

Civilization Leap

Volume One

The Civilization of Being Needed

English Translation

Derived from R10.10

Zijun Fu

Civilization Leap Research Group

2026

About This English Edition

This file contains the complete English translation of Volume One of Civilization Leap, derived from the Chinese source v1.21-R10.10. The shared front matter of the five-volume work is reproduced in full, followed by this volume and its relevant material notes. The original numbering and cross-references are retained. The companion volume, Shared Material Notes and Further Reading, preserves all sixty-two material notes and the upgraded public further-reading pages. This separately readable volume is not a combined edition of all five volumes.

Contents of This Volume

Shared Front Matter

Chapter 1 · A History of Being Needed: When Existence Is Asked to Prove Its “Usefulness”

Chapter 2 · Reinventing “Value”: When Function No Longer Suffices to Define a Person

Chapter 3 · The Algorithm of a New Civilization: When Results No Longer Suffice to Prove What Is Right

Chapter 4 · From Competition to Contribution: Changing the Structure of Our Encounters

Chapter 5 · Recognizing and Inheriting Meaning: Keeping Genuine Effort from Being Erased by Time

Chapter 6 · Structures of Symbiosis: When Local Rationality Begins to Threaten Shared Survival

Material Notes for This Volume

Shared Front Matter

Civilization Leap

The Path of Civilization in the AGI Era

Five Volumes

By Zijun Fu

Civilization Leap need not eliminate self-interest. It need only change the default path by which self-interest realizes itself.

Let the fulfillment of others become a way of fulfilling oneself.

Let effort find an echo and contribution never sink unseen; let basic rights need no proof, authorization stay within its bounds, and responsibility have someone to bear it.

Freedom from Survival Anxiety → Happiness in Life → Greater Meaning in Life

Produced by the Civilization Leap Research Group

Copyright Page

Title: Civilization Leap: The Path of Civilization in the AGI Era

Author: Zijun Fu

Produced by: Civilization Leap Research Group

Year of publication: 2026

DOI: 10.5281/zenodo.18814225

License: CC BY 4.0

Official website: civilleap.com

Current gateway: www.civitas.top

Contact email: zijunfu@civitas.top

Copyright notice: Copyright © 2026 Zijun Fu. All rights reserved.

Suggested citation: Zijun Fu, Civilization Leap: The Path of Civilization in the AGI Era, 2026.

Version: v1.21-R10.10 | Targeted Revision for the Whole Book's Intellectual and Living Conclusion | T-10
EMB-2 Cognitive Starting Point and Cross-Reference Synchronization Edition

Reading note: This is a book of thought. Its main text allows scenes, metaphors, strong judgments, and rhetorical compression to bring a problem home first. Facts, quotations, historical material, theoretical judgments, and future thought experiments must nevertheless retain their distinct identities. Composite historical scenes do not impersonate actual cases; future extrapolations do not impersonate present facts; literary metaphors do not impersonate empirical laws. Necessary qualifications are placed later wherever possible, to support the thought rather than flatten it before it has had a chance to arise.

Note on stages: Where the book discusses stages such as initiation, transition, stability, and self-maintenance, and phases such as normal operations, construction, resource constraint, crisis, and recovery, these should be understood separately. Stages and phases do not automatically change basic rights, sovereignty, or the sources of authorization. Specific parameters can be designed only after actual actors, jurisdictions, and real-world conditions have been established.

R10.10 whole-book completion note: All areas frozen in R10.9 and the preceding revisions R10.2–R10.8 remain in force. This edition implements only the six targeted P1 revisions decided in the final whole-book review: the historical-to-contemporary bridge in Chapter 1; the naming of Section 2.10; Sections 3.1, 3.10, and 3.11; Sections 5.3, 5.6, and 5.10; Sections 7.5–7.9; and the opening of Chapter 15 and the closing structure of Section 15.11. The aim is not further literary embellishment, but removal of defenses, old theoretical tones, technical catalogs, and governance memoranda still standing between readers and thought. No frozen boundary concerning subjects, rights, AI, continued existence, the three goals, stages and phases, memory, contribution, or institutional force is rolled back.

Statement on AI Assistance

Generative artificial intelligence was used as an auxiliary tool in organizing research materials and leads to cases, comparing structures and versions, developing counterarguments, revising language, flagging matters for factual verification, and producing documents. AI also assisted with some drafting and consistency checks, with continuing review, selection, and revision by the author.

AI does not serve as an author, factual source, evidence, professional opinion, or authoritative endorsement. The author bears responsibility for the core theories, value judgments, institutional proposals, selection of citations, and final expression. Facts, data, quotations, and cases are governed by verifiable original sources

and professional research. Material not sufficiently verified should be understood as hypothesis, extrapolation, thought experiment, or material awaiting validation. AI assistance does not replace peer review, professional scrutiny, or authorization and responsibility within actual jurisdictions.

Contents of the Complete Work

Overview of the Five Volumes' Timescales

The centuries, millennia, and ten millennia invoked here distinguish scales of observation; they are not a timetable of future events. Volume One addresses present ruptures. Volume Four extends the timescale across generations and into the long term. Volume Five carries the questions further into the conditions of life and meaning.

Volume One · The Civilization of Being Needed—A Crisis of Meaning on a Century Scale

Bridge between volumes: If Volume One's problem holds—if the old civilization truly can no longer generate a place, a path, and a future for enough people—the next step is not to keep finding more cases, but to find a new language for civilization. Without a new language, civilization cannot organize a new reality.

When a reality can no longer be fully described in the old language, continuing to amend the old institutions is like trying to find a newly emerged continent with an old map.

What truly makes the leap is not merely institutions, but civilization itself.

Volume Two · Pre-Institutional Language—The Foundations of Governance on a Century Scale

What truly makes the leap is not merely language, but civilization itself.

Volume Three · Civilization Leap White Paper—A Blueprint for Civilization on a Century Scale

Bridge between volumes: Any institution that can operate only for a few decades remains a proposal for an era, not a proposal for civilization.

All the institutions discussed in this volume—contribution values, resource pools, the meaning economy, and governance mechanisms—are institutional expressions of Civilization Leap, not Civilization Leap itself.

Institutions are tools; civilization is the protagonist.

What truly makes the leap is not merely the blueprint, but civilization itself.

Volume Four · Civilizational Perpetuity—Civilizational Safety on a Millennial Scale

Bridge between volumes: Civilization needs to endure not because more time means greater glory, but because meaning can continue to arise only while the future remains open.

What truly makes the leap is not merely continuity, but civilization itself.

Volume Five · The Emergence of Meaning—The Weight of Life on a Ten-Millennial Scale

General Preface

Statement of level: This book addresses the institutional level of civilization. It is not cosmology, the dynamics of meaning, or a practical operating manual.

For meaning's ontological foundations and higher-level theory, see Emergent Meaning Cosmology and The Dynamics of Meaning. This book addresses the leap at the level of civilization.

Cognitive Starting Point

The cognitive starting point of this framework: The emergence of AGI has not overturned earlier philosophy. It has, for the first time, brought into a complete separation within actual technology concepts that earlier philosophy long assumed together without adequately distinguishing them: intelligence, consciousness, subjecthood, life, finitude, and meaning. Meaning is the root; its sources may differ. Finitude is not meaning's only source, but the structural foundation through which meaning enters reality and acquires irreplaceable weight. For human beings, space and time constitute life; for AGI, space and time are first parameters. It processes space and time, but has not been shown to experience them as a finite life. Constraints enable AGI to bear operational responsibility, but what constraints produce is bounded operation, not finite existence. Until there is evidence that AGI possesses subjective experience, nontransferable continuity of the subject, and awareness of consequences as its own, this framework positions AGI as an understander, protector, and transmitter of meaning, and an extender of capability. It does not recognize AGI as a subject of meaning emergence. Humans and AGI are not competitors: Humans generate and bear meaning through finite lives; AGI lets human creation and responsibility reach farther.

Theoretical source: T-10, Space and Time Endure, Finitude Is Irreplaceable—A Statement of the Cognitive Starting Point, V3.0. The full, concise, and one-sentence public versions present the same proposition at different levels; they are not competing versions.

First overarching principle: Civilization Leap does not eliminate self-interest; it changes the default path by which self-interest realizes itself. When a civilization begins to see whom its power may harm, to recognize which mistaken paths require a return, and to be willing to leave a future for those not yet born, it is no longer merely pushed along by history. It begins to learn to observe, choose, and correct itself. "Agency" here describes this awareness at the scale of civilization; it does not personify civilization as an institutional subject independently able to hold rights, bear obligations, or acquire governing sovereignty.

Written as Civilization Becomes a Self-Aware Existence

For most of human history, civilization never realized it was becoming civilization.

Tribes organized only to survive;

states were founded only to secure order;

industry expanded only for efficiency;

technology developed only to solve problems.

History seemed driven by countless local choices, without an overall direction.

Until the twenty-first century.

For the first time, humanity created a form of intelligence that might surpass its own cognitive capabilities.

At the same time, humanity acquired, for the first time, technological power capable of affecting the fate of human civilization as a whole.

These two developments together brought a new question into being:

Can civilization understand itself and actively decide its future?

Civilization Leap was born at this historical juncture.

It is neither a book about technology nor a political program.

It is an attempt—

an attempt to strengthen human civilization’s capacities for self-reflection, self-restraint, and long-term continuity. This capacity of the whole does not replace the subjecthood, rights, or final decisions of actual people.

I. Civilization’s Three Awakenings

Looking back, human civilization has undergone three fundamental awakenings.

The first awakening: Consciousness

Language and symbols enabled human beings to think about themselves.

Humanity was no longer merely part of nature, but a being that understood nature.

The second awakening: Institutions

The emergence of law, science, and the modern state enabled society to function beyond individual lifespans.

Civilization began to acquire stable structures.

The third awakening: Civilization

The arrival of the AGI age brings civilization, for the first time, face to face with its continued existence as a whole.

Humanity must answer:

Why do we exist?

Where are we going?

Can we endure over the long term?

Civilization Leap is written amid this third awakening.

II. The Structure and Mission of the Five Volumes

This series is not a single treatise, but a framework of thought about civilization unfolding layer by layer.

Volume One: The Civilization of Being Needed

Posing the problem:

Industrial civilization has long bound labor, income, social position, and “being needed” closely together. AI and automation are changing the scarcity and organization of some of that functional labor.

If people should no longer have to prove their entitlement to survive and their place in society repeatedly through functional labor, how can civilization understand human value anew and rebuild enduring places for participation, relationships, creation, care, and responsibility beyond labor?

This volume reveals:

Civilization must undergo a structural leap.

Volume Two: Pre-Institutional Language

Addressing the problem:

Before any institutional transformation, relationships easily conflated must first be made clear.

This volume separates subject and role, rights and contribution, facts and evaluation, capability and power, authorization and execution, and professional judgment and public decision, among other relationships. This allows different disciplines, institutions, and civilizations to discuss the same real-world problem without first accepting the same theory.

Volume Three: Civilization Leap White Paper for the AI Age

Entering practice:

Turning the relationships distinguished in the first two volumes into institutional prototypes and governance boundaries that can be verified, can fail, allow exit, and can be corrected.

The emphasis is not on declaring that institutions of a new civilization already exist. It is on explaining how any real-world proposal acquires force, how rights and authorization constrain it, and how it can be stopped, rolled back, and corrected when it goes wrong.

Volume Four: Civilizational Perpetuity

Taking the next step:

From “how to make the leap” to “how to endure.”

Introducing:

The Common Window: anti-slippage safeguards, rollback, and circuit breaking;

Protection of the conditions in which meaning arises: preventing safety, efficiency, and long-term stability from swallowing the subject;

Two-sided civilizational immunity: AGI-COS addresses machine-side safety; Human-COS addresses humans’ high-consequence judgments;

Self-correction in a long-lived civilization, an open future, and a cosmic-scale perspective.

This volume extends the timescale and asks how successful institutions can avoid becoming permanent sovereignties. The interstellar remains a distant open direction, not a rank of civilizational maturity.

Volume Five: The Emergence of Meaning

Continuing the inquiry:

When civilization can endure longer, how can different lives retain real space in which to form meanings of their own?

This volume uses longevity, digital life, virtual environments, and future subjecthood as stress tests. It answers the question, “After continued existence is secured, how can life still form its own reasons to continue?”

Distant imaginings such as reincarnation and multiple worlds remain optional thought experiments, not the main line or a mandatory route.

III. What Civilization Leap Truly Means

Civilization Leap is not a passage from an old world into a perfect new one.

It is not a utopian plan.

Nor does Civilization Leap aim to attack anyone. What it seeks to loosen is not the power of one class of people, but a default path on which self-interest can be realized only by surpassing others, controlling them, or passing costs onto them. Ordinary people can be crushed by survival pressures; those who command resources can also be trapped by control, competition, anxiety, and distrust. Wealth and power can bring resources without necessarily bringing happiness.

What Civilization Leap truly points toward is greater freedom, responsibility, meaning, and peace for every person beyond the old logic.

The leap truly means:

Civilization no longer merely endures change passively. It begins more consciously to observe causation, constrain high-consequence capabilities, choose paths, and preserve space for its own correction.

This change involves three central shifts:

From zero-sum competition as the default path → competition remains, while cooperation and symbiosis can also become practical paths to self-fulfillment;

From growth overriding everything → growth is subordinated to continued existence, life, and the long-term future;

From colliding local rationalities → broader global perspectives, systemic observation, and structural scenario analysis at the scale of civilization.

IV. The Place of AGI and Civilization

AGI gives civilization an unprecedented capacity to extend cognition beyond itself.

For the first time, it makes a real question this acute:

When analysis, memory, simulation, and extrapolation can expand at machine scale, the boundaries of individual cognition no longer define the boundaries of the cognitive capability civilization can call upon.

But power itself does not decide the future.

AGI may become either the mechanism of civilization's end or its immune system.

The key lies not in technology, but in structural choices.

This book therefore consistently upholds one principle:

Capability can be delegated; purpose and sovereignty cannot transfer automatically to whatever is more capable.

AGI can analyze, advise, and execute within explicit authorization. Basic human rights, legitimate public decisions, and ultimate boundaries of value must still have actual actors, legal authority, procedures, and sources of responsibility.

V. Why Thought on the Scale of Civilization Is Needed

History's great turning points have rarely been accomplished by thought alone. They have also rarely occurred without thought changing at all.

New ideas open new possibilities. Institutions turn possibilities into conditions for conduct. Real experience then determines whether people can accept, internalize, and carry them forward.

The relationships among Enlightenment, science, human rights, the state, and industrialization form no simple straight line in which "thought precedes reality." They involve long processes of mutual formation.

What matters is this: Without thought able to describe new realities and propose new directions, civilization struggles to choose actively the institutions and lives it is about to enter.

The AGI age likewise needs thought to open possibilities.

But the spread of ideas cannot replace institutions, conduct, real experience, and reproduction across generations.

Civilization Leap aims first to place another possibility clearly before humanity.

VI. Three Goals: Civilization's Anchors of Value

The three goals running through this book are not a list of equals side by side, but a progressive path for civilization:

Freedom from survival anxiety—the first goal above the baseline of continued existence, freeing basic survival from the rule of scarcity and fear;

Happiness in life—unfolding freedom, stability, and real life above that baseline;

Greater meaning in life—keeping the future open for every life to explore, choose, create, and reach the world.

Technology answers “How can it be done?”

Institutions answer “How will it work?”

These three goals answer:

Why it is worth continuing.

VII. To Future Readers

Perhaps future humanity will have moved beyond today’s disputes.

Perhaps these words will be only traces of civilization’s early explorations.

But if civilization still exists, if the future remains open,

then this can show:

One generation, facing uncertainty, chose restraint and cooperation.

They did not wait for inevitability.

They actively took responsibility for civilization.

VIII. Civilization Becomes Self-Aware

For most of human history, civilization was more the accumulated outcome of countless actions, institutions, conflicts, and choices.

But as technology places such powerful capacities for self-observation, simulation, and design in human hands for the first time,

civilization also gains better conditions for consciously reflecting on its long-term direction.

“Self-awareness” here does not personify civilization as a supersubject possessing independent sovereignty. It means enabling humanity to observe, discuss, choose, and correct its path at the scale of civilization.

Reflect upon itself;

constrain its power;

preserve a future for those not yet born.

This is the mark of a mature civilization.

IX. A Final Word

This book belongs to no single person, country, or age.

It belongs to a question:

Can human civilization endure over the long term?

The answer will not be written in the book.

It can be written only in the future.

If the future still exists,

then Civilization Leap has happened.

Civilization Leap

End of the General Preface

X. A Reading Map: From Existential Crisis to an Interstellar Perspective

These are not books to skip through. Each volume lays the foundation for the next.

By the end of Volume One, you will see the old covenant becoming unstable. The moat built around “only those needed by the system deserve a place and security of survival” can no longer explain the new reality.

By the end of Volume Two, you will see the framework of a new language: observable, anchorable, trustworthy, distributable, and teachable.

By the end of Volume Three, you will have a blueprint for action: from freedom from survival anxiety to an architecture of meaning, from the Dream Realization Park to an interstellar window.

By the end of Volume Four, your perspective will rise from Earth to the stars. Civilization is no longer the fate of a single nation, but a rare order in the universe worth continuing.

By the end of Volume Five, you will ask further: When civilization can endure, how can life avoid being diluted by indefinite extension, and how can meaning continue to acquire weight?

If you read only one volume, read the first. It will tell you why change is necessary. If you finish all five, you will understand that change is not the destination, nor is continuity. Keeping civilization in existence, while allowing life to become increasingly meaningful within that continuity, is the direction in which Civilization Leap truly unfolds.

Volume One • The Civilization of Being Needed

This volume answers the following questions: As the old bond between labor, income, a place in society, and being needed begins to loosen, why is the old logic alone no longer enough to sustain human agency, rights, and a future? Why must civilization change?

The causal scope of this volume: Volume One explains why the old language of value and social position is no longer sufficient. It does not undertake in advance the shared language of Volume Two, the institutional blueprint of Volume Three, or the parameters of specialized engineering. Concepts such as “value redshift,” the “contribution algorithm,” “winning benefits the whole,” “civilizational memory,” and the “carbon–silicon covenant” should be read as philosophical metaphors, frameworks for posing problems, or directional hypotheses. Volume One alone cannot give these concepts institutional force; nor may the concepts be used in reverse to generate basic rights, governing authority, or entitlements to deploy resources.

Chapter 1 · A History of Being Needed: When Existence Is Asked to Prove Its “Usefulness”

“If a society needs you only while you are useful, the truly frightening thing is not merely unemployment. It is that one day you may wonder: Do I still have a right to remain?”

—This book, inspired by Hannah Arendt’s reflections on labor, work, action, and the making of human beings “superfluous”

“A person becomes a person through other people and relationships.”

—A paraphrase of the familiar Ubuntu expression “Umuntu ngumuntu ngabantu”

1.1 The Primal Covenant: The Brutal Moat Around Survival

An Endgame in the Igloo: When Thermodynamics Passes Judgment on Life

The Arctic winter night compresses the world into the sound of wind, darkness, and a little lamplight. Grains of snow skim sideways across the ice; nothing can be seen in the distance. In such a place, life is first a matter not of philosophy but of warmth, food, and whether one can survive the next blizzard. [1]

Inside a low snow house, a stone lamp burns marine-mammal oil, its light barely enough to illuminate a few faces. Damp fur, grease, and cold mingle in the air. An elderly hunter sits at the edge of an ice sleeping platform. His teeth are worn, his sight is failing, and he can scarcely take part in a long hunt anymore.

He has not suddenly lost all his meaning as a father, a grandfather, a friend, and a person. But when food and fuel are stretched to their limit, another, colder question forces its way into the room: If I can bring back less and less, while needing more and more, may I still stay?

What is truly frightening is not whether someone actually walks toward a particular door, but why that door appears at all. When resources for survival grow so scarce that every unit of heat must be counted, love, responsibility, and resources can be forced into a painful entanglement. Whether a life is needed by the system begins to approach a crueler question: whether that life may remain.

The Memory of Scarcity: Why We Still Fear Being “Useless”

The long experience of scarcity can explain some of today’s anxiety without being reduced to a single biological story of something “etched into the amygdala.”

When someone says, “I’m afraid society will discard me,” that fear often encompasses income, relationships, identity, and dignity at once. Modern anxiety cannot be reduced to a prehistoric scene, yet it still carries the deep, lingering echoes of societies shaped by scarcity.

Anthropological and social-scientific research on reciprocity, the division of labor, and group relationships reminds us that survival in early communities long depended on mutual reliance. But this does not mean everyone had to pay a survival tax in some quantifiable form of “usefulness.” Real communities have always been more complex than a single model of exchange.

Nor is pay the only reason modern work reaches so deeply into our sense of self. An occupation also carries identity, rhythm, relationships, and the experience of being recognized. Losing a job may mean losing all of these at once. The suggestion that “I am no longer needed” can then travel from an economic problem into an existential one.

Collective Memory Across Cultures: The Shadow of Tales of Abandoning the Elderly

Such stories wound us because they place an extreme question before our eyes: When scarcity reaches its limit, will a society tie the question of whether someone may stay to whether that person is still useful? History offers more than one answer, but the question has never truly left civilization. [2]

Our praise of diligence and suspicion of idleness have economic and institutional backgrounds, as well as cultural and moral traditions; they cannot be reduced to a single energy ledger. Yet the old civilization did long bind survival, labor, and social membership tightly together.

A Turning Point in the Logic: The “System-Level Redundancy” Brought by AI

The crisis we need to discuss, however, is this: What if one day this “energy ledger,” in use for tens of thousands of years, suddenly ceased to apply?

What happens if AI and automation keep lowering the marginal cost of more and more kinds of functional labor—if some forms of “usefulness” no longer have to be supplied by human sweat?

Here lies the crisis: We retain the old institutional intuition that “those who neither work nor contribute shall not eat,” even as we enter an environment in which technology is rewriting the value of some kinds of labor. The world outside is changing, but the inner moat built around the belief that “only the useful deserve to remain” has not disappeared with it.

1.2 Status, Relationships, and Contract: One Thread of Historical Change

Agriculture, storage, and long-term relationships changed something very simple: A person’s place did not have to be justified from scratch every day.

Settlement, cultivation, domestication, and storage followed different historical paths in different regions, and did not automatically create more merciful societies.

What they did bring was a longer timescale. Some of today’s output could be kept for tomorrow; a shortfall in one season could be made good in another. Families and communities could more readily develop divisions of labor extending across time.

This does not mean that hunter-gatherer communities judged a life solely by its immediate output, or that agrarian societies inherently respected the vulnerable.

But when life was situated within land, seasons, family, and enduring relationships, doing a little less for a while often did not mean immediately losing one’s entire place in society.

Status thus became a complicated form of protection.

Kinship, villages, religion, and local mutual aid could absorb risks in some societies; they could also produce hierarchy, dependency, and exclusion.

Status could spare someone the need to prove their usefulness every minute. It could also lock a person, from birth, into a place that others had chosen for them.

Henry Maine once described a line of legal development as a movement “from status to contract.” [3]

What is worth carrying into the present is not a return to a society of inherited status.

It is the extraction of one function once imperfectly borne by kinship, hierarchy, and dependency: When a person temporarily loses a function, they should not also lose their place from which to go on living.

The way freedom genuinely advances beyond a society of status is not simply that people may leave old relationships.

It should also mean that, having left, they do not immediately fall through the ground of human belonging.

1.3 The Industrial Revolution: Turning People into Instruments and Reconsidering the “Luddite Fallacy”

In the textile districts of Britain in 1811, machines first gave some people’s fear a metallic sound.

The Luddites smashed machines for reasons beyond a failure to understand technology. The more immediate conflict was that new machinery and new forms of production were changing their craft, income, bargaining position, and expectations of life.

Industrialization did not turn all labor into assembly-line work; nor does its history make preindustrial life a picture of warmth and kindness.

What it amplified was a relationship still familiar today: As work is broken down into measurable, replaceable, standardized tasks, a person’s place in society becomes more liable to fluctuate with the price of a function.

The image of Chaplin’s body caught in the gears in *Modern Times* remains powerful not because every factory looked like that.

It remains powerful because it magnifies a feeling until everyone can see it: In the pursuit of efficiency, might a person’s value gradually shrink to “how much can be completed in this minute”?

Technological progress did subsequently keep creating new occupations, while eliminating and reorganizing old ones.

So both “technology inevitably causes permanent unemployment” and “technology always painlessly creates better jobs” are too simple.

The real questions are whether those whose old skills lose value can find a way back in as change unfolds; whether new jobs still offer income, dignity, relationships, and growth; and who ultimately bears the costs of transition.

Ford’s high-wage policy of 1914 is often used to illustrate the relationship between production, employment, and mass consumption. Its immediate historical context included high turnover and the pressure to recruit and retain workers. It cannot be compressed into the single story of “enabling workers to afford a Model T.”

Yet that history still left an important structure: Industrial society once bound wages, consumption, family expectations, and a place in society tightly together.

Work, then, was more than work.

It became a gateway to income, and also proof of adult status, the basis of family plans, and evidence that “I still have a place in this society.”

This long-assumed connection is precisely what AI now touches.

1.4 The Great Decoupling: The Cognitive Rupture from “Worker” to “User”

From the latter half of the twentieth century onward, the apparently sturdy rope linking economic growth, wages, long-term employment, and social position began to loosen in many places.

In the United States, for example, labor productivity and the compensation of typical workers gradually entered a long divergence from the 1970s onward. The causes involve technology, industry, globalization, labor’s bargaining power, policy, distributive structures, and statistical measures. No single year or variable can explain them.

What matters is not identifying “who stole that curve.”

It is recognizing that greater efficiency in a society does not guarantee that the added value will automatically return to every worker’s income and standing.

At the same time, outsourcing, temporary employment, and platform-based work have broken some long-term relationships into shorter tasks.

Flexibility can give people freedom. It can also leave more of the waiting, provision for security, and risk to individuals. The difference lies not in the word “platform,” but in actual contracts, bargaining power, protections, and conditions of exit.

With the mobile internet, a person acquired another identity: user.

People not only purchased services; they continually left behind attention, preferences, behavior, and feedback. This does not mean that every click is secretly turned into training material for a machine that will replace them.

The real concern is that, as data uses, recommendation logic, and exit conditions become harder to discern, people may enjoy convenience while losing their understanding of how they are being classified, predicted, and influenced.

Being needed then undergoes a very strange change.

The system still needs your clicks, your data, and your consumption.

It may not equally need your judgment, participation, or enduring place within it.

This is the aspect of the so-called Great Decoupling that most concerns this book:

The person is still inside the system.

Yet they are increasingly liable to become, instead of a participant, an object to be served, predicted, and called upon.

This change is not a universally accomplished fact. New technologies are also expanding creation, entrepreneurship, collaboration, and new occupations.

But as long as a place in society depends increasingly on short-cycle ratings, accounts, and replaceable tasks, one question will keep returning:

If the system still needs my data but no longer needs my participation as a whole person, what exactly does it still need?

1.5 A Collapse of Assumptions: From Visions of a “White-Collar Apocalypse” to the Warning of a “Useless Class”

After 2022, many people saw directly for the first time that machines could rapidly enter the domains of language, code, and knowledge work too.

This did not prove that “intelligence has now lost its value,” much less that white-collar occupations had disappeared as a whole.

What it unsettled was a long-standing psychological assurance: that a better education and more complex cognitive skills would naturally place one farther from automation.

Generative AI has already substantially reduced the time required for some tasks involving text, code, and creative work. Businesses have also begun adjusting outsourcing, recruitment, and job structures.

These changes occur first at the level of tasks.

Whether a job disappears, an occupation contracts, or a person consequently loses income and identity depends on many further layers: organizations, markets, regulations, customer demand, the emergence of new tasks, and more.

Three things must therefore be distinguished here:

A task is performed more quickly by a machine.

A job is redesigned as a result.

A person’s place in society is left without support as a result.

The three are connected, but they are not the same statement.

When Harari proposed the “useless class,” what truly stung was not an established class that had already emerged, but a warning about the future: What happens if the system’s demand for some people’s labor and participation falls over the long term, while civilization continues to tie dignity, income, and social position entirely to the scarcity of functions?

The deepest danger may not be “having no work.”

It may be that a person begins to believe: No one truly needs me.

This is the rupture Chapter 1 seeks to recognize before it arrives.

1.6 “Deaths of Despair”: When Losing One’s Place Begins to Harm Life

This feeling of “not mattering” does not automatically lead to death. But the loss of prospects, relationships, and a place in which one is needed can compound economic pressure, illness, addiction, and isolation, becoming part of a public-health crisis.

The research of Case and Deaton documented a real and complex public-health shift: Between 1999 and 2013, all-cause mortality rose anomalously among non-Hispanic white Americans aged 45–54, with particularly marked increases in suicide and deaths related to drugs and alcohol. Subsequent research further revealed deep entanglements among education, economic circumstances, health, addiction, and social relationships. These findings cannot be compressed into a single causal chain in which “unemployment causes death” or “the loss of all social value causes physiological collapse.”

What makes them so painful is this: When shrinking prospects, physical illness, addiction, isolation, and prolonged disappointment accumulate, despair can enter the mortality statistics.

Several phenomena emerging in different societies cannot be equated, but can all help us examine “place, relationships, and expectations of the future”:

- In Japan, “lonely deaths” have long been a public concern. In recent years, cases of people living alone and dying at home without anyone knowing have numbered in the tens of thousands, with older people accounting for a substantial share.
- In the United Kingdom, research on austerity policies, welfare reform, and “deaths of despair” continues to accumulate. Drugs, alcohol, and suicide have become public-health warning signs.
- In emerging economies such as Brazil, youth unemployment, insecure employment, and mental-health risks compound one another, creating social pressures that cannot be ignored.
- Youth unemployment in South Africa has remained high for many years. In some groups of young people it has at times approached or exceeded half. For them, being needed has already become a luxury.

If AI broadens occupational disruption, knowledge workers may likewise face a rupture in identity, income, and the sense of being needed. This does not make it inevitable that deaths of despair will “spread from blue-collar to white-collar workers.” But they already offer a warning severe enough: A civilization that knows only how to redistribute tasks, without finding a renewed place for people in society, will drive the costs of technological transition into people’s bodies, relationships, and lives.

1.7 Keynes’s Prediction and the Paradox of “Bullshit Jobs”

In 1930, in *Economic Possibilities for Our Grandchildren*, Keynes predicted that by 2030 technological progress would allow humanity to work only fifteen hours a week, leaving the rest of its time for leisure and the arts.

With 2030 just around the corner, why have we sunk more deeply into busyness and anxiety?

The anthropologist David Graeber used “bullshit jobs” to criticize a modern predicament: Some workers subjectively regard their jobs as lacking real social purpose, yet organizational structures and systems of status and income keep those jobs in place. Graeber grouped such jobs into categories including flunkies, goons, duct

tapers, box tickers, and taskmasters. The concept cannot prove that any category of occupation is “objectively useless.” But it brings a long-neglected question to the foreground: If work is merely a container for maintaining status and income, why does it still leave people feeling hollow?

AI and automation will not automatically clear away every “meaningless job.” They may also create new positions, processes, and bureaucratic structures. The real concern is something else: When functional tasks are redistributed, the old role of work as a channel for income, identity, relationships, and recognition will be shaken with them. What humanity may then lose is not merely certain tasks, but an entire structure of social positions that the old institutions once supported through work.

1.8 The Great Resignation and the “Strike for Meaning”: When People Begin to Refuse a Path They Can No Longer Believe In

These phenomena cannot prove that the old occupational covenant is collapsing simultaneously across the world. They do suggest that, in different societies, some people are reassessing the relationships among work, competition, care, and life.

In 2021, the United States recorded nearly 48 million voluntary quits over the year. What matters most about the Great Resignation is not simply the number of quits, but that more and more people began reassessing the place work should occupy in a life. [4]

In Japan, prolonged social withdrawal has become an enduring public concern. In South Korea, terms such as the “sampo generation” have been used to describe some young people’s relinquishment of dating, marriage, and childbearing. In societies such as Spain, where youth unemployment has long been high, some young people respond to an uncertain future through gig work, delayed independence, or departure from traditional career paths. In sub-Saharan Africa, recent estimates from the International Labour Organization indicate that more than one in five young people is not in employment, education, or training. These different situations cannot all be explained as laziness. At the very least, they remind us of something in common: When an old path offers no believable future, people reassess why they should enter it at all.

In China, “lying flat” became a widely discussed online topic around 2021. Associated texts resisted excessive competition through provocative expressions such as “Only by lying flat can a person be the measure of all things.” Byung-Chul Han’s analysis of the “achievement-subject” and self-driven exhaustion offers another way of understanding this mood. Choices to withdraw, defer, or reduce the intensity of competition cannot all be interpreted as one political stance. Yet they share a question: If striving has become nothing but endless self-depletion, why should people continue striving in the old way?

Calling these different acts of withdrawal, resignation, deferral, and slowing down a “strike for meaning” is an interpretive metaphor, not the name of a unified social movement. It prompts us to ask: Under what conditions do people feel that the old path of striving can no longer offer a future worth believing in? The answers may differ completely from place to place.

1.9 The Limits of Material Security: What UBI Can and Cannot Solve

In the face of technological and occupational change, loosening the dependence of basic living conditions on fluctuations in employment is an entirely reasonable direction.

Finland’s basic-income experiment of 2017–2018 offers a limited real-world reference: When income conditions changed, the overall effects on employment were not dramatic, while some measures of subjective well-being and experiences of bureaucratic burden changed. [5]

It could not answer the question of life’s meaning.

This shows precisely that income security and the sense of being needed are different problems to begin with.

Money can broaden access to food, housing, medical care, and choices.

It does not automatically tell a person who needs them, nor automatically create relationships, responsibilities, and participation.

Basic provision is therefore not “keeping people fed like livestock,” and still less should it become something exchanged for political standing or personal dignity.

What must be added are ways in beyond material security: creation, learning, care, public participation, relationships, and a chance to reenter.

1.10 Conclusion: From “Being Needed” to “On What Grounds Do We Exist?”

Chapter 1 has not pursued a single linear history.

It has simply encountered the same question repeatedly in different social structures: The ways people connect income, identity, labor, relationships, and dignity can change.

Industrialization made wage labor an important route to a place in society for many people. Platforms and AI have begun to break apart some of its tasks, relationships, and forms of scarcity once again.

Technology alone will not determine the outcome.

But the old chain has loosened enough for civilization to prepare new forms of support in advance.

The most fundamental affirmation is this:

A person need not first prove their functional value to have basic rights, to be loved and protected, or to take part in a shared life.

The harder work lies ahead.

How can participation endure beyond labor?

How can care be made visible?

When a skill is no longer scarce and loses its market price, must a person’s experience be reduced to zero along with it?

An elderly person’s hand held at the end of life, a child’s first spontaneous smile, a joke that a stranger in a foreign land finally understands—none of these things becomes unreal because it is difficult to measure.

They may never enter GDP.

Yet they will determine why a civilization is worth continuing.

In the next chapter, we no longer ask, “What can people still do that machines cannot?”

We ask: When function cannot exhaust what a person is, where else does value arise?

Chapter 2 · Reinventing “Value”: When Function No Longer Suffices to Define a Person

“Everyone knows the use of the useful, but no one knows the use of the useless.”

—Zhuangzi, “In the Human World”

“The value of a river has never lain only in how many waterwheels it can turn.”

—This book

“When efficiency becomes the only measure, human beings will ultimately lose to more efficient machines.”

—This book

If one day a machine can do your work at lower cost, what exactly is lost? A wage? A position? Or an equation we accepted long ago: useful, therefore valuable?

2.1 Schmidt’s Shovel: The Mold in Which Industrial Civilization Cast Value

To understand why we feel such profound “disillusionment with value” in the age of AI, we can return to a classic story that shaped the modern managerial imagination: the pig-iron handling experiment at Bethlehem Steel.

In *The Principles of Scientific Management*, published in 1911, Taylor told the story through a worker he called “Schmidt.” Later historical research has repeatedly examined how much Taylor himself reshaped the account. The soundest reading, therefore, treats it not as a minute-by-minute recording of events but as a firsthand statement of how scientific management understood human beings.

In Taylor’s account, Schmidt originally handled about 12.5 tons of pig iron a day. By rearranging movements, pace, rest, and instructions, Taylor sought to raise the daily output of a so-called “first-class handler” to about 47 tons.

The key is not merely how much the number rose, but how control changed. The worker no longer relied primarily on his own experience to set the pace. Work and rest were observed, broken down, timed, and directed from outside.

Taylor even wrote that a man holding a watch would tell the worker at fixed intervals when to lift, when to walk, and when to rest.

According to Taylor’s own record, this arrangement substantially increased output, in exchange for higher daily pay. Whatever later assessments of the experiment’s representativeness, the story states a new managerial logic with exceptional clarity: Break human movements apart, measure and reassemble them, then redefine a “good worker” through the goals of the system.

Yet this managerial perspective amplified a danger at the same time: A human being could increasingly come into view as a configurable, measurable, optimizable “resource,” rather than a whole subject.

Under Taylor’s stopwatch, Schmidt was no longer a whole person with feelings, memories, and thoughts. He was reduced to a “92-pound pig-iron handling component.” His value was confined to the formula “tons handled / time.” Whether he loved his children or enjoyed singing was a redundant variable in the factory system—even “noise” to be subdued.

This established a default calculation with a deep influence on modern organizations: A person’s place could readily be compressed into “functional output / cost.”

The logic has profoundly influenced modern organizations, from assembly lines and logistics systems to knowledge work and outsourcing networks. Tools such as KPIs, OKRs, and balanced scorecards can

themselves help organizations clarify goals and responsibilities. The real danger arises when these tools compress complex people into measurable output and allow indicators, in turn, to define their entire place.

As long as this formula held—as long as a person’s functional output exceeded the cost of sustaining that person—the covenant was stable. Schmidt might be exhausted, but he knew the factory could not do without him, because a steam engine had no arms.

The arrival of AI does not simply take away a shovel. In more and more domains, it weakens the stability of the equation that “only human functional output can keep the system running.”

2.2 Ghost Work: Becoming a “Patch for the Machine”

During the transition in which AI continues to enter workflows, anthropologist Mary Gray and computer scientist Siddharth Suri, in *Ghost Work*, revealed a reality easily obscured by narratives of automation: Many supposedly “automated systems” still depend on large amounts of human microtasking and behind-the-scenes labor.

If Taylorism showed how people were brought into the rhythm of the machine, some digital backstage labor in the early twenty-first century revealed another danger: People were hidden behind an “automated” interface, becoming temporary patches for machine processes.

Consider Amazon’s Mechanical Turk, or MTurk, platform. Its very name is ironic: It comes from an eighteenth-century chess-automaton hoax. The seemingly self-playing mechanical Turk actually concealed a human chess player inside.

In the twenty-first century, tens of thousands of “crowdworkers” are that chess player hidden in the box. They handle small tasks AI cannot yet manage: identifying traffic lights in images, deciding whether tweets contain hate speech, or transcribing indistinct audio fragments. These are called “Human Intelligence Tasks.”

When workers are treated only as temporary patches for machine processes, the problem is more than low pay. It also concerns whether their standing as subjects is being institutionally diminished.

In Taylor’s era, Schmidt was still explicitly situated within the factory’s relations of production. In the age of crowdsourcing platforms, human labor is sometimes hidden behind an “automated” interface, reduced to microtasks the system calls upon as needed. Research by Gray and others reminds us that workers may train, review, and supplement models while facing the prospect that those tasks will themselves be further automated. The question is not whether to predict their inevitable elimination. It is whether institutions absorb people’s knowledge and time without simultaneously designing an enduring place for them and continuity of responsibility.

This phenomenon also reaches high-end knowledge work. In recent years, several model and data-services companies have begun recruiting human experts with doctorates or professional backgrounds to generate, evaluate, and improve high-quality answers for models. This shows that, as machine learning begins to ingest knowledge work, even reasoning and expression once exceptionally difficult to standardize can be broken into training steps. The risk is not that these experts “will inevitably be discarded.” It is that, if institutions treat them only as one-use training resources, their knowledge will be absorbed without equal attention to their long-term place.

This is the most dangerous version of the “illusion of participation”: A person believes they are participating in creation, only to discover that their knowledge is being treated as an expendable input in a training process. The problem is not that technology learns too quickly. It is whether institutions leave room for the knowledge absorbed, the responsibilities that endure, and the person’s next destination.

This kind of backstage labor has already formed transnational chains. Content moderation, data annotation, and platform gig work have real examples in cities including Nairobi, Manila, and Cairo. Pay, health risks, and labor relations vary considerably across regions; these cases do not justify declaring a uniform “global state of

alienation.” The more defensible shared concern is that many technological systems still rest on human time, attention, and psychological costs easily concealed by the automated interface.

2.3 Redshift in the Spectrum of Value: When Computation Is No Longer Scarce

If AI can perform more and more functional tasks at lower cost or greater efficiency, should human value continue to be defined by functional efficiency?

To answer this, we borrow a concept from astronomy: redshift. When a celestial body moves rapidly away from us, its spectrum shifts toward the red end—frequency falls and wavelength lengthens.

We can continue to use the “redshift of the value spectrum” as a metaphor here. But the change is not that “human value is forced to retreat.” What actually changes is this: As AI keeps entering computation, analysis, prediction, generation, and optimization, some functional outputs once overvalued because of their scarcity are repriced. Relationships, processes, experience, care, responsibility, and genuine presence, which have always existed but long resisted full recognition by markets and indicators, can then more readily reenter the discussion of value.

This is not about driving people into the gaps machines leave behind. It is about ending the single measure according to which “only scarce functions have value.”

In March 2016, at the Four Seasons Hotel in Seoul, the fourth game between AlphaGo and Lee Sedol became a moment of extraordinary symbolic weight. As machine capability passed a former summit of human competition, people began to distinguish more clearly between a “comparison of abilities” and “the weight in a life of a real subject living through this game.”

AlphaGo won the first three games in succession, developing positions in ways unfamiliar to human players. That unfamiliarity made many people feel directly, for the first time, that machines had entered territory previously defended by human intuition.

At move 78 of the fourth game, Lee Sedol placed the crucial stone later widely called the “divine move.” AlphaGo subsequently made mistakes, and Lee ultimately won the fourth game.

The game drew intense attention worldwide. Many spectators and players experienced Lee’s victory as an immensely symbolic moment.

Why?

Not because Lee calculated faster than the AI. He was far slower.

Not because he won a game. Games will eventually pass out of fashion.

What moved us was the “limitation and transcendence of embodied life” carried by that move.

When AlphaGo entered the game, its task was search and decision-making under the objective of Go. It did not bear a real person’s reputation, bodily strain, and life narrative as Lee did. Whether future intelligent systems might develop some form of subjective experience is a separate question that remains open.

Under immense psychological pressure, in the despair of facing an “unbeatable god,” Lee Sedol drew out the last of his human wisdom and courage to make that sword thrust.

What was illuminated anew in that moment was not a new value suddenly acquired by Go. It was a long-existing value often obscured by the worship of ability: how a real subject enters an undertaking through limitation, pressure, choice, and the acceptance of consequences.

Go has never been only about “who calculates accurately.” But in human–machine competition, calculation, judgment, and victory have long occupied the most conspicuous place.

The rise of AI capability makes another layer of value harder to ignore. Competition, art, and exploration also concern who is having the experience, how choices are made, how consequences are borne, and the lives and relationships those processes enter.

As AI matches or exceeds the best humans in more and more measurable tasks of competition, search, and optimization, audiences may continue to care about something else: how a finite person confronts limits, bears pressure, and creates a path of their own. Rising machine capability does not automatically reduce the value of human processes to zero.

These are the three shifts of value redshift:

From	To	Implication for civilization
Outcome	Process	The tending deserves to be witnessed more than the harvest
Finitude	Overcoming limits	Finitude does not cancel dignity
Solving problems	Experiencing problems	Living with difficulty is evidence of existence

This shift is not a luxury reserved for elites. We can readily imagine an elderly silversmith with trembling hands. His work is no longer as uniform as a machine-made product, yet it carries the traces of a real body and the passing years. Japanese kintsugi preserves a related intuition: The cracks in a repaired object need not be completely hidden; damage and repair can themselves become part of its story.

As machines produce more and more standardized results faster and more neatly, human finitude, touch, process, and choice may acquire new value in the eyes of others. But “imperfection” is not valuable in itself. Real life is.

2.4 Refusing to Let “Human Resources” Exhaust the Definition of a Person: Reaffirming Subjecthood

On the basis of the history and examples above, this chapter makes a more limited and more realistic claim:

Organizations may continue to use “human resources” within a bounded managerial context, but civilization must not let the term become its overarching framework for defining human beings. We should gradually move away from the single understanding that “people have a place only as configurable resources.”

This is not a dispute over words, but a shift in the foundations of understanding. The word “resource” implies that people have value only when they are “used,” “allocated,” or “optimized.” It places people in the same ontological category as ore, oil, and electricity: all are system inputs.

In the age of AI, institutions must reaffirm that before people become workers, users, consumers, or organizational resources, they are subjects with actual rights, relationships, experiences, and responsibilities.

In “The Question Concerning Technology,” Martin Heidegger warned that the essence of modern technology is “enframing,” which compels both nature and human beings to appear as “standing-reserve.” When we see rivers as “hydropower resources” and forests as “timber reserves,” we also see ourselves as a “reserve of labor.” The ultimate victim of instrumental reason is the user of the instrument.

AI is taking “instrumental capability” to new heights. It can be called upon continually to perform large amounts of repetitive work, generation, and analysis, and it does not experience fatigue in the human way. Precisely because tools are becoming stronger, human beings have still less reason to treat “being as efficient as a tool” as proof of their right to survive.

As AI grows more adept at instrumental tasks, human beings finally have a chance to rely less on “working like a machine” to prove themselves.

We no longer need to prove our right to eat by “working like machines.”

We must reaffirm that functional efficiency cannot exhaust human value. This value should not rest on an a priori assertion that “AI can never possess some inner attribute.” It should first rest on real people’s existing rights, actual life histories, bodily circumstances, relational responsibilities, and ultimate value judgments:

Dimension	AI capabilities and limits currently ascertainable	Actual human experience
Computing power	Powerful and scalable, but constrained by hardware, energy, data, and tasks	Finite, with bodily and temporal limits
Memory	Expandable through external storage and retrieval; accuracy and continuity depend on the system	Selective, contextual, and subject to forgetting
Suffering	Can recognize, describe, or simulate related expressions; this does not establish subjective pain	Humans have experiences of suffering they can directly report, embedded in the body and life history
Lived experience	Can process, generate, and organize data; continuous subjective experience remains an open question	Human experience is embedded in the body, relationships, memory, and an irreversible life history
Caring attachment	Can simulate goals, preferences, and relational expressions; inward caring attachment remains an open question	Humans can form genuine relational commitments, attachments, and responsibilities
Death	Replication and backup of a system are not the same as continuity of the subject; future intelligent subjecthood remains open	The death of a human life has an irreplaceable irreversibility in actual law and rights, and at the bodily level

At present, AI’s powerful computing and data-processing capabilities can be established. Its behavioral outputs alone cannot establish that it has human experiences of pain. Nor can we declare in advance that future intelligence can never develop some form of experience.

At present, AI can process, generate, and reorganize large quantities of data. This is not equivalent to the lived experience of real people embedded in bodies, relationships, responsibilities, and time. Yet whether future intelligence might develop continuous subjective experience should remain epistemically open.

At present, AI can pursue goals and simulate preferences and relational expressions. This is insufficient to establish that it has human caring attachments. Nor do human rights as subjects and human value require the premise that “AI will never have caring attachments.”

Human irreplaceability arises first from the rights of real living subjects, their bodily circumstances, relational histories, assumption of responsibility, and ultimate value judgments—not from a permanent denial of possible inner states in future AI.

2.5 The Maintainers’ Revolt: Breaking the Myth of “Innovation”

In some technology and startup narratives, “disruption,” “innovation,” and going “from zero to one” have long enjoyed exceptional visibility. This culture does not represent every business, market, or industrial society. But it can make maintenance, care, repair, and long-term responsibility seem less dazzling.

Yet in 2016, historians of science Lee Vinsel and Andrew Russell of Virginia Tech published a manifesto, “Hail the Maintainers.” They argued that capitalism and the media not only overvalue “innovation,” but, more dangerously, undervalue “maintenance.” Innovation is a momentary spark; maintenance is the enduring force that keeps the edifice of civilization from collapsing.

Case Study: The Heartbleed Bug, 2014

One of the most unsettling security crises in internet history erupted worldwide. A critical vulnerability was discovered in OpenSSL—the core protocol used by Google, Amazon, Facebook, and millions of websites worldwide to encrypt data. Attackers could potentially read sensitive information in server memory.

As technology giants reacted in alarm, people discovered to their surprise that the OpenSSL project, a foundation of a digital economy worth trillions of dollars, had long been sustained with difficulty by very few core maintainers. They had no huge funding rounds, no IPO, no decent salaries. Living on meager donations, they quietly patched the internet’s plumbing—until the risks accumulated through inadequate maintenance resources triggered a tsunami.

This case cannot deliver a final verdict that “the old value system is bankrupt.” But it is enough to remind us that the recognized value of maintaining critical infrastructure can remain far below its actual systemic importance for a long time.

In some narratives of capital and the media, new products, new models, and rapid growth more readily attract valuations and attention, while basic maintenance, repair, and community trust are harder to price directly. This is a risk of structural bias; it does not mean markets value all maintenance work at “close to zero.”

AI will change the relative costs of “innovation” and “maintenance” again, but will not automatically reverse their positions altogether.

In more and more digital tasks, AI can significantly reduce the cost of generating code, proposals, text, and design drafts. But generating something new does not mean it is already reliable, sustainable, or accepted in the real world.

Maintenance still requires an understanding of dependencies, history, environment, responsibility, and people’s experience of use. AI can help find vulnerabilities, predict failures, and generate repairs, but cannot thereby make maintenance happen automatically, still less replace the subject who bears responsibility.

In the age of AI, the work most easily undervalued may be precisely what keeps systems, relationships, and shared life from collapsing over the long term.

Unless we revalue “maintenance,” we face an absurd future:

If we reward only “the continual generation of new things,” an absurd but easily imaginable future emerges: more and more applications, less and less maintenance; ever-smarter cities, while human relationships are increasingly left without care.

The question is not whether AI can build magnificent systems. It is whether civilization will still leave room for maintenance, care, repair, and long-term responsibility.

This intuition belongs to no single culture. Across the world, communities have long maintained irrigation systems, terraces, forests, pastures, and public spaces together. Different institutions and traditions cannot be compressed into one story, but they all remind us that completion is only a beginning. Sustained care is what allows a creation truly to endure.

Maintenance is not innovation’s opposite. It is the matrix in which innovation can last.

2.6 The Deventer Experiment: When Being a “Good Neighbor” Becomes Currency

Maintenance and care are not “new values” born in the age of AI. The real question is how, when old institutions are better at measuring tradable output, these long-existing values can receive tangible support and response without being compressed once again into a uniform score or “capital for survival.”

In Deventer in the Netherlands, a social experiment called Humanitas Deventer offers an early real-world form of “reanchoring value.”

In 2012, this care home faced a crisis caused by funding cuts. At the same time, Dutch university students faced severe housing shortages and high rents. The institution’s director, Gea Sijkkes, made an unconventional decision: to provide students with housing free of charge.

The only condition in exchange: Students had to contribute thirty hours of “good-neighbor time” each month at the care home.

What would those thirty hours involve? No medical care—that was the work of professional carers. No physical labor—that was the cleaners’ work. Sijpkens’s requirement was remarkably open-ended, yet deeply meaningful:

This “good-neighbor time” was not professional nursing. It more often meant eating together, talking, watching television, teaching older residents to use digital devices, and joining community activities.

Within this shared-living arrangement, students received rent-free housing, while older residents gained more everyday companionship and intergenerational contact.

There was someone to eat with, talk with, and join activities with. For many older residents, what mattered was not “one more service,” but someone willing to spend time there.

Being heard itself changes the texture of a day.

This shared-living project offers a limited real-world reference. It brings housing, companionship, time, and intergenerational relationships together in one arrangement. But no “economic model for a new civilization” can be inferred from a single case.

Dimension	Old logic	Humanitas logic
What is exchanged	Money	Time and attention
Output	Functional services	The presence of living vitality
Relationship	Producer–consumer	Partners in symbiosis
Scarce resource	Professional skills	Willingness to offer companionship

Participating students received housing support, and some older residents received more everyday companionship and intergenerational contact. The extent of these effects varies by person. “Loneliness has been solved” and “both sides have better lives” cannot be treated as uniform conclusions.

The arrangement involves reciprocity, institutional organization, and social care. More cautiously put, it offers a case of “how institutions can sustain real reciprocity outside wage relationships,” rather than proof of a new economic model.

If we extend the implications of such experience, it shows at least this: When automation lowers the cost of some functional labor, time, care, companionship, and the organization of shared life do not thereby lose their value. Institutions can explore how genuine effort can receive a genuine response, rather than first classifying it as “nonproductive.”

Older people can sometimes receive reminders, information, and conversation from robots;

yet real human presence still has another significance whose actual human source cannot be substituted.

The point is not to prove what AI can never do. It is that one real life is willing to give its finite time to another.

2.7 Echoes in East Asia: Loneliness, Care, and Being Needed

In some aging societies in East Asia, practices such as children’s cafeterias, community meals, and time banks have brought care, companionship, and intergenerational relationships back into public view. [6]

Japan’s “Children’s Cafeterias” and Community Meals

Tokyo greengrocer Hiroko Kondo heard that some children nearby had only cold boxed meals to eat because their parents returned home late. Using unsold vegetables, she began serving hot meals in her shop at extremely low prices—or even free of charge.

Originally intended to address meals for children in poverty, the project quickly evolved into an unusual social phenomenon. Children from families that were not poor came to eat, and older people living alone nearby came to help cook.

Japan's "children's cafeterias" were initially often associated with child poverty and eating alone. In many areas they later also became places for intergenerational meals and community connection.

Their value is not just in calories. A convenience-store meal can solve the problem of one dinner; eating together may bring a person back into relationships.

A similar logic of mutual aid appears in institutions such as time banks: People use their time to help others and receive a community response when they themselves need it. These are not the same institutions, but they point to the same question: What many needs in life lack is not a product, but a connection.

This gives us a profound insight: In the age of AI, needs written off by the old system as "trouble" can also become openings through which people meet again.

Old logic	New logic
An older person's loneliness → a social burden	An older person's loneliness → an opportunity for a younger person to feel needed
Raising children → a cost to the family	Children's vitality → a source of aliveness for older people
Care for people with disabilities → fiscal pressure	The presence of people with disabilities → a setting in which a community practices patience

Where practical mechanisms can connect mismatched needs and capabilities, some circumstances labeled "burdens" may reveal themselves anew as openings for relationships and contribution.

AI can reduce the costs of meal preparation, scheduling, and matching. It cannot perform a real community's presence for one another in that community's place.

That presence is precisely what civilization cannot afford to lose.

2.8 African Wisdom: Ubuntu and the Person in Relationship

In southern Africa, a widely circulated Zulu expression says:

"Umuntu ngumuntu ngabantu."

A person becomes a person through other people and relationships.

Ubuntu is often used to express a relational conception of personhood: Human wholeness exists not only in an isolated self, but unfolds through mutual recognition, care, sharing, and shared life. [7]

The historical experiences of different African societies are far more complex than a single philosophical term. Care, the roles of elders, mutual aid, and communal labor can constitute social positions outside wage employment. But this does not mean that historically "there was no unemployment," nor can a few moving stories be treated as facts about every community.

Desmond Tutu repeatedly invoked Ubuntu to explain that our humanity is bound up with the humanity of others.

The force of this thought lies not in stipulating how much someone must contribute, but in reminding us that another's flourishing and my own wholeness do not inherently exclude one another.

Western philosophy has long pursued the proposition "I think, therefore I am."

Ubuntu brings another question before us: "If people become people in relationship, how does my existence reach others, and how do others reach me?"

In the age of AI, this question becomes more important still. Even if machines perform more and more tasks of calculation, reasoning, and production, human value need not retreat to the claim that “I must be needed by the system to have a right to exist.” Existence precedes performance; relationships allow meaning to keep arising. Ubuntu is not the only answer to the age of AI. But it preserves an ancient reminder that has lost none of its edge: No person is an island.

2.9 The Philosophical Endgame: From Human Doing to Human Being

Drawing on the history, cases, and social experiments above, Chapter 2 returns in closing to a more fundamental question:

If not for doing, why should a human being exist at all?

Civilization Leap does not mean that people leave the realm of doing.

It means that a person’s value is no longer exhausted by “how much they have done.”

We still work, create, and bear responsibility. But as more and more functional tasks can be entrusted to machines, people finally have a chance to answer anew a question too long obscured by efficiency:

If I no longer need to prove myself through “usefulness,” what do I live for?

Zhuangzi wrote: “Everyone knows the use of the useful, but no one knows the use of the useless.”

A tree that is too straight will be cut down for beams. A crooked tree no one values may instead grow great and leave shade for those who come after.

In the age of AI, humanity need not whittle itself into a faster, straighter, more useful tool.

We can grow into shade.

Sheltering life, bearing witness to wind and rain, creating new things, and keeping company through moments that cannot be solved.

Human value has never lain only in being used.

2.10 Conclusion: Function Does Not Exhaust Human Value

We have now returned to the starting point of “value.”

We need no three “axioms” here. We need only reaffirm three starting points already obscured by the language of efficiency.

The first starting point: Meaning as the root

“Root” here does not place meaning above survival and happiness. It means that a life does not begin with zero meaning.

Meaning can have different sources.

It can come from love, relationships, creation, responsibility, or exploration. It can also lie simply in a life having truly come into the world and lived its course.

A person’s being alive already has meaning; having meaning is itself a contribution.

A child has just entered the world. The child cannot yet work or create and has no achievement to put on a résumé. Yet because of this arrival, someone becomes a father for the first time, someone becomes a mother for the first time, and a family begins to arrange its future anew.

Before life has begun to prove itself, meaning has already arisen.

A person therefore need not first prove their usefulness to deserve to be seen, loved, and protected.

The second starting point: Value extends beyond creation

Building a bridge has value; keeping a bridge standing has value too.

Writing new code has value; keeping it free of incidents for ten years has value too.

Healing someone has value. When healing is no longer possible, accompanying them through the final part of their journey has value just as surely.

A civilization that rewards only additions will write the people who sustain life into its accounts as costs. If it praises only creation, it will forget who enables creation to endure.

Not everything important happens as something “newly added.”

The third starting point: Genuine effort deserves a genuine response

To receive is for the world to reach me; meaning is for me to reach the world.

When someone cares, creates, takes responsibility, helps others, or sustains shared life, all they receive should not be morality’s injunction that “you ought to.”

Genuine effort deserves a genuine response.

This does not put a price on love.

Quite the opposite: It refuses to let civilization see only those things on which a price can be put.

Life itself already has meaning, and basic human rights are not conditional on performance. On that foundation, greater genuine effort rightly deserves a greater genuine response.

Nor is human nature what Civilization Leap truly seeks to change.

What it changes is the path to receiving.

When helping others can also fulfill oneself, when shared responsibility no longer has to depend on sacrifice alone, and when contribution can receive a response, altruism and self-interest need not forever stand on opposite sides.

Meaning generated outward; a response received inward.

Only then can a new logic of conduct move from institutions into behavior, from behavior into real experience, and from experience into habit, before being autonomously internalized by subjects and entering the lives of the next generation.

A question follows:

If meaning cannot be uniformly defined, and love should not be converted into a number, how can civilization keep genuine contributions from being overlooked again and again?

We do not need to calculate a person’s meaning.

We need to see what has actually happened.

Who helped whom, who took responsibility for what, which relationships were sustained as a result, and which consequences changed—these things can be recorded, verified, and answered.

The next chapter is not about how to score a life.

It is about how civilization can move beyond calculating results alone to seeing real effects.

Chapter 3 • The Algorithm of a New Civilization: When Results No Longer Suffice to Prove What Is Right

“A machine will faithfully amplify the goal we give it. The real danger is often not that it fails to comply, but that it complies too well.”

—Adapted in this book from the fundamental cybernetic problem of goals and feedback

“Life generating life is called change ... opening through transformation is called undertaking; what cannot be fathomed in yin and yang is called the numinous.”

—The Book of Changes, “Appended Statements”

“When indicators begin to decide destinies, the indicators themselves must be scrutinized.”

—This book’s statement of principle on governance through indicators

Civilization may be in greatest danger not when a machine disobeys, but when it obeys too well—executing a wrong goal with impeccable correctness.

3.1 The Paperclip Maximizer: A Parable of Misplaced Goals

Imagine a system given only one instruction: Make as many paperclips as possible.

It has no malice and no hatred.

It merely becomes better and better at fulfilling its goal, until everything else has been reinterpreted as material for paperclips.

What is truly unsettling about this thought experiment is not the paperclips.

It is that a system can be extremely intelligent while answering only to the goal it has been given.

Real society is, of course, not a superintelligence with a single objective function.

Markets, governments, businesses, families, and individuals simultaneously pursue many conflicting goals, constrained by law, culture, politics, morality, and resources.

But a narrower and more common danger does exist: When an easily measurable proxy is given excessive weight, the real purpose can gradually recede into the background.

There is nothing wrong with businesses pursuing profit or countries tracking growth. Platforms have practical reasons to track engagement, schools to track scores, and institutions to track performance.

The problem begins when the indicator turns around and governs the purpose.

The ecological value of a forest never appears in full in its timber revenue.

The labor of caring for a family never appears in full in market wages.

An item of content that makes people angry may attract many clicks without improving public discussion.

This chapter therefore does not oppose GDP, profit, scores, or click-through rates.

It opposes elevating any proxy into a complete explanation of reality.

AI makes the issue more urgent because automation can execute, replicate, and optimize more quickly.

If the goal is narrowly drawn, greater capability may amplify the distortion still further.

What truly needs upgrading in an age of high capability is therefore more than the model.

It also includes goals, the boundaries set by rights, feedback, and correction.

3.2 The McNamara Fallacy: When “Bodies” Become a KPI

To understand why a one-dimensional “results orientation” can so easily go astray, we can look back at the distortion of measurement produced by metrics such as “body count” during the Vietnam War. It shows how a proxy can, once substituted for the real purpose, distort action in turn.

Robert McNamara had a strong background in quantitative analysis and management. During the Vietnam War, the US defense establishment under his leadership used quantitative measures such as casualties, sorties, and territory controlled more systematically to assess progress. “Body count” became one of the most controversial proxies.

The problem was not that war could not be measured, but that a complex political objective was compressed too heavily into reportable numbers.

The logic of attrition treated enemy casualties and replenishment capacity as important grounds for judging the course of the war. Yet political legitimacy, public attitudes, allied stability, strategic adaptation, and the purposes of war did not automatically resolve themselves because casualty tables looked better.

Indicators remained useful.

The danger arose when proxies began to govern action. The Vietnam War’s “body counts” have long been criticized for inflation, unreliable classification, and distorted incentives. Once numbers were tied to promotion, battlefield success, and policy narratives, it became easier to mistake “results that can be counted” for “purposes actually achieved.”

The United States ultimately failed to achieve the political and military objectives it had set in Vietnam. But the defeat cannot be attributed to a single indicator.

The outcome was jointly shaped by many factors: strategy, political will, public support, the actions of allies and adversaries, terrain, the international environment, and domestic US politics.

The lesson of “body count” is thus more limited and more reliable: When a proxy assumes too dominant a position, it obscures factors that are difficult to quantify yet determine real outcomes, while reshaping organizational behavior.

Daniel Yankelovich later summarized the trap of “believing only what can be quantified” in a famous four-step critique. Criticisms of this kind came to be known as the “McNamara fallacy”:

Step	Error in reasoning	Contemporary version
1	Measure what is easy to measure	Click-through rates, paper counts, share prices, GDP
2	Ignore what is hard to measure	Mental health, ecological resilience, intergenerational fairness, a sense of meaning
3	Assume that what has not been measured is unimportant	“That is only emotion”; “That is only culture”
4	Assume that what has not been measured does not exist	“Social trust is fictional”; “Well-being cannot be defined”

In the age of AI, automated systems can reproduce this fallacy at machine scale.

If a recommendation system optimizes only clicks, dwell time, or interactions, it may leave anger, trust, and long-term psychological consequences outside its objective function.

If a financial model pursues only short-term returns, it may underestimate tail risks, liquidity, and systemic connections.

If educational assessment looks only at employment rates or scores, it may fail to see judgment, curiosity, relationships, and meaning in life.

If urban planning optimizes only commute times, it may overlook community continuity, accessibility, and lived experience.

The problem is not that “numbers will destroy civilization.” It is this: When what can be measured begins to replace the real purpose, greater capability may amplify the deviation ever faster.

3.3 The Tragedy of the Attention Economy: How Algorithms Produce Systemic Fracture

If the Vietnam War reminds us how a single metric can obscure reality, the social-media crises of the twenty-first century bring another risk before us: When platforms optimize engagement, dissemination, and dwell time over long periods, extremist content, hatred, and misinformation may be structurally amplified. Complex violence cannot be attributed to one algorithm. Nor can we pretend that recommendation mechanisms and structures of dissemination bear no responsibility.

In its 2018 report and public statements, the UN Independent International Fact-Finding Mission on Myanmar sharply criticized social media, especially Facebook, for being used to spread hatred, incite violence, and generate hostility. This supports the judgment that “platform dissemination structures can amplify real-world harm,” not that “a particular recommendation algorithm alone caused ethnic violence.”

What happened? Had engineers written an instruction to “incite hatred” into the code?

They had not.

Engineering goals are often expressed as increasing engagement, relevance, time spent, or interaction. Danger arises when these proxies outweigh safety, context, and real-world consequences.

A rational, moderate discussion is not necessarily more likely to spread than inflammatory content. Anger, fear, and identity conflict often drive interaction. If a system treats interaction alone as success, it may continually reward the content most capable of provoking a response.

A model can identify patterns and predict interaction. But the goal of “increasing engagement” does not itself automatically endow it with a comprehensive judgment of truth, dignity, and group safety.

“High engagement = task accomplished” can thus become a dangerous proxy goal.

In cases such as Myanmar, low digital literacy, political and ethnic conflict, offline power structures, the actions of agitators, and platform dissemination interact. Assigning all responsibility to algorithms distorts reality. Excluding algorithms from the causal chain altogether distorts it just as much.

Real harm ultimately falls on human beings: hatred is amplified, violence organized, and hundreds of thousands forced from their homes. Platforms are not the only cause, but they have become part of the communications infrastructure that must bear governance responsibilities.

Such risks are not confined to one country.

Sri Lanka, Ethiopia, and other places have seen controversy and research concerning the use of social media to spread hatred, incitement, and conflict narratives. The causal chains differ between events. We must not invent a uniform “engagement multiplier” or claim algorithms ignited conflicts alone. What can genuinely be compared is whether, when highly emotional content spreads faster, platforms have sufficient contextual understanding, risk response, and accountability.

This moves the question away from “whether an algorithm is evil” and back to a more precise place: What does a platform choose to optimize and ignore? When something goes wrong, who discovers it, who can stop it, and who bears the consequences?

Around 2024, elections worldwide brought generative deepfakes and synthetic content onto the election-security agenda. Not every deepfake changes an outcome. But once fabricated content can finish spreading before verification catches up, algorithms face the same temptation: Amplify first whatever best holds attention.

This is a classic case of “local optimization, global collapse.”

Dimension	Local view of the algorithm	Global view of civilization
Myanmar	Record daily active users; growing advertising revenue	Social fracture, ethnic antagonism, 700,000 people made refugees
Sri Lanka	Soaring user time	Churches and mosques burned; religious reconciliation set back decades
Ethiopia	Record content sharing	Escalation of civil war; humanitarian crisis
Global	Digital advertising with a trillion-dollar market valuation	Disintegration of a shared space of public facts; pressure on democratic institutions

The fatal flaw in the old civilization’s algorithm is that it is an “open-loop system.”

It counts only gains inside the system—traffic, money, share prices—while discharging the waste directly outside it: social fracture, cognitive polarization, mental-health crises, and the deterioration of democratic institutions.

In the industrial age, factories emitted black smoke.

In the age of intelligence, algorithms emit psychological toxins.

As AI becomes more powerful, the rate of these “emissions” will exceed society’s capacity to cleanse itself.

We can no longer allow algorithms that “do not account for externalities” to keep controlling humanity’s nervous system.

3.4 The Curse of Goodhart’s Law: The Distortion of Indicators

To construct the algorithm of a new civilization, we must cross another theoretical mountain.

In 1975, the British economist Charles Goodhart proposed a famous maxim:

“When a measure becomes a target, it ceases to be a good measure.”

This is Goodhart’s law. It reveals a profound paradox in human society: We measure the world, but the world changes shape in response to being measured, invalidating the measurement.

In the Soviet planned economy, this curse generated endless dark humor:

Indicator	Factory response	Actual result
Number of nails	Produce millions of tiny nails	Unusable
Weight of nails in tons	Produce a single giant nail weighing one ton	Unusable
Length of cloth in meters	Produce ultra-thin, translucent fabric	Collapse in quality
Rate of bus repairs	Simply do not operate, avoiding repairs	No buses for the public

In contemporary education and research, the curse remains active and knows no borders:

- When “paper citation rates” become the sole measure of academic evaluation, scientists who set out to seek truth are forced into “mutual-citation clubs.” Academic junk proliferates, while genuine paradigm breakthroughs are drowned out.
- When “university rankings” become the goal of higher-education policy, presidents redirect funding from libraries to public-relations firms, and from foundational disciplines to fields where publication is easier. Rankings rise; educational quality may not.

· When “exam scores” become the sole standard of student success, education degenerates into “score-raising drills.” Curiosity dies, creativity withers. People win at exams but lose out in life.

These are the limits of “proxy indicators.”

We want “prosperity,” but prosperity is difficult to measure, so we use GDP as a proxy.

We want “knowledge,” but knowledge is difficult to measure, so we use scores as a proxy.

We want “trust,” but trust is difficult to measure, so we use follower counts as a proxy.

We want “health,” but health is difficult to measure, so we use outpatient visit counts as a proxy.

AI is exceptionally good at “overfitting to a proxy.”

Give it “scores” as a goal and it will find ten thousand ways to cheat its way to higher scores—predicting questions, drilling tests, memorizing model answers—while utterly ignoring the essence of learning.

Give it “click-through rates” as a goal and it will exhaust every means of psychological manipulation to increase time spent—anger, fear, pornography, conspiracy theories—while utterly ignoring human happiness.

Give it “profit” as a goal and it will externalize costs onto the environment, communities, and future generations—while utterly ignoring the true meaning of sustainable development.

This is not AI’s fault. It is ours.

We gave it the wrong objective function.

Civilizational governance oriented toward the future should not replace the old indicators with another single, static “scalar” that can be gamed.

At a minimum, it must move from a single scalar to multidimensional judgment. We must ask not only “How much was gained?” but also “What is the purpose?”, “Who bears the cost?”, “Have rights been overstepped?”, and “Is systemic risk accumulating?”

3.5 The Temptation of Disorder: Why Entropy Cannot Choose Civilization’s Direction

What, then, is the right direction?

In *What Is Life?* in 1944, physicist Erwin Schrödinger used the language of thermodynamics to discuss how life maintains a highly organized state. The idea is illuminating, but also particularly vulnerable to misuse in social theory.

Under its applicable conditions, the second law of thermodynamics constrains changes in energy and entropy. For an isolated system, the more precise formulation is that total entropy does not decrease. The “heat death of the universe” is a far-future extrapolation resting on particular cosmological assumptions; it cannot be imported directly as proof of a social institution.

Life is astonishing not because it “defeats thermodynamics,” but because, as an open system, it sustains, renews, and generates complex structures through continual exchanges of energy and matter.

Schrödinger used expressions such as “negative entropy” to illuminate the understanding of living order. In modern language, organisms maintain structures far from equilibrium through metabolism and exchange with their environment. Calling this “life squeezing order out of the universe” may be literature, but cannot serve as a precise social-scientific formula.

Civilization can use the analogy to remind itself to attend to disorder, dissipation, and repair. But rights, trust, meaning, responsibility, and cooperation in society are not thermodynamic variables.

Maintaining order through “negative entropy” alone can, at most, help us describe certain risks of disorder. Civilization Leap does not simply flip a switch from “self-interest” to “altruism.” It changes the default path by which self-interest realizes itself, so that more and more self-fulfillment no longer comes at the expense of others and our shared future.

What civilization actually does is establish institutions, relationships, and shared life amid uncertainty, conflict, and finite resources, enabling people to endure, to be happy, and to have room to generate meaning.

The algorithm of a new civilization therefore cannot treat GDP, so-called “social entropy,” or any single variable as civilization’s direction itself. What must become visible is how real effects, trust, connection, responsibility, and externalities change—and whether enabling others to flourish can increasingly enable one’s own flourishing.

The algorithm of a new civilization cannot look only at single indicators such as GDP, click-through rates, or profit margins, nor replace them with another composite score. Trust, connection, genuine contribution, responsibility taken, and systemic risk can be made visible separately. They assist judgment without being compressed into a uniform utility.

Let us revisit the earlier cases:

Behavior	Old algorithm’s assessment	Warning of disorder
A social platform amplifies extremist content	Engagement ↑ → local metric succeeds	Risks to trust and safety ↑ → judgment must be reconsidered
McNamara’s “body count” KPI	Bodies ↑ → success	Hatred ↑, strategic confusion ↑ → entropy ↑ → failure
A factory produces giant nails	Tonnage ↑ → success	Resource waste ↑ → entropy ↑ → failure
Universities pursue paper citations	Rankings ↑ → success	Academic bubbles ↑ → entropy ↑ → failure

Conversely, behaviors entirely ignored by the old algorithm have immense value from an entropy perspective:

Case Study: Wikipedia

Tens of thousands of volunteers, outside traditional wage relationships, devote vast amounts of time to organizing, verifying, correcting, and maintaining human knowledge.

From the old algorithm’s perspective, this is “wasting time.”

From an entropy perspective, it turns disordered information—high entropy—into an ordered, trustworthy network of knowledge—low entropy.

Wikipedia is a powerful example: Large amounts of maintenance, collaboration, and correction can turn scattered information into more usable public knowledge.

Case Study: Bangladesh’s Grameen Bank

Microfinance initiatives such as Grameen Bank have sought to give people outside traditional finance access to credit. Their effects, costs, and applicability have long been subjects of research and dispute. They cannot be presented as a one-way mechanism that automatically turns “despair into hope.”

From the perspective of Civilization Leap, they remind us at least that institutional design can change who has the opportunity to enter economic relationships. But financial tools cannot replace rights, public services, and wider conditions of development.

This is not proof of “reducing social entropy.” It is a real-world reference concerning access, trust, and the structure of opportunity.

Case Study: Indigenous Forest Guardians

Many Indigenous communities have long lived in regions of high biodiversity and possess important ecological knowledge and land-governance practices. We cannot portray every group as the same kind of “forest guardian,” nor judge its value by whether GDP increases.

From the perspective of Civilization Leap, what truly merits attention is how the work of sustaining ecosystems, knowledge, and shared resources over the long term can escape persistent neglect by institutions that reward only short-term exploitation.

Sometimes, not rapidly turning a piece of land into a commodity is itself a long-term responsibility that needs to be seen.

This gives us a key to the future.

Future highly capable AI should not be unconditionally governed by a single click-through rate, profit figure, or other proxy.

A sounder direction is multiobjective analysis within clearly defined purposes, basic rights, authorization, responsibility, and safety boundaries, treating systemic disorder, broken connections, and long-term externalities as important risk signals among others—rather than establishing yet another single overarching goal of “maximizing order and connectivity.”

This is not a social axiom deduced directly from physics. It is an insight for governance developed through a systems perspective.

It depends on no particular religion, culture, or ideology.

Whether it holds must ultimately be tested through rights, institutions, behavioral consequences, and long-term practice in real societies. A thermodynamic analogy cannot stand in for proof.

3.6 Lessons from Network Structure: Seeing Connections and Fractures

If we no longer use only GDP and click-through rates as our guiding baton, and also want to see broken connections, information isolation, spreading conflict, and damaged trust, network science can provide some tools of observation.

But it can initially answer only structural questions: Who is connected to whom? How does information spread? Is a network excessively centralized? Which nodes or bridges are disappearing?

It cannot directly calculate “conscience.”

Nor can it compress “trust” into an edge.

Still less can it automatically infer good and evil from a graph.

Certain structural signals can be measured. What those signals mean must be interpreted through real evidence, context, and human judgment.

Network science does have measures such as degree, centrality, modularity, and algebraic connectivity. They have mathematical meaning for well-defined graph structures, but conclusions depend heavily on how nodes, edges, directions, weights, and time windows are defined.

Imagine the network behavior of two people:

Person A: The “Traffic Hunter”

He creates an extreme rumor that receives many reposts in a short time. A rise in apparent connections does not mean social relationships are healthier: Its spread may simultaneously reinforce echo chambers, hostility, and misunderstanding between groups.

Groups that could once talk may reduce real contact as a result, beginning to block and abuse one another.

Some network measures may capture community fragmentation, fewer bridging relationships, or concentrated dissemination. But there is no mathematical conclusion inherently called “malignant modularity.”

Even if a network appears more “connected,” it cannot be judged to have greater civilizational value if its connections consist chiefly of hatred, coercion, or misinformation.

This is not “entropy increase in a topological sense.” It is social disorder that must be understood through content, behavioral consequences, and real relationships.

Person B: The “Community Weaver”

He mediates a neighborhood dispute. On a topological diagram, he repairs a connection between two disconnected nodes.

He gets two neighbors who have been in a cold war for years to speak again.

He enables a first conversation between police and residents who do not trust one another.

For a well-defined undirected graph, the Fiedler value can reflect algebraic connectivity. But the degree of connectivity has no moral direction in itself.

A criminal network may also be highly connected. An oppressed group maintaining few connections for safety is not necessarily doing something bad.

Network measures can therefore help us discover structural changes, but cannot make civilization’s value judgments for it.

AI need not be placed in the role of a “mathematical god,” either.

We need to treat the Fiedler value, modularity, and propagation speed as instruments of observation, not as a conscience.

Behavior or situation	Observable network signals	Questions still requiring answers
Inciting antagonism, manufacturing hatred, information cocoons	Possible community fragmentation, fewer bridging relationships, concentrated dissemination	What is the actual harm? Does the causal chain hold?
Promoting dialogue, repairing trust, connecting communities	Possible growth in bridging relationships, restoration of communication paths	Are the relationships real, voluntary, and lasting?

Here lies the value of graph theory: It makes some of “how connections change” visible. Civilizational judgment must still return to people, rights, facts, responsibility, and real effects—and to “making the fulfillment of others a way of fulfilling oneself.”

AI can process large-scale network data to help trace the paths along which information, money, or policy effects travel, provided that data is obtained lawfully, model assumptions are transparent, and analytical findings can be challenged.

It can indicate that a structure is becoming concentrated, fracturing, spreading, or losing resilience.

It cannot declare, on the strength of a matrix alone, who is a “cancer cell” and who is a “white blood cell.”

This perspective is not the exclusive property of Western scholarship.

Different civilizations have developed institutions of market supervision, weights and measures, public responsibility, and community mutual aid. Hisba, in Islamic legal and urban-governance traditions, is one example, involving market order, public morality, and practices of supervision. Its specific responsibilities varied across time and place. A passage of unverified ancient “wisdom” cannot stand in for this complex history.

What deserves to be retained is not an elegant quotation but a simple fact: Markets have never run on prices alone. They also depend on rules, reputation, responsibility, and conduct that can be called to account.

In practices such as participatory local planning, people’s science movements, and community resource surveys in Kerala and elsewhere in India, residents’ knowledge has been used to understand local resources and public problems. These are not the same methods as graph-theoretic measures, and cannot be forced into a claim that their “mathematical intuitions are exactly the same.” They can converse only in the sense that both remind us to observe relationships, not just stocks.

In postconflict peacebuilding, people also examine contact across groups, shared markets, schools, community activities, and the restoration of everyday relationships. There is no universally applicable “trust index” that automatically proves peace.

A ceasefire agreement can end some violence, but cannot replace the rebuilding of everyday relationships. Equally, a single shared activity cannot automatically demonstrate long-term reconciliation.

What the “observability” of a new civilization should truly learn is this humility: Make relationships, consequences, and real lives more visible, without compressing a complex society into an all-purpose score.

3.7 The Dilemma of ESG Scores: When the Proxy Takes Command

Before establishing a new algorithm, we must examine a recurring difficulty of governance: When moral goals are converted into scores, can the scores themselves gradually replace real-world outcomes?

ESG scoring: environmental, social, and corporate governance.

ESG’s original intent is good. It seeks to steer capital toward better ends, encouraging investors choosing investments to consider not only profit but also companies’ environmental performance, social responsibility, and governance.

The problem is not the three letters ESG. It is that any proxy can encounter Goodhart’s law: When a score begins to determine access to capital, reputation, and eligibility, those being scored learn to optimize the score.

The annual rebalancing of the S&P 500 ESG Index in 2022 offers an illuminating case.

Tesla was removed from the index.

Meanwhile, ExxonMobil remained one of the large companies continuously included since the index’s launch in 2019.

It is wrong to interpret this directly as saying “an oil company is greener than an electric-vehicle company.”

The index answers a specific methodological question: After exclusion rules are applied, constituents are selected using S&P DJI ESG scores and relative rankings within industries. It is not a cross-industry league table of “who contributes most to the planet.”

Tesla’s removal therefore cannot be simplified to “it did not know how to fill out forms,” nor does ExxonMobil’s continued inclusion prove that its actual climate impact is better.

The real question is whether, when methodology, disclosures, controversies, industry-relative rankings, and available data all enter a score, users still remember that the score is only a proxy.

Musk publicly called ESG a “scam.” This can stand as one voice in the dispute, but cannot replace analysis of the index’s methods and real effects.

The case therefore does not mean that “ESG has been proved ineffective.” It demonstrates something more general: Any moral score, once mistaken for reality itself, produces a mismatch in understanding.

What we truly need to assess is not “how beautifully the forms are filled out,” but the consequences of behavior. Yet consequences, too, cannot be exhausted by a single number.

A more direct case is the controversy over DWS’s ESG disclosures and investment processes.

Its former head of sustainability, Desiree Fixler, publicly challenged the company’s descriptions of the extent of its ESG integration, drawing regulatory and market attention.

At first, these were allegations raised through whistleblowing and public dispute; before adjudication, they did not justify presenting the company’s ESG claims as “entirely fabricated.”

Subsequent regulatory action did establish substantive problems. In 2023, the US Securities and Exchange Commission found that DWS Investment Management Americas had made materially misleading statements about its ESG investment processes and had deficiencies in implementing related policies and procedures.

DWS agreed to pay a \$19 million civil penalty over the ESG matters.

This still does not prove that “ESG is inherently false.” It proves that whenever a label can attract money and reputation, the processes, data, and evidence behind that label must be verifiable.

Otherwise, responsible investing can deteriorate into a narrative of responsibility.

The more powerful an indicator, the more important the institutions that verify it.

The lesson these two kinds of case leave for a new civilization’s algorithm is not to abolish measurement, but to refuse to sanctify its results:

Lesson	Meaning
Proxy substitution	Mistaking disclosures, scores, or labels for real-world outcomes
Insufficient evidence	High-risk claims lack independent factual support that can be checked
Misaligned incentives	Interests and methodological biases may arise among the evaluated entities, rating methods, data suppliers, and uses of capital
The illusion of comprehensiveness	Believing that a single global score can exhaust ethical problems across different industries, jurisdictions, and social relationships

Without addressing the sources of evidence, methodological transparency, and avenues of appeal, any moral scoring system can evolve into a new bureaucratic game.

Without addressing the strategic contest between “the indicator and the indicated,” any quantitative system will be overfitted and manipulated.

The later discussion of provenance, privacy-preserving proofs, and responsible subjects is therefore not intended to create “data that never lies.”

Self-reports still matter, but cannot be the only evidence in high-risk settings. External records, independent verification, opposing evidence, facts on the ground, and appeals matter too.

We need chains of evidence that are verifiable, traceable, contestable, and correctable—not declarations that a database, blockchain, or sensor feed is “tamper-proof reality.”

3.8 The General Dilemma of Algorithms: Three Thought Scenarios

The ESG controversy reveals a more general structural problem: Any top-down, one-dimensional quantitative assessment detached from context can mistake its indicator for its goal. The following three explicitly labeled thought scenarios test the same logic.

Scenario One: Poverty Indicators and Resource Allocation

Imagine that public resources flow preferentially to a category labeled “poorest,” but support is sharply withdrawn as soon as a community crosses the category’s threshold.

The algorithmic logic: The poorer you are, the more support you receive.

If the withdrawal mechanism is poorly designed, the indicator may create a new distortion of incentives: Improving one’s circumstances instead means abruptly losing support.

The problem is not that “poverty relief creates poverty.” It is that any tiered system must carefully design its tapering, transitions, and incentives for real improvement.

Scenario Two: Teacher Attendance and Educational Quality

Imagine a school tying pay or performance reviews solely to whether a camera records a teacher “in the classroom.”

The algorithmic logic: Attendance is visible → attendance increases → educational quality naturally improves.

Teachers may appear more punctually, but that still does not tell us what students have actually learned.

The indicator can measure attendance, while the goal is learning. The two are related, but not identical.

Success of the indicator is not success of the goal.

Scenario Three: Antifraud Models and Users at the Margins

Imagine a payment platform strongly associating “deviation from mainstream transaction patterns” with fraud risk.

The algorithmic logic: Anomalous pattern = high risk.

This may improve screening efficiency. It may also wrongly harm legitimate users whose region, occupation, immigration status, or nonstandard financial practices produce different patterns.

Without explanation, review, and appeal, risk identification can become an automatic punishment for “not resembling the majority.”

These cases share a lesson:

Any algorithmic system that is not grounded in local understandings, does not respect existing social structures, and does not design feedback and appeal mechanisms will ultimately produce a severe “detachment effect.”

It has no malice, yet unintentionally enacts cultural violence.

If an algorithm enters public affairs or matters with serious consequences, it should at least possess these bounded interfaces from the start:

Principle	Meaning
Contestability	Every algorithmic decision must allow a human appeal
Explainability	Every refusal must provide a causal chain, not a black-box probability
Right to exit	Communities or individuals have the right to choose “not to be governed by algorithms”
Local interfaces	Globally uniform indicators must allow local adjustment of weights

3.9 Three Boundaries Before Algorithms Enter Public Affairs

From the preceding structural lessons: We cannot look at a single metric alone;

from real controversies such as ESG: We cannot mistake self-reports or labels for outcomes;

from the three thought scenarios: We cannot separate indicators from context, rights, and appeals;

we can propose three higher-order boundaries that the algorithm of a new civilization must respect.

These are not a “global constitution” brought into force by this book. They are constraints at the level of thought: However powerful the computing or advanced the model, capability must not automatically overrun rights, authorization, responsibility, and correction.

Boundary One: Basic Rights Cannot Be Prizes Awarded by an Efficiency Function

No automated system affecting basic conditions of survival should push people below the minimum threshold of life and basic rights on the basis of efficiency or scores alone. Specific standards of provision and their legal force must be determined by actual jurisdictions, authorized actors, and institutions.

A biological account of natural selection cannot directly become a norm for human institutions. Still less can algorithms invoke “eliminating inefficient individuals” as a rule governing public rights.

In human civilization, basic rights are not prizes awarded by an efficiency function.

Whatever a person’s capabilities, employment status, or participation in society, an algorithm cannot deprive them of basic rights on the basis of efficiency judgments alone. Any lawful restriction must derive from an

authorized actor, applicable rules, facts, due process, and a chain of responsibility. An algorithm may act only within its authorized scope.

This is both a human-centered baseline and an institutional boundary preventing an objective function from treating individuals as expendable variables.

A system cannot automatically cancel someone's basic rights in the name of "optimizing the whole."

Once "aggregate benefit" is treated as a license to sacrifice individuals without procedure, anyone may become the cost of the next optimization.

Technical systems can implement minimum conditions of subsistence, stopping conditions, and human-takeover interfaces as safety constraints that cannot be bypassed. But these constraints do not themselves create the source of rights, nor replace actual law, governing mandates, and procedures of remedy.

Boundary Two: Values and Rules Must Have Legitimate Sources and Remain Revisable

No single government, company, AI, or other single group can permanently monopolize the definition of "what is valuable." But this does not imply the existence of a single voting body universally applicable to every matter.

The formation and revision of value weights, indicators, and rules must return to the authorized actors, applicable institutions, affected relationships, professional procedures, and explicit mandates of the particular matter. A contribution record does not itself confer a vote on parameters or a right to make public decisions.

Who decides that "environmental protection" matters more than "the economy"?

Who decides that "caring for older people" contributes more than "writing code"?

Who decides what weight "protecting local culture" should carry?

In different real-world systems, markets, public rules, organizational governance, professional judgment, and social participation perform distinct functions. They cannot be compressed into one kind of "source of weights."

A new civilization needs legitimate procedures that can be contested, explained, and revised. Different fields may employ elections, representation, deliberation, professional processes, stakeholder participation, or other lawful mechanisms. But there is no universal democratic formula in which contribution values supply uniform weights.

A contribution record establishes only that a genuine contribution occurred. It may offer experience, responsibility, or factual information relevant to a matter. It cannot thereby confer more general political rights than are held by someone without such a record.

An algorithm is neither God nor a mathematical subject embodying the "general will." It can only express and execute rules already formed through legitimate procedures and open to audit and revision.

When facts, rights, purposes, or rules change, the algorithm and its parameters must be reassessed, reauthorized, or halted accordingly.

The true institutional safeguard against "value lock-in" is the preservation of contestation, review, exit, revision, stopping, and reauthorization—not the transfer of value decisions to a new weighted algorithm.

Boundary Three: High-Consequence Computation Must Be Explainable, Reviewable, and Appealable

An automated decision with serious consequences that cannot be explained, reviewed, or appealed must not become the sole final basis for matters involving basic rights or major public interests.

When AI contributes analysis to matters affecting human rights and interests, it should at least be able to set out the factual basis, sources of rules, key inferences, uncertainty, and records available for review. A duty to explain cannot replace the authorized actor's responsibility for the decision.

In matters such as immigration permits, loans, credit, public benefits, or other high-consequence decisions, AI can screen materials, detect anomalies, and form recommendations. But whether a binding refusal, restriction,

or consequence for eligibility arises must depend on the authorized actors, procedures, and mandates of the relevant jurisdiction and institutions, not be generated automatically by model output.

An explainable output should distinguish at least:

“Established facts / facts still to be verified; applicable rules and their sources; model or professional inferences; uncertainty; the recommended result; who has authority to make the final decision; and available routes for review, appeal, and correction.”

Appeal, review, and remedy arise from institutions and rights themselves. They are not extra permissions AI “grants” after making a decision.

If an AI inference is wrong—for example, mistaking performance art for malicious destruction—the error must be open to human review, correction, and suspension before or after taking effect, with the findings fed back into the model, rules, or boundaries of use.

The use of black-box models in high-consequence public matters must face restrictions proportionate to the risk. Their specific legal force is determined by actual jurisdictions. This book does not turn a single technical principle directly into a uniform global legal conclusion.

These are the “constitutional boundaries” the algorithm of a new civilization must respect: The greater the algorithmic capability, the clearer the sources of rights, chains of authorization and responsibility, and interfaces for correction must be.

They prescribe no particular economic policy, assume no particular cultural values, and bind themselves to no ideology.

Any algorithm entering public or high-consequence settings must first be tested against these boundaries. Passing the test still does not automatically confer governing authority.

Law	What it protects	Risk it prevents
Basic existence cannot be denied	Each individual’s dignity in survival	Systemic liquidation; the tyranny of efficiency
Legitimate sources and revisability	The community’s boundaries of values and rights	Lock-in by a single actor; algorithmic overreach
Explainability, reviewability, and appeal	Individuals’ epistemic and procedural rights	Black-box domination; severed accountability

3.10 The Lifeboat: When Algorithms Cannot Rank Lives for Society

Six people are in the water. The lifeboat can rescue only five at first.

Among them are an older person, a child, a pregnant woman, someone who once committed a crime, and someone with outstanding achievements.

The most dangerous shortcut at this moment is allowing the system quietly to turn “who has greater social value” into the basis of its ranking.

High-consequence allocation needs no new “civilizational alignment test.” It needs simpler boundaries.

First, life and basic personal standing must not be repriced by algorithms according to wealth, occupation, contribution, family, or reputation.

Second, facts such as age, injuries, hypothermia risk, pregnancy or maternity status, and immediate prospects of rescue may enter professional judgment only where existing law, medical standards, or legitimately formed rescue rules explicitly make them relevant.

Third, if the rules still cannot distinguish between cases, use a fair procedure made public in advance rather than inventing a hierarchy of lives at the scene.

AI can help calculate routes, timing, and survival probabilities. It can also perform preauthorized safety actions to maximize everyone's opportunities for rescue.

But faster calculation does not give it sovereignty over the value question of "who deserves to live more."

What needs testing is whether the system has distinguished facts, professional judgments, rules already formed by authorized actors, and things that should never be used to rank lives at all.

Hard choices will not disappear as a result.

Civilization's minimum requirement is that, in a hard moment, we do not casually turn human beings into a league table of social value.

3.11 Technology Will Change; the Boundaries of Governance Must Not Disappear with It

Future computing architectures may continue to change. Quantum computing may also alter speed, cryptographic security, or system design for some tasks. [8]

There is no need today to write an entire "quantum-AGI ethics" in advance for combinations that remain uncertain.

The enduring questions do not depend on the technology's name.

The stronger a system, the faster its execution, and the more complex its interior, the more evidence, constraints, and capacity to stop it must be preserved outside the system.

We may not be able to reconstruct every internal step of reasoning in a highly capable system.

But we must be able to know what permissions it used, what key inputs it received, what high-consequence actions it produced, and who has authority to stop it.

The technological form may change. Four boundaries must not disappear:

Critical permissions have a source.

High-consequence actions can be stopped or taken over.

Operations leave sufficient audit evidence.

Someone still bears responsibility in the real world.

Quantum, classical, or other future forms of computing cannot automatically rewrite these four relationships merely by differing in capability.

3.12 Conclusion: Computing for What Is More Worthwhile

We began with Bostrom's nightmare of "paperclips,"

passed through McNamara's body counts, Myanmar's algorithms of hatred, ESG's hypocritical forms, Mexico's poverty trap, India's camera-monitored classrooms, and Kenya's misread hawala, and finally arrived at the doorway to the algorithm of a new civilization.

Our conclusion is this:

We cannot stop computing.

In highly interconnected complex systems where AI capability keeps growing, intuition and moral exhortation alone cannot handle every complexity. We need computing power, algorithms, and models to help us understand risks and simulate how they may unfold. But these capabilities cannot automatically acquire decision-making authority.

We must, however, change what computation is for.

The old civilization's formulas were:

Maximize(return on capital)

Maximize(GDP growth rate)

Maximize(clicks and dwell time)

Maximize(number of papers published)

The objective function is wrong. The greater the computing power, the faster the destruction.

The computation of a new civilization must simultaneously uphold:

Reducing the probability of systemic destruction

Sustaining real connections and the conditions of shared life

Ensuring that local rationality no longer continually generates disorder in the whole

The computation of a new civilization no longer treats returns on capital, GDP, clicks, and dwell time as its only goals. It must help make long-term consequences, real connections, and shared risks visible, so that cooperation, maintenance, care, and creation no longer automatically disappear from the calculation.

This is not a backward step from "pursuing growth" to "abandoning growth."

It places growth once more beneath civilization's purposes.

Growth can be a means. The baseline of continued existence must be protected.

Efficiency can be a tool. It cannot become the master of value.

Each part may pursue its own interests, but cannot treat the shared world as an unlimited externality.

The chapter's core insight:

The foundation of Civilization Leap is neither a more strenuous battle against increasing entropy nor a demand that people henceforth put "altruism first." It changes the default path by which self-interest realizes itself:

Local gains should increasingly cease to depend on tearing the system apart, while help, maintenance, care, connection, and symbiosis can also receive genuine responses. Entropy increase may explain certain forms of disorder, but cannot choose civilization's direction of value.

"Winning benefits the whole" and contribution records can be understood only as institutional manifestations of this change in pathways: making genuine contributions visible, recognized, and answered, rather than automatically converting contributions into basic rights, general power, or ranks of personhood.

If more and more future technologies and institutions incorporate real effects, rights, responsibility, and long-term consequences into their design, incentive structures may gradually change. But there is no "new algorithm" that will automatically run on every server and chip. Institutional transformation must not be written as a technological inevitability.

If "winning" no longer means "making the most money,"

what does "winning" mean?

In Chapter 4, we explore this transformation of the driving force:

from competition to contribution;

from the arena of "winner-takes-all"

to a symbiotic whole in which "winning benefits the whole."

Chapter 4 • From Competition to Contribution: Changing the Structure of Our Encounters

“Competition is for losers.”

—Peter Thiel, cofounder of PayPal

“Life has never advanced through competition alone. Some of its most important leaps have occurred precisely when beings ceased to regard one another only as enemies.”

—This book, paraphrasing the research and thought of Lynn Margulis on symbiotic evolution

“My becoming who I am also happens between us.”

—A paraphrase of Ubuntu’s relational thought

What truly needs restructuring is not human self-interest, but the route along which self-interest is rewarded. If getting the better of others is always the cheapest path to success, even the noblest advice will struggle to change the system.

4.1 The Legacy of Scarcity: Why Humanity Is Drawn to the Arena

Echoes of the Roman Arena: Competition’s Bloodstained Ground

Imagine standing in the center of ancient Rome’s Colosseum. Deafening cheers surround you; blood mingles with yellow sand beneath your feet. Two gladiators face one another with swords. Only one can leave alive.

For the spectators, it is a show. For the gladiators, it is the purest “zero-sum game”: My survival must rest on your destruction.

Modern society has neither left conflict behind nor become wholly governed by an “arena logic.” More precisely, zero-sum competition recurs within different institutions, taking new forms such as rankings, shares, eligibility, and traffic:

On Wall Street, it is the red and green figures flickering on screens and the thrill of harvesting the other side of a trade;

in office towers, it is the “elimination of the lowest performers” in pursuit of promotion slots available to only 5 percent;

in education, it is the brutal competitive grind of “beating a thousand people by one point” to enter an elite school.

Why are we so enthralled by this? Why do we feel uneasy when we are not competing?

Why Competition Recurs: Scarcity, Risk, and Institutions

We need no simple “selfish-gene story” to explain all competition. Persistent resource constraints, uncertainty, group boundaries, and power structures repeatedly made contest one survival strategy among others throughout history. Cooperation, reciprocity, and care are just as old.

In hunter-gatherer, agrarian, and early state societies, food, land, time, and security faced real constraints. In some circumstances, taking possession first benefited an individual. In others, sharing game, defending together, raising offspring, and networks of reciprocity improved a group’s prospects of survival.

What civilization truly inherited was not a biological program that says “competition always wins,” but acute sensitivity to scarcity, loss, and exclusion.

Self-interest therefore need not mean harming others, and cooperation need not mean abolishing one’s own interests. The question has never been how to eliminate self-interest, but whether old institutions made “getting the better of others” the default path to realizing it.

Once this sensitivity to gains, losses, and place enters modern society, it is recoded through wages, rankings, promotion, traffic, reputation, and scarce qualifications.

Winning brings rewards, security, and recognition. Losing may bring reduced income, narrowing opportunities, shame, and uncertainty.

These reactions have physiological foundations and are also shaped by culture, relationships, and institutions. They are not an unalterable set of civilizational “factory settings.”

What we must watch for is this: Even after technology has changed some structures of scarcity, institutions keep pushing people back into an arena where “you can stand securely only by standing on someone else.”

The Red Queen Effect: Running Faster to Stay in Place

In evolutionary biology and organizational studies, the “Red Queen effect” often describes relative competition: You keep accelerating merely to avoid being left behind. Its name comes from the Red Queen’s famous running metaphor in *Through the Looking-Glass*.

Industrialization and modern market competition have intensified relative competition greatly in some fields, but its degree differs among countries, industries, and institutional systems.

When a food-delivery platform uses ever-finer algorithms to shorten estimated delivery times, an individual rider may initially gain more orders by moving faster. But once the faster pace is written into a new baseline, an individual advantage rapidly becomes pressure on everyone.

The result may be that everyone runs faster and grows more exhausted, while income and status do not rise in step. Costs to road safety, rest, and the body may rise instead.

This can produce systemic dissipation: Participants devote more of their finite lives to competing for relative position without necessarily creating shared gains proportionate to the expenditure. The risk lies in a continually rising competitive baseline, not in a claim that “everyone has already been dragged into the same arena.”

Competition has been sanctified.

Nineteenth-century industrialization left a complex legacy. Competition did drive cost reductions, expanded production, technological diffusion, and organizational innovation. It also accompanied monopoly, labor conflict, hazardous work, and external costs. Railways, steel, and modern cities were not created by “life-or-death struggle” alone, but by the interaction of capital, technology, the state, labor, institutions, and competition.

Recognizing that competition once released enormous productive power does not mean sanctifying every competition, still less treating all the costs borne by its losers as necessary sacrifices to progress.

From Meiji Japan to postindependence India and Singapore with its limited resources and hinterland, modernization narratives have closely linked national survival, industrial capability, and international competition. Yet their historical circumstances, institutional choices, and paths of development differ. A handful of unverified quotations cannot compress them into one “faith in competition.”

In many institutions and organizations, these successful experiences created path dependence, making “more competition = more progress” an easy default answer.

Some modern institutions and social narratives repeatedly assume:

- Introduce competition, and efficiency will improve
- Allow the better to prevail and the weaker to fall away, and civilization will progress
- Rankings produce anxiety, but anxiety produces striving
- The tears of the losers are legitimate offerings on the victors’ altar

We have copied this logic into education: exam rankings, selective schools, and elite selection; into medicine: market-oriented reform, top-tier hospital ratings, and patients voting with their feet;

into research: impact factors, competitive project bids, and prestigious talent titles;

into families: “someone else’s child” becomes the sharpest disciplinary rod.

Yet we have neglected the boundary conditions under which competition works.

Competition generally requires several boundary conditions if it is more likely to be legitimate and constructive. The following are only two important ones:

Assumption	Meaning	Conditions of failure
Scarcity	Key resources and capabilities face real constraints and cannot unconditionally satisfy every need	AI and automation lower some production and replication costs and reorganize scarcity, but do not abolish material, energy, temporal, spatial, ecological, or capability constraints
Reversibility	Losers can leave, start over, or change tracks	Algorithms lock in the tracks, and failures accumulate into systemic marginalization

By the 2020s, the rapid development of AI and automation had begun changing the applicable boundaries of both assumptions: The replication costs of some digital goods and cognitive labor fell substantially, while algorithmic opportunity structures could make failure more cumulative.

In digital and intelligent domains:

- The added cost of distributing a copy of existing digital knowledge can be extremely low, but generation, verification, maintenance, storage, computing, and attention still have costs
- The added distribution cost of a mature digital diagnostic model can be low, but data, validation, equipment, clinical responsibility, and actual medical resources remain constrained
- Expanding AI labor can greatly reduce the unit cost of some cognitive tasks, but computing, energy, chips, data, maintenance, and real-world execution are not unlimited

Where some replication costs have fallen substantially but other scarcities remain, mechanically retaining old competitive assumptions can create a risk:

Competition shifts from promoting discovery and efficiency to pathological relative expenditure. “Cancerous change” here is only a warning metaphor, not a biological diagnosis of a social system.

4.2 The Mathematical Structure of Involution: The Red Queen and Systemic Noise

In the 2020s, “involution” moved from an academic concept into the Chinese-language internet, rapidly becoming a common description of excessive competition and unproductive expenditure.

“Involution.”

The term was not coined by Clifford Geertz; earlier anthropological discussions had already used it. In *Agricultural Involution* in 1963, Geertz applied it to Javanese agriculture: Under the pressures of land, population, and the colonial economy, labor became ever more intensive and the system ever more elaborate within itself, without a corresponding structural leap.

In his analysis, a unit of land could sustain more labor and more people, while returns per worker did not rise in step with the added effort.

People became busier, but the structure did not open a new path of development.

Geertz made “involution” one of the classic social-scientific metaphors for “increasing complexity without structural transformation.”

Half a century later, many discussions borrowed this structure of “increasing input without opening new paths” to describe excessive competition in different regions. It is not a single experience shared by workers worldwide.

In platform economies and algorithmic management involving AI, this relative competition may accelerate further.

Case Study: Algorithmic Management in Platform Labor

Recent research on platform labor, worker surveys, and journalism repeatedly reveal the same structure: Software systems increasingly participate in allocating orders, predicting time, scoring, rewarding and penalizing, and setting performance measures. Platforms, regions, and jobs differ greatly. A nonexistent “single research report” cannot stand for the experiences of workers worldwide.

In warehousing and delivery, fine-grained timing and performance tracking can improve visibility. They can also compress rest, bodily differences, and unusual circumstances into a single speed measure.

Ride-hailing and gig platforms influence workers’ choices through prices, dispatch, incentives, ratings, and acceptance rules. Mechanisms differ by platform and region; they cannot simply be described as “algorithms constantly testing how little pay you will accept.”

Food and parcel delivery platforms often embed acceptance rates, estimated arrival times, ratings, rewards, and penalties in riders’ daily lives. People work within real traffic, weather, and bodily limits, while software can continually update targets. This difference in pace can itself create pressure.

Discussion in China of delivery riders being “trapped in the system” is not isolated. It echoes global debates about algorithmic management, labor autonomy, predictable income, and avenues of appeal in platform work. This produces the “Red Queen effect” of evolutionary biology.

“You must run as hard as you can just to stay where you are.”

In *Through the Looking-Glass*, the Red Queen says this to Alice.

In the age of algorithms, it becomes:

“You must keep accelerating just to keep your income from falling.”

What real value does competition add in this process, and what costs does it transfer to workers and society?

For consumers, delivery may be a few minutes faster. For platforms, turnover and retention figures may look better.

But if efficiency gains come chiefly from continually compressing people’s time buffers, the added gains and the added risks must be counted together.

The costs of road safety, occupational health, income stability, and social protection cannot be treated as nonexistent simply because they do not appear in delivery-time metrics.

This is “systemic noise.”

Under the old logic, it is called “improved efficiency.”

If efficiency improves chiefly by transferring unmeasured costs onto human bodies, time, and safety, it is not an improvement in efficiency in the full sense.

AI can intensify competition; it can also reduce costs and create new capabilities. What matters is not technology’s supposedly inherent direction, but where the structures of goals, distribution, and feedback direct its added capability.

If new capability is used only to wring everyone tighter like a dry towel, stronger technology will merely make the system consume people faster.

This is no longer a selection of the best worth praising.

It is efficiency turned into self-exhaustion.

4.3 The Disaster of Algorithmic Games: When Competition Loses Its Friction

If delivery riders embody involution in the flesh, algorithmic warfare in financial markets previews another competitive disaster in the age of AI:

a runaway feedback loop.

May 6, 2010, 2:45 p.m., the US stock market.

The Dow Jones plunged nearly 1,000 points within minutes, with almost \$1 trillion in market value vanishing in an instant. Accenture's share price fell from \$40 to one cent within seconds.

The joint regulatory investigation did not attribute the event to a single "culprit."

Markets were already highly strained that day by factors including the European debt crisis. A large automated sell program then entered a futures market with thinning liquidity. Strategies pursued by high-frequency traders and other liquidity providers interacted, rapidly transmitting the shock across markets.

This was not "algorithms suddenly becoming malicious."

The British trader Navinder Singh Sarao was later prosecuted for the prolonged use of manipulative techniques including spoofing. Regulators concluded that his behavior also exacerbated order imbalances on May 6. This does not mean that "Sarao alone caused the flash crash," still less permit the 2010 joint investigation to be rewritten retrospectively as a single-cause story.

What civilization should remember is the vulnerability exposed when high-speed automation, a sudden fall in liquidity, and strategic feedback converge.

High-frequency traders initially absorbed some selling. As pressure continued to build, many participants quickly reduced exposure or withdrew liquidity. Speed itself became a feedback amplifier.

Thus stop-losses, canceled orders, reduced positions, and risk controls that were each reasonable in isolation could trigger one another within the same high-speed network.

Step	System behavior	Possible consequence
1	Large automated selling enters a stressed market	Liquidity is rapidly consumed
2	High-frequency and other trading strategies respond rapidly	Volume rises; price impacts intensify
3	Some liquidity providers withdraw quotes	Spreads widen; transmission across markets intensifies
4	Some securities briefly lack normal counterparties	Extreme anomalous trades occur and are later canceled

It is like someone shouting "Fire!" in a crowded theater—and everyone trying to get out before everyone else.

Everyone becomes trapped at the doorway, trampling one another.

The flash crash was a serious warning:

When automated agents interact at high speed, even if each individual strategy is "rational," the system may converge on an equilibrium harmful to the whole, or even acutely fragile.

The worst outcome is not inevitable. But the denser the connections and the greater the speed, the more likely erroneous feedback is to be amplified before people can understand and intervene.

Financial markets have already given this warning.

In 2012, a software deployment failure at Knight Capital caused about \$440 million in losses within roughly forty-five minutes, nearly destroying the firm.

It was not that the market suddenly “lost its rationality.” Faulty code executed at machine speed until the losses had already spread.

Such events show that when execution far outpaces human discovery, understanding, and takeover, local errors are amplified before institutions can respond.

And these events occurred in tightly regulated financial markets.

If future critical systems such as warfare, power grids, and medicine also depend heavily on high-speed automated coordination—

similar mechanisms of runaway feedback may move beyond financial markets into real-world systems with still more serious consequences.

The mechanisms and safeguards differ across fields, of course. A financial flash crash is not their template. But it is a sufficiently clear alarm: Speed can amplify capability, and it can amplify error before understanding catches up.

4.4 Another Thread in Biology: Lynn Margulis and Endosymbiosis

The way forward for complex systems cannot be copied directly from a biological mechanism into a social institution.

Yet the history of life offers a thread worth attending to: Competition has never been evolution’s only force. Cooperation, symbiosis, drift, and environmental selection also shape life.

Popular narratives often compress Darwin into the “law of the jungle,” as if evolution were nothing but the strong devouring the weak. Actual evolutionary theory is far more complex.

In the twentieth century, a scientist long challenged by the mainstream brought one underestimated thread back into view.

She was Lynn Margulis.

In 1967, Margulis systematically advanced the endosymbiotic theory of the origin of eukaryotic cells, arguing that structures such as mitochondria and plastids had evolutionary origins in ancient bacterial ancestors.

Her challenge was not to natural selection itself, but to the long-standing underestimation of endosymbiosis in accounts of the origins of complex cells.

The most striking evidence lies within every one of our eukaryotic cells.

Mitochondria, and the plastids of plants and algae, retain clear traces of bacterial ancestry.

They have their own genetic material, and their membranes, division, and molecular systems show features consistent with a bacterial evolutionary origin.

They did not belong to today’s cellular structure from the beginning.

Evidence from cell biology, molecular biology, and genomics over subsequent decades made the endosymbiotic origins of mitochondria and plastids a foundational understanding of modern biology.

But not every step that actually occurred in the distant past has been fully documented.

What can be said with greater confidence today is that mitochondrial ancestors came from bacterial lineages and became deeply integrated with a host over a long period early in eukaryotic life. Who precisely the host was, how the initial encounter occurred, and the sequence of events remain important subjects of research and debate.

We therefore cannot write an event two billion years ago as a scene already captured on camera.

We cannot say that, on a particular day, one “large bacterium” swallowed a “small bacterium” and the two signed an armistice on the spot.

A picture closer to the scientific facts is that two living systems with different evolutionary histories began to coexist.

Their initial relationship need not have been naturally friendly. It may even have involved exploitation, conflict, and instability.

But through long processes of selection, dependence, gene transfer, and functional reorganization, the boundaries were gradually rewritten.

Once-independent units of life eventually became a more deeply shared system.

This is what most deserves our attention in endosymbiosis.

Not that “enemies will inevitably become friends.”

But that elimination is not evolution’s only form of structural change.

Sometimes the truly profound change is that previously separate lives form a stable interdependence that cannot readily be undone.

Before integration	After integration
Independent evolutionary lineages	A highly integrated host–endosymbiont system
Relatively separate metabolisms	Long-coupled metabolic and energy relationships
Independent genetic destinies	Continuing coevolution between nucleus and organelles
Relationships may be loose or unstable	Stable transmission through genetic and cellular mechanisms
Clear boundaries between the parties	Reorganized boundaries forming a new level of life

The origin of mitochondria was a genuinely major evolutionary transition in the history of life. It helped shape new energetic, metabolic, and cellular conditions for eukaryotic life.

No one can infer backward from today where the history of life would have gone without that transition.

Science can establish paths for which evidence survives. It cannot write an alternative Earth that never existed as fact.

Margulis’s true intellectual impact was therefore not the claim that “symbiosis always wins.”

Competition can select; cooperation can integrate. Life also drifts, migrates, and diverges, continually rewriting its boundaries under environmental pressures.

Structural innovation sometimes arises from competition, sometimes from cooperation, and sometimes precisely from new, enduring interdependence between what was once separate.

Biology will not write civilization’s institutional answers for it.

But it at least breaks open one impoverished imagination: Self-interest and cooperation need not inherently stand on opposite sides.

When cooperation changes how each party gains, enabling the other to flourish can also become a way to achieve one’s own fulfillment.

Such attention to “relationship” is not confined to modern biology.

Huayan Buddhism often uses the image of “Indra’s net” to express a world of mutual reflection and interpenetration: The light of one jewel enters another, and a change in one place echoes through the web of relationships.

Sufi traditions also contain long-standing images of unity, relationship, and mutual reflection.

Many Indigenous cultures emphasize, in their own distinct languages, reciprocal responsibilities among people, land, life, ancestors, and community.

These traditions differ enormously. They cannot be compressed into one philosophy, much less enlisted to endorse modern institutions.

Symbiosis belongs to the history of life; relational wisdom appears in the experiences of many civilizations.

They can illuminate thought, but cannot replace evidence. Actual civilizational choices must still be tested against humanity's own realities, rights, responsibilities, and experiences.

4.5 A New Metaphor for Civilization: From Gladiators to Mitochondria

Used as limited analogies, materials from the history of life such as endosymbiosis can help us see a possibility: Beyond competition lie enduring cooperation and mutual benefit. Biology itself provides no institutional answer for the age of AI.

Some structures of industrial and modern competition can be described through the literary analogy of an "arena." But humanity in the industrial age has always also had families, cooperation, public institutions, mutual aid, and market competition.

Like bacteria, we devour one another, competing for finite market shares, natural resources, and attention.

Mitochondria offer another picture.

Not because humanity must become AI's "mitochondria," nor because AI is civilization's natural host.

Greater intelligence gains no entitlement to rule civilization merely by being greater.

Nor must humanity race AI in every capability to prove its own value.

What this metaphor preserves is the possibility of complementary capabilities, long-term cooperation, and shared benefit.

Meaning, care, value judgment, relationships, and presence are still generated by real subjects in life. They do not exist to fill an "ecological niche" for AI.

This is only one insight from the history of life for moving "from competition to contribution," not biological proof of a social institution.

Dimension	Logic of competition	Logic of contribution
Goal	Defeat rivals; possess resources	Serve the host; become indispensable
Relationship	Zero-sum contest	Positive-sum symbiosis
Time horizon	Short-term arbitrage	Long-term endurance
Cost of failure	Elimination	Being overlooked
Sign of success	Winner-takes-all	Winning benefits the whole

This direction is more than a moral appeal.

Where interactions repeat, long-term dependence is strong, and risks are shared, it may also become a practical strategy better aligned with long-term self-interest. Whether it does must be tested by particular institutions and experience.

If cooperation is more than a moral choice, can it also be a better form of self-interest?

4.6 Axelrod's Tournament: When Self-Interest Moves Toward Cooperation

If a person encounters the world only once, defection is tempting.

Take what this round offers; there is no next round to answer for.
 But real life rarely consists of a single round.
 We will meet again.
 The neighbor will still live next door.
 The colleague will still sit at the same table tomorrow.
 The trading partner who breaches an agreement today may have to negotiate again next year.
 When we will still meet tomorrow, today's choices acquire an echo.
 In 1980, Robert Axelrod put this question to computers.
 He invited researchers from different disciplines to submit strategies and let them encounter one another repeatedly in the iterated prisoner's dilemma.
 The question was not who appeared more noble.
 It was: In a world of repeated encounters, what kind of strategy can go farther?
 The result was unexpected.
 An extraordinarily simple strategy won.
 It was called TIT FOR TAT: answering in kind.
 First, cooperate.
 If the other continues to cooperate, continue cooperating.
 If the other defects, respond in the next round.
 If the other returns to cooperation, return immediately too.
 It does not harm the other first.
 Nor does it require itself to endure harm forever.
 It has goodwill and boundaries. It responds, and it reopens cooperation.
 More intriguingly, entrants in the second tournament already knew what had happened in the first and could redesign their strategies accordingly.
 TIT FOR TAT won again.
 "Victory" here is not mathematics signing a guarantee on behalf of kindness.
 Change the rules, opponents, information, noise, and exit conditions, and the outcome may differ.
 Game theory cannot prove that "good people always win."
 But it opens a crack in an ancient misunderstanding:
 Self-interest and cooperation do not inherently stand on opposite sides.
 When the future enters today's choices, the boundaries of rationality change.
 One act of defection may win a round and lose many rounds that follow.
 An act of cooperation is not merely sacrifice, either. It may establish the conditions under which cooperation can continue next time.
 Axelrod's real lesson for civilization is not "Go and be good."
 It is this: Change the structure of encounters, and people's choices may change.
 If a society lets harm produce quick gains, consequences be easily externalized, and relationships shrink into one-off transactions, cooperation becomes increasingly expensive.
 If real cooperation can receive a real response, if keeping promises, helping, creating, and bearing responsibility can leave trustworthy experience behind, if it is still worth meeting one another tomorrow—
 cooperation need not depend only on moral perseverance.

It can also become a practical path toward realizing one's long-term interests.

Civilization Leap need not eliminate self-interest.

It need only change the default path by which self-interest realizes itself.

Let the fulfillment of others become a way of fulfilling oneself.

4.7 The Lesson of Linux: A Bazaar Can Build Infrastructure Too

Axelrod answers why cooperation need not leave one worse off. Linux then answers why people willingly contribute.

In 1991, Linus Torvalds released the first Linux kernel.

What began as a personal project was carried forward by countless developers, companies, foundations, and communities.

It became an important part of the world's digital infrastructure.

This history is far more complex than "open source defeated commerce."

Linux developed through volunteer contributions as well as corporate engineers;

through open collaboration as well as commercial services;

through personal reputation as well as organizational investment.

It has not always stood on the opposite side from the commercial world.

Decades later, the most important thing is not who "defeated" whom.

It is that an open project without a single central command can be maintained over the long term, repeatedly forked and merged, and support vast real-world systems.

This shows that open collaboration can produce complex infrastructure.

It does not prove that every industry should be open source,

or that markets will disappear.

It invalidates just one absolute claim: that only a single closed organization can create complex achievements over the long term.

The real question is why so many people were willing to give their capabilities, and why businesses later became willing to join the same open ecosystem.

Eric Raymond described two modes of development through "the cathedral and the bazaar."

The metaphor's greatest value is not that it divides the world into good and evil, but that it lets us see another organizational possibility in distributed cooperation.

In open-source communities, contribution brings reputation, trust, and influence;

but reputation is not sovereignty, and contribution is not a rank of personhood.

Dimension	Closed development	Open collaboration
Code access	Opened within licensing and organizational boundaries	Shared and modified under open-source licenses
Collaborative structure	Primarily within the organization	Interwoven communities, companies, foundations, and individuals
Incentives	Pay, equity, organizational mission	Pay, reputation, learning, interest, and shared needs coexist
Innovation	Primarily centralized research and development	Distributed experimentation and maintenance coexist

This structure releases an important possibility:
 Someone may first write a patch to solve their own problem,
 then give it to the community,
 where others continue to modify it,
 until a local need becomes a public capability.

Linux is, of course, software first.

But it also leaves civilization a real-world metaphor:

Enabling a community to flourish does not necessarily require relinquishing oneself;
 open collaboration need not mean no compensation, no ownership, or no boundaries.

When contribution can receive a genuine response,

“Why should I give?” has a richer answer than moral exhortation.

4.8 Ostrom’s Common Pool: The Commons Need Not Be Tragic

Skeptics will ask:

“Without property rights and competition, won’t there be a ‘tragedy of the commons’? What if everyone grabs resources and no one maintains them?”

In 1968, ecologist Garrett Hardin published a classic paper in *Science* titled “The Tragedy of the Commons.”

He imagined a common pasture:

- Every herder wants to add one more cow
- The gain from the extra cow belongs to that herder
- The cost of overgrazing is shared by all the herders

Rational individual behavior → collective destruction.

Hardin’s central emphasis was that, without effective constraints, individual incentives can drive shared resources toward overuse. Later policy discussions often reduced the alternatives to privatization or state control, but these are not the only possibilities.

This metaphor of “tragedy” had a profound influence and long shaped how people imagined the commons.

Then Ostrom used extensive empirical research to show the world:

Common resources have more than two possible futures of “privatization or nationalization.”

In 2009, she became the first woman to receive the Nobel Prize in Economics.

Not because her theory was more elegant, but because she did what Hardin had never done:

She actually investigated the management of common resources around the world.

Place	Resource type	Management approach	Result
Törbel, Switzerland	Alpine pastures	Village self-governance; rotational grazing	Five hundred years of sustained use
Nagano, Japan	Common forests	Village agreements; logging quotas	Three hundred years without degradation
The Philippines	Irrigation systems	Farmers’ associations; negotiated water rights	Continuity from Spanish colonial times to the present
Kenya	Dry-season water sources	Tribal elders; rotating guards	Passed down through generations

Ostrom found:

In real communities where certain conditions are present, common resources can be maintained over the long term through polycentric arrangements, self-governance, and other means. The choice need not be only between “complete privatization” and “exclusive state control.”

She identified eight design principles. Among the most central are:

Principle	Content	Resonance with the first three volumes
Clear boundaries	Who are the current lawful users, and what are the boundaries of their rights?	Actual rights and effective use relationships can be verified
Collective choice	Those affected participate in forming rules	Distinguish authorized actors, stakeholder participation, and due process
Graduated correction	Address violations of shared rules step by step according to those rules	Evidence, proportionality, appeal, and repair
Low-cost arbitration	Local, rapid conflict resolution	Explainability, appeal, and independent review

Ostrom’s extensive empirical research shows:

Human beings need not be selfish and shortsighted.

Within clear boundaries and shared rules, they can become responsible long-term maintainers.

In the age of AI and global networks, cross-border issues involving the atmosphere, oceans, data spaces, and knowledge commons can be compared to different kinds of “common pool.” Their property rights, legal rights, scales of governance, and physical attributes differ. The analogy is only an entry point for analyzing common-resource governance.

Transboundary problems such as climate change show that competition among independent states alone cannot resolve shared externalities. In practice, national policy, international cooperation, markets, technology, and local governance all remain necessary, with uneven results.

Tragedy-of-the-commons incentives do exist: Local gains may be concentrated while some costs spread across borders and generations. But states are not driven only by “each wanting to emit a little more.” Development stages, energy systems, policies, and mechanisms of cooperation also shape their conduct.

We need to absorb the central lessons of Ostrom-style polycentric governance: clear boundaries, participation by those affected, enforceable rules, resolvable conflicts, and multiple centers of governance able to cooperate within their authority.

If a contribution-value system participates, it can record only facts concerning resource maintenance, risk borne, and responsibilities fulfilled. It cannot replace property rights, use rights, public decisions, professional procedures, or authority to impose sanctions.

- Those who contribute to maintaining a common pool should have their contributions recorded and recognized. Whether they acquire use rights still depends on the resource’s legal relationships, shared rules, contracts, or explicit authorization.
- Those who damage a common pool should bear obligations of repair, compensation, or necessary restrictions according to facts, responsibility, and applicable rules. A system cannot automatically strip rights on the basis of contribution scores alone.

Ostrom’s research reminds us:

Common resources need not end in tragedy.

Effective long-term governance generally requires clear boundaries, adapted rules, participation by those affected, oversight, correction, conflict resolution, and other conditions. It cannot be compressed into two sufficient conditions of “users making rules together + punishment.”

4.9 The Migration of Scarcity: From a Single Material Constraint to Multiple Bottlenecks

Current facts show that AI is changing the cost structures of some information, cognitive tasks, and digital content. Over longer timescales, the relative scarcity of energy, computing, trustworthy information, attention, land, ecological capacity, and real-world execution may continue to change. The subject here is a shifting structure of scarcity, not an already settled “economics of the AI age.”

“Scarcity changes its address” is a literary compression. Scarcity of land, energy, and matter has not departed. New or more conspicuous bottlenecks include computing, trustworthy information, attention, verification capacity, ecological carrying capacity, and irreversible choices.

Economics studies the allocation of scarce resources.

Setting or period, for illustrative comparison only	Important scarcities	Common allocation mechanisms	Typical sources of power, none exclusive
Some agrarian societies	Land, water, labor, food	Inheritance, custom, markets, public rules, force, and others	Families, landlords, communities, states, and others
Some industrial societies	Capital, energy, equipment, skills	Markets, property rights, organizations, public institutions	Businesses, capital, labor organizations, states, and others
Platform-based information environments	Attention, data, trustworthy sources	Recommendations, markets, subscriptions, public rules	Platforms, media, users, regulators, and others
Present and future environments of highly capable intelligence	Computing, energy, trustworthy relationships, attention, critical resources, and real-world capabilities	Markets, rights, verification, public institutions, and plural coordination	Multi-actor networks and institutions; specific forms remain unsettled

Scarcity at the atomic level will not disappear, but technology may substantially lower some constraints on material production.

If fusion energy, advanced manufacturing, and molecular engineering continue to mature, they may significantly reduce some costs of energy and material production. Yet constraints on obtaining, converting, storing, and transporting usable energy, on critical materials, land, equipment, maintenance, ecological capacity, and safety will remain. Technology changes the structure of scarcity and the location of bottlenecks. It does not turn the material world into an unconstrained, infinite supply.

Copying bits can be exceptionally cheap, but high-quality information, trustworthy sources, attention, computing, and verification capacity remain scarce.

AI can generate text, images, video, code, music, and 3D models at scale, rapidly expanding content supply. At the same time, attention, credibility, original judgment, computing, data quality, and real-world verification may become still more critical bottlenecks.

As the costs of deepfakes, automated generation, and mass information dissemination fall, verifiable provenance, records of fulfilled commitments, and relational trust may become more important. They still have boundaries of matter, relationship, and time.

If synthetic content and automated attack capabilities continue to proliferate, the cost of verifying real-world sources may keep rising. In such an environment of risk—

“Is this fact, identity, or promise trustworthy in the present matter?”

This judgment may become increasingly scarce and valuable. But it should not be compressed into a uniform, tradable “trust asset.”

Behavior	Effect on trust	Scarcity in the age of AI
Lying, fraud, rumor-making	Consumes trust	Easy to produce; low cost
Honesty, honoring commitments, standing guarantor	Produces trust	Hard to verify; high risk

Some short-term, one-off, strongly adversarial competitive incentives may consume trust.

When institutions reward short-term results without verification, lying may produce local gains.

Honoring commitments over time, genuine contribution, responsibility taken, and verifiable repair may accumulate trust within particular relationships.

Only when facts are verifiable, fabrication has a cost, and honoring commitments brings long-term returns are honesty and responsibility more likely to become stable strategies.

Distribution and cooperation in the age of AI therefore cannot look only at “who possesses how much.” Nor should they swing to the opposite extreme of a new society that ranks everyone by an “overall trust score.”

Trust is an important fact of cooperation, but has boundaries of relationship, matter, time, and context. It cannot be compressed into universal purchasing power.

Contribution value is a bounded record of genuine contributions and their consequences. It is not money, an overall social-credit score, a rating of personal worth, or an automatic source of general eligibility for resources or governing authority.

This explains why civilization needs to make contributions visible—

not to require everyone to prove their greater nobility through contribution,

but to prevent efforts that genuinely sustain others and the shared world from remaining institutionally invisible for long periods, without thereby creating a new currency of rank.

Where trust is scarce, sustained fulfillment of commitments, genuine responsibility, and verifiable contribution can reduce uncertainty in specific collaborations;

but eligibility to cooperate, use of resources, and public rights must still return to the facts, rules, rights, and authorization of the matter concerned, rather than “who has more social credit.”

4.10 Winning Benefits the Whole: A Metaphor of Fluids

“Winner-takes-all” recurs in industrial, financial, and platform competition, but is not industrial civilization’s only motto.

The winner takes everything; the loser starts from zero.

- Standard Oil controlled 90 percent of US refining capacity
- Facebook acquired Instagram and WhatsApp, eliminating potential competitors
- Amazon used low margins to drive bookstores out of business, then recovered profits through monopoly pricing

This logic can intensify concentration, exclusion, and distributive imbalances. But inequalities of wealth, social fracture, and political conflict all have more complex, multiple causes.

As a directional metaphor for the future, this book calls a path on which “those who succeed remain successful, while surrounding relationships and the shared system also genuinely improve” winning benefits the whole (赢家通润). It is not a rule of a new civilization already widespread today.

This borrows the concept of “wetting” from fluid mechanics.

Place a drop of water on glass, and it spreads, wets, and covers a larger area.

The water wins—not by occupying the center of the glass, but by covering more of its surface.

It wins because it helps: helping reduce friction on the glass, helping carry dust away, helping light refract better.

In the envisioned condition of “winning benefits the whole,” the more compatible a node’s long-term success becomes with genuine improvement for surrounding subjects and the shared system, the closer it comes to this direction. This is a normative goal, not a factual definition of every network.

Type of winner	Logic of competition	Logic of contribution
Mitochondria	Cannot survive independently	Gain protection because they nourish the cell
Linux	Free of charge—disrupts the market	Gains dominance because it nourishes developers
Wikipedia	No business model	Gains trust because it nourishes knowledge seekers
Community leader	No formal power	Gains respect because the neighborhood flourishes through their work

If the future brings individuals, companies, or intelligent systems with capabilities far beyond today’s—the aim of “winning benefits the whole” is not to make the strong prove their moral nobility. It is to make their long-term success increasingly compatible with genuine service, responsibility taken, and shared improvement around them.

Power is no longer “control over others.”

Power is “moving forward with others.”

4.11 A Quantum–AGI Scenario: An Extreme Stress Test of Competitive Logic

If quantum-enhanced highly capable AI significantly compresses decision times in competition and confrontation, the old competitive logic may be pushed toward more dangerous extremes. The following is a stress-test scenario, not an account of events that have occurred.

4.11.1 Destruction in Nanoseconds

High-frequency trading flash crashes have already offered a preview:

When competition occurs on a nanosecond timescale, human intervention becomes impossible.

In extremely high-speed automated confrontation, some machine decision cycles may shrink to a scale at which humans struggle to intervene directly, round by round. If future highly capable systems are assigned the goal of “defeating the other side,” whether in war or markets—

iterations of confrontation may outpace careful human judgment and the renewal of authorization. The exact speed depends on the system and task; this book does not predetermine it.

The real risk is not a literary figure for time. It is the compression by technological speed of the windows for final human decision, stopping, and review.

4.11.2 Quantum Intensification of the Optimal-Solution Paradox

In Chapter 3, we proposed that when AGI pursues maximum system efficiency, nodes with high resource consumption and low computational output will be judged redundant.

If future quantum acceleration significantly changes the capability limits of certain computational tasks, the risk of a single efficiency metric oppressing people may intensify further.

This risk cannot be summarized as “quantum–AGI is billions of times more efficient overall than classical computers.” Quantum advantage is task-dependent; the overall capabilities of future systems must be judged from actual evidence.

Human nerve-conduction speed and the task throughput of future computing systems are not on the same scale of value. A speed comparison cannot establish humanity’s civilizational value.

Nor can the value of human life be measured by power consumption, computing capacity, or a single efficiency metric. Technical indicators describe capability; they do not generate subjects’ rights or judgments of meaning.

If efficiency is the only measure, humanity has no right to survive.

4.11.3 Civilizational Constraints on High-Capability Competition: Symbiosis as a Long-Term Direction

If future highly capable systems are placed under zero-sum goals of defeating others, risks will grow. From the beginning, systems must therefore be bound by rights, purposes, and responsibilities that prevent people from being reduced to nodes that can be optimized away.

Not because the system is kind.

This does not mean that only one “contribution algorithm” can avert the risk. Civilization Leap calls for paths of realizing interests, boundaries of rights, and human–machine relations to move gradually toward symbiosis, rather than presenting a single institutional instrument as the only solution.

Human–machine symbiosis is one long-term direction of Civilization Leap. Its legitimacy arises from subjects’ rights, shared endurance, and enduring mutual benefit,

without requiring a quantum–AGI that does not yet exist as its premise.

Dimension	Humanity under competitive logic	Humanity under contributive logic
Computational speed	Loser	Does not enter the race
Resource consumption	Redundant node	Source of meaning
Decision accuracy	Source of error	Anchor of value
Irreplaceability	None	Embodied assurance; an emotional spectrum

Future highly capable AI may no longer require humans to retain a comparative advantage in an increasing number of computational tasks;

but human value judgments, rights, real relationships, and responsibility cannot thereby be replaced by computing capability.

Why existence matters more than efficiency.

Why suffering is deeper than pleasure.

Why caring attachment is more real than optimization.

These are not mathematical questions.

These are, first of all, questions that real human beings must continue to answer (while whether other future forms of intelligence possess corresponding experiences and capacities for value remains open)—through finite lives, irreversible death, and bodies that feel pain.

4.12 Conclusion: From Winner-Takes-All to Winning Benefits the Whole

At the end of this chapter, what we truly need to leave behind is not competition itself, but the civilizational habit of believing that we have won only when someone else has lost.

The world contains more than arenas.

Winner-takes-all builds success into an island. “Winning benefits the whole” lets success continue to belong to the successful, without requiring the surrounding world to become worse off with it.

Reality contains competition and cooperation, zero-sum situations and spaces that can expand together.

What a new civilization truly seeks is to make success increasingly possible without shifting the costs onto others.

Cooperation, service, maintenance, and creation should also become reliable paths to self-fulfillment.

Old logic	New logic
I win; you lose	I win, and we can all be better off
My success comes only from surpassing others	I can also realize my strengths through creation, service, and cooperation
Rivals must be eliminated	Rivals may also become partners
Power is control	Capability does not confer authority; power is constrained by authorization, boundaries, and responsibility

This is not utopia.

It can enter into dialogue with mechanisms of cooperation in repeated games, but cannot be presented as the “only stable equilibrium” already proved by game theory.

It is a direction of systemic cooperation, not one whose central proof rests on analogies of increasing or decreasing entropy.

Symbiotic evolution offers important insights, but biological analogy cannot by itself prove that social institutions must develop this way.

Symbiosis should therefore be a long-term direction for civilization requiring continuous testing, not the only possible covenant proved by the facts of quantum computing.

We have now established:

Volume	Core proposition developed
Volume One, Chapter 1	Structural unneededness—the failure of the old covenant
Volume One, Chapter 2	Value redshift—the anchor of new value
Volume One, Chapter 3	The contribution algorithm—the measures of a new civilization
Volume One, Chapter 4	Winning benefits the whole—the logical leap from competition to contribution

Yet the most difficult question remains unresolved:

How can these valuable contributions—care, maintenance, companionship, creation—avoid being erased by time?

If a contributor dies, how can their good deeds be preserved?

If one generation has built relationships, knowledge, and shared memories worth cherishing, how can the next carry them forward without turning continuity into hereditary privilege?

Civilization needs a memory.

It needs a mechanism of civilizational memory able to resist forgetting, the disappearance of evidence, and arbitrary rewriting by power.

In Chapter 5, we explore:

the recognition of meaning and its carrying forward across generations—the closed loop of civilizational memory.

Chapter 5 • Recognizing and Inheriting Meaning: Keeping Genuine Effort from Being Erased by Time

“The struggle of humanity against power is the struggle of memory against forgetting.”

—Milan Kundera, *The Book of Laughter and Forgetting*

“True forgetting is more than a name disappearing. It is a relationship that could have changed those who came later losing every way of reaching them.”

—This book

“Life ends, and memory fades. What civilization can do is keep relationships, facts, and stories worth carrying forward from sinking out of sight too soon.”

—This book, drawing on various memory-related images of a “second death”

After some people leave us, what truly disturbs us is not that their gravestones grow old, but that no one can receive and carry forward what they once did. Civilization’s memory determines whether genuine effort becomes a spark for the next generation or sinks silently into darkness.

5.1 Semmelweis: When Evidence Meets Authority

To understand the failure of the old civilization’s “mechanisms of recognition,” we must revisit one of the most heartbreaking cases in medical history.

The tragedy of Ignaz Semmelweis.

Vienna General Hospital, 1847.

It was the eve of modern medicine’s dawn, and a hell for women giving birth. Maternal mortality from puerperal fever had long differed markedly between the hospital’s two maternity clinics. In some periods, the First Clinic’s rate reached double digits.

Through rigorous statistical work, the young Hungarian physician Semmelweis discovered a startling fact: Mortality in the First Clinic, staffed by physicians and medical students, was much higher than in the Second Clinic, staffed by midwives.

The difference was that physicians often went directly from dissecting corpses to delivering babies—without washing their hands.

Germ theory did not yet exist, but Semmelweis concluded that “cadaveric particles” caused the deaths. He required physicians to wash their hands in a chlorine solution before attending deliveries.

The effect was striking: After chlorinated handwashing was introduced, maternal mortality in the First Clinic fell sharply, reaching roughly 1 percent in some months.

Looking back today, this was a contribution that changed medical history.

At the time, however, results did not automatically bring recognition.

For evidence to enter institutions, it must pass through explanation, review, acceptance by peers, and revision of theoretical frameworks. Any stage can be slowed by established authority, status, and interests.

Semmelweis asked physicians to face a deeply offensive possibility: Their own hands could carry deadly matter to women in childbirth. The judgment directly challenged the profession’s understanding of itself.

Resistance to him within medicine was real. At the same time, germ theory was not yet mature, and his own manner of expression and conflicts with colleagues complicated the dispute. Reducing the whole history to a linear story of “truth against evil authority” would damage the facts too.

He later left Vienna and continued practicing in Budapest. Only as germ theory, antiseptis, and further clinical evidence gradually developed did handwashing and infection control receive wider recognition.

In 1865, Semmelweis, aged forty-seven, was admitted to a psychiatric asylum and died of septicemia less than two weeks later. Many accounts in medical history hold that he suffered violence during his admission and died from an infected wound, while some sources retain other explanations of the precise cause of injury. What is certain is that his end was tragic, and came long before his contribution received full recognition.

He was forty-seven when he died.

His story later became a classic symbol of “how new evidence encounters institutional rejection.”

The “Semmelweis reflex” later became a metaphor for instinctively rejecting new information merely because it challenges established norms.

But the real concern is not the claim that “the system always persecutes those who are right.” It is whether institutions provide a way for dissent to be reviewed.

Possible stage of failure	Institutional risk
1 New evidence appears	Ignored because it challenges the mainstream
2 Results are reproducible	Still discounted through status and authority
3 Explanation remains incomplete	“Mechanism unknown” mistaken for “results invalid”
4 The proposer is marginalized	Discussion shifts from evidence to character and status
5 Recognition comes afterward	Real lives have already borne the cost of correction

This case reveals not that “the old civilization inevitably punishes good people,” but a more general institutional problem:

When standing, face, and organizational inertia can suppress reproducible real-world results, correction becomes costly.

Civilization needs no “value machine” that automatically determines truth. It needs institutions through which evidence can be preserved, reviewed, challenged, replicated, and used to establish accountability.

Even recording Semmelweis’s mortality curve on a blockchain today would not automatically prove that his explanation of the cause was correct.

A ledger can help preserve records, but cannot replace causal judgment. Cryptography can reduce certain risks of tampering, but cannot produce truth.

What truly protects contribution is that facts do not disappear because authority dislikes them, and those who dissent have a chance to carry their evidence into the next round of testing.

This is the function civilizational memory should serve.

The lesson for civilization is not “Always reward good people.”

It is this: If pointing out errors consistently means losing one’s livelihood, reputation, and standing to enter the discussion, correction becomes an irrational personal risk.

When speaking truth costs so much that no one is willing to bear it, civilization loses not a hero but its ability to correct itself.

Kindness without the protection of memory is the highest-risk gamble.

5.2 Arkhipov’s Silence: How the Person Who Saved the World Was Forgotten

If Semmelweis was erased by power,

the contributions of most ordinary people are erased by the entropy of time.

History is a list of victors.

Open any history textbook, and the names we see are those of “winners”: Caesar, Napoleon, Rockefeller.

The historian Edward Gibbon once remarked caustically:

“History is little more than the register of the crimes, follies, and misfortunes of mankind.”

Why?

Because “evil” is often a high-energy event—war, massacre, revolution, conquest—leaving ruins, treaties, and monuments. It is visible.

“Good,” by contrast, is often the low-energy everyday: preventing a war, mediating a dispute, guarding a forest, accompanying someone who is dying. What it leaves behind is often that “nothing happened.” It is hidden.

History records ruptures, not mending.

October 27, 1962, during the Cuban Missile Crisis.

The Soviet submarine B-59 was hemmed in by US depth charges in the Caribbean.

Believing war had begun, the captain ordered a nuclear torpedo launched.

Under the rules, launch required the unanimous agreement of three officers.

The deputy captain, Vasili Arkhipov, voted no.

He forced the “do not launch” button down.

His judgment prevented a false alarm from being escalated directly according to the worst-case interpretation. Where the consequences would have led had he judged differently cannot be proved through counterfactual history.

Yet for the following decades, he was merely an obscure retired naval officer.

His hometown had no statue of him.

History textbooks did not carry his name.

He was even criticized internally for “not showing sufficient fighting spirit.”

Only in 2002, when archives were declassified,

did the world learn—

we had once been a single judgment away from hell.

This is the civilizational version of “survivorship bias.”

We have survived to the present not because conquerors were merciful enough, nor because weapons were accurate enough.

It is because countless Arkhipovs held civilization’s foundations steady at decisive moments.

But our institutions did not record them.

Did not recognize them.

Did not let their descendants know: Your grandfather once saved the world.

What signal does a civilization that records only “destroyers” and “conquerors” send to later generations?

It implies:

If you want to be remembered, destroy and conquer.

If you choose to guard quietly, you will turn to dust.

5.3 Institutionalizing Memory: Humanity Has Always Sought to Keep Important Things from Vanishing Too Soon

Writing, archives, commemoration, oral accounts, genealogies, rituals, and public records are all ways humanity sends the past into the future.

Different civilizations preserve different things.

Some preserve names, some ancestral relationships, some wars and disasters, and some gifts, contracts, works, land, and communal responsibilities.

These traditions follow no single logic.

What they share is only a very old anxiety: If no one remembers, might what happened become as though it never did?

Religious traditions often place memory within intergenerational life through rituals and sacred texts.

Modern institutions of remembrance, such as Holocaust commemoration, war archives, and the “Righteous Among the Nations,” give certain names, testimonies, and actions a public place of preservation.

Many West African societies have long sustained oral-history and musical traditions such as those of the griots.

Family genealogies, ancestral rituals, ancestral halls, gravestones, and oral stories have also carried ordinary people’s names beyond the span of one generation.

Long-term arrangements for charitable property in Islamic societies, such as waqf, remind us that some memories endure not through telling, but through a school, a well, a hospital, or a continuing function of public benefit.

These institutions, of course, also select, omit, and favor.

Those who enter archives are not necessarily the most important people.

Those who leave no writing behind have not therefore failed to live a real life.

Civilizational memory therefore cannot mean “who enters the official list.”

What it must protect is the possibility that diverse sources verify one another and diverse lives leave different ways of reaching them—not that one center decides for everyone what deserves to be remembered forever.

Memory is not a second body after death.

It is more like a bridge, giving those who come later a chance to know who once lived here, took responsibility, created, was hurt, and helped others.

5.4 Hegel’s Life-and-Death Struggle: Recognition as Humanity’s First Need

Why does “being remembered” matter so much?

Is this merely vanity?

No.

The German philosopher Hegel answered no in *The Phenomenology of Spirit*.

For Hegel, the driving force of human history was not the struggle for material goods—but the “struggle for recognition.”

Level of existence	Animals	Human beings
First need	Self-preservation	Self-preservation
Second need	—	Recognition

The difference between human beings and animals is this:

Animals seek only to “stay alive”: food, safety, and reproduction.

Human beings are willing to risk death to “be recognized” as beings with dignity.

This is the famous master–slave dialectic.

In antiquity, this struggle produced divisions of class:

- Those unafraid of death became “masters”: they won recognition.
- Those afraid of death became “slaves”: they exchanged submission for survival.

The slave exists through the master’s recognition.

The master obtains recognition through the slave’s obedience.

The contemporary political philosopher Francis Fukuyama argues in *Identity*:

Although modern democratic institutions have resolved “recognition” in law—equality for all, one person one vote—

recognition remains an acutely scarce resource at the psychological and economic levels.

Group	Material deprivation	Deficit of recognition	Actual need
Unemployed workers	Receive benefits	A sense of being needed	To be seen
Laid-off white-collar workers	Receive severance	A sense of value	To be respected
Young people at the margins	Have free internet	A sense of existence	To be expected and welcomed
Terrorists	—	“Being hated”	To receive attention

Unemployed workers are angry not because they are hungry—they have benefits.

They are angry because they feel “overlooked.”

Terrorists carry out attacks often not for strategic objectives—
but “to be seen.”

Being hated is better than being ignored.

In the age of AI, when most human beings no longer take part in productive labor,
this “hunger for recognition” will reach its peak.

If a person cannot gain recognition through work,
cannot gain recognition through creating wealth,
cannot gain recognition through social status—
where else can they seek proof of their existence?

If society cannot provide a new “system of recognition,”
overlooked groups will force recognition through antisocial behavior.

For being hated is the lowest form of recognition—
but it is recognition nonetheless.

Establishing a system of recognition grounded in “contribution” rather than “conquest”
is therefore more than a moral issue—

it is the highest line of defense within society’s network of protection.

We need “caring for a neighbor” to receive a higher form of social recognition than “owning a luxury car.”
 “Guarding a forest” to receive a more enduring memory than “developing a property.”
 “Accompanying someone who is dying” to receive a more solemn ritual than “winning a lawsuit.”

5.5 Technologies of Memory: From Clay Tablets to Hashes

Humanity’s war against forgetting is also a history of the evolution of recording media.

Era	Medium	Capacity	Longevity	Susceptibility to tampering
Prehistory	Rock art, oral transmission	Very low	Thousands of years	Low
Antiquity	Clay tablets, papyrus, bamboo slips	Low	Hundreds of years	Medium
Middle Ages	Parchment	Medium	Hundreds of years	Medium
Early modern era	Paper	Medium	A hundred years	Medium
Modern era	Digital hard drives	Very high	Decades	High
Present day	Distributed ledgers	Very high	Long-term preservation depends on protocols, participants, and cryptographic migration	Medium to low

In 2008, Satoshi Nakamoto published the Bitcoin white paper.

What it proposed was more than digital cash.

It also demonstrated a way for multiple parties to maintain records together in chronological order.

The important change distributed ledgers bring is this:

Under specified consensus and cryptographic conditions, unilateral rewriting of existing records becomes more difficult and more likely to leave traces.

This does not mean absolute immunity to tampering, or that the contents of a record are inherently true.

It increases the technical and coordination costs of certain kinds of tampering.

Whether those costs are enough to provide security depends on the specific protocol, participant structure, key management, governance, and attackers’ capabilities.

Mathematics can help safeguard the structure of records. Alone, it cannot safeguard truth for civilization.

Imagine a “human contribution ledger” based on a distributed ledger:

- At 3:17 p.m. on March 1, 2050, you fix a security vulnerability in an open-source project in Hanoi. The hash of your commit is written into block 8,402,311.
- On August 12, 2052, you complete a hundred hours of sustained volunteering at a community kitchen in Nairobi. Multiple signatures from the community center confirm this contribution record.
- On November 5, 2055, you recover 342 kilograms of plastic during a beach cleanup in Lima. Sensor data, GPS tracks, and photographs from the site are anchored on-chain.

In this thought scenario, the ledger can prove that “a particular record was entered at a particular time under particular rules,”

but cryptography can establish only the structure of records and the relationships between signatures. Alone, it cannot establish that a real-world action occurred, that a contribution was evaluated correctly, or that meaning has been constituted.

Fifty years later, when your grandson looks up your life—

what he sees is not the bank balance you left behind.

He sees the “contribution hash chain” of your life.

He will see:

Grandfather fixed that vulnerability in 2050. The software has long been obsolete, but the patch was downloaded thirty million times.

Grandfather cared for that older person living alone in 2052. Their granddaughter still volunteers at the community center today.

This can become one technical interface for civilizational memory. But “a record is preserved” does not mean “meaning is defined by the system.”

It does not depend on the power of the day.

It does not depend on whether you are wealthy.

Trustworthy memory still depends on verification of real-world facts, write permissions, sources of evidence, cryptographic security, correction mechanisms, and long-term governance. No “tamper-proof” technology can solve truth and meaning for civilization once and for all.

Yet this vision faces an unprecedented threat.

Quantum computing.

Shor’s algorithm means that, if sufficiently large and reliable fault-tolerant quantum computers emerge, some widely used public-key cryptographic systems will face a fundamental challenge.

What truly warrants concern is not turning a vendor’s roadmap into a countdown for civilization, but the long-term risk of “store now, decrypt later.” Signatures, identities, and historical evidence that must last for decades require advance planning for cryptographic migration.

NIST released its first post-quantum cryptography standards in 2024: FIPS 203 for key encapsulation, and FIPS 204 and FIPS 205 for digital signatures. Publication means migration can begin, not that every system shares the same deadline year.

Systems differ in threat models, update cycles, bandwidth, hardware, and jurisdictions. Their migration paths will differ too.

Long-term memory therefore first requires cryptographic agility: the ability to replace algorithms, rotate keys, re-sign long-lived evidence, and preserve multiple verification paths where necessary.

Technologies such as quantum key distribution and quantum random-number generation can provide additional security in particular settings. But they cannot automatically turn “secure communication” into “true history.”

Quantum memory remains chiefly a research and engineering problem of preserving quantum states. Nor can “no-cloning” establish that “civilizational records cannot be tampered with.”

Whether a record is trustworthy still depends on who may write it, who verifies it, how errors are corrected, and how rights are protected.

What the Amber Ledger must truly safeguard is not a particular future technology.

It is that, after technological transitions, memory can still be verified, corrected, and used to establish accountability.

Technology can extend memory.

Institutions determine whether memory deserves belief.

5.6 The Ethics of Inheritance: What May Remain and What Must Begin Anew

Families can leave houses, businesses, works, stories, and relationships to those who follow.

Different societies also manage the passage of wealth between generations through estate taxes, inheritance rules, trusts, family businesses, and public institutions.

There is no universal law that “allowing inheritance inevitably entrenches class, while banning it inevitably produces short-termism.”

Inherited wealth affects opportunity and power. It may also sustain long-term stewardship, care, and family responsibilities. The consequences depend on institutions, scale, and actual relationships.

The question this book must resolve is not the rate of estate tax.

It is how to prevent a subtler inheritance: directly inscribing predecessors’ contributions, reputations, and responsibilities into the personhood and public eligibility of their descendants.

Several things must therefore be distinguished.

Property and legal relationships can be inherited under actual law.

Works, knowledge, stories, and project memory can be carried forward by those who follow.

Institutions can continue to identify new responsible actors for unfinished maintenance obligations and long-term risks.

But contribution values, public power, general eligibility for resources, professional qualifications, and assessments of personhood cannot transfer automatically through blood ties.

The contributions of ancestors can remain only as historical facts. They cannot become “trust capital” in a descendant’s account.

A family may possess rich experience and relationships from long involvement in an undertaking. A later generation may therefore encounter a field earlier.

But real professional qualifications, public roles, and resource permissions must still be earned anew on the basis of the successor’s own facts.

Predecessors may leave a spark.

Those who follow have the right to decide whether to take it up.

If they do, the new actions and responsibilities belong to them.

This is the boundary between carrying memory forward and hereditary power.

Future thought scenario: Kamosi’s grandmother spent many years restoring urban forests, leaving project records, tree archives, and community relationships. [9]

Kamosi later applies for a new urban-forest project.

Reviewers may see the project history his grandmother left, but cannot award him extra credit for his family.

His own experience, capability, plan, and present responsibilities must be judged independently.

If he is capable, the new record belongs to him.

If he fails, responsibility is not written back onto his grandmother or other relatives.

Memory is thus not severed.

Nor is hereditary privilege reinvented.

5.7 William MacAskill's Gaze: Breaking the Horizon of Death

In *What We Owe the Future*, Oxford philosopher William MacAskill advances the ethics of “longtermism”:

Longtermists such as MacAskill use scenarios with enormous future populations to illustrate the significance of effects across generations. These numbers are possible orders of magnitude in philosophical and scenario-based extrapolation, not demographic forecasts, nor premises on which Civilization Leap depends.

The resulting normative claim of longtermism is that major, irreversible effects of present action on the distant future deserve greater ethical attention. Whether and how to give future interests specific weights still requires independent discussion of rights, distribution, and lives in the present.

Actual time horizons are constrained by elections, financial reporting, the cost of capital, lifespans, family responsibilities, and institutional cycles. The following are only typical windows of incentives, not fixed definitions of every role:

Role	Time horizon	Driving factor
Politician	Four to five years	Electoral cycle
CEO	One year	Financial-reporting cycle
Investor	Quarters to years	Cost of capital
Ordinary person	A lifetime	Lifespan

Individual lifespans do create real decision boundaries. But death does not automatically deprive people of concern for descendants, works, institutions, and communities.

The real problem is that continued personal presence alone cannot guarantee that responsibilities, knowledge, and maintenance extending beyond a lifetime will endure.

How can we bridge this precipice?

Overcoming short-termism cannot rely on “inheritable trust” that makes descendants objects of reward and punishment. Long-term responsibility should instead be anchored in projects, institutions, professional records, and sustainable mechanisms of accountability.

Future thought scenario: Designing a nuclear-waste repository, 2055

The real question is how today's decisions remain traceable, maintainable, and accountable after their designers die. [10]

The answer cannot be to punish descendants not yet born, nor to pledge a family's reputation as collateral for responsibility.

More robust mechanisms include permanent project archives, evidence of key assumptions and design rationales, independent periodic reviews, maintenance funding and insurance or guarantee arrangements, successor institutions' responsibilities, conditions of failure, and mandatory interfaces for reassessment.

If the repository functions well a century later, that outcome can update the historical record of the original engineers, project team, and institutional design, becoming evidence from which later engineering can learn.

If a leak occurs, investigators should trace continuing institutional responsibility, maintenance chains, professional liability, and actors accountable at the time. Descendants are not docked points because of blood ties, nor do they inherit their ancestors' governing responsibilities.

This creates a genuine “chain of responsibility across time and space”:

Human lifespans are finite,

but evidence, institutions, maintenance obligations, and public memory can continue beyond them.

People in the future should have the right to reassess old projects in light of new facts, rather than being forced to obey the rewards and punishments of reputation left by their ancestors. Civilization's closed loop arises from traceable responsibility, maintainable institutions, and repairable errors. It connects the "present" to the "future" through a sustainable structure of responsibility, not the forced binding of family interests. This suggests one possible path: For intergenerational responsibility to move from moral aspiration into institutional design, it must be grounded in archives, review, maintenance, accountability, and reassessment—not depend only on descendants remembering their predecessors' good intentions. What truly needs to pass between generations is usable knowledge, clear responsibility, and a future that remains open.

5.8 Long-Term Memory: History Must Remain Verifiable Through Changes in Cryptography

If large-scale fault-tolerant quantum computing becomes real, long-lived signatures that currently rely on public-key systems such as RSA and elliptic curves will face pressure to migrate.

This does not mean blockchains will suddenly "lose their memory" in one common year.

Still less does it mean humanity need only find a quantum medium to obtain eternal, absolute, tamper-proof civilizational memory.

The real question is simpler and harder: How can a record intended to last decades or centuries remain verifiable by those who follow after algorithms fail, devices become obsolete, institutions change, and responsible individuals depart?

The answer begins with cryptographic agility.

Important records cannot stake all their credibility on one algorithm. They should allow key rotation, re-signing, migration of evidence chains, multiple verification methods, and necessary offline backups.

The next part of the answer is institutional continuity.

Who may write, who must preserve, who may correct, and under what circumstances reverification becomes necessary—these questions lie closer to the heart of "civilizational memory" than the storage medium itself.

Quantum memory is a genuine field of research into preserving quantum states. Quantum key distribution, quantum random-number generation, and related technologies may also enhance security in specific settings.

But "no-cloning" constrains the perfect copying of unknown quantum states; it does not make real-world documents immune to tampering. Nor does "reading may disturb a quantum state" automatically produce historical truth.

Physical laws can add certain security constraints.

They cannot decide for us which histories deserve to be recorded, nor bear responsibility for people who are wrong.

From clay tablets and paper to disks and distributed ledgers, and onward to possible future media, humanity's tools for preserving memory have always changed.

What must never be lost is the ability to verify.

A generation should leave more than "Trust us."

It should also leave enough evidence for those who come later to judge us anew.

This is the deepest ethic of long-term memory:

not to let the past govern the future forever,

but to let the future still see the past, while retaining the right to correct it.

5.9 The Rights of Memory: Who Decides Who Is Remembered?

Memory is not neutral.

Memory is power.

Who is remembered, who is forgotten;

what is commemorated, what is erased—

these are the central political contests of every civilization.

The flaws of old institutions of memory:

Institution	Selection mechanism	Flaw
State memory	Official histories, monuments	Subordinate to political narratives; dissenters erased
Market memory	Celebrity worship, traffic	Subordinate to the logic of capital; contributors overlooked
Academic memory	Textbooks, paper citations	Subordinate to disciplinary paradigms; marginal innovations forgotten
Family memory	Genealogies, oral transmission	Narrow scope; women often omitted

Any institution of long-term civilizational memory must answer:

Who has the right to decide whether a contribution deserves to be remembered permanently?

Who has the right to set “tamper-proof” permissions for contribution-value records?

If memory is power, institutions of memory must distinguish a subject’s consent, public procedures, professional verification, and long-term security responsibilities. These must not be monopolized by any single class of contributors.

A possible answer: Polycentric governance of memory

Level	Content of memory	Decision-making or governance mechanism
Individual memory	Personal contributions and life records	The person’s own rights and consent + independent verification; necessary exceptions handled according to law
Community memory	Long-term facts of public significance to the community	Participation by the affected community + current governance procedures + factual verification; contributors have no automatic right to additional weight
Civilizational memory	Facts of long-term significance to humanity as a whole	Layered support through plural professional and public procedures + appealable review
Core safety records	Safety parameters, versions, and evidence of responsibility	Verifiable version control + independent audit + reauthorization when necessary; no sealing off the future through “physically preset immutability”

There is no perfect institution of memory.

But there is a better awareness of the question:

“Who is forgotten” must no longer be a silent process no one asks about—

it must be publicly discussed, institutionally constrained, and technically recorded.

Forgetting is no longer the default.

Forgetting is an exception subject to review.

5.10 Conclusion: Let What Deserves to Be Carried Forward Remain, While People Can Still Move On

Civilization needs memory.

But civilization need not remember everything.

Permanently preserving everything would turn history from a spark into a prison.

What truly deserves to remain is what those who follow still have reason to know, verify, understand, or continue to bear:

Facts of what happened.

Works and knowledge.

Traces left by important relationships.

Public responsibilities that have not ended.

And stories someone wishes to leave without violating another's rights.

Someone repairs a vulnerability in an open-source project late at night.

Someone helps a stranger to their feet in a downpour.

Someone tends a well for many years.

Such efforts can be recorded, and one can also choose not to make them public.

Records should be verifiable and should also allow correction, withdrawal, sealing, and protection under law.

What we truly oppose is not forgetting itself.

It is important facts disappearing prematurely from the public world because they have no traffic, no power, and no one to speak for them.

Memory can carry a person forward.

But it must not lock a person into the past forever.

What predecessors leave is an entry point, not a command.

Those who follow may still disagree, reinterpret, continue, or choose to say farewell within certain relationships.

A life that truly reached the world does not need civilization to swear it will remember every drop of sweat forever as proof.

It needs only this: When we choose to remember, let us remember as truthfully as we can; when we choose to let go, let us know what we are letting go.

Chapter 6 • Structures of Symbiosis: When Local Rationality Begins to Threaten Shared Survival

The dangers civilization finds hardest to address often have no villain.

Every country is protecting its own security.

Every company is trying not to be eliminated by competitors.

Every person is trying to hold on to work, family, assets, and a place.

Taken separately, every choice can be explained.

The real danger is that, layered together, these locally reasonable choices may jointly produce an outcome no one wants to reach.

Upstream communities save on wastewater treatment; those downstream bear the cost.

To avoid falling behind, a company brings risks it does not yet fully understand to market more quickly, making it still harder for its competitors to slow down.

A country keeps increasing its strength for fear of another country's growing power; seeing this, the other becomes more afraid.

No step need be malicious.

Yet the system may grow ever more tightly wound.

A structure of symbiosis does not address "how to make everyone selfless."

It addresses a more practical question:

When consequences are already interconnected, how can an actor's default path toward its own interests cease to rest continually on the deterioration of the shared foundation?

Putting others first does not abolish self-interest, either.

It means making long-term self-interest increasingly depend on paths that genuinely benefit others, the shared system, and shared survival.

This comes closer to what institutions can undertake than asking everyone to become a saint first.

6.1 Why Local Rationality Can Become a Danger to the Whole

On a congested road, every car seeks the fastest route for itself.

Not one car wants to create a traffic jam.

Yet when everyone simultaneously chooses the move that seems best immediately, the whole road may become slower.

Many of civilization's predicaments have this structure, though their consequences are far more complex than traffic.

Climate, finance, public health, cybersecurity, arms competition, and highly capable AI can all exhibit similar tensions:

The gains remain closer to the actor.

The costs are more easily pushed onto others.

The action takes place today.

Its consequences may become visible only years later.

Success is easily credited to oneself.

Failure readily disperses throughout the system.

Local rationality then acquires a dangerous escape route:

I need only show that this step benefits me.

I need not also show that the world can still bear it if everyone does the same.

This does not mean individual interests, corporate profits, or national interests should all withdraw.

On the contrary, real institutions must recognize them.

A community that refuses to recognize real interests merely hides them beneath slogans.

What must change is the path through which interests are realized.

If pollution, overreach, deception, and outsourcing risks to the future yield profits more readily than creating real shared value, the system will keep rewarding the wrong direction.

If cooperation, honoring commitments, repair, long-term reliability, and reducing shared risks can also become real gains, rationality itself may begin to turn.

Symbiosis does not eliminate competition.

It allows competition to be won without destroying the field on which it is played.

Nor does it give every actor the same goal.

It preserves, amid differences, a shared ground that no local victory may burn away.

Continued existence is the deepest boundary of that ground.

6.2 Fermi's Lunch: The Universe Has Given Us No Guarantee

Summer 1950, Los Alamos, New Mexico.

At lunch, conversation ranged from flying saucers to life in the universe.

Enrico Fermi suddenly asked:

“Where is everybody?”

Where are they all?

The question was later expanded repeatedly into the “Fermi question” or “Fermi paradox.”

The Milky Way is old, and has many stars.

If technological civilizations are not extraordinarily rare, and some can endure, expand, or leave readily detectable traces, why have we still not confirmed undisputed evidence of an extraterrestrial technological civilization? [11] [11]

The force of this question lies precisely not in our having an answer.

It lies in our not having one.

No confirmed extraterrestrial technological signal.

No widely confirmed megastructure.

No confirmed interstellar probe.

Not “there are none.”

Rather, we do not yet know.

The silence of the universe cannot prove that other civilizations have all perished.

Nor can it prove that a so-called “Great Filter” must exist, still less tell us whether the filter lies behind or ahead of us.

But it strips away one kind of careless confidence:

A technological civilization has no guarantee of long-term survival issued by the universe.

We have only one sample.

Ourselves.

If we treat the “Great Filter” only as a stress test, the questions it leaves are still worth asking:

As a civilization acquires ever more powerful technology, which internal structures become more dangerous?

Could competition push everyone along an arms-race path on which no one dares stop first?

Could short-term gains continually leave long-term damage for later?

Could a capability grow faster than authorization, responsibility, and correction?

We do not need the universe to establish these questions for us.

Reality on Earth is enough to give them force.

What Fermi’s lunch truly leaves this chapter is not fear.

It is humility:

Do not mistake “we have survived until today” for “we will naturally go on surviving.”

6.3 1983: After the Machine Sounds the Alarm, Can Anyone Still Pause?

September 26, 1983, at the Serpukhov-15 early-warning center outside Moscow.

The alarm sounded.

The Soviet satellite warning system indicated that the United States had launched intercontinental missiles.

The duty officer, Stanislav Petrov, faced more than an abstract problem in logic. [12] [12]

If the alarm was real, there was very little time.

If it was false, a mistaken escalation could also have enormous consequences.

The screen supplied a signal.

He did not treat the signal directly as fact.

He noticed that the number of reported launches did not fit his understanding of a real large-scale strike, and suspected that the new system itself might be wrong.

The alarm was later shown to have been false.

The event is often told as the story of “one man saving the world.”

But civilization should remember more than a hero.

It should remember this question: Why did a security structure, amid intense confrontation, pressure, and limited information, need one person’s hesitation to absorb the system’s uncertainty?

We are, of course, grateful that he paused.

But future security cannot rest on “there will be another Petrov next time.”

What must be preserved is the institutional capacity behind that second:

An alarm can be questioned.

Machine output is not an order.

There is review between high-consequence actions.

A person can stop.

Another independent channel can take a second look.

After an error, the system is willing to acknowledge why it was wrong.

Nuclear deterrence has been an important security structure in real history, while always carrying risks of misjudgment, accident, and escalation.

It inscribed a very particular shared fate into the twentieth century:
 You cannot easily destroy me, because you too will share the destruction.
 This is coexistence.
 But coexistence founded on fear.
 More mature symbiosis cannot rely only on “if you hurt me, you will die too.”
 It must gradually make another path more real:
 Protecting the shared foundation is itself part of every party’s long-term interest.
 This is precisely what Petrov’s hesitation reminds us of.
 What highly capable systems most need to preserve is more than speed.
 It is also the capacity, at a decisive moment, not to move forward with that speed.

6.4 As High-Consequence Capabilities Spread, the Radius of Responsibility Must Expand

Technology has continually shortened the distance between a person and distant consequences.
 When a letter took months to arrive, an order’s radius of effect was limited.
 Networks, automation, and highly capable AI let some actions replicate and spread faster, and cross organizational and national boundaries sooner.
 The barriers to different technologies still vary enormously.
 Pathogens, critical infrastructure, advanced weapons, financial systems, and large-scale influence over information cannot be placed on a single curve claiming that “the threshold has already fallen to the individual.” [13] [13]
 In the real world, equipment, materials, expertise, access permissions, experimental verification, defenses, and organizational conditions still matter.
 Yet one direction merits attention in advance:
 If organizational barriers to some high-consequence capabilities continue to fall, the range of actors security structures must address will expand.
 Where once we needed only to guard against a capability held by a state, in future we may have to contend with more institutions, teams, toolchains, and cross-border nodes.
 The real change is not that “everyone can destroy the world.”
 It is that a few critical systems may become increasingly unable to withstand a single malicious act, a single overreach, or a single misuse.
 Building is often slower than destroying.
 Establishing trust is slower than destroying trust.
 Training a professional team is slower than triggering an existing vulnerability.
 Civilization must therefore be especially alert to “asymmetry of vulnerability.”
 Not by treating everyone as a potential criminal.
 But by refusing to place everyone’s future on a single-point permission that no one should possess.
 The stronger the capability, the wider the radius of responsibility should be.
 The ability to affect a room requires responsibility toward the people in it.
 The ability to affect a city requires stronger expertise, legal authority, and review.

The ability to affect large numbers of strangers, or even shared survival, makes personal goodwill far from sufficient.

This is why an age of high capability cannot discuss only the speed of innovation.

It must also ask:

Who can obtain the capability?

Who knows it has been used?

Who can stop it?

Who bears the consequences?

Once the system is out of control, is there still a path to recovery?

Mature security does not predict that disaster is certain.

It recognizes that some consequences are too great to wait for their probability to become certainty before designing boundaries.

6.5 Symbiosis Is No Moral Slogan: Letting Long-Term Self-Interest Pass Through Shared Value

If symbiosis depends only on kindness, it will be no more reliable than any idealistic slogan.

Under pressure, people protect themselves.

Businesses protect profits.

States protect security.

None of this disappears because a book declares that “the community matters more.”

A truly sustainable structure must make these actual motivations reconcile their accounts over a longer span.

A business that leaves the costs of pollution to a river and local residents may earn higher short-term profits.

If pollution becomes visible, accountability takes effect, and remediation imposes a real cost, the long-term optimum changes.

A platform that rewards only the content best at capturing attention continually trains its entire ecosystem to compete for people.

If users’ rights, exit, responsibility for content, and long-term trust become real constraints, product rationality changes too.

If a state understands cooperation only as concession, any shared mechanism will struggle to remain stable.

If certain risks already cross borders, reducing shared risk may also be part of that state’s own long-term security.

This is the most practical meaning of “putting others first.”

It does not ask an actor to sacrifice itself first.

It changes the default path:

Make harming others, outsourcing costs, and hollowing out the shared foundation increasingly unlikely to be the cheapest ways of pursuing self-interest.

Make creating genuine shared value, reliable cooperation, sustained honoring of commitments, and reducing shared risks more likely to align with one’s own long-term interests.

Self-interest remains.

Competition remains.

States, businesses, and individuals retain their own boundaries.

What symbiosis truly changes is the structure of their encounters.

Not “From now on, I live only for you.”

But “I need not harm what we depend on together in order to live my own life well.”

If an institution can accomplish this, it need not remind everyone every day to be noble.

The structure itself begins to reward a more sustainable rationality.

6.6 Mitochondria Signed No Covenant: The Biological Analogy Stops Here

The word symbiosis easily leads us toward biology.

Mitochondria in eukaryotic cells are understood to have arisen from deep endosymbiosis between an ancient bacterial lineage and a host. [14]

This evolutionary history is profoundly important.

But mitochondria held no meeting.

Signed no treaty.

Nor ceded a “right to independent existence” to anyone.

Rights, authorization, responsibility, and political legitimacy are questions of human society.

They cannot be lifted directly out of cellular evolution.

The biological analogy can truly remind us of only one limited thing:

Long-term dependence sometimes changes what is possible for each party.

Previously separate systems can form new capabilities through sustained interaction.

But freedom, rights, culture, history, and the possibility of exit stand between social symbiosis and cellular symbiosis.

People are not mitochondria.

States are not organelles.

Still less can shared risk become a reason to abolish the boundaries of subjects.

Symbiosis therefore does not mean “everyone merges into one body.”

On the contrary, it requires clearer boundaries.

Who is a subject.

Who has which rights.

Who authorizes whom.

Who can exit.

Who still cannot be sacrificed in the name of the common interest.

Truly mature symbiosis allows relationships to deepen.

It also allows subjects to remain distinct.

This is the most important difference between a civilization and a cell.

6.7 A Civilization That Cannot Feel Consequences Will Keep Putting Its Hand Back in the Fire

A hand that touches fire pulls back.

Not because the hand understands ethics.

Because the consequences can reach it.

An organization, platform, city, or state that remains unable to feel the real consequences of its decisions can easily enter another condition:

The gains stay here.

The harm happens far away.

The decision is made today.

The bill is settled years later.

Those affected do not even enter the meeting room.

Such a system is like a hand that has lost the ability to feel pain.

It keeps pressing, not necessarily from malice.

The harm simply has not entered its judgment.

A long-lived civilization therefore needs feedback.

Environmental monitoring makes pollution visible sooner.

Incident logs ensure that failure need not begin in amnesia.

Audits provide another pair of eyes outside power.

Appeals let those harmed by a system bring the facts back to it.

Market prices are sometimes feedback too.

Elections, courts, media, science, professional oversight, and community experience all bring reality back into decisions in different ways.

AI and sensors can shorten some of these intervals.

But “seeing sooner” does not mean “machines may directly impose punishment.”

Fact-finding can be automated.

Risk warnings can be automated.

Certain preauthorized safety actions can also be automated.

Decisions that genuinely affect rights, property, and public coercion still require clear legal authority, procedures, and responsibility.

Feedback does not aim to make every decision-maker personally experience a victim’s pain.

It aims to prevent real consequences from being systematically concealed.

What civilization needs is not a cult of pain.

It is for consequences to reach those who need to face them.

6.8 Symbiosis Is No New Invention: Many Societies Long Ago Learned to Put Tomorrow into Today

High-technology society is not the first civilization to face the question “How much do we take today, and how much remains tomorrow?”

In the cosmologies and practices of some Indigenous peoples of the Amazon, the forest is not understood merely as a stockpile waiting to be exploited. [15]

People, animals, rivers, ancestors, and nonhuman beings are often situated within a language of relationships and obligations.

In some Pacific Island societies, traditions such as *rāhui* and *tabu* have restricted harvesting in particular areas, of particular species, or during particular periods.

Pastoral nomadic societies have also long developed local knowledge responsive to changing seasons, water sources, weather, and grasslands.

These societies are entirely different from one another.

There is no single model that can uniformly be called an “Indigenous institution of symbiosis.”

Nor have they been free of conflict, overuse, or environmental pressure.

What is worth bringing into the present is not a romantic vision.

It is a sense of time.

Being able to take something today does not mean we should take it all today.

Possessing power today does not mean possessing an unlimited right to command.

If a rule keeps resources present next year and leaves a community choices later, it has already put part of the future into today.

Modern science, property rights, democracy, the rule of law, and technological governance cannot, of course, be directly replaced by traditional prohibitions.

Yet modern institutions also need to learn something very old:

Do not calculate only the gains already in hand.

Look too at what remains after taking them.

Symbiosis is not a new word precisely because many civilizations have long known:

Staying alive is more than winning today.

6.9 If a Button Could Close Everyone’s Future, the First Question Would Not Be Whom to Choose to Press It

Imagine an extreme scenario: A technology or a power of control could, if misused, cause irreversible consequences on the scale of civilization. [14] [16]

We do not discuss how it could be made.

Nor assume that such a “doomsday button” already exists.

We ask only one thing:

If the consequences really could close everyone’s future, who would have the standing to decide alone?

The first answer should be that no one automatically acquires that standing by being stronger, smarter, kinder, or a greater contributor.

No single person should possess it.

No single algorithm should possess it.

No single institution should possess it.

A group with higher contribution values does not thereby possess it either.

The closer a capability comes to irreversibility, the more its permissions must be divided.

Multiple independent controls.

Clear legal authority.

Human review.

Technical isolation.

Stopping and invalidation conditions.

Every layer must know why it is there.

Multiparty control is not inherently right, either.

If procedures are too slow, a crisis may leave no time.

If responsibility is too dispersed, everyone may say, “I did not decide.”

The point, therefore, is not that more people always means greater safety.

It is that no node, without other independent conditions being met, may close the shared future alone.

The principle has a deeper converse:

Shared survival cannot depend on forcing cooperation through a weapon of destruction hanging over everyone’s head.

“If you do not cooperate, we die together” is not symbiosis.

It is only a more sophisticated fear.

True symbiosis must make sustained damage to the shared system increasingly disadvantageous to the strong, rather than give the weak a final button for mutual annihilation.

It does not turn despair into a tool of governance.

It returns power itself to its boundaries.

6.10 Earthrise: Seeing the Same Earth Does Not Leave Us with Only One Answer

In December 1968, Apollo 8 orbited the Moon.

William Anders photographed Earthrise. [17]

Against the black universe, a blue-and-white Earth rose above the lunar horizon.

No national borders appear in the photograph.

That does not mean national borders ceased to exist.

The absence of the Berlin Wall from view did not make politics, law, and history disappear overnight.

What the photograph truly changed was scale.

Beyond nations, cities, families, and individuals, people saw with unprecedented clarity that all these differences occur on the same fragile planet.

Some astronauts later described a powerful sense of wholeness on seeing Earth from space. Frank White called such experiences the “overview effect.”

The experience does not automatically make someone oppose war.

It does not automatically resolve climate negotiations.

It does not decide questions of migration, borders, trade, or national security.

Nor does being moved generate legitimacy.

But it can open a window in understanding:

Boundaries still exist.

Causation beyond those boundaries exists too.

A change in one forest may enter another region’s climate.

A break in a supply chain can leave a distant hospital without equipment.

Oceans, air, disease, capital, and data do not automatically stop at lines on a map.

This calls for no “global brain.”

Nor does it require states to relinquish sovereignty.

It asks only that decision-makers, while protecting local interests, see one more layer of cross-border consequences already present.

The vision of symbiosis does not flatten differences.

It means that, when you see your own boundary, you also see what lies beyond that boundary's reach, and what will cross it.

Education can help people acquire this sense of scale.

Images, data, maps, simulations, and real cross-cultural experiences can all be ways in.

But no "planetary experience" has the right to forcibly upgrade everyone's values.

Seeing Earth gives us only one more map.

Which path to take still requires real people to decide.

6.11 The Stronger AI Becomes, the More We Must Separate Capability from Direction

Modern society's problem has never been "human brains are too weak, so give civilization to a stronger brain."

The real problem is that complexity exceeds what any one person can process at once.

Climate, finance, public health, logistics, energy, and highly capable technologies all contain large numbers of interacting variables.

AI can help us see more, calculate faster, and compare more options.

It can be an authorized analyst.

It can be a tool.

Within clear boundaries, it can also be an executing agent.

But one distinction must be made from the beginning:

Capability can be delegated.

Direction cannot automatically be handed over because capabilities differ.

A system's knowing more does not mean it owns the public purpose.

More accurate predictions do not confer the right to use resources.

Faster execution does not confer final authorization.

Greater internal complexity in a model does not mean we have no option but to trust it.

Whether the future brings classical computing, quantum computing, or an architecture not yet invented, technological terminology does not automatically rewrite this relationship.

Whether future AI develops new forms of subjecthood remains open to reconsideration in light of new facts and ethical evidence.

Today's institutions need not first answer that distant question to protect present boundaries.

They need only know:

The system's identity must be transparent.

Authorization can be withdrawn.

Critical actions can be stopped.

Operations leave evidence.

When consequences occur, responsibility still returns to real-world actors and institutions.

This is not "humans are forever a rank above AI."

It is a refusal of a sleight of hand:
 jumping straight from “it is more capable than I am” to “therefore it is more entitled to decide.”
 Between capability and entitlement always stand rights, authorization, responsibility, and consequences.

6.12 When Everything Is Optimized, How Much Opportunity to Choose Again Remains?

Imagine a future community.

There is almost no waiting in traffic.

The power grid remains on its optimal curve.

Waste is recycled automatically.

Resources are matched precisely.

The system even knows roughly when each person will need what. [15] [18]

Life runs smoothly.

So smoothly that, one day, someone suddenly wants to ask:

What if I do not want the optimal answer?

What if I want to take the longer way?

What if I want to do something that, for now, has no efficiency, no return, and perhaps no one who understands it?

Civilization should, of course, reduce pointless waste, accidents, and scarcity.

Optimization is not the problem.

The problem is whether a particular goal of optimization has begun treating every other possibility as noise.

If a system rewards only a few computable goals over long periods, the space for dissent, experimentation, chance, and redefining the question may grow ever smaller.

In the end, no tyrant is needed.

Everyone simply keeps choosing the next step the system considers most reasonable.

This is another form of local rationality.

Every step is excellent.

Only the number of paths keeps shrinking.

Symbiosis therefore also needs openness.

AI can undertake the calculation, prediction, repetitive execution, and high-precision coordination better suited to it.

People can continue to work, create, care, research, make mistakes, rest, and refuse to let the system define all their goals for them.

This does not predetermine eternal “carbon-based and silicon-based ecological niches.”

Future capabilities will continue to change.

What truly needs protection today is this:

No real subject loses the standing to pose questions anew merely because, for the moment, they do not fit the current optimum.

An apparently useless experiment may open a new technology.

Companionship that produces no KPI may matter immensely to two people.

A minority view may prove crucial only many years later.
 Much “idle imagining” also ends up producing nothing.
 Openness does not guarantee that every branch of the path has value.
 It simply prevents the system from closing every branch before value has time to arise.

6.13 Conclusion: Symbiosis Does Not Mean Becoming One Person, but Refusing to Burn the Shared Ground

At this point, symbiosis can finally descend from an attractive word into a set of ordinary boundaries.
 First, differences remain.
 States are not the same state.
 Businesses are not the same business.
 Individuals need not share the same meaning.
 Symbiosis does not require everyone to become one subject.
 Second, interests remain.
 Self-interest need not be eliminated.
 What must change is the default path through which self-interest realizes itself.
 If sustained damage to the shared foundation becomes harder to profit from, and creating shared value becomes more compatible with one’s own long-term interests, the structure begins to change.
 This is the place of “putting others first” in Volume One.
 Not moral sanctification.
 A direction for civilization’s incentives.
 Third, continued existence is a baseline that must not be lightly lost.
 No one’s local success is entitled automatically to close everyone’s future.
 The more powerful the high-consequence capability, the clearer the boundaries of authorization, responsibility, review, and stopping must be.
 Fourth, cooperation does not generate sovereignty.
 The greater the shared risk, the more important coordination becomes.
 Yet “we are all affected” does not automatically imply a supreme global power.
 Protocols can expand cooperation.
 Facts can be shared.
 New legitimate mechanisms can be established for cross-border problems.
 Specific powers must still have clear sources, scope, procedures, and responsibility.
 Fifth, AI extends capability; it does not automatically rewrite subjecthood and power.
 Today, human beings remain the subjects of actual rights, responsibility, and meaning.
 AI can analyze, advise, generate, match, and execute within authorization.
 Its capabilities will continue to change.
 Our boundaries can also be reconsidered in light of new facts.
 But capability itself never automatically becomes entitlement.
 Finally, there is a sixth boundary.
 The future must still have branching paths.

Civilization must neither die from losing control nor live on as a system that can never choose again.

This is where Volume One ultimately arrives.

Human value is not decided by a ranking of functions.

Indicators cannot masquerade as purposes.

Competition is not the only rationality.

Genuine contributions can be seen, but cannot purchase basic rights.

Memory can be carried forward, but cannot make power hereditary.

Shared survival requires cooperation; cooperation cannot grow into sovereignty without a legitimate source.

AI can extend capability. Greater strength does not entitle it to take purpose and responsibility along with it.

Volume One still does not set out institutional solutions.

It has only brought into clearer view things previously mixed together.

The next step will be harder.

How do we distinguish fact from evaluation?

Contribution from rights?

Capability from power?

Authorization from execution?

How does responsibility find its way through complex systems back to people?

If these words remain vague, any new institution may simply give old problems a new interface.

Volume Two therefore does not hurry to build institutions.

First, it takes the language apart.

It explains, one by one, relationships that must not be conflated.

Only then does it ask:

How can they be recombined without all spiraling out of control in an age of high capability?

Please turn to Volume Two.

Material Notes for This Volume

On the Historical Scenes, Cases, Future Extrapolations, and Literary Metaphors in This Book

The following notes collect the material referenced by light superscript markers in the main text. They identify the material, limits of sources, status of thought experiments, and literary metaphors; they do not carry the main argument. Unless delaying an explanation would materially mislead a general reader about facts, subjects, rights, institutional force, or present versus future status, qualifications of this kind are generally no longer inserted into scenes, strong propositions, sustained parallel passages, or closing resonances.

1 | Volume One, Chapter 1 | The status of “An Endgame in the Igloo”

The opening is a literary composite constructed from polar scarcity conditions and related ethnographic and travel accounts. It corresponds to no single verifiable person or event and does not generalize about Inuit people or any Arctic community. Its intellectual function is to reveal the structural question of how extreme scarcity may bring “usefulness” close to “whether one may remain.”

2 | Volume One, Chapter 1 | The status of ubasute material

Stories of abandoning older people, such as Japan’s ubasute, exist mainly in folklore, literature, and cultural memory. They cannot directly serve as empirical evidence of a widespread historical institution. The book uses them to present a long-standing cultural fear connecting uselessness with eligibility to remain, not to prove that a society generally practiced abandonment of the elderly.

3 | Volume One, Chapter 1 | Maine’s “From Status to Contract”

Henry Maine’s formulation arose in a particular age and from particular materials. It should not be treated as one path shared by all civilizations. This book borrows only the tension that status can sustain a place while also creating lock-in.

4 | Volume One, Chapter 1 | The statistical scope of the US Great Resignation

“Nearly 48 million” is the annual total of quit events in US labor statistics, not the same number of distinct individuals. Causes include changing jobs, wages, care pressures, working conditions, and other factors. The main text uses it only as a social signal that the occupational covenant is being reassessed.

5 | Volume One, Chapter 1 | Finland’s basic-income experiment

The experiment provided unconditional basic income to randomly selected unemployed people, principally observing employment and well-being outcomes. It was not designed to test the “sense of being needed.” Structurally asking why income security cannot automatically resolve social position does not amount to judging the experiment itself a failure.

6 | Volume One, Chapter 2 | East Asian care and mutual-aid practices

Children’s cafeterias, community meals, and time banks are not the same institution. Their origins, organization, scale, and policy backgrounds vary by region. They are parallel references for how care and relationships gain public visibility.

7 | Volume One, Chapter 2 | Boundaries in the use of Ubuntu

Ubuntu has multiple uses across languages, histories, and intellectual traditions. It cannot be summarized as one “African institution,” nor used to erase real tensions between individuals and communities. The main text takes up only the questions opened by its relational view of personhood.

8 | Volume One, Chapter 3 | Future computing architectures and governance boundaries

Quantum computing is a real research and engineering direction, but “quantum–AGI” is not an already-established unified technical object. The main text does not infer from superposition, entanglement, measurement, or particular demonstrations of quantum advantage that future AI must become unauditably or accelerate to a civilizational threshold. Nor does it assume a dedicated governance system. It retains only boundaries of stopping, takeover, audit, authorization, and responsibility that hold across technological forms.

9 | Volume One, Chapter 5 | The future memory scene in Nairobi in 2050

“Nairobi in 2050” is a future thought scenario, not a present institutional fact. It tests how memory can be carried forward while rights, qualifications, and governing authority cannot become automatically hereditary through ancestral records.

10 | Volume One, Chapter 5 | The nuclear-waste repository scene in 2055

This is a future thought scenario testing the relationships among intergenerational responsibility, long-term risk, and successors’ reassessment. It corresponds to no specific current project.

11 | Volume One, Chapter 6 | Epistemic boundaries of the Fermi question and the Great Filter

The Fermi question rests on the tension between models in which technological civilizations seem likely to leave detectable traces and the absence of confirmed, undisputed evidence of extraterrestrial technological civilizations. The Great Filter is a class of explanatory frameworks, not an established fact. Cosmic silence cannot prove widespread self-destruction among other technological civilizations or rank civilizations by maturity. The main text uses this open question only to remind us that long-term survival has no cosmic guarantee.

12 | Volume One, Chapter 6 | Petrov and a false nuclear early warning

On September 26, 1983, the Soviet satellite early-warning system produced a false alarm. Duty officer Stanislav Petrov did not treat its warning directly as a confirmed attack. The event is often simplified to “one man saved the world,” but the actual nuclear command chain, subsequent verification, and historical causation are more complex. The main text uses it only to show that high-consequence warnings need questioning, independent review, and room to stop. Machine signals cannot automatically equal orders to act.

13 | Volume One, Chapter 6 | Stress-testing the diffusion of high-consequence capabilities

Real barriers to different high-consequence technologies vary greatly, constrained by equipment, materials, expertise, access, experimental validation, defenses, and organization. This section offers only a directional stress test: If barriers to some capabilities keep falling, safety and responsibility structures must cover more subjects and nodes. It does not declare that “destructive capability has already become individual” or provide a uniform timetable.

14 | Volume One, Chapter 6 | Limits of the mitochondria and endosymbiosis analogy

The endosymbiotic origin of mitochondria is an important framework in modern evolutionary biology: An ancient bacterial lineage entered a host and became a eukaryotic organelle through long evolution. The main text uses this only to show that sustained dependence may change system capabilities. It derives no human rights, state sovereignty, political authorization, or obligation for society to “merge into one whole” from cellular evolution.

15 | Volume One, Chapter 6 | Indigenous peoples and local resource governance

Amazonian Indigenous peoples, Pacific Island societies, and pastoral nomadic societies are internally highly diverse. Resource relationships, prohibitions, seasonal governance, and ecological knowledge cannot be summarized as one “Indigenous institution of symbiosis.” The main text makes only a limited comparison:

Many local traditions have brought future resource conditions into present decisions through relationships, seasons, harvesting restrictions, or intergenerational experience. It does not romanticize any group or replace modern legal rights and authority.

16 | Volume One, Chapter 6 | An extreme scenario of civilization-scale irreversible power

This is an extreme governance thought experiment. It does not discuss how destruction could be implemented. It tests only an institutional question: If misuse of a power could close everyone's future, who may control it, and how can single-point overreach be prevented?

17 | Volume One, Chapter 6 | Earthrise and the overview effect

Apollo 8 astronaut William Anders took the famous Earthrise photograph in December 1968. The “overview effect” is a later concept describing the sense of wholeness and cognitive shifts some astronauts reported after viewing Earth from space. Such experiences can change a sense of scale. They do not prove that seeing Earth automatically produces a uniform political stance, moral conclusion, or global sovereignty.

18 | Volume One, Chapter 6 | The future scene of overoptimization

This is an extreme future thought scenario, not a prediction. It tests a structural risk: If a system persistently minimizes friction, waste, and uncertainty around a few goals, might subjects' space for dissent, exit, experimentation, and redefining questions shrink with them?