4
5 The Honesty Mark	6
6 Can Corruption Be Engineered Away?	7
7 Starting Small	8
8 One People, Two Systems	9
9 When Amendment Is a Patch	10
10 Justice in a Single Transaction	11
11 What Code Cannot Save	12
12 An Invitation to the Age of Repair	13
Sources and Inspirations	14

Preface

On the third of July, 2026, a small state was founded inside a computer. It had a registry, a senate, a treasury, a budget, and a timelock. In the final step of its founding, the engineer who built it, that is, I, signed a single transaction: I transferred my own powers into the system. A few minutes later, out of habit, I tried to write an appropriation directly into the treasury. The system answered: *Only admin*. The administrator was no longer me; the administrator was the timelock, fed by the senate. The thing I had built did not recognize me.

What I felt in that moment was not defeat; it was a strange, weightless relief. Because over every promise in human history hangs the shadow of the one who made it: “I will keep my word, and if I don’t, who exactly will stop me?” That day, for the first time, I had made a promise I could not take back even from myself.

This is a book about that weightlessness. Its claim is large: we can rebuild our institutions, budgets, assemblies, courts, cooperatives, out of promises that cannot be broken, and doing so touches the root of a surprising share of humanity’s great problems. But it is not a salvation recipe; it is a *repair manual*. I will describe what code can hold, and, in the later chapters, what it never will. Because the first rule of the system this book defends is honesty: every promise must carry a mark that says whether it is physics or merely testimony.

Nothing in these pages is theory. Every mechanism described here, the budget that cannot be overspent, the system that does not recognize its own founder, the ruling that enforces itself in a single transaction, exists today in running code. Small, yes. In a laboratory, yes. But the Wright brothers’ first flight lasted twelve seconds.

Chapter 1

The Age of Broken Promises

A city budget is blown; nobody goes to prison. An aid convoy's cargo melts away before it reaches the village it was meant for. A pension fund is drained by the very people sworn to protect it. A lawsuit is won, and cannot be collected. An election is held, and the losers doubt its result while the winners forget its limits. A climate accord is signed; targets are announced, applauded, missed; nobody answers to anybody, because there is no mechanism for answering.

We imagine these events have separate names: corruption, incompetence, bad luck. Underneath, the same broken gear turns in all of them: *the promise is entrusted to the goodwill of the very people who are supposed to keep it*. The rule does not bind the powerful; the powerful administer the rule. A budget is an intention, a bylaw is a wish, a verdict is a request, if the hand that must execute it may choose not to.

Look one by one at humanity's great problems; a startling majority are not failures of knowledge but failures of *coordination*. We know how to cool the climate; we cannot make each other keep the promises we exchange about it. We know which transfers reduce poverty; we cannot stop the money from evaporating along the way. We know what to do in a pandemic; the commitments about who does what and when dissolve at the first panic. Coordination runs on promises. And our promises are broken.

Trust is the scarcest resource of our age, scarcer than oil, than lithium, than clean water. Trust in institutions is measured everywhere and falling almost everywhere. Economists politely call its price "transaction costs"; on the street it has plainer names: a lawyer for every contract, a kickback margin on every tender, a doubt attached to every receipt. Where there is no trust, everyone pays for everything twice.

This book proposes to look at that picture with an engineer's eyes. What is broken is not our morality, morality was never sturdier than it is today. What is broken is *infrastructure*: the technology that carries our promises has not changed in three thousand years. Writing, seal, signature, witness, threat of punishment, all of it hanging from the same weak link: somebody must still *choose* to obey the rule even when the rule touches them.

What if they didn't have to choose? What if a promise, once made, left the hands of its maker entirely?

Chapter 2

Institutions: Humanity's Operating System

Writing was not invented for poetry. The oldest tablets we possess are the grain ledgers of Sumer: who brought what, who owes what. Humanity's first information technology was born to hold up an *institution*, the temple granary. From that day on the story has flowed in one direction: institutions dragged technology forward. Property gave us surveying; taxation gave us statistics; trade gave us double-entry bookkeeping; the state gave us the census.

When Hammurabi had his laws carved in stone, what he built, in today's terms, was a *read-only broadcast*: the rule stood in public view, unalterable. But stone solves only the *readability* problem; it does not solve *execution*. Whatever the stone says, the tax collector can still take what he likes. Constitutions were invented for exactly this: the attempt of power to bind itself. Odysseus lashing himself to the mast, the mind that knows it will hear the sirens ties its own hands in advance. But the ropes of a constitution are made of paper, and the knot is held, again, by people.

Once you begin to see institutions as *technology*, a strange asymmetry leaps out. We replace our hardware every two years. Our software updates weekly. But civilization's operating system, the layer that decides how choices are made and carried out, how value is collected and distributed, how disputes end, still runs on paper, seals, and fear. We handed the airplane to the autopilot; we still entrust the budget to conscience.

There was a reason for the asymmetry: until recently, a rule could be executed only by human hands. To give a rule to a machine, the objects the rule works on, money, membership, authority, records, must live where the machine can reach them. That is precisely what changed in the last decade. An infrastructure appeared in which money, membership, authority, and records can live on a shared ledger that no single operator holds. Its first great application was currency; but currency was the prologue. The real promise is this: *for the first time, the hand that executes the rule and the hand that is subject to it can be separated.*

The rest of this book is about what can be built on that separation.

Chapter 3

The Notary and the Mechanism

Let us begin with a distinction; the rest of the book stands on it.

A *notary* records what happened. A good notary is priceless: who decided what, who signed what, in which order, a tamper-proof history. The overwhelming majority of what you see today under the names “digital government” or “blockchain governance” are accelerated, beautified notaries. A decision is made, the decision is *recorded*... and then? Then, as always: carrying it out is somebody’s job. And somebody can forget, postpone, “interpret,” or change their mind.

A *mechanism* is a different species. A mechanism binds the decision to its consequence: what is approved *happens*; what is not approved *cannot happen*. The difference is the difference between a request and a rail.

Let us make it concrete. In a paper state, a budget is an instruction: “spend at most this much on roads this year.” The instruction is addressed to officials with signing authority, and someone with signing authority can, by definition, sign outside the instruction too. Overspending the budget is a *crime*; which is to say, it is possible, merely forbidden. Now build the same budget as a mechanism: before every outflow, the treasury asks the budget contract; a payment with no appropriation line, or beyond its line’s remainder, does not happen, no matter whose signature carries it, including the treasury’s own owner. The transaction reverts. In a paper state, overspending is a crime; in a coded state, it is an *impossibility*. The distance between a crime and an impossibility is the distance filled today by armies of auditors, inspection boards, corruption trials, and lost decades.

Once you see this distinction, you look at every institution around you differently. When you hear “we have decided,” the question becomes: is this decision a record, or a rail? Could failing to carry it out be *someone’s choice*? If yes, you are looking at a notary. And notaries, however digital, cannot close the age of broken promises.

Chapter 4

How a Promise Becomes Unbreakable

So how does a decision get onto rails? The road has four stops.

First stop: the decision carries its own action. The first condition of an unbreakable promise is that it not be vague. In our system, a proposal that says “appropriate five hundred units to schools” is not just a title and a justification; inside the proposal sits the action itself, which contract, which call, which parameters, ready to fire if approved. What is voted on is not an intention but a loaded trigger.

Second stop: the vote is arithmetic. Who may vote is determined by a membership registry; the quorum and the approval threshold live inside the contract. Ordinary decisions may pass by simple majority while existential ones, dissolution, constitutional change, are bound to higher bars. What we call constitutional rigidity is a single line of configuration: which category of decision passes at what percentage.

Third stop: the appeal window. An approved action does not run at once; it is queued into a timelock where it waits, in public view, for a set time. The waiting is not ornamental: while the window is open, a wired arbiter, a court chip, can suspend the operation; the case is heard; the ruling either releases the act or annuls it. The appeal window has teeth. A hasty decision can be caught before its execution.

Fourth stop: execution belongs to everyone. When the window closes, *anyone* may execute the queued action, the founder, a member, a stranger, a robot. This odd-looking detail is in fact the keystone: if execution is nobody’s monopoly, then *blocking* execution is nobody’s power either. Keeping the promise no longer requires anyone; breaking it now requires defeating everyone.

And then, sealing it all, the ceremony: **the handover**. While a system is being built, every string is in the founder’s hands, there is no technical way around that. When construction ends, the founder signs one last signature: administrative power over every chip passes to the timelock fed by the assembly. The answer to “who controls what”, the control graph, is written into the system’s deed. From that moment the founder holds exactly what any member holds: the right to propose and the right to vote. That is the scene from the preface: *Only admin*. The first law of the system is that its maker obeys it.

Four stops and a ceremony. None of it is new science; all of it is running, tested pattern. What is new is that these have been made into an *institutional language*: assembly, treasury,

budget, registry, court, each a building block, and the blocks wired to one another. Founding an institution begins to feel like building a circuit.

Chapter 5

The Honesty Mark

Now comes the least glamorous and most important idea in this book. Brilliant ideas build systems; honest ideas keep them *standing*.

Every one of our building blocks carries a mark. The black diamond (◆): this block's rule is *physics*, violating it is not forbidden but impossible. A treasury's balance cannot go negative; a proposal without quorum cannot pass; an expired work order's bounty refunds by construction. The white diamond (◇): what this block says is *testimony*, a claim about the world beyond the chain, vouched for by someone appointed to vouch. "The goods were delivered." "The transfer arrived." "The school was actually built." The chain cannot know these things; it can only notarize someone's word for them.

Why does this matter so much? Because the most destructive lie in the history of technology is hiding the difference. Most sentences that begin "the blockchain guarantees..." sell a white diamond painted black, and at the first scandal, a justified wave of distrust takes everything down with it. A system that states its own limits, by contrast, can be trusted. Our rule is simple: every block's mark stays visible from the design table to the deployment document. When you found your institution you can *count* which of your rules are steel and which are word. In today's library, sixteen of the twenty-five blocks are black diamonds and nine are white, and that ratio is itself a design instrument.

The most instructive blocks are the ones that carry both marks at once. Consider the redemption desk that converts company scrip into real money: the moment a redemption is requested, the scrip is *burned*, supply shrinks, irreversibly; black diamond. The corresponding bank transfer closes with an officer's "sent", white diamond. Inside a single contract, physics ends and trust begins, and the borderline is drawn in code.

The mark must also be applied to this book's own claims. "Code will save humanity" is an unmarked assertion, and it should not be sold in that form. What this book defends is humbler and sturdier: we can now draw the *boundary* of what can be saved, what can be turned to steel, and what will remain human forever. The next chapters tour the black diamond's territory; chapter eleven is an honest journey into the country of the white.

Chapter 6

Can Corruption Be Engineered Away?

The short answer: no. The long answer: no, but the *arithmetic* part of it, yes; and the arithmetic part is larger than you think.

Some corruption is the darkness of the human heart; engineering cannot go there. But a startling majority of what we call corruption in daily life is not a matter of the soul but of *accounting*: money, authority, or a record has traveled a path it should not have. And paths are exactly what engineering is about.

Embezzlement is a conservation violation: more left the vault than entered it. On a conserving treasury, that transaction *cannot be written*. Overspending, as we saw, is a payment without a line: it reverts. The ghost vendor is an unregistered payee: no tender payment can flow to an address absent from the charter registry. The uncollectable fine is a broken wire between courtroom and vault: if the ruling carries its own enforcement, there is no wire to break. The rigged tender is an unsealed comparison: when bids live on-chain, the arithmetic of ranking is public.

A thought experiment: you donate to a relief organization. In today's world your money enters an account and disappears; at year's end you read a glossy report. In a coded institution your donation enters as an appropriation with its purpose written on the line: the school-construction line, the vendor registry, milestone work orders, escrows that release on delivery approval. The set of paths your money *can* take is fixed at the moment of the gift; there are no others. Your question stops being "did they steal my money?" and narrows to "who signed the delivery attestation?"

And that is precisely the honest point. Corruption does not vanish; it *migrates*. Lay steel pipes and the leak moves to where the pipes end, to the white diamond, the point of testimony. The inspector who certifies that the school really stands becomes the system's new weak link. But the migration bears two gifts. First, the attack surface *shrinks*: risks once scattered across a thousand clerks and ten thousand daily choices collect at a few countable points of attestation. Second, it becomes visible: who attested what, and when, is sealed forever; whoever lies, lies over their own signature. You cannot delete corruption. But you can corner it, and there is light in the corner.

Chapter 7

Starting Small

“Saving humanity” conjures great stages: parliaments, world summits, constitutional referendums. The road this book proposes runs the other way: it starts at a forty-person cooperative.

The first reason is moral: nobody should entrust millions of lives to an untested mechanism. The second is strategic: institutional technologies have always walked from small to large. Double-entry bookkeeping ran first on the merchant counters of Venice, and only later in imperial treasuries. The joint-stock company financed spice voyages before it financed the world. The technology of promises will take the same road: first the apartment building that wants its dues collected transparently, the farm cooperative that wants the harvest shared fairly, the three villages with a common irrigation fund, the software studio that wants milestones paid the second they are approved.

Two things happen in these small places. First, the mechanism is *tested*: quorums meet real people, budgets meet real spending, appeal windows meet real disputes. Second, and deeper: people acquire the *daily taste* of the unbreakable promise. The member who watches on her phone where the dues went; the shopkeeper who sees the tender’s arithmetic; the craftsman whose payment lands in the same second as the delivery approval, these people do not easily return to the old way. Trust is not built by decree from above; it is built by habit from below.

The third gift of starting small is *forkability*. An institution made of code can be copied and adapted like open-source software. The constitution a fishing cooperative in Chile distilled through four hard years arrives at an Anatolian village as a template: take it, rename it, set your own thresholds, run it. Good institutions stop being the secret of nations and become the library of humanity. Today that library holds some two dozen templates, village federation to software company, association to micro-state. Tomorrow’s library will be written by those who fork them.

And should the infrastructure ever collapse, war, disaster, a long darkness, the library takes on another meaning: portable civilization. In a solar-powered box, a registry, a treasury, an assembly, and a court: the skeleton a forty-person camp needs to begin again without surrendering to the strongest. Institutions are infrastructure, like water and power; the difference is that they now fit in a suitcase.

Chapter 8

One People, Two Systems

The oldest dead end in political argument is incomparability. “This policy reduces poverty!”, “No, it increases it!” Both sides may be right, because both sides are describing a future that has never been lived. History does not run experiments; you cannot make the same country live twice under two arrangements. Until you can.

One of the quiet gifts of coded institutions is this: because an institution is now *executable*, it is also *rehearsable*. In the laboratory we built the following experiment: twenty-four named lives, their births, talents, strokes of luck and misfortune identical to the die roll, were placed inside two different orders. The same people, the same luck, two systems. Eighty simulated years later the two worlds were photographed side by side: who stands where on the ladder of needs, which order lifted whom, which dropped whom. In another experiment, ten citizens lived by their own strategies inside a small state whose land registry, taxes, and rolls ran on genuinely binding rules.

The lesson to draw is not “this system won”, such models are shadows of shadows of reality, and must be presented as such; remember the white diamond. The lesson is that *the ground of the argument can change*. “In my opinion” can give way to “here is what the model shows, here are my assumptions, change them and see.” It is like the weather forecast: nobody knows tomorrow’s rain for certain, but we carry umbrellas by the forecast, and where the forecast fails we fix the *forecast*, we do not go back to the fortune-teller.

To rehearse a constitution before living inside it, to try a tax rate, a quorum, a budget rule first in simulation, then in a forty-person pilot, then in a town: that is the step that moves politics from a contest of conviction toward a discipline of engineering. It will not end the wars of ideology; but it gives them a shared laboratory bench. Losing at the bench is cheaper than losing in the square.

Chapter 9

When Amendment Is a Patch

The most load-bearing rule of any institution is the rule for changing the rules. An institution that cannot change dies; an institution that changes without restraint becomes the toy of the strong. In the age of paper this tension was never resolved: constitutions either hardened into stone or were rewritten every morning by power.

In a coded institution, change has a natural form: the *patch*. Programmers have lived with patches for half a century: a signed, versioned, reversible description of what is to be added, removed, and altered. An institutional patch is the same: “add a central bank to this system, these two blocks, these three wires, these parameters.” The genius of the patch is that it is *typed*: it is written not against one named system but against kinds of blocks, “applicable to any system with a treasury and an assembly.” One community’s invention becomes, overnight, an option for every institution that speaks the same language.

And who applies the patch? Here the two worlds meet: the patch is adopted *by vote*. The amendment proposal is the patch itself; the assembly approves, the appeal window runs, the patch rolls onto the rails. Constitutional amendment and software upgrade become two names for one mechanism. And humanity gains, for the first time, a shared shelf where inventions of governance are versioned and exchanged, a GitHub for constitutions. An island cooperative’s recall procedure, a mountain village’s schedule of graduated sanctions, a city council’s participatory budgeting pattern: fork it, adapt it, vote it, run it.

It would be naive not to see the danger: the patch is also the shortest road to a constitutional coup. An invisible power hidden on line forty of a patch that promises school funding, a silent back door. That is why the patch regime has one non-negotiable companion: the *verifier*. Every patch, before application, is checked against institutional invariants. Does quorum remain reachable? Does any single actor, founder included, become able to drain the vault? Does every power keep its appeal path? Is an ownerless administrator being born? In the age of artificial intelligence these questions burn hotter, because soon machines will draft patches as often as people do. Whoever writes the draft, mathematics must hold the sieve: *unaudited intelligence does not touch the constitution*.

Chapter 10

Justice in a Single Transaction

The least discussed scandal of the world's justice systems is this: winning the case is not the same as getting your due. Everywhere on earth, courthouse corridors are full of people holding favorable judgments and facing emptied accounts. The verdict is a piece of paper; collection is a second war, and that war is usually won by whoever can afford the longer siege. Inequality of access to justice grows widest at the gate of enforcement.

In a coded institution the two acts collapse into one scene: *the judgment is the settlement*. Within the very transaction in which the court chip rules, damages transfer, the license suspends, the record is annulled. The moment a case is admitted, the disputed sum is frozen, the defendant cannot drain the vault while the hearing runs; the judgment is born collectible *by construction*. The appeal period is our familiar timelock: the ruling queues, the window stays open, the higher court can act, and when the window closes enforcement runs of itself. The poor plaintiff's judgment and the rich one's execute in the same second, with the same certainty.

The price of this beauty is the heaviest warning in the book: such a court is the most powerful contract in the world. Whoever captures it captures every vault, every license, every record within its jurisdiction. In the paper world a court's power is softened by delay, objection, friction; in the coded world there is no friction. The counterweights are therefore not ornaments but conditions of existence: judicial terms and elections; layered appeal; jurisdiction granted only by consent, at registration, by declaring "I recognize this court as an authorized executor over these assets", and scoped to those assets alone; and changing the court itself set as the decision with the highest threshold and the longest window in the whole constitution. Power that code concentrates, code must also dilute.

And one boundary must be drawn in honesty: single-transaction justice covers the consequences the chain *holds*. "The defendant shall publicly apologize" remains in the country of the white diamond. Physics can transfer the damages; only a human can offer the apology.

Chapter 11

What Code Cannot Save

This is the most important chapter in the book. Because those who believe in a tool's power sink with the tool on the day they cannot state its limits.

Code does not know persons. What the chain recognizes is a keypair, and a keypair is not a human being. Does the citizen who loses her key lose her citizenship? Where do a dead man's shares go? Guardianship, inheritance, and the Sybil attack that hollows out one-person-one-vote with a hundred wallets, none of this can be solved purely on-chain; all of it requires an identity layer with recovery and attestation, which is to say, *trusted people*. The system's most human question, "who are you?", is a white diamond forever.

Code does not know mercy. Remember the scene in *The Merchant of Venice*: the bond is clear, the penalty written, a pound of flesh. Perfect enforcement can be perfect injustice. Life always arrives in forms the strict rule did not foresee: a payment late but righteous, a debtor broken but worth saving. This is why we deliberately built human valves into the mechanisms: the appeal window, the power to suspend, the arbiter. These are not defects of the system; they are its *ports of conscience*. You do not want a justice made entirely of steel; be wary of those who do.

Code cannot manufacture legitimacy. A perfect system that no one believes in governs nothing. The movie scene is wrong: people do not obey a machine; people obey *what they decided together*, the machine merely holds the decision. The last line of our film belongs here: the system stands not because no one *can* break it, but because everyone *knows* that no one can. Knowing is social; code cannot compel knowledge.

And code cannot create the wanting. You can pour the rules into code; you cannot pour in the will that wants them. The desire for justice, the sense of a common good, the worry for those not yet born, these are the mechanism's inputs, never its outputs. That was Ostrom's lifelong lesson: institutions live not on paper but in the way people look at one another. Code can carry that look, protect it, make it cheap to extend; it cannot produce it.

So let us put this book's claim in its final form. Code will not save humanity. Humanity saves itself, as it always has, with its institutions. What code does is multiply the *holding strength* of those institutions a thousandfold: a ground on which a given word cannot be broken, power can bind itself, and trust leans on mechanism rather than on men. A foundation does not save a building. But no building stands on rotten ground.

Chapter 12

An Invitation to the Age of Repair

Hammurabi wrote his laws in stone so they could not be altered. We can write ours in code so they cannot be broken, and so that we may alter them only together. In the three thousand years between the stone and the code, humanity renewed everything else; the technology of the promise is what remains.

The road map is unglamorous. First, *simulation*: draw your institution, rehearse it, break it, mend it, with no one harmed. Then the *pilot*: a forty-person cooperative, real money, real rules, low stakes. Then *federation*: pilots finding one another; a common ground where every community runs its own validator and no one can write history alone. And only then, public scale. The bar at every stage is the same, and it is not up for negotiation: *Is the founder provably unable to override the system? Does a vote provably cause the voted action? Is the budget provably impossible to exceed? Does a ruling provably carry its consequences? And are the rules for changing all of the above themselves on the chain?* A system that cannot answer five yeses is, whatever its name, a notary.

Who is this invitation for? For the engineer: come cut stone, every missing block in the library is a broken promise in somebody's life. For the community leader: be a pilot, your building, your cooperative, your association; small but real. For the researcher: verify, mechanisms deserve trust only after they have been attacked in earnest. For the lawyer: build the bridge, let the decision on the chain become, through your translation, a judgment in a courtroom. And for the citizen, for all of us: learn to ask the one question. When an institution says "we have decided," ask: *is that a record, or a rail?*

The age of broken promises is not a fate; it is an infrastructure problem. And infrastructure gets repaired, it always has: we laid the sewers and the plagues receded; we strung the grid and the darkness receded. Now comes the invisible infrastructure. I do not know everything that will recede when we lay a ground that can carry our promises; but one needs no prophet to imagine what humanity might do in a world where a word, once given, is kept.

The promise can now be unbreakable. The rest is up to us.

Sources and Inspirations

This book stands on the shoulders of a library rather than of any single scientist. Elinor Ostrom’s lifework on the commons (*Governing the Commons*, 1990) and the grammar of institutions she wrote with Sue Crawford; Douglass North’s economics of institutions as “the rules of the game”; Lawrence Lessig’s *Code and Other Laws of Cyberspace* (1999) and its warning that code is law; Primavera De Filippi and Aaron Wright’s *Blockchain and the Law* (2018); Nick Szabo’s anticipation of smart contracts (1997); Satoshi Nakamoto’s nine-page paper (2008); Amy X. Zhang and colleagues’ PolicyKit (2020); and Nathan Schneider’s *Governable Spaces* (2024).

Every mechanism in this book, the budget that cannot be overspent, the governance handover, the appeal window, the honesty mark, amendment as a typed patch, exists in the running code of the STOA project; the technical account is given in the companion paper (*The Budget That Cannot Be Overspent*) and the project documents. The errors here are the author’s alone; their mark is $\diamond$, and they await correction.