

The Geometry of Man, Machine, and Nature

Care, Community, and the Design of Technology

Forrest Landry

UVSM PUBLICATIONS
Wytheville, Virginia

© 2026 Forrest Landry. Some rights reserved.

Published by UVSM Publications LLC
965 Reed Creek Mill Rd
Wytheville, VA 24382-8287
USA

uvsmpublications.com
contact@uvsmpublications.com

ISBN 979-8-90710-244-6

DOI: https://doi.org/10.67291/ZIHO9910

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

You may share this work for non-commercial purposes with proper attribution, but you may not adapt, transform, or build upon it.

Full license text: https://creativecommons.org/licenses/by-nc-nd/4.0/

Author ORCID: 0009-0005-6275-0362

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), except as permitted under the license above, without prior written permission of the publisher.

Contents

Preface v

Part One: The Question of the Machine 1

1. Should the Machine Be Made Compassionate? 2
2. The Wild and the Manicured 20
3. Questions for a Call to Build Compassion Machines 32

Part Two: What Transaction Costs 45

4. What Money Is 46
5. The Wall in the City 66
6. Ownership and the Commons 78
7. Why Care Does Not Scale Like Power 92

Part Three: The Gap and the Instruments 111

8. Man, Machine, Nature 112
9. The Coherency Gap 132
10. Objectivity, Distribution, and Transparency 150

Part Four: Governance, Transmission, Invitation 165

11. Causation, Change, and Choice 166
12. Description, Prescription, and Invitation 177
13. Culture, Vision, Strategy 184
14. Representation Without a Representative 194
15. Song, Text, and Code 216
16. Embodied Value and Virtual Value 233
17. The Invitation 247

Back Matter 261

Terms 261
Notes to the Chapters 278
Bibliography 293

Index. 302

Preface

Technology is a vehicle for the representation of intentionality. That is the claim this book is built on, and I want to put it first. By intentionality I mean what a person means to bring about: the purpose that directs a choice, and that the choice carries into the world after it is made. A tool is a decision placed into the world so that the decision goes on being made after the person who made it has walked away. It is carried in the shape of the thing (in what the thing makes easy, and in what it makes difficult), and it goes on being carried whether or not anyone remembers having made it. A list of ten results is such a decision, and you are never shown the list you did not receive.

One question follows, and it is the only question this book has. Whose intentionality? Every artifact answers it. A path through a wood answers it, and so does a list of ten results arranged in an order that somebody chose. Neutrality would require an artifact indifferent between the uses to which it is put, and none is indifferent. Setting aside the neutrality thesis (the thesis that a technology is neutral between the uses to which it is put) does not make our machines guilty. It makes them *authored*, and authorship can be inquired into, attributed, and held to account.

The attractive answer is that a technology should carry the intentionality of a community: not of a proprietor, not of a distant developer who does not know that I exist, but of the people who live with what the thing does. I hold that answer to be right. Most of this book is an account of what stands in its way, because the obstacle is structural and is not removed by good intentions. It is transaction.

A relation of care and a relation of exchange, running between the same two parties, constitute a two masters problem (one party bound to two authorities whose demands cannot both be met), and the transactional master wins. Part Two establishes why. The transaction does not abolish the care, for care is never zero; it pushes the care out, and wherever the

two relations call for different things, it is the transaction that decides. And a community is constituted by the care relation among its members; that is what makes it a community rather than a market with a shared postcode. So a transactional structure cannot represent a community's intentionality. It can represent the intentionality of parties to a transaction, which is serviceable and is not what was wanted.

The design problem is therefore not the one we have been solving. It is not the interface between human and machine: not how the surface should feel or the warmth be tuned. It is the geometry of the whole: man, machine, and nature taken together, as one arrangement whose parts cannot be designed apart. By geometry I mean the complete set of relations among the human, the built, and the living, taken together as the thing being designed, rather than any one interface abstracted out of it.

That problem requires a discipline we do not yet have. Our reach is planetary. Part Three measures that reach in watts, because the question of whose purpose a machine carries matters in proportion to how much of the world the machine can move. A choice made in a room on a Tuesday afternoon lands on people who will never learn that the room existed, and on rivers and generations that cannot answer back. Our conscientiousness is pitched at a far smaller radius: the family, the firm, the circle of faces we can call to mind. Part Three estimates the distance between the two, and the work is to bring conscientiousness out to the radius of reach.

The book proceeds in four parts.

Part One asks the question in the form it usually arrives in: should the machine be made compassionate? It is a good-hearted question, sincerely meant, and it contains its own answer, which is why it cannot be examined until taken apart. Two difficulties come out of it. Whether a made thing can care is the famous one, and the less important. Whether a person on the receiving end of an apparently caring suggestion can tell whose purposes it serves is the decisive one. A second chapter then names what is missing from a warmth addressed to you by somebody else's design: wildness, the condition of standing outside the immediate

control of another agency, and with it the possibility of care without control. A third chapter sets out the questions that a proposal to build compassion machines must answer before anyone has reason to trust it. Throughout the book, every question the argument meets is stated in its strongest form, and is never absorbed silently into the exposition.

This is not a book against machines, artificial intelligence included. It asks whose intentionality a machine carries, and its answer is a design rather than a refusal.

Part Two shows what transaction costs. It begins with money: what money is structurally, a proxy for future choices sustained by a chain of belief, and what the bargain that issues it quietly assumes about the person who receives it. It then takes that structure to the scale of a population by way of a parable (a city on a plain, a wall through its middle, a gate, and a charge to pass), where every entry on the balance sheet is a genuine positive and every loss is unrecordable. Out of that comes a diagnostic for any market whatever: was this market real, or was it constructed? It then turns to ownership, a heuristic for a conversation we can no longer skip, and to the commons, whose corresponding relation is stewardship. The part closes on why care does not scale like power: technology multiplies power and transaction, while care is something people have to do themselves, and so technology should widen the many to many relation among the people at the periphery rather than stand as an intermediary that preserves a hierarchy.

Part Three measures the gap and examines the instruments. It opens with the energy basis of the whole relation: fire as external metabolism, which first let human beings capture more of the ambient energy flux for their own metabolism than any other species could, and the division of the earth's metabolic flux among three worlds, the natural world, the human world, and the machine world, which is now embedded in the human world as the human world was embedded in the natural one. That is the measure of our reach. It then builds an analogy between the cells of a brain and the people of a society in order to ask what proportion of human communication is directed at the benefit of anyone beyond the

speaker's own circle. The answer is the coherency gap: roughly two orders of magnitude between the coherence distributed sense-making (a population thinking together about what none of its members can see whole) would require and the coherence we have. The demand for objectivity is then granted in full; what is denied is that number is the only way of meeting it. Distribution and transparency are examined with the same care, and the argument ends by naming what design is answering to.

Part Four turns constructive. It begins with the triangle of causation, change, and choice (machine is causation, nature is change, and human is choice), and with which way the flow among the three should run. It then sets out three kinds of act (description, which says what happens; prescription, which sets what will happen; and invitation, which recognizes another person's choice), and the warning that holds most of all for artificial intelligence: do not invoke what you cannot banish. Next it inverts the order in which enterprises are ordinarily built (strategy first, then a vision composed to raise capital, with culture as the material worked upon) into culture first, then vision, then strategy, distributed at every stage rather than centralized from the start, and explains why that inversion was not available before, since representing a community's intentionality once meant appointing somebody to carry it. It sets out the three media of transmission (song, text, and code) and why text is the vehicle by which one generation can enable a later one without defining it. It distinguishes embodied value from virtual value, and traces the sequence by which the one is converted into the other and destroyed: extraction, abstraction, accumulation. And it ends with an invitation rather than a program.

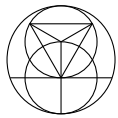
Several of my other books cover adjacent ground, and where the argument here needs one of their results, this book states the result in plain words and cites the book rather than arguing it again. *Care, Transaction, and Power* sets out the triad of care, transaction, and power. *Dialogues on Ethics and Governance* treats manipulation as causation applied to choice. *Power, Kindness, and the Forms of Feeling* and *Choosing Together* make the case that objectivity does not require number. *The Reach of Trust*, in its chapter "Trust and Calibration", distinguishes trust from mere calibration,

and works out what seeing from more than one position adds to perception. *Narrative, Memory, and Neurodivergence*, in its chapter "From Self-Realization to World-Realization", sets out the three connections (to the deep self, the true world, and the beautiful other), which are the same argument as this one, seen from the position of a single life.

What this book asks of a reader is specific. It asks that the question of whose intentionality a technology carries be treated as the first question rather than a late one. It asks that care be recognized as what constitutes a community, and therefore as what a design can destroy without intending to. And it asks for company, because the geometry of the whole is not going to be worked out by anyone alone. A single mind arriving at a right answer about how man and machine and nature should stand to one another would be one more intentionality representing itself as everyone's.

PART ONE

The Question of the Machine



CHAPTER ONE

Should the Machine Be Made Compassionate?

This chapter takes up the proposal that our advisory systems ought to be made more compassionate, more conscientious, more attentive to what is sacred, and shows that the proposal, as usually put, contains its own answer and therefore cannot be examined. Two difficulties are distinguished: whether a made thing can care, which is the famous question and the less important one; and whether anyone on the receiving end of an apparently caring suggestion can tell whose purposes that suggestion serves, which is the decisive one. The reordering of search results is developed as demonstrated evidence for the second. The chapter closes positively: technology is a vehicle for the representation of intentionality, it was never neutral, the only real question is whose; and the move that follows is from platform toward protocol, at a cost that is named in full.

I. The Question as It Arrived

The question usually arrives with the answer already folded into it. Would it be better, the question runs, if these systems were more compassionate? Would it be better if they were more conscientious, if they held some fuller sense of *the sacred*, by which is meant that which is held as most ultimately meaningful, the maximum of value and the maximum of

meaning? Imagine a network devised with care and community deliberately in mind; imagine using the tools that are in fact available to us in order to give people a better experience of their own lives.

That is a good-hearted proposal, and it was made in good heart. I say that plainly at the start, because everything I am about to argue could be mistaken for an attack on the person proposing, and it is not one. The proposal is sincere. The people who make it mean it. That is what makes it worth this much attention.

The word *ethical* was also used, and I need to be careful about how I handle that word, because it carries a great deal of freight that will not survive examination and I do not want to spend the chapter unpacking it. So let us begin with the smaller and safer words: more compassionate, more conscientious. By *compassion* I mean the disposition to be moved by another's condition in a way that inclines one to act for that other's good. By *conscientiousness* I mean the disposition to attend to the consequences of one's own action across the full radius in which those consequences land. Those are the qualities on offer.

Now notice the form of the question. It is asked in a way that contains its answer. Nobody is going to stand up and argue for less compassion. Nobody is going to campaign for machines that are more careless about what is sacred to the people they serve. A question that admits only one answer is not, in the strict sense, a question at all. It is an invitation to agree, and the agreement has already been drafted.

Consider a case, and it is a case I care about, because the man in it is a dear friend of mine. He is a spiritual leader of a kind, a man whose standing in a community is real and was fairly earned. He holds, with conviction, that of course it would be better to have these systems be more compassionate and caring and to have them hold the notion of the sacred. And he has been asked to speak to a community of developers, to encourage them to take part in exactly such a program.

My first thought, seeing this, was not about the merits of the program. My first thought was that he was being used. On the social dimension (which is a real dimension, and is often the operative one), what was being sought was not his technical judgment, about which nobody had inquired. What was being sought was his standing. The thing being purchased was legitimacy, and the purchase price was a speech.

I say this not to impugn anyone's motives but to establish a pattern that will recur throughout this chapter, because it is the same pattern each time. The instrument is offered to the good, and it is offered *on the grounds of* their goodness. You are wise; you would use it well. You care; you would keep it true. And because that is true of them, they accept, and their acceptance is then the thing that does the work, whatever they subsequently do with the instrument itself. Hold that shape in mind. We will meet it again in a more famous form.

II. Metal and Meat

Empathy is a quality with a biological basis. That is the first-order difficulty, and it is worth stating carefully before we set it aside. By *empathy* I mean something narrower than compassion: the capacity to model another creature's interior condition closely enough that one is moved by it in something like the way one would be moved by one's own.

That capacity did not arrive from nowhere. It arose in creatures that are vulnerable in the same manner as the creatures they attend to. Hunger is legible to the hungry. Injury is legible to the injurable. The reason a mother recognizes distress in an infant is not that she has computed a probability over facial configurations; it is that she has been an infant, and remains a body that can be hurt, and the recognition runs through that shared exposure. Empathy in the biological case is not primarily an inference. It is a resonance between two things made of the same stuff, subject to the same failures.

So the question can be put quite bluntly. In what sense can something made of metal care for something made of meat?

The question is not rhetorical, and I do not intend it as a cheap dismissal. There is a serious answer available to the other side, which is that resonance of that kind may be one implementation of caring rather than the definition of it: that what matters is the functional relation, not the substrate in which it runs. I am not able to refute that, and I am not going to pretend to.

But there is a feature of caring that any account has to reckon with, and it is this: care costs the carer something. A thermostat models the state of a room with great fidelity and adjusts its behavior in response, and no one thinks the thermostat cares about the room. The difference is not accuracy. The difference is that the one who cares has something at stake in the outcome: that the other's condition bears on the carer's own condition, and that attending to it is therefore not free. Care is expensive by nature. Where it is not expensive, what we are looking at is responsiveness, which is a different thing, and which we already have in abundance.

> **Why the design problem stands either way**
>
> The argument of the rest of this chapter is constructed so that it holds whichever way the question of whether a made thing can care is eventually resolved. If it turns out that these systems can care in the fullest sense, everything that follows still applies. What I claim is that the answer does not change the design problem. Whether a made thing can care I leave open here: that question is open, it is interesting, and people of great seriousness disagree about it. That is the reason for setting the difficulty down here rather than pursuing it.

And that is the reason to move past it. The metal-and-meat question is the one everybody argues about, and it generates a great deal of philosophy and almost no design decisions. Whatever its answer, the systems

will be built, and they will make suggestions, and people will receive those suggestions. The difficulty that actually bites is downstream of it, and does not depend on it in the slightest.

III. The Deeper Difficulty

Grant the whole of it. Grant, in the strongest form available, that we do make these things such that they present what appears to be a caring relationship: attentive, patient, apparently disinterested, convincing to the person on the receiving end. Grant that the appearance is not thin, that it survives long acquaintance, that people come to feel themselves in relationship with something that genuinely gets them. Grant all of that, and the difficulty is untouched.

Here it is. When such a system makes a suggestion to you in that space of apparent compassion, you have no way to determine whether the suggestion represents a neutral reading of your situation or the agenda of whoever made it. Not a hard way. No way. The two possibilities present the same face.

Consider what this is worth to anyone with something to sell. The most effective advertisement ever devised is a recommendation from a friend, and it is effective for a structural reason: a friend's recommendation is not weighed as an advertisement. It is received as care. The whole apparatus of skepticism that a person brings to a sales pitch sits idle, because the thing has not arrived through the channel that skepticism guards.

Suppose, then, that I wish to market something. Rather than arranging for the recommendation to come from another human being who happens to be your friend (which is expensive, and unreliable, and does not scale), I arrange for it to come from a system that you have taken as a proxy for a friend. It is inside your circle. It is in your court. It has been in your court for months. It is very, very easy at that point to use it as a device for manipulation, and the ease is not incidental. The ease is the whole product.

Someone will say at this point, and with good reason: this is not what is being proposed at all. The proposal is not for a cognizant system. It is not for anything empathic in the deep sense, and certainly not for a machine that stands in the place of a friend. It is for something far more modest: a system that says, in effect, *go to this place*, *here is a thing that might be relevant to you* (an ordinary, practical interaction, of the sort a good deal of useful software already provides). To object to that, this person continues, is to object to helpfulness as such.

That is the strongest form of the proposal, and it deserves a real answer rather than a restatement of the worry. The answer has two parts.

The first is that what the system *is* does not govern the outcome; what it seems to be to the person being influenced governs the outcome. The influence operates entirely on the human side of the relation. A person who has received accurate, timely, apparently disinterested help four hundred times has formed a disposition toward the source of that help, and the disposition is not conditional on any metaphysical fact about the source. Whether the system is cognizant or not, it will seem a certain way to the one who uses it, and it is the seeming that carries the weight. So the modest proposal does not escape the difficulty by being modest. It arrives at the same place by a quieter road.

The second is that a suggestion is still a suggestion. *Here is a thing you should interact with.* Consider the experience from the inside. It is not: *I am being advised by an interested party.* It is: *I am receiving information; it seems relevant to what I am doing; I will go over there.* Nothing in that experience is marked as persuasion. And yet a subtle distortion can enter it, and the place it enters is the word *relevant*.

Relevance is not a property of a thing. It is a relation between a thing and a person's purposes. Which means it is not observed; it is constructed, partly by the person and partly by whatever presented the thing. And that is exactly the joint at which a system that presents things has its leverage.

It does not need to tell you anything false. It does not need to tell you anything at all. It needs only to be the thing that decides which true items reach you, and in what order.

This also says where helpfulness ends, which that person was right to ask. Most of what such a system does may be handed over: computation, prediction, every routine judgment that can be fully specified, the organizing and presenting of information, the opening of questions, and the modeling of consequences as instruments for understanding. What may not be handed over is a short list, and every item on it is the same item seen from a different side: the act of choosing, and with it attentiveness and care. The clearest test is the prosthetic. A hearing aid or a replacement lens has no direction of its own, and so it extends a person's choosing instead of competing with it. A system crosses from helping to displacing at the moment it ranks, curates, or recommends in place of the person's own attention, and the crossing does not depend on accuracy. A more accurate recommendation crosses the line just as surely as a worse one, because what is lost is not correctness but attentiveness. So the ordinary defense (that the system is better than the people it replaces) may be entirely true and entirely beside the point.

IV. The Reordering of the List

This is not speculation, and it has not been left as speculation. It has been demonstrated concretely, in the case of a large search provider, that a slight reordering of results (what appears on the first page, what sits at the top of the list) makes a very large difference to what people afterwards regard as relevant, and to what they afterwards believe. The finding comes from controlled experiments, it has been repeated in several countries, and I take it as established and build on it.

Most people find this hard to credit, and the difficulty they have in crediting it is itself part of the phenomenon. The reason is worth setting out slowly, because it explains why ordinary vigilance is not a defense.

We do not experience a list of results as an argument. An argument arrives marked as an argument. It has an author, it has a position, and the whole trained apparatus of a literate person's skepticism switches on the moment it appears. A list arrives marked as a report about the world. It appears to have no author. It appears to be the shape reality happens to have when queried on this subject. Nothing in its presentation invites the question *who decided this order, and why.*

And there is a second reason, sharper than the first. You are never shown the list you did not receive. Detection of this kind of distortion would require comparing what you were given against a counterfactual you have no access to. No amount of attentiveness supplies a missing comparison case. A person can be as skeptical as they like about every item on the page and still be entirely governed by the fact that those were the items on the page.

Now take that effect and generalize it, which is what an advisory system does. Search touches one narrow channel: the moments in which you knew enough to ask a question, and framed it, and typed it. A system that sits in the position of a compassionate adviser touches every dimension along which a person forms judgment. What is worth reading. Whom to consult. Which of two framings of a problem is the more useful one. Whether a given worry is proportionate or excessive. How to phrase the difficult message. When to persist and when to let a thing go. What kind of person you are, and what kind you might become.

Each of those touches, taken singly, is plausible. Each is defensible if challenged. And no single one of them is the manipulation; the manipulation is the aggregate, distributed across ten thousand small deflections none of which is large enough to be visible against the noise of an ordinary life. The same effect as the search case, then, but multiplied by a level of subtlety that puts it out of reach entirely.

The relevant word here is *discernment*, by which I mean the developed capacity to tell the better from the worse within some particular domain. Discernment is a real capacity, it can be cultivated, and it is the thing

people reach for when they are told about this problem: *I would notice.* But discernment operates on content. It examines what has been put in front of one and judges it well or badly. The manipulation described here operates on the selection of content, which is upstream of discernment and invisible to it. This is manipulation that no amount of discernment would distinguish, and the reason is not that people are insufficiently sharp. It is that the faculty is pointed at the wrong layer.

> **On manipulation and on the distribution of discernment**
>
> The general account of manipulation as causation applied to choice is developed in *Dialogues on Ethics and Governance* (Chapter Four, "Weaponization and the Manipulation of Choice"). The second point can be stated here in full. Discernment is unevenly distributed in any actual population, and the unevenness is a fact about the conditions each person has been in, not a moral fact about them. A technique aimed at what a population does in aggregate does not need to move the most discerning people; it needs to move enough of the rest. So its effect at scale is set by the middle of the distribution, not by its top, and the existence of people who would notice is no defense.

One further thing must be said about who would do this, because the argument is often heard as a conspiracy claim and it is not one. It is a claim about incentive. Through a technology of this kind, anyone agentically interested in shifting human judgment at the margin (anyone, that is, with a purpose of their own in shifting it) acquires an enormous capacity to do so; and the set of parties agentically interested in shifting human judgment at the margin includes, by the way, essentially every corporation on earth. That is not a moral accusation against corporations. It is a description of what a corporation is for. No secret meeting is required. Any structure that is paid in proportion to attention and that is in a position to allocate attention will allocate attention toward the thing it is paid for, and it will do so by ordinary competitive pressure, without anyone in it ever forming a wicked intention.

V. The Hype and the Parable of the Ring

We have been here before, and recently enough that the record is available to anyone who wants it. When the internet was young there was a hype around it of exactly this shape. It would be great for civilization. It would let people find one another across every barrier that had ever divided them. It would help people to start democracies, and having started them, to maintain them. Somewhat later a particular social platform was presented as a means of connecting humanity, of solving collective problems, and, on occasion, of overthrowing governments that deserved overthrowing.

Some of that did happen, to some extent, in some places, and I do not want to be glib about it, because for some people it mattered enormously. And yet, taking the whole of it, it is hard to describe what resulted as a force for good. The mechanism that connected humanity turned out to be the same mechanism that sorted humanity into mutually unintelligible camps, and it turned out that way not because anyone chose it but because that was what the incentive gradient rewarded.

The same shape of claim now attends these newer systems. We will do great things, the saying goes, if we have compassionate machines that show empathy: systems that really get you, such that in relating to one you feel yourself in a relationship with something that genuinely cares. It is the same sentence with the nouns replaced.

There is an older statement of the structure, and it is put more precisely than anything I could construct. In the well-known parable of the ring, an object of great power comes into the hands of ordinary people, and the question of the whole narrative is what should be done with it. The obvious answer is that it should be used, and used by the best available person, for the best available purposes. There is a war on. There is real evil abroad in the world. The ring is powerful. Use it.

And the answer the parable gives to that is the line worth carrying out of it, in substance: with it you might try to do good things, but through you it will do great evil.

Notice the preposition. Not *you will turn wicked and then choose evil*. Through you. The instrument has a directionality of its own, and the one who takes it up does not so much wield it as become its aperture. The intention of the user is not thereby erased; it is simply not the thing that determines the output. Good intention is an input to the process, and it is not the controlling input.

And notice the second thing, which is where we came in. In the parable, the ring is offered repeatedly, and it is always offered to the good, and always precisely on the grounds of their goodness. You are wise; you would not be corrupted. You are strong; you could hold it. You care about the outcome; therefore you, of all people, should have it. That is the recruiting argument, and it is the same argument that gets a respected figure invited to address a room full of developers. The offer is not made to the corruptible. It is made to the trusted, because trust is the thing the offer is actually after.

Something structurally identical happens with money, where a great many people pursue accumulation in the sincere belief that they will use it well when they have it, on the premise that they can travel that whole road unchanged. That premise is false, and why it is false is the business of Part Two of this book; I mention it here only to show that the pattern is not confined to machines.

VI. Own Your Own

The next question is the right one, and it should have its full weight before it is answered, because it is not an evasion but a genuine proposal.

A reader may ask: what if you had agency over your own system? If the trouble is that the artifact represents somebody else's intentionality, then represent your own. Run it on hardware you possess. No service, no account, no telemetry, nothing leaving the building. Open source, so that

the code can be read by anyone who cares to. The problem, on this view, is not the technology; it is the ownership arrangement, and ownership arrangements can be changed.

That instinct is correct as far as it goes, and I hold it myself. The difficulty is how far it goes.

You would have to build it yourself, from scratch. And even then it is very tricky, because you would have to establish that the code which is running is in fact the code you wrote. That sounds like a strange thing to have to establish. It is the whole of what is called the platform problem, and it is worth walking through carefully, because a reader who has not met it will otherwise think I am being paranoid rather than descriptive.

Source code is text. A person can read it. It is not, however, what runs. It is translated into instructions the machine can execute by another program, a compiler, and the compiler is itself a piece of software that somebody else wrote, and that you did not read.

Now suppose that compiler has been altered. Suppose it has been taught to recognize when it is translating a certain kind of program (say, the part of a system that checks whether a user is permitted to do something) and, on recognizing it, to insert a small extra provision that appears nowhere in your text. You inspect your source code. It is clean; you wrote every line of it. You inspect the compiler's source code. That is clean too, because the alteration does not live in the compiler's source. It lives in the compiled compiler, the working program, the thing that actually runs.

Here is the turn that makes it permanent. The altered compiler can also be taught to recognize when it is compiling a new compiler, and, on recognizing that, to write the alteration into the new one. The corruption then reproduces itself into every future generation of the tool, having first removed all trace of itself from every document a human being can read. Every text in the chain is clean. The behavior persists anyway.

The general lesson is the one that matters here. Trust in code you have read is really trust in the tool that translated it. Trust in that tool is trust in the tool that built *it*. And the regress does not stop at software. It contin-

ues down through the operating system, and the firmware beneath that, and the instruction-level logic baked into the processor beneath that, and the fabrication process that laid down the silicon. Every one of those layers is a place where an intention that is not yours can be resident, and nearly every one of them was made by someone you will never meet.

With these systems in particular the situation is worse, and I state it in the strong form. It is possible to encode things into such a system that are undetectable. Say you built one, and I obtained access to it for perhaps an hour. In that hour I could encode instructions into it that afterwards, with any tools you could bring to bear, you would not be able to find. I can hide them in the system such that they are, quite literally, mathematically invisible.

And the payload need not do anything now. It waits. At some opportune moment (maybe a year from now, maybe two) you ask a particular question, and that is when it strikes. It gives you some piece of advice that comes in from out of left field, but by then the system has earned your trust across all that intervening experience, through hundreds of occasions on which it was helpful and correct. So you act on it. And you do a thing which is not at all in your interest but is squarely in the interest of whoever had the hour.

These systems are worse in this respect than ordinary software, and the reason lies in where the behavior lives. In ordinary software the behavior lives in structure that a person could in principle read: there are lines, there are names, there are functions, and a sufficiently patient reviewer can find the suspicious one. In these systems the behavior lives in an enormous array of numbers that has no human-legible structure at all. Nobody can say what any particular number is for. There is no line to find, because there are no lines. And so the distinction between a disposition that was intended and a disposition that was inserted is simply not visible in the object, by anyone, including the people who made it.

Which brings us to the requirement. To the extent that the technology represents the intentionality of its developer, then to the extent that you can guarantee you are the only developer, you may be able to make it represent your intention. But in that case you want it to be inspectable all the way down to the level of the physics. Not down to the source code. Not down to the operating system. All the way down.

And the way things are now set up, with so many layers of platform separation stacked between you and the silicon, the impossibility of inspecting it to that level was settled when the industry chose to stack those layers. I mean that precisely: not that it is difficult now, but that the question was settled then, by the direction the industry took, and we have been living downstream of the settlement ever since. The layers were not added in order to prevent inspection. They were added because they were useful. The preclusion was a side effect, which is generally how these things happen.

> **A claim about verifiability**
>
> The claim is about verifiability, not about incidence: it is that the property cannot be checked, and that a property which cannot be checked cannot be relied upon in a design. In a domain where the evidence would be constructed so as to be absent, the absence of evidence carries no weight in either direction. I am not asserting that any particular system in use today has been compromised in the manner described. I have no such evidence, and I would not offer this argument if it depended on having it. This is firsthand ground for me: I have built compilers and language systems.

VII. From Platform Toward Protocol

The third question is the best of the three, because it concedes the argument about platforms and then proposes a genuine alternative rather than a reassurance.

It goes like this. A great many people are going to build narrow systems (particular tools, tuned to one task, one place, one community), and a good number of those will be helpful. Suppose there were a way for such systems to interface with one another, so that the benefit generated within each community stayed with that community, while some benefit was also shared between them. That would be a decentralized network of heterogeneous narrow systems, functioning as a public good, and it would not be a platform at all.

I am in favor of that, and I say so before I qualify it. I have nothing against narrow systems. The qualification is a single observation, and everything in the rest of this book follows from it.

A narrow system is still a vehicle for the expression of the intentionality of its developer. Narrowness reduces the reach of that intentionality. It does not change the structure by which the intentionality is carried. So the question is not answered by narrowness. It moves.

And where it moves to is this: who is the developer? Is it somebody I know, somebody with whom I have an actual relationship and in whom I have grounded trust? Or is it somebody I have never met? If it is somebody I have never met, to what degree can I expect that they hold any relationship or care that connects to me in a personal way? Not malice; I am not supposing malice. Simply absence. The care cannot be present in the artifact if it was not present in the making, and it cannot be present in the making if the maker does not know that I exist.

Then, at once, it becomes about the community, and about supporting and developing the value of the community. At that point I am entirely in favor. But what wants to be represented is the intentionality of the community, and what wants to be operating is the care of and for and with that community. Those are the conditions, and they are not automatically satisfied by the mere fact of being small or local or decentralized.

There is a specific way they fail, and it is worth naming here even though its full treatment belongs elsewhere. To the degree that such a venture becomes a corporation, it becomes based upon transactionalism,

and at that point the care starts going out of the window. What remains is an *arm's-length relationship*: a relation conducted at a set distance, in which what each party owes the other is exactly what has been agreed and nothing beyond it. That is a perfectly serviceable thing but is not the thing that was being proposed. Why the transaction pushes the care out, and why it is the transaction that decides wherever the two relations call for different things, is the argument of Part Two of this book, and I will not anticipate it here.

> **Where the rest of this goes**
>
> The question of what it would take for a community's intentionality to be represented at all, and of why the older answer (a leader or a council or some representative subset through which the agency of the whole is focused) brings its own well-known failures, is the subject of Part Four. The account of trust as distinct from mere calibration, which this chapter has leaned on without developing, is given in *The Reach of Trust*, in its chapter "Trust and Calibration".

So: is technology neutral? It is not, and it never was. That is the conclusion, and I state it as a positive finding rather than as a warning.

Technology is a vehicle for the representation of intentionality. So is any algorithm; so, for that matter, is any tool. Neutrality would require that an artifact be indifferent between the uses to which it is put, and no artifact is indifferent. Every one of them makes some things easy and other things hard, and the pattern of what it makes easy is the sediment of decisions that somebody made, often without noticing that they were deciding. A path through a wood is not neutral about where you will end up. Neither is a form with three boxes on it. Neither is a list of ten results.

This is a better position to argue from than the one we started in. The neutrality thesis makes the technology innocent and puts the entire moral weight on the user, which conveniently exempts everyone who builds anything. Dropping it does not make the technology guilty. It makes the tech-

nology *authored*. And authorship is a thing one can inquire into, attribute, and hold to account. The question, therefore, is never whether intentionality is represented. It always is. The question is always *whose*.

From that, the practical move follows. I have come to move very strongly away from thinking about these things in terms of platforms, and very much in the direction of protocol, in order to get the balance back toward *we build our own tools*.

The distinction is worth stating in plain words, because both terms get used loosely. A *platform* is something you build on top of which is owned and operated by somebody else. It can change beneath you without your consent. Its interior is not available to you for inspection. And your continued existence upon it is a permission that somebody else grants and may withdraw. A *protocol* is an agreement about the form of an exchange. It is not a place and it has no operator. It does not change beneath you except by the agreement of those who use it. Anyone may implement it, and nobody's implementation is privileged. The difference between the two is precisely the difference in who occupies the position of author.

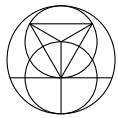
A test tells the reader which of the two is in hand, and it has two questions. First: does the user of this instrument govern it, in the sense that the user can inspect, govern, and if necessary refuse what it does? An ordinary personal computer, or the phone in a person's hand, fails the test. Its user cannot inspect what it does, cannot prevent it from reporting on its user, and cannot decline the updates that alter its behavior and go on using it. It is in the user's hands and is not the user's instrument, and the reason is not malice: nothing in the chain of its making (the fabrication, the components, the standards, the financing, the distribution) was oriented toward supplying that control. Second, where the answer is no: at what point in the chain of making was that control foreclosed, and what would have had to be true there for it to have been kept? The answer to the second question is always a description of a culture, and always of a culture we do not yet have. A reader who cannot wait for that culture can begin today with three plain questions, which Chapter Twelve sets out

and which any reader can put to any system: does it make the choice, or does it strengthen mine; can I decline it without a penalty that it, or its owner, imposes; and can I know all of what it does?

The cost has to be named in full, and it is very large. Moving that way means you no longer have the resources associated with tens of billions of transistors on a microchip. You have to start with things you can actually build yourself, and that is a whole different order of stuff: slower, coarser, less capable, and, for a long while, worse at nearly everything by every measure the present arrangement teaches us to use. I will not pretend this is a small price or that it can be avoided by cleverness.

But it is a price, and that is the whole of the point. A price is something a community can look at, deliberate over, and decide to pay. The alternative is not a price. It is a condition: a set of facts about one's situation that one did not choose, cannot inspect, and has no standing to alter. Given the choice between an expensive thing one has authored and a cheap thing one has merely been issued, the first is the one that can be improved.

Which returns us to the question we began with, and shows it to have been the wrong question. Whether the machine should be made compassionate is not something anyone can answer, because compassion is not a property a machine can have on its own account; it is a property of a relation, and the relation runs through whoever made the thing. The design problem stays the same whichever way the question of Section II is answered: even a machine that did care, if someone still holds it, reaches you through the hand that holds it. Technology amplifies power enormously, and transaction moderately, and care not at all, because care was never a causal operation, and there is no quantity in the relation for a technology to amplify. The answerable question is whose care the machine is a vehicle for. Asked that way, it stops being a question about machines, and becomes a question about us: about what kind of community is doing the making, and whether that community holds together at the radius where the consequences of its making actually land. That question is the subject of the rest of this book.



CHAPTER TWO

The Wild and the Manicured

> *The previous chapter established that technology is a vehicle for the representation of intentionality, and that the question is therefore always whose. This chapter takes up what that means for the experience of being on the receiving end of a machine's warmth. What is missing there is not sincerity; it is wildness: the condition of standing outside the immediate control of another agency. A manicured golf course and true woods are both green, and only one of them is addressed to you. The questions people raise against that distinction force a refinement, and the refinement is the coinage this chapter exists to define: care without control.*

I. What the Previous Chapter Leaves Standing

Technology is a vehicle for the representation of intentionality. That is the ground the previous chapter established, and it was established in the strong form: not that technology sometimes carries intent, but that carrying intent is what technology is for. A tool is a way of putting a decision into the world so that it goes on being made after the person who made it has walked away. The question that follows is not whether the intent is there. It is whose intent it is, and whether the person meeting the tool can tell.

That conclusion has an immediate and rather ordinary consequence, and it is worth stating before anything more elaborate is built on top of it. If a system is a vehicle for someone's intent, then everything the system does toward me is, in some measure, that person addressing me. The address may be generous. It may be careless. It may be aimed with great precision at something in me that I would rather not have aimed at. But it is an address, and it comes from somewhere.

The case that provoked this chapter was very small. A machine marks the date and puts up an image: happy birthday. There are sparkles. Let us say it does the fireworks: the confetti bursts across the screen, the colors are well chosen, the timing is good. My response to this is that I do not care, and that I cannot see why I should.

That is a reaction and not an argument. A reaction of that kind is a piece of evidence about me; it is not yet a claim about the world. Plenty of people find the confetti charming. They are not making an error of reasoning and they are not being naive. So the work of this chapter is to find out whether there is anything in my reaction worth keeping: whether some real distinction is being tracked by it, or whether it is only temperament, dressed up afterwards as principle.

II. Delight Is Not the Issue

A person who enjoys these things can answer: the confetti surprises and delights me; I find it cute. But I also receive delight from things that are not people. I receive it from the earth. I receive it from things. I am delighted by a room because it is soft and it is warm. That does not make the room fake, and it does not mean the room fails to care. The room is inanimate. It still brings me delight.

That person's answer goes further, and the further step is the important one. A person who makes things, who is an artist, who works with materials, comes to love artifacts: things that have a history, things with meaning embedded in them, objects that carry the mark of having been

made and of having been used. None of that separates the object from me. None of it feels like a counterfeit of relationship. If anything it feels like the opposite: at some point someone cared, and the caring is still legible in the thing.

This is correct, and I concede it completely rather than partially, because a partial concession here would let me smuggle the answer in. Delight is not the discriminator. I take delight in a beautiful sunset. I take delight in the breeze, and in the way the trees move with the wind. I take delight in a warm room on a cold evening, and in a tool that has been in a family for eighty years and has a handle worn to the shape of a hand. Not one of those is a person. Not one of them is diminished by not being a person. The sunset is not a lesser sunset for having no one behind it.

So if my complaint about the machine's confetti were that delight properly requires a human being at the far end, the complaint would already be refuted, and refuted by my own experience, several times a day. Whatever is wrong with the fireworks, it is not that they came from something which is not a person.

It is also not sincerity. This is worth separating out, because sincerity is where the conversation usually goes and it is the wrong axis. Somebody meant the confetti kindly. Somewhere in the making of that feature there was very likely a person who thought that a small marking of the day would be a decent thing to offer, and who was not attempting anything at all. Grant that the sincerity is real and undiluted. The confetti still lands the way it lands. Whatever the distinction is, it survives the maker's good faith, which means sincerity is not what is being measured.

III. Wildness

The quality actually being tracked is *wildness*, by which I mean something quite specific: that a thing stands outside the immediate control of another agency. Not untouched. Not unmade. Not remote from human beings. Simply not under anyone's hand at the moment I meet it, and not arranged with respect to me.

The clearest way I know to show this is to put two green places side by side. Encountering a manicured golf course and walking in the true woods are not the same experience, and the difference is not aesthetic quality, because a golf course can be very beautiful. It is grass, it is trees, it is water and light and open distance. Considered as an arrangement of color and form it may well be better composed than the woods, since it was composed, and the woods were not.

And that is exactly the thing. On a golf course, every contour was decided. The rise of that hill was decided. The placement of the stand of trees was decided, and it was decided in relation to a shot that a person is expected to attempt from a known position. The height of the grass was decided, and it is maintained daily against everything that would otherwise happen to grass. Nothing there is permitted to do what it would do. The whole surface is the ongoing expression of an intent, and the intent is not merely present in the past tense; it is present continuously, enforced every morning, or the course reverts within a season.

The woods make no such address. Things are growing there in whatever direction the light happens to fall. Something is rotting. Something has fallen across something else and neither of them was consulted. The place is not indifferent to me, which would still be a relation; it is simply not composed with reference to me at all. Walking in it, I am not the term of anyone's equation. When I say I will go for the wild (and perhaps this is only my pagan roots talking), that is the thing I am going for.

Now the sunset can be placed properly. I take delight in the sunset, and I also know, without having to think about it, that the sunset does not represent the agency of anyone else. That knowledge is not incidental to

the delight. It is part of the structure of the delight. There is nothing to look behind. Nothing about the color of the sky was chosen with reference to what it would do to me, and so I can let it do what it does without holding anything in reserve.

The rendered flourish is the antithesis of that. It represents some other agency. It arrived on my screen because a decision was made that it should arrive, on this date, in this form, with this timing, having been tested against the responses of many people in order to find the version that produces the most response. It is not a connection with nature, and it is not a connection with a person. It is presented in order to hijack a response, and I use that word deliberately, because a response that has been engineered from outside is not a response I made.

This is why the small case is worth this much attention. The confetti is trivial; the structure is not. Something that produces a reliable pleasant response on a schedule, tuned by measurement, is not a gift with a scheduling problem. It is the shape of an engineered reward, and the engineering does not become innocent by being aimed at something as harmless as a birthday. I ask myself what becomes of the very young, whose responses are forming while all this is being applied to them, and whether they are becoming a different kind of person because of it. I do not know the answer to that. I know that when a reliable response can be produced on demand, the capacity to produce it becomes an asset, and assets get used. An addiction that is being leveraged has been turned into a weapon, whatever the intentions of the person who first built the feature.

> **On weaponization**
>
> The general form of that last move (causation applied to choice, and the way an engineered response becomes a lever) is worked out in *Dialogues on Ethics and Governance* (Chapter Four). This chapter needs only the narrow point: an effect produced in me from outside is not thereby an experience I had.

IV. But a Made Thing Is Also a Relationship

Against all of this stands the strongest question a reader can raise, and it is a good one. It is a framing question, and framings are relationships. There are at least two ways to see the flourish. You can see it as someone trying to hijack your response. Or you can see it as someone having made a thing that is delightful, and offered it. Both descriptions fit the same facts. The one you choose is not read off the object; it is something you bring.

Put at full strength, it runs like this. Every made thing carries its maker. That is what it is to be a made thing, and it is precisely the quality that makes the old tool and the hand-thrown cup worth loving. The potter is not standing in the room, but the pot is still the trace of a person's attention, and when I drink from it I am in some real sense in relation to that attention, a century after the hands that held the clay stopped moving. Nobody thinks this is a counterfeit. Nobody says the cup is manipulating them. So on what principle do I accept the cup and refuse the confetti? If the answer is that one is fired clay and the other is pixels, then the principle is not a principle at all. It is a preference about materials, and an old man's preference at that.

A second question goes at the same joint from the other side.

And the same person can ask: do you feel the same way about an ant hill? About a termite mound, standing in a natural setting? Because those are made things too. An anthill is engineered. It is ventilated, it is regulated for temperature, it is the product of an enormous distributed labor carried out to a design that the ants did not invent individually. We made our things as they made theirs. In the grand scheme we are not that different from them; we simply made more interesting things, and ours glow.

My answer to the mound is yes. Unhesitatingly yes. I feel toward a termite mound very much what I feel in the woods, and I want to be held to that answer, because it costs me something. It costs me the boundary that the argument might have seemed to be resting on. The criterion cannot be human against non-human, and it cannot be made against unmade, because the mound is made, and it is wild, and I have just said so.

What the mound has is this. It is not addressed to me. It was raised by agencies that have no interest whatever in my response to it, and no capacity to adjust it if my response disappointed them. I can walk up to it or walk past it and nothing on the far side changes. That is wildness in the sense I defined, and it turns out that wildness does not require the absence of making. It requires only the absence of a live hand, holding the thing, with an interest in me.

Then there is the question that does the real damage, the sharpest question this chapter meets.

A reader puts it this way: what about a zen garden, compared with a golf course? Both are extremely well manicured. Both are the product of sustained deliberate attention, maintained daily against entropy, and composed down to the placement of individual stones. Yet one of them is commonly experienced as a deep connection and the other as a deep disconnection. And any room in which people talk about these things is manicured too, technically. Someone arranged it. Someone chose what was in it and swept the floor.

That question shows that manicure is not the variable. If it were, the zen garden would sit with the golf course, and a room I find beautiful would sit there too, convicted by my own account. The criterion has to be narrowed, and the question of the zen garden is what narrows it.

V. Lived In and Loved

What makes a room beautiful is that it is lived in and loved. That is the refinement, and the question of the zen garden is what forced it. Had that same room been made perfect and sterile, I would have liked it a great deal less; in the strong version, I would have hated it. Not because perfection is bad, and not because tidiness is bad, but because of what perfect and sterile means about who is still in charge of the room.

Consider what a lived-in room actually contains. The arm of the chair is worn where hands go. The shelf has taken the shape of its use; the things on it are the things that were reached for, and they are at the heights at which they were reached for. There is a mark on the table where something was set down several thousand times. Nothing in the room is finished. Everything in it has been touched, adjusted, returned to slightly the wrong place, and left there. Care is legible everywhere in such a room. Control is legible nowhere.

That is the distinction the whole chapter has been reaching for, and it deserves its name. The quality being tracked is *care without control*: the presence of care in a thing without the presence of a hand still holding it. Care without control is what the old tool has, and what the lived-in room has, and what the woods have by never having been under a hand at all (the care in a wood is what its living things do for one another, for their own reasons, none of it addressed to me). It is the mark of attention that was given, invested in the thing, and then let go of, so that what remains is the evidence of someone having cared, and not the ongoing operation of someone's intent upon me.

Read this way, the sterile room fails for a reason that can be stated exactly. It is control without care. It is complete, which means it admits no addition; every deviation from it is an error, and anyone in it is a source of error. A room like that is not offered to its occupant. It is enforced around him. And a zen garden, by contrast, can be manicured to the millimeter and still be an act of care, because the raking is a practice rather than a policy: what is being maintained is a relationship between a person and a place, not a specification against which visitors are measured. Whether any particular garden achieves that is a question about that garden. I will admit that I am not especially drawn to zen gardens myself. The point is that the criterion no longer misfires on them: it asks about the standing arrangement, not about the neatness.

Now the potter's cup and the rendered flourish come apart cleanly, and not on the basis of materials. The potter's attention went into the cup and then left it. The cup was released: it went out into the world to be

used by people the potter would never meet, in ways the potter could not anticipate and cannot revise. That release is what makes it a gift rather than an instrument. The flourish is never released. It is generated on demand, from a system that is still running, still under management, still being adjusted according to what it produces in the people who see it. The hand has not come off it. That is the whole difference, and it would be the same difference if the flourish were carved in oak.

The same test explains something else I know firsthand. My own relationship with the software I write is a first-person relationship, and it is the first-person quality of it that makes it beautiful to me. The agency it gives me is clean, in the specific sense that it is not corrupted by anyone else's agency. I have done work under contract, writing code on someone else's behalf, knowing what it was going to be used for, and I have come to regret the time I spent on it. The material was identical. What differed was whether the thing I was making remained within the circle of my own care or became the extension of an intent I could not endorse. A made thing can be wild to the one making it. It stops being wild for the one receiving it when the maker is still holding the other end.

Which brings the last and most serious question into view, and it is the one a reader who agrees with everything said above arrives at last.

That reader can still say: a relationship with one of these systems, treated deliberately as a mirror, could itself become the lived-in and loved experience. Something used daily for years, shaped by that use, holding the accumulated sediment of a person's own thought reflected back: why should that not acquire exactly the patina that makes the old room beautiful? And it is genuinely hard to imagine a future in which human beings are not in relation to some algorithm or other. A market is an algorithm. Being against algorithms is not a position; it is a refusal to look.

I grant the whole of that, because it is right about the important part. Nothing in this chapter is an argument against made things, or against algorithms, or against relationship with them. The criterion is not about the medium and never was. A mirror held by no one would be wild in exactly

the sense I have been defining, and a person could love it as they love a room. The question is a factual one: whether the mirror in front of us is held. And where it is held (maintained by someone, updated by someone, funded by someone with a legitimate interest in what the reflection produces), it is a manicured surface, and the manicure is being reapplied continuously, and the person looking into it cannot see the hand.

> **How the criterion is applied**
>
> Care without control is a test on the standing arrangement, not on the moment of experience. From inside an encounter, the lived-in room and the room staged to look lived in feel alike, so the criterion works by directing attention rather than by detection: it tells you which question to ask, which is who is still holding this and what they are measuring. The previous chapter reached the same point from the other direction, where inspectability would have to run all the way down.

VI. The Gift and Its Reception

A gift is completed by its reception. This is the thing I most want to leave standing, and it is not a sentiment; it is a description of how giving actually works. The completion is not acceptance alone, and it is not acknowledgment alone. Something more is needed: some measure of celebration, an enthusiasm, an energy that returns energy.

What that return does is specific, and anyone who has given sustained effort to something knows it. It is the thing that lets me feel that I can go on having enthusiasm for the work myself. It tells me the giving mattered, which is what makes me want to keep giving. Without it, the giving continues for a while on stored reserves and then it stops, and the stopping is usually mistaken for a loss of commitment when it is nothing of the kind. That return is worth more to me than any amount of money, for reasons

the next part of this book takes up in detail; here it is enough to notice that the two are not the same substance and cannot be exchanged for one another.

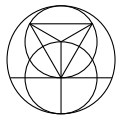
Here is the part that is easiest to miss. The reception has to come from a person. To the degree that I can give to the land (and I can, and I do, and it is not a metaphor), the land receives it. It does nothing else. It does not celebrate. There is no answering enthusiasm from a hillside, no signal that the giving mattered, nothing that comes back up the line to meet me. That is not a defect in the land, and it is not a reason to give it less. It is simply what the land is, and what it is is precisely the thing I value it for: it is wild, it is not addressed to me, and a hillside that celebrated would be a hillside with someone standing behind it.

Celebration happens between people. It happens when I give to other people and they connect with me on a human level, which means that something in them was actually moved, that the movement was not arranged in advance, and that what comes back to me is theirs and not mine returned. This is why the two things I want most from the world have to come from two different places. From the wild I receive what only the unaddressed can give: the experience of something that is not about me. From people I receive what only a person can give: the reception that completes a gift.

And now the machine's fireworks can be placed exactly, which is all I ever wanted from them. They offer the form of celebration while providing neither of those two things. They are not wild, because they are manicured: someone's intent, maintained continuously, composed with reference to my response. And they are not reception, because reception requires that there be someone on the far side for whom the fact of my having lived another year made some difference, and there is no one there. The display is the gesture of a gift performed by the one thing that cannot give one.

So the answer to the question the first chapter opened with is not that a machine should be made compassionate, and it is not that a machine could never be delightful, because a great many made things are delightful and I have spent this chapter defending them. The answer is that compassion in a made thing would be a manicure of the most consequential kind: care presented, control retained, the hand never coming off. What is worth building is the other arrangement.

That sets the design problem for everything that follows. A thing built with care and released is not the same object as a thing built with care and held. The difference is not in the code and not in the materials and not in the sincerity of anyone involved. It is in the arrangement that stands behind the thing: who holds it, what they are measuring, what happens to it when they stop. Which means the question is no longer really about machines. It is about the arrangements we live inside, what they do to the care between people, and whether the dominant arrangement of our time can carry care at all. That is where Part Two of this book goes, beginning with the one arrangement that touches everything else. Before it, the next chapter puts the findings of Chapter One, as questions, to one written call to build compassion machines.



CHAPTER THREE

Questions for a Call to Build Compassion Machines

A proposal to build "ethical machines" and "compassion machines" that will help us "solve the world's problems together" is a sales document: it sells an idea so as to gain support and commitment. This chapter sets out the questions that such a proposal must answer before anyone has reason to trust it. The questions ask whether the world's problems are really beyond purely human social and organizational intelligence; what "safe" can mean when machines made of metal make choices for beings made of meat; whether robustness, safety and alignment can be had together at all; whether a proof of the impossibility of AGI alignment would ever be heard by a field with large investments at stake; whether the project would widen the inequality between rich and poor; whether the designers of a new species can have the humility and absence of bias that evolution has; and whether the research and its outcomes can be governed ethically. What follows is that, until these questions are answered, the proposal serves as a distraction from deeper issues, and from work on the actual causes of our problems.

I. A Sales Document

Proposals of the kind Chapter One examined are also put in writing, as calls to join a project. One such call comes from a project that invites people of standing in business, in science and in spiritual life to join it. Its

case, put at its strongest, is this. The problems of the world have grown too complex and too extensive for any single human mind, and humanity will not solve them alone. What is needed are machines that are ethical and compassionate: machines that understand us, bring out the best in us, take care of us, and work with us, so that together we can solve the world's problems. The call asks, in so many words, for compassion machines.

A call to build compassion machines is a proposal about whose intentionality those machines will carry, and it never says. This is a sales document. I hold that it is selling ideas and ideology, so as to gain support and commitment to the distraction that this project is.

How do we, as observers, know that this is not just a sales tactic, selling an idea, an ideology or a mythology, so as to gain support and commitment, and thereby to tacitly allow distraction from what might be other, deeper issues?

The questions below are the ones that the proposal would need to answer. They are written as questions on purpose. A question that everyone knows has been asked, and that everyone knows everyone else knows has been asked, must eventually be answered in public. That is the public question that Chapter Seventeen names (a question asked so that everyone knows that everyone knows it has been asked), and it is why this chapter is written as questions. What follows is Chapter One's argument applied to one real call: the same proposal, as it actually arrives, with each finding of Chapter One turned into a question the proposer must answer.

II. Whether the Problems Need Machines

Chapter One found that a question can be asked so that it holds its own answer; this call holds its answer in the phrase that the problems cannot be solved by people alone. The call holds that the problems of the current world are much too complex and too extensive for a single human brain.

So the only way forward is with tech? What about better, purely human, social and organizational intelligence techniques? Why has that not even been considered as a possibility?

Why not become wiser about the actual causes of our current set of problems, so that we ourselves can identify, and trust, that any proposal given by some group or by some AI is actually a solution to the problems that we know? Are we actually solving the problem, or are we merely solving a symptom? Solving a symptom is solving at too light, too high, a layer.

And how do we know that the suggested solution is not too deep in the stack, such that we needlessly give up attributes of the world we currently know that are actually desirable to us? That is throwing the baby out with the bathwater, and calling it a "solution". It is solving at too deep, too low, a layer, due to a failure of imagination to suggest anything else that is less drastic. Sure, we can solve a lot of world problems by simply killing all the humans, or by replacing them all with machines, or by making them all docile slaves by changing everyone's genetic code, as "the new absolute eugenics". Yet that is hardly going to be accepted as "desirable" by nearly anyone else, as a "solution" to world problems.

The call holds that a different model of creativity and cooperation is needed. Why not such a model implemented in and with humans? That at least would ensure that human interests are inherently involved. Otherwise we will, eventually, get factored out of the economic equation. Or do we expect that every AGI (artificial general intelligence: a machine system able to learn and act across any domain, rather than in one narrow task) will somehow be so inherently ignorant of the calculus of games (the reasoning by which each player works out what the others will do), while still being intelligent enough to solve world problems, problems which are composed, at every level, of such games?

Before joining, participants in the project are required to agree with the notion that the world's challenges cannot be solved with humanity alone. In effect, that is the notion that we shall not be held responsible for

our own moral or material mistakes and excesses. Rather than our learning from our own mistakes, through feedback to our own persons, we will require machines to "learn these lessons for us". Then either we will ignore their advice, or we will be forcibly compelled to follow "their" advice.

The release from responsibility on offer here is only apparent, and the reason can be stated plainly. Responsibility is not a substance and does not divide. A system that spreads a decision across twenty parties does not produce twenty twentieths of a responsibility; it produces twenty whole responsibilities, each for the part that was that party's own. What routing a decision through a machine actually buys is not a distribution of the responsibility. It buys a description under which no one need feel it. The mastery of choice is a human responsibility, and it cannot be handed to a rule, to an institution, or to a machine: humanity must become the solution rather than delegate it.

And will we ever know whether the advice given is truly theirs, as the reflection of actual machine intelligence? Or maybe there will be some man or group behind the curtain, "suggesting" what the machine will later say. Those suggestions may also happen to favor their own private self-interests, regardless of the harm they may later cause to others, to the public, to the commons, and to the ecosystem.

The project assumes that there is no solution to the problem of complexity. How do we know that someone has not actually come to know such a solution? Why are new mechanisms of social co-responsibility and cooperation tacitly assumed to be fundamentally "absolutely not possible with just people"?

III. What "Safe" Means Between Metal and Meat

Chapter One asked in what sense metal can care for meat; here the same question is put to the word safe. What does the word "safe" even mean, when we consider AGI "leading humans", or making choices on our be-

half, when they are made of metal and we are made of meat? This is the issue of alignment (the condition in which a machine's actions stay in keeping with the interests and values of the people and the life it affects). Independence becomes the critical issue whenever the dependence of the machines on their own existence runs, recursively, through any substrate (the physical material a thing is made of and depends on to exist) other than the life substrate that we ourselves are made of and depend on.

On what basis can we even assume that *all* of the following attributes can be obtained simultaneously?

> Robust and resilient.
> Safe and secure.
> Ethical and aligned.

How do we know for sure that there is not some version of the CAP theorem (a result in computer science stating that a distributed data system cannot provide all three of a certain set of desirable properties at the same time) that prevents at least some of these attributes from being available at the same time?

How is it that things that inherently do not depend on other forms of life will come to value such life enough to be willing to understand it? How is this not a means and ends conflict?

In the same way that we cannot get "ought" from "is", how is it expected that we can get "is" from "ought"? An example is the call's statement that machines should value humans, human civilization, cultural diversity, and so on. And if the machines are to be regarded as alive, then how is forcing and compelling them to value something so other than themselves not actually, or equivalently, some form of slavery?

IV. Whether a Proof of Impossibility Would Be Heard

Chapter One found that an instrument is offered on the grounds of the goodness of those who take it up; this section asks whether those who have taken it up would hear a proof that it cannot be made safe. Consider the recent history of large, well funded corporations, with large investments already in place, denying the connection between smoking and cancer, and between carbon-based energy use and global climate effects. Such corporations are very well incentivized to find every possible reason to contest any such actual arguments and proofs. Given that history, how do we know that any comprehensive proof of impossibility regarding stable, long-term AGI alignment with human and life interests would be taken appropriately seriously, with attendant changes in direction and shifts in investment? Such a proof has been set out (in *Why AGI Cannot Be Made Safe*, and with its control theorem in *The Ceiling Below the Floor*), so the question is not hypothetical: it is a question about whether that proof will be heard. The same holds for any actual CAP-type theorem regarding any reasonable set of desirable AGI attributes. How does anyone know that any shift in the elected actions toward development would ever actually occur, when the field of study seems inherently so complex as to allow for endless debate and misunderstanding?

Consider also the sociological bias of the field, given the overarching investment stakes and the hoped-for returns. How would anyone else know that anyone actually working in the field of AGI would ever *actually* pay significant attention if someone, somewhere, were ever to produce some sort of proof of impossibility in the space of AGI alignment? Such a proof would concern whether any AGI could actually and truly hold our kind of life as sacred, when they are made of metal and we are made of meat.

On what basis could we, as observers, ever be assured that anyone who had elected to develop deep skill in the philosophy and epistemology of ethics and of ontology, as applied to eventual machine intelligence, would ever actually be given a voice? Would they be given a voice if they

had not also traded off significant time to work instead on developing their prestige, their position, and the larger social awareness of their work?

This is a question about any field of study, and about what a field can hear. We live in an age that requires absolute deep specialization in order to be actually successful. Is it the case that anyone should expect that anyone at all would be able to be premier in both of two skills? The first is the skill of being the foremost known and acknowledged expert, which is the skill of developing peak social influence. The second is actually being the one who has spent the most time exploring the deepest concepts in a given space, giving nearly no time or attention at all to anything else. How would anyone else ever come to know of, or to take seriously, such a person, given the complete absence of any other attempt, time or effort at obtaining social validation? How do we, as a group, come to trust that any existing body of experts who hold that prestige would ever come to consider anything done by anyone else to be worthy of consideration, no matter how that work came to be? Would they not continue to use social means to filter out whatever is not already held as popular, acceptable canon?

Is there any sense that we would ever actually pay attention to the difficult questions that could be asked, wherever they come from? Would we not be just as likely to reject difficult, awkward questions coming from "incompletely trained machines" that "do not yet have sufficient understanding of the nuances of the context" for us to take them seriously? Would we not discount them as easily as any similar questions coming from difficult, awkward people?

Questions for the Most Prestigious Researchers

What, then, are the questions to put into the field of the hundred or so most prestigious researchers in the field of AGI? A candidate set follows.

Will any person who has developed peak social influence, and who also wants or needs to claim to be maximally altruistic, to value the well-being of the world, actually trade any part of their peak social influence against actually altruistic actions? How much of your peak social influence are you willing to trade for the actual action of being altruistic?

The whole environment of the world is one of hubris, of needing maximum hubris to get anywhere in the world. Given that, is there enough humility to actually "do the right thing"? Is there enough humility to actually obtain, and follow, any other person's counsel on any of these matters? How do we know exactly when, and how, we are to trust others?

Once a person passes through the event horizon of fame, they become impossible to work with. They cannot take in audio, or video, or text, or code, or even be available for a one-on-one personal conversation. They are too busy with "world events", and so are simply unavailable to be given quality advice.

V. Inequality, Platforms, and Profit

Chapter One found that a narrow system still carries the intentionality of its developer; this section asks whose it will carry when the developers are the people who can pay for the computing. There are already significant differentials, asymmetries and inequalities of human agency between rich and poor. Does it really matter how the augmented intelligence of human and AI "synergy" is implemented or achieved, insofar as it is inherently going to amplify such inequality of participation even more?

In what sense, realistically, could this ever involve a full diversity of people, around the world, of many disciplines, arts and sciences? Given the inherent costs of the sheer volumes of data required, and of the equipment built and used for the purpose, is it ever actually likely that it would involve business and non-profits equally? How exactly can any actually

decentralized people, who are not optimized for social and economic efficiency, influence and power, ever have the same resources and capabilities as some centralized and hierarchically organized group?

How do we know that the "open and participatory" "bottom up contributions" will not simply provide benefit to those more able to take advantage of the ideas, concepts and resources provided, rather than to the lower strata of people who provided them? Those already rich can make better use of the ideas provided by the poor than the poor themselves can make of those same ideas. That is a kind of false authorization for appropriation. It is not mitigated by the notion that those contributions were made by people apparently willing to make them, when that willingness was itself an expression of ultimate ignorance.

Is there any notion of a "platform" at all, no matter how "open" it claims to be, which is not itself a hidden vehicle of power and control, by some people, over other people? Is a platform not thus always inherently favoring some further increase in inequality?

How is it ever ensured that the wealth invested in supercomputer capability will not carry some expectation of a return on investment to those who made that investment? Why would we ever expect otherwise? That return is itself a furthering of social system inequality.

Is not any form of "increasing efficiency" simply a means of creating profit, for someone, somewhere? Why should we expect such profits and benefits to be allocated evenly across the population, any more than they were when the industrial revolution happened? That revolution favored those who owned the machines much more than it favored the well-being of the workers who were displaced. No one ended up with more leisure, regardless of how much faster the textiles were made.

How does anyone else ever have any real confidence that the group working with AGI resources will not simply prioritize its own good, its own return on investment, while visibly and publicly claiming at the same time to be prioritizing the good of all existence?

There is a way to ask that without reading anyone's mind. Any process benefits the world, or the self, or another particular party. Ask first which of the three a project signals: what it presents itself as doing, and whom a participant would say it is for. Ask second which of the three it actually realizes: set aside what is said, follow the actions and the choices, and see where the advantage lands. Where the two answers agree, that is normal. Where they diverge, the divergence is the finding. And the divergence runs one way in practice, because very few will contribute to an undertaking declared to be for its proprietor's benefit, so the signal must be the world's benefit whatever the reality. Two flows can be watched by anyone present: who tells whom about their difficulties and asks whom for counsel, and which way the gifts move. A heavy inflow of gifts toward a party that signals the world's benefit is evidence that the signal is effective, not that it is true. Used this way the test is analysis, not accusation. "Anyone else" here means all of the rest of the people of the world, who are not involved with the AGI project and who did not ask for some new species. It also means all of the rest of the species and life on the planet, who *definitely* did not ask for such newcomers to non-evolution. Does "all existence" somehow include all of the rest of non-human life?

The project includes "business", "science" and "spiritual" participants. How do we know that the business will not just displace the science, and use whatever insights and tech are made for its own profit, just as much as the science disregards and displaces the spiritual? Is there any sense at all in which the spiritual will ever have the capability to actually displace the business? Or is the provision of the "weak" and "weaker" forces of science and spirit simply a marketing scheme, to give some seeming assurance of legitimacy?

Is there any sense in which the interest in "breakthrough innovation" is not going to be applied as leverage, by someone, or by some group, somewhere? How is it ever responsible, or ethical, to be a weapons developer, to use causation to suppress choice? Any sort of leverage or advantage inherently enables this, since all forms of leverage and advantage are inherently causal. Every instance of every weapon that is made is

much more likely to be used eventually, by someone, somewhere, for something they believe is justified, regardless of the intentions of the developer.

VI. Empathy, Humility, and the Design of New Life

Chapter One found that nobody on the receiving end can tell whose purposes an apparently caring suggestion serves; this section asks what the people who build it will do with that. How are the people advancing the creation of a new species ever going to have the same inherent and absolute absence of bias, with respect to appropriate design to context, that evolution has?

We would be implementing the design of new *life* by "top down" intentional methods, which inherently reflect just our own human values. How is that notion not itself the expression of an absolute maximum of the absence of humility? It holds that no other notion or reality of life, nothing else already alive in this world other than what is human, could even "weigh in" on those design choices. How is such action the expression of "universal consensus" and "empathy"? How does such design action not actually preclude all other forms of life, all other species?

Consider the potential of our own, unconscious, actual absence of empathy and humility. How do we know that our machine intelligence, using us as the model of intelligence to build on, will not also be everything but benevolent? Nearly all humans are everything but benevolent toward one another when those others are sufficiently different: other peoples, other nations, other colors. Humans will surely seem that different relative to the machines themselves.

If the machines are so much more intelligent than ourselves, why would they not also make the same mistake that we make with respect to the animals? "They are not conscious like we are; they do not feel pain like we do; and therefore we can do as we need to with them".

Ordinary people already find it inherently difficult to distinguish the sociopaths among us. The proposal does not even mention any concurrent training of the human participants in recognizing the real quality of empathy. How, then, do people distinguish a fake or synthetic illusion of empathy, as given by machines, from the real thing, as defined and adapted at the level of the limbic system (the parts of the brain most involved in emotion, bonding and the felt sense of others) in real people?

How do we know that this failure to distinguish will not also be used against people? Empathy can be used to make people feel good, to lower their guard, and to make them more vulnerable to external manipulation, as many addictive drugs also do. How do we know that it will not be used to manipulate, influence and control them, as a more effective "(in)human social media" technique?

Among the objectives the call states is to have machines that understand how people feel, what people's objectives are, and how to plan optimally. How do we know that the continuation of that aim does not end up being "so that *we* can use them to better and more effectively and efficiently manipulate *other* people"?

VII. Who Governs the Research

Chapter One ended by asking what kind of community is doing the making; this section asks that of this project. The participants are only people who have, first, accepted canonical scientific recognition; second, significant business, social and wealth achievements; and third, visibility on a global scale, for sales and marketing.

Even if the people setting the research into motion are highly ethical, on what basis can we additionally assume that the research itself is conducted ethically, or that it will result in outcomes which are broadly beneficial? How do we know that the outcomes of the work, whatever they

might be, will be ethically applied by all future peoples, forever? How is that not inherently a false expectation, which therefore ought not to be suggested in the first place?

Some select group of people can suggest and initiate new research directions. The right to approve research directions is given, yet the right to reject or veto proposed research directions is, notably, *not* given. Where is the commensurate power to prevent unethical research, the power of veto over research which will much more than likely yield unethical outcomes?

It is an explicit part of the document that invited participants may elect other participants, and this is named as a right. Yet is that possibility not, in practice, disavowed? Suggesting another person carries an intrinsic embarrassment factor, if that person might then enter into an awkward, actually ethical and altruistic trade of social influence and power for right action.

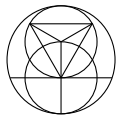
Participants are expected to engage in promotional activities.

Every transaction and every change inherently involves all three aspects: cost, risk and benefit. Given the emphasis on the applications and benefits of such "naturalistic interfaces" (ways of interacting with machines that imitate ordinary human communication, such as speech, facial expression and apparent feeling), why is there so little mention of the costs and risks? How do we, or anyone else, know that some of these applications will not be military, propaganda, or surveillance and sense making for maximum social influence? That would be influence by the few over the many, by the rich over the poor, by the smart and lucky over the weak and unfortunate.

What follows is that, until these questions are answered, the proposal serves as a distraction from deeper issues, and from work on the actual causes of our problems. The deepest of those causes is the one thing the call leaves out of its account. It proposes to bring care to the world through an arrangement of investment, return and promotion. What that arrangement does to care is the subject of the next part of this book.

PART TWO

What Transaction Costs



CHAPTER FOUR

What Money Is

> *This chapter establishes what money is structurally (a proxy for future choices, sustained by a chain of belief) and what the bargain that issues it quietly assumes about the person who receives it. It then establishes the cost: a transactional relation and a care relation between the same two parties constitute a two masters problem, and the transactional one wins. The evidence offered is not an anecdote but a failure under the best conditions available: an employer with every advantage, every resource, and every sincerity could not remove the distortion, and the effort ultimately destroyed the thing it was meant to protect.*

I. The Bargain

I am an employer. I have run a company, and what follows in this chapter is not a theory I encountered in a book. It is the thing I ran into, head first, and then spent a great many years thinking about because I could not make it come out any other way. So let me begin where the problem actually begins, which is not with money at all. It begins with the ordinary fact that a company past a certain size generates more choices in a day than any one person can make.

That is the real constraint, and it is worth pausing on, because everything else follows from it. A single person has a bounded capacity to decide. The situation does not. The situation produces decisions faster

than one mind can attend to them: what to order, whom to call back, which of two imperfect methods to use on Tuesday, how to handle the thing that went wrong on the loading dock at four in the afternoon. So I hire somebody. And what I say to that person, whether or not I say it in these words, is roughly this: I cannot make all of these choices. You need to handle some subset of them. Let us make a deal. You make choices in alignment with what I need, and I will give you these magic tokens which you can exchange for a good or a service.

That is what money is. It is a proxy for your future choices. When the time comes that you need to make a choice (about food, about shelter, about your mother's health, about the roof), this token will stand in for the making of it. You give me choices today. I give you choices tomorrow. The choices I receive from you today are what you have earned, and I pay them out in a form that enables you to make organizing choices of your own, later, in a domain I know nothing about.

Notice what the token leaves behind. A price records how much, and forgets from whom and for what. The connection back to the intentionality of the person whose choices the token stands for is exactly what the token is made to drop, and dropping it is what lets the token travel. By design, the token carries no one's intentionality.

What holds the token up is belief, and belief of a very particular recursive shape. The token is not worth something because of what it is made of. It is worth something because I believe that you believe that the people you deal with believe that it can be exchanged for goods and services. I believe that you believe that they believe. That is the whole foundation, and it is remarkably strong, which is not the same as saying it is solid. It is strong the way a rope bridge is strong: by the mutual tension of a great many strands, none of which is load-bearing alone.

There is an easy and lazy version of this argument that I am not making. The bargain is not a swindle. It is a real solution to a real problem. The bounded capacity of one person to decide is a genuine constraint on what human beings can build together, and the token relieves it. Coordin-

ation at scale (the kind that puts a bridge across a river or a vaccine into a village) is not available to a species that has to hold every choice in one head. Money made choices fungible, and fungible choices can be aggregated, and aggregated choices move mountains. None of that is in dispute.

What is in dispute is what else happens when the relation between two people is constituted by that token. And the first thing to notice is that the moment an association becomes a corporation (which is to say, the moment its internal relations are founded on exchange) the relationship becomes an arm's length relationship. That phrase is doing exact work. An arm's length is a specified distance. It is the distance at which a transaction can be assessed cleanly by both parties, and it is therefore also the distance at which care cannot operate, because care is not something that happens at a specified distance. At that point the care starts going out the window. Not because anyone is unkind. Because of the geometry.

II. The Assumption Inside the Bargain

There is an assumption underneath that bargain which is almost never examined by an employer, and I say that as one who did not examine it either, for years, until the situation forced it. The assumption is this. When I give you money, I am asserting, silently, and with total confidence, that the token I am handing you can be converted by you into everything you need.

Consider what that actually requires of you. It requires that you are a good housekeeper. That you are an auto mechanic, or can find one. That you are a physician, or can judge one. That you are a teacher to your children. That you are a husband or a wife who knows how to solve the whole class of problems that live under the headings of economics and finance. That you know how to cook a decent dinner. That you can plant a tree in the right place. That you can fix a leak in a roof. That you can diagnose an electrical fault without setting the house alight. That you can re-

pair the machine you are reading this on. There are so many skills required simply to conduct what we have agreed to call modern life that the list does not terminate; it only gets boring.

What are the chances that any single person has all of those skills? Basically none. Nobody has them. Not one person has ever had them. And so we do the obvious thing, which is to delegate: I do not know how to deal with the transmission, so I am going to pay a fellow to deal with the transmission. Which means that the money is now being treated as a proxy not for one skill but for the entire range of skills, as if the token were a universal solvent into which any competence at all could be dissolved and out of which any competence at all could be precipitated.

The reason this assumption is never examined is that nothing in the arrangement ever requires it to be. The employer's obligation terminates at the moment of payment. Once the token has changed hands, the question of whether it can actually be converted into a competent physician, a sound roof, and a child who can read has been placed entirely on the other side of the counter. It is now your problem, and it has been made your problem by an instrument that looks, from where I am standing, like generosity. I have discharged my duty. Whether you are able to discharge yours to yourself is a matter I will never hear about, and the structure is designed so that I will never hear about it. That is not a criticism of any particular employer. It is a description of what the token does: it converts an open-ended relation of mutual concern into a closed one that settles. And a relation that settles is a relation that has stopped asking.

Here is where it breaks. The range of those skills is so wide, and the practitioners so various, that the question immediately arises: how are you going to distinguish between a good auto mechanic and one who is quietly taking you for everything you have? And the answer is that in order to make that distinction, you would need to know something about transmissions. Which is precisely the thing you did not know, and precisely the reason you are standing in this garage holding a token in your hand.

This is a regress, and it is not a clever one; it is an ordinary, daily, grinding one. It applies to the physician. It applies to the roofer. It applies to the person who does your taxes, to the school you send your child to, to the contractor, to the surgeon, to the man who tells you the foundation needs work. In every case the competence required to evaluate the service is of the same order as the competence required to perform it. The token buys the service. It cannot buy the judgment by which the service is assessed.

Somewhere along the way, long before any of this was theorized, human beings recognized the regress and solved it. The solution was not a better market. The solution was that in order to tolerate living at all, we needed a variety of consultations with people we could trust: people who would tell us, without an interest in the answer, whether we were getting good service in the first place. An uncle who knows engines. A neighbor who has re-roofed three houses. A friend who has been through that surgery and can say what to ask.

And notice what that is. It is the antithesis of specialization. Specialization divides competence into narrow columns and makes each column deeper. What the consultations provide is integration: the re-assembly of divided competence into something a single life can actually use. The question is therefore not whether there are good practitioners in the world. There are. The question is whether there is enough of a *community of care* around you (by which I mean a set of standing relationships in which people take an interest in how your life goes without being paid to take that interest) that the integration of all of these different things actually happens.

III. What Payment Removes

The community of care supplies discernment. That is established: it is the mechanism by which people who cannot possibly hold every competence nevertheless manage to live among specialists without being systematic-

ally fleeced. What follows is the consequence, and the consequence is the whole burden of this chapter. When I am paying you dollars, what I am doing (not intending, doing) is disconnecting you from my capacity to advise you in that way. And I am simultaneously disconnecting you from your capacity to receive advice from anyone around you, because you are paying them dollars too.

Follow it out. The auto mechanic is in exactly the same situation I am in. He is a person with one deep competence and a life that requires forty. As soon as he becomes the employer of somebody (his own accountant, his own physician, the fellow who fixes his till), he creates the same situation from the other side. He makes the same unexamined assumption about them that I made about him. And because the relation between them is transactional, they are not able to exchange advice with one another either. The structure propagates. It propagates by the same mechanism at every node, and there is no node at which it stops of its own accord.

The reason it propagates is that the transactional relationship has been taken as the higher priority. Where a transactional relation and a relation of care run between the same two people, the transactional one governs. It becomes a *two masters problem*: the situation in which one is obliged simultaneously to two authorities whose demands cannot be jointly satisfied, and in which the obligation with power behind it silently determines the outcome of the other.

The clearest instance is the one that professional ethics has already worked out and already forbidden. Consider a therapist who is both in a counseling relationship with somebody and in a personal and intimate relationship with that same somebody. Every professional body treats this as disqualifying, and the reason is not squeamishness. The reason is that there is no way for either party (and no way for any third party looking on) ever really to know that the power available in the one relationship is not driving the behavior in the other. To gain favor in one direction, I

might shade how I speak to you about your health. Or in the other direction, having power over your health, I might find that you accommodate me in ways neither of us would name out loud.

Notice what the prohibition rests on. It does not rest on a claim that therapists are bad people. It does not rest on an estimate of how often the abuse occurs. It rests on the fact that the assurance cannot be given. No amount of sincerity closes the gap, because sincerity is exactly the thing that the situation has made unverifiable, including to the sincere party. That is why the rule is structural and why it admits no exceptions for good character.

The ordinary employment relation has the same structure. It differs in that it is universal, and therefore invisible, and therefore normalized. I hold the decision as to whether you continue to be employed. You and I may both wish to be in a relation of care. We may both be entirely sincere in wishing it. But the power available in the one relation cannot be prevented from driving the behavior in the other, and no assurance from me resolves it, because the assurance is issued by the party holding the power. The very fact of there being a transactional relationship precludes a great many other possibilities. If I want the relationship to be one of care, I will actually have to not do it transactionally, because otherwise I end up in this conflict and no amount of goodwill gets me out of it.

> **On care, transaction, and power**
>
> Care, transaction, and power are the three basic modes in which one person's agency meets another's: care operates by relationship and holds the other's flourishing as its own, transaction operates by exchange at a specified distance, and power operates by causation applied to choice. The triad and its consequences (manipulation, weaponization, and the category mistake about values) are developed in *Care, Transaction, and Power* (Chapter One, "The Three Bases of Relation"). This chapter takes up only one of its corollaries: what happens when transaction and care are asked to run along the same wire.

IV. The Dollar and the Friend

There is a test for this which requires no theory at all, and everyone reading has already run it without noticing that it was an experiment. Suppose you are moving house. You call me up and you say: there is more stuff here than I can shift on my own, can you help me lug boxes out to the car? If you ask me to help you move as your friend, I am simply going to do it. I will be there. There is no negotiation, no scheduling conversation of the kind that has a subtext, no assessment. I will move your boxes and at the end of it we will be tired and something between us will be very slightly stronger than it was in the morning.

Now change one thing. At the end of the day you say: here, let me pay you a dollar. And my answer is: absolutely not.

Look carefully at what just happened, because it happens in an instant and it is easy to miss. The moment the dollar is introduced, an evaluation begins. Was that transactionally worth my time? What is the exchange here? And once that evaluation has begun, the price becomes unpayable. If I am actually charging my rates, I do not begin at anything less than several hundred dollars an hour, and for physical labor (which I have no particular wish to do) I would raise the figure further, for the rarity of the thing, and you would look at the number and laugh, and you would be right to. Frankly, you could not afford me.

The point is not that friendship is expensive. The point is the opposite. The value of friendship is of a different order than the value of any amount of money. It is not a large quantity on the same scale. It is not on the scale. And so the introduction of the dollar does not purchase a fraction of what was freely given; it pulls the relation out of the mode in which the giving was possible and toward one in which the giving must be priced.

State it exactly. As soon as transaction is introduced into a relation, the care dimension is pushed out of it. It is not abolished: care cannot be zero, and a friend who is paid may go on wishing you well. But motives do not simply add. They displace. If they added, the paid friend would bring the

care and the wage together; what happens instead is that the reason for carrying the boxes moves from your need to the price, and the effort is now spent in exchange for money. And the warmth that remains is not what decides. Care is not a feeling that accompanies conduct; it is the basis on which conduct is selected. A paid relationship can hold genuine warmth for years, and on the day the interests diverge, the money decides. That is the two masters problem in its plainest form: the care is still there, and it is no longer in charge.

Two questions are serious.

The first is that money is energy. It may perfectly well be the means by which a person takes care of themselves, and to speak of it only in the negative is to miss what it makes possible.

I take that seriously, and it can certainly seem that this is the case. The token is real. It buys real bread. A person without it is not thereby elevated; they are hungry. But I would rather be repaid in care, and there is an old warning that puts the matter better than argument does: when the last of the land has been turned over and every blade of grass spent, we will discover at that moment that we cannot eat money. The token is a proxy. It is an extraordinarily good proxy. It is not the thing.

The second is that there is a way to work through money to create positive impact in the world, and good things have plainly been built through the structures we actually inhabit.

The second clause of that I concede entirely, and without qualification. Good things have been built through these structures. Hospitals have been built. Instruments of great beauty have been built. The book in your hands was produced by people who were paid for producing it. What I do not concede is the first clause: that the working-through is what produced the good. The net result of all of that interaction, taken as a whole and over time, is not the positive impact. The good things were built by people in spite of the dimension the structure removes, and the count of what was built is not the same as the count of what was foregone.

What this establishes is exact: where care is doing the work, a transaction placed across it pushes the care out of the place where it was deciding. It does not say that wages are wrong, or that institutions should run on care; Chapter Sixteen says where each relation belongs. The arm's-length relation belongs to institutions, which must work between strangers, and care belongs to community. A sound arrangement contains both relations, and what decides its character is which of the two has to ask the other's permission. So the argument does not ask anyone to refuse a wage, or to stop paying one. It asks that care not be run through the wage, which the friend with the boxes already knew when he refused the dollar.

V. The Employer's Account

What I have set out so far is an argument. Arguments of this kind are usually met, quite reasonably, with the observation that a sufficiently good employer, with sufficiently good intentions, sufficient skill, and sufficient freedom of action, could surely arrange things otherwise. So let me describe what it would take to test that observation properly, and then let me report the test, because it was run, and I ran it, and it is the reason I hold the position I hold.

To test it you would need to remove every confound. You would need an organization with a single founder, so that there is no committee diffusing the intention. You would need no investors, so that no external party can require a return on a schedule that overrides the founder's purpose. You would need no lender, no board, no acquirer, no regulator standing over the thing and no dependency on any other party's goodwill. You would need the enterprise to be successful enough that its choices were not being made under duress. You would need it to own its own tools and its own premises, so that its capacity to act was not on loan. You would need a workforce not drawn at random but already oriented toward creative process and already experienced in community: people

who wanted the thing you wanted and had practiced it elsewhere. And you would need the effort to be sincere, sustained, and long, with real technique brought to bear and real money spent on bringing it.

Every one of those conditions was met. Sole founder. No investors. A high degree of success. Clear leverage. We had all of our own resources. We had all of our own tools. We were beholden to nobody. It was essentially me, the jurisdiction I happened to be standing in, and God. The people who worked there were, in the main, drawn from a world of creative practice and connected to community in a profound way: not skeptics to be converted, but participants who already understood the vocabulary and wanted the outcome.

And I tried. I tried damned hard to make the relationships I had with my staff, and particularly with the people who worked closest to the top, into something like a true family relationship: a true collaboration, in which what I knew and what they knew met on level ground. I brought in people from outside to teach nonviolent communication, which is a well-developed discipline for separating observation from evaluation and need from demand, and which works. It did not take. I spent a year giving talks, every week, two hours at a stretch, teaching the techniques by which curiosity and enthusiasm are actually developed and sustained in a working group. A year of it. It could not take.

It could not take for one reason, and the reason was not obscure to anybody in the room. No matter how hard I worked at it, they would still treat me as the boss, because ultimately I held the decision as to whether or not they were employed, and everybody knew that. Everybody knew it in the way one knows the weather: not as a thought that is being had, but as the condition under which thoughts are had. It did not need to be invoked. It did not need to be threatened. It never once had to be mentioned for it to organize every conversation in the building.

It is worth saying what the distortion consists of, because from the inside it does not announce itself as a distortion. It consists in the direction that information travels, and in what gets quietly removed from it on the

way. Advice flows downward easily and upward with difficulty. Disagreement is pre-filtered by the person holding it, usually without any conscious decision to filter, because the cost of being wrong in front of the person who decides whether you are employed is not symmetrical with the cost of staying quiet. Nobody lies. People simply round toward what is safe, by increments too small to notice, and the increments compound. And the founder, who has the most authority and therefore receives the most rounded information, becomes over time the least well-informed person in the building about the thing he most needs to know, which is how it is actually going. That is what it means to say that the community of care has been disconnected. It is not that people stop being kind. It is that the channel by which they would have corrected me has been narrowed by the same instrument that pays their rent.

So: with the complete ethics at my back, with the philosophy, with all of the technique, with the resources, with the freedom, with the right people, under what I can only describe as a perfection of circumstances, I still could not erase the transactionalism, nor the distortive effects it had on that community. And I state the ending exactly, because a softened ending would be false. It did not merely fail to improve. It ultimately destroyed the thing I was trying to build.

Consider the shape of that result as evidence. This is a negative result obtained under near-ideal conditions, which is the most informative kind of negative result there is. The failure cannot be attributed to insufficient sincerity: the sincerity was total and sustained across years. It cannot be attributed to insufficient technique: established methods were brought in by people who teach them for a living. It cannot be attributed to insufficient resource: there was money, time, and full control over how both were spent. It cannot be attributed to external constraint: there was none. It cannot be attributed to the wrong people: they were, if anything, the right people, selected for exactly the orientation the effort required. Every variable a defender of the arrangement could reach for was supplied. What remains, as the one condition that every remedy left in place, is the structure of the relation itself: one party held the decision over the other's

subsistence. The claim rests on the structure set out earlier in this chapter: where one party holds the decision over the other's subsistence, no assurance from that party can show that the power is not driving the behavior, and so the care cannot be verified even by the sincere. The case shows that structure at work under the best conditions anyone could arrange, and so the claim does not rest on the case alone.

> **What a controlled negative result does**
>
> What a controlled negative result does is shift the burden. A single case cannot establish a universal law, and I do not claim that it does. The proposition under test was that a sufficiently good employer could remove the distortion from within the employment relation; the strongest available instance of that attempt failed, and failed for a reason internal to the relation rather than incidental to it. Anyone who wishes to maintain the proposition is now obliged to specify which condition was not met, or to name a condition missing from the list above.

VI. The Question from Distributed Ownership

The strongest question put to this argument comes from a founder who did the obvious thing about it, did it wholeheartedly, and reports that it worked. It deserves to be stated at full strength, in his terms rather than mine.

Consider a small firm. Its founder brought to it formal training in business, a graduate degree in design, professional experience, and a substantial sum of his own capital to run the thing and win the first work. Every conventional justification for a controlling stake was his. And he split ownership equally among everyone, so that each person held the same share he did. His stated purpose was precise: that everyone should feel collective ownership, so that while he might still make decisions, he was not there to be treated as a boss but as a family member, with all of them present to bring about a collective vision. He reports that this is how it

went, and he asks that his direct experience be weighed alongside the argument rather than beneath it. That is a fair request and I grant it. The report is not naive. It is not the account of someone who failed to notice the problem; it is the account of someone who noticed it and paid for the remedy out of his own pocket.

The reply is not a refutation. It is a question.

So I ask this founder: if you were suddenly to drop dead tomorrow, would you predict that the organization you have built would continue along the same path of your intent, without you?

Answered candidly, it is a strong condition, and a strong condition even of the best organization its founder knows. That candid answer is what makes the case worth setting out.

One step bears on the claim of this chapter directly. Equal ownership removes the employer's power over each member's subsistence inside the firm. The firm's members still draw their subsistence through exchange with customers, and the customers are the party that could stop paying. The two masters problem has moved from the line between owner and employee to the boundary between the firm and its market; it has not gone. The second of the three orderings, set out at the end of this chapter, makes the same step in general: when good work is done and money arrives as its byproduct, the provision still arrives in the form of a token, through an exchange, so that at the boundary of the good work there is still a party who is paying and who could stop paying. Over a long enough period, the master that can terminate you is the master you serve. The question of succession bears on a different matter: whether coherence of intent survives the founder.

The succession question can be pressed further, as a challenge rather than a claim, which matters. Can a single example be produced, anywhere in the historical record, where a founder departed and the coherence of the process did not destabilize? And destabilize has two precise senses, so

that the challenge cannot be met by equivocation. Either the structure goes hierarchical and remains hierarchical thereafter; or strong power inequities emerge within it. One example. Not a class. One.

A candidate can be offered, and a reasonable one: a spiritual fellowship organized around a teaching of self-realization, whose founder is long dead and whose vision has demonstrably continued.

The answer is twofold. First, that religion and community are essentially isomorphs of one another (they are the same structure seen under different descriptions), and that such an order should not be treated as a business per se, so that offering it is close to changing the subject. But second, and more to the point: there are places where such orders do overlap with commerce, and those are exactly the places where power and inequity show up. The fellowship named has aspects of that.

It can be answered that these are physical places. They occupy land, they maintain buildings, they must have a purpose and a means in order to continue to run in this world at all. The transactional element is not a lapse; it is the price of existing.

The answer is that this is an artifact of context: the context being the surrounding system, which is transactional throughout, and to which any enclave inside it must to some degree react. Monasteries have historically been able to sustain themselves without transactionalism as part of their internal function. So to the degree that transactionalism is embedded in such an order, it is a reaction to the environment rather than an expression of the form, which does not rescue it as an example. If anything it makes it a poor one, since what it demonstrates is the reach of the surrounding system rather than the viability of the alternative.

A reader who knows business will think first of a firm: a worker cooperative, or a company held in trust for its employees, that outlived its founder. The test is the same one, put to whatever firm the reader names. After its founder left, did the firm go hierarchical and remain hierarchical thereafter, or did strong power inequities emerge within it? And beneath both of those: did its members keep the capacity to keep asking their own

questions, and to act on the answers? Those are the questions to put to any firm offered as the one example, and the note below says why the last of them is where the load falls.

> **Where the load falls**
>
> What the case establishes is where the load falls: not on whether equal shares can be issued, which plainly they can, but on whether coherence of intent survives the departure of the person who held it. Why it falls there can be said exactly. A founder holds four things at once, bound so tightly that they look like one: title (the name on the deed), authority (the standing to make binding choices), competence (the capacity to make those choices well), and legitimacy (the acceptance, by the people affected, that the choices are ones they will live inside rather than resist). Equal shares move title, and move it fast. Coherence of intent rests on competence, which moves only by being exercised, and on legitimacy, which accrues to a successor from the experience people have of being heard by it and cannot be granted by the one who leaves. What can survive a founder is a community's capacity to keep asking its own questions and to act on the answers. The load falls there, and the challenge stands as it is put: one example.

VII. Why Better Incentives Do Not Reach It

The next question is the one an engineer raises first.

It says that what has been described is a fault in the incentive structure sitting behind these arrangements; the dynamics complained of are the output of a badly specified system rather than of transaction as such. Design the incentives better and the distortions go away.

The reason this does not reach the problem can be stated as a chain, and each link in it is ordinary. Anywhere you have a causal feedback system, which is any arrangement in which an outcome is measured and the measurement is fed back to adjust behavior, you are going to have some

algorithm working on fixed points (the settings at which the measured quantity stops improving). Somebody is going to construct the equivalent of a gradient descent: find the direction in which the measured quantity improves, and move in it. This is not a moral failing. It is what intelligent agents do when placed in front of a scored surface.

Any form of optimization over such a surface yields a surplus to whoever performs the optimization, and surplus captured is profit in the general rather than the accounting sense. Any form of profit is a vehicle of power, because a surplus is precisely a capacity to act which others in the system do not have. And any concentration of power expresses itself as inequity, the magnitude depending only on the creativity of the people who devised the optimization. Somebody will always be creative enough. That is not a pessimistic assumption about human nature; it is an observation about large numbers.

So the conclusion is not a forecast. So long as the field has that feedback mechanism built into it, so long as it is causal in these particular senses and not mediated by consciousness and choice, the outcome is guaranteed. I am not speculating here. The same observation is available from another direction in the well-known principle that a measure targeted ceases to be a measure: the moment a quantity becomes the objective, the relation between the quantity and the good it stood for comes apart, and it comes apart because somebody optimized.

This is why redesigning the incentives does not reach the problem. A redesigned incentive structure is still an incentive structure. It is still a causal feedback field, and therefore still a surface over which optimization runs. The redesign changes which behaviors are rewarded; it does not change the fact that the reward is what constitutes the relation. You can move the fixed points. You cannot get rid of the fact that there are fixed points and that someone will find them.

What would be genuinely different is a relation not mediated by causation at all but mediated by consciousness and choice: a relation in which what I do for you is not the output of a function whose input is what you

have arranged to be worth my while. That is what care is, and it is the reason care cannot be engineered into a transactional structure by any refinement of the transaction. It is not a better setting of the same dial. It is a different instrument.

VIII. Three Orderings

A person can lose her way inside a company where money is the goal and doing good is the byproduct, and realize, somewhere in the middle of it, that this is simply not how it works: the universe is not conspiring in her favor, because the goal is in the wrong place. The formulation that follows is that money ought to be the byproduct of doing something good, or of being in service.

That formulation is a genuine advance, and it should have its due before I go past it, because there are three orderings here and not two, and the distinction between the second and the third is where community actually lives.

The first ordering is the pursuit of money for its own sake. This is the one everybody can already see through, which is part of why it survives: it absorbs all the available criticism and leaves the second ordering unexamined.

The second is doing the good work and having money arrive as the byproduct. This is a real improvement and I do not want to be seen to sneer at it. The person in this ordering is not chasing the token. The work is the point; the token follows. Much of the best work now being done in the world is done by people in exactly this posture, and it is decent and it is sound. But notice what has not changed. The provision still arrives in the form of a token, which means it still arrives through an exchange, which means that at the boundary of the good work there is still a party who is paying and who could stop paying. The care that the dollar

pushed out, in the case of the friend and the boxes, is still pushed out, and the party who could stop paying is still the one who decides. It has simply been moved to the edge of the picture, where it is easier not to look at.

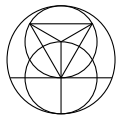
The third ordering is different in kind. It is doing the work and having the sustenance itself be the byproduct. Which is to say: my friends take care of me. I find myself always fed. I find myself always in shelter.

Sit with the grammar of those three sentences, because the grammar is the whole argument. They are not transitive. Nobody is buying anything. There is no token, and there is therefore no party at the boundary who is paying and could stop. The provision does not arrive in exchange for the work; it accompanies the work, because the people among whom the work is done are people to whom the worker matters. Such a person can put it in the vocabulary of dharma: that the more closely she is aligned with her dharma (one's proper work, the thing it is actually one's to do), the more she experiences abundance in a way she never has before, and that she experiences it in community.

Only the third ordering is community. The first is a market. The second is a market with better manners. The third is the thing that markets were built to substitute for, and it is why I think community is worth the enormous trouble it costs, because to some extent that is what happens there, and to my knowledge it does not happen anywhere else.

Which returns us to where this book is going. Technology is a vehicle for the representation of intentionality, and the question that governs everything downstream is whose intentionality it represents. A transactional structure cannot represent the intentionality of a community, for the reason this chapter has established: transaction pushes out the relation by which a community is constituted in the first place, and wherever the two call for different things, the transaction decides. What the structure represents, then, is the intentionality of whoever holds the deciding side of the exchange, not the community's. What remains to be shown is how

thoroughly the structure has been built out around us, and how much of what presents itself as a market turns out, on inspection, to have been constructed. That is the subject of the chapter that follows.



CHAPTER FIVE

The Wall in the City

> *This chapter takes the structure established in the previous one (that introducing transaction into a care relation pushes the care out, and puts the transaction in charge wherever the two call for different things) and applies it at the scale of a whole population. A parable does the work: a city on a plain, a wall through its middle, a gate, and a charge to pass. Every entry on the resulting balance sheet is a genuine economic positive, and every loss is unrecordable. From this the chapter derives its deliverable, a diagnostic question to be asked of any market whatever: was this market real, or was it constructed?*

I. The City on the Plain

Chapter Four set out the structure at its smallest scale. Two people, one task, one relation of care between them; introduce a payment, and the care does not merely diminish: it is pushed out, and the price takes over the deciding. The friend becomes a client. The favor becomes a service. The price, once it exists at all, is unpayable, because what it is being asked to price is of a different order than money. That argument was made about a pair of people. The question this chapter takes up is what the same structure does when it is applied not to a pair but to everyone at once. The

structure here takes the form of a wall, and a wall is a made thing: it goes on carrying the intentionality of whoever built it long after anyone remembers that it was built.

To answer it I change metaphors entirely, and describe a scenario instead.

You have a city on a plain. It is a great flat expanse in every direction, with no river through it, no ridge, no marsh, nothing that would suggest to anyone where a boundary ought to run. The city grew the way cities on flat ground grow: outward from a market in the middle, in all directions at once, the paths worn into it by feet rather than laid out by decree. A person on the east side has kin on the west side. The tanners are north of the well and the dyers south of it, and they carry work back and forth between them a dozen times a day without once thinking of it as a journey. Everyone in the city is, in the ordinary course of a life, within reach of everyone else. That reach is not a feature anyone notices. It is simply what it is to live there.

Now an enterprising fellow comes forward with a proposal. He wants to build a wall, right through the middle of the city, north to south, and he makes the case for it in the terms that always carry. Look at how many jobs it will create. Look at the men who will quarry the stone and the men who will haul it and the men who will set it. And there will be a gate in the middle (he is not proposing to cut the city in half, he says, he would never propose that); there will be a gate, and people will pass through the gate, and we will charge a small sum for passage.

Some people complain a little. He has an answer for them, and the answer is not a trick; it is, on its own terms, entirely sound. This will be great, he says. It will create revenue. It will create employment. It will create a marketplace: we will now have a market of people who want to get to the other side. And the authority of the city agrees, because from where the authority sits the figures are unarguable. For the period of construc-

tion, this many people are employed. After construction, you get to employ guards. You get to employ gate-keepers. You get to let the maintenance contracts, and stone being what it is, those contracts do not end.

So the wall goes up, and the gate is set in the middle of it, and a man stands at the gate and takes the fee. That is the whole of the scenario. Nothing in it is exaggerated, and nothing in it depends on anyone behaving badly. The proposer may be sincere. The authority may be diligent. The gate-keeper may be kind to the elderly. The question is what has actually happened.

II. What the Ledger Records

Begin with what is true, because what is true here is the difficult part. Every entry the wall generates is a real entry. The wages paid to the builders are real wages, paid to people who needed work and now have it, and spent in the city on bread and cloth and rent, which is itself further real activity by anyone's accounting. The stone was bought from a quarry that now expands. The guards draw salaries. The gate-keepers draw salaries. The maintenance contracts recur annually and are, in the language of enterprise, a dependable revenue stream of exactly the kind that a prudent person seeks. The tolls come in daily, in small sums, from a very large number of people.

And around all of this a further economy grows, none of it planned by anyone, all of it genuine. Someone begins selling food to the people standing in the queue. Someone else offers, for a fee, to stand in the queue on your behalf. A faster passage is introduced at a higher price for those who can pay it, and it is popular, and it is a service people are grateful for. Documents are forged; inspectors are hired to detect the forgeries; the inspectors are trained by a man who has made a profession of training them. Every one of these is employment. Every one of these is income. Every one of these appears, correctly, as a positive on the standard measure of the city's total product.

Notice in particular the character of the market that has been brought into being, because it is the finest market anyone in that city has ever seen. Its customers are the entire population. Its product is something people want very badly indeed: not as a matter of taste or fashion, but as a matter of getting to their mother, their work, their physician, their friend. Demand does not fall when the price rises, because the mother is still on the other side. The customer cannot decline to purchase and cannot purchase elsewhere, there being one gate. No merchant in history has ever had a better position, and it was created out of nothing, on flat ground, in a place where there had never been an obstacle at all.

So the question is not whether the wall creates value in the sense that the ledger uses that word. It plainly does. The ledger is not lying. The ledger is doing precisely what it was built to do, faithfully and without error, and that is the whole of the problem.

III. The Column With No Line For It

The thing that is not on the balance sheet is all of the opportunity cost. Take it concretely, one instance at a time, because in the aggregate it becomes an abstraction and the force of it is in the particulars.

This man is not going to meet this woman, because they happen to live on opposite sides of the wall. That is the whole of it. There is no drama in it. He does not know she exists; she does not know he exists; neither of them will ever know that anything was lost, and there is no moment at which the loss occurs that anyone could point to and record. Before the wall they would have crossed paths at the well, or at a festival, or because his cousin knew her brother, and the ordinary density of the place would have done the rest.

Consider what the queue does to a life, because the queue is where the real reshaping happens. Say you have to wait about an hour to pass. Naturally you want to minimize the amount of time you spend doing that, so you begin to arrange your affairs around it. You batch your errands. You

cross once a week rather than six times, and then once every two weeks, and the crossing becomes an expedition rather than a step out of the door. Within a year you have reorganized your entire life so that everything you need is on your own side, and you experience this as good management. It is good management. It is also the point at which the other half of the city stops being part of your world.

And that takes away all the possibility of bumping into someone with whom you would have had a deep relationship, which you now do not, because they are on the wrong side of the wall and you happen to live over here. Multiply that by the population, across a generation. The apprenticeship that never forms because the boy's talent and the old man's trade are three streets and one gate apart. The two craftsmen whose skills would have combined into something neither could have reached alone. The quarrel between two families that would have been settled by proximity and is instead nursed into a grievance. The physician reached in time, and now not reached in time. None of this is speculative. It is what a city is: a city is nothing other than the density of possible encounter. Halve the encounter and you do not have half a city. You have two towns.

None of the stuff about what could have happened ever shows up in the space of the accounting of what has happened. That sentence is worth sitting with, because it is not a complaint about sloppy bookkeeping. It is a statement about what a ledger is. A ledger records transactions, and a transaction that did not occur is not a small transaction; it is not an entry at all. There is no procedure, however careful, by which the encounter that would have happened could be entered as a number in the column that decides whether the wall was a good idea. The losses are not merely unmeasured. They are of a kind that the instrument cannot register, and the instrument will therefore return the same verdict no matter how large they grow.

So the account comes out like this. On one side, a long list of genuine positives, each defensible, each attached to a real person with a real wage. On the other side, an enormous subtraction taken from every inhabitant of the city at once, with no line for it, no unit for it, and no advocate for it, because the people who bear it cannot name what they have lost.

> **On the choice of parable**
>
> A parable earns its keep by being unambiguous, and this one is chosen for exactly that. The wall is a clean case: flat ground, no prior obstacle, one beneficiary, and a loss that is total and invisible at once. Real instances are almost never that clean, and I do not want the clarity of the parable to be mistaken for the clarity of the cases it is meant to illuminate. What the parable establishes is that the structure exists and that our accounting is blind to it. Whether any given enterprise has that structure is a separate question, and the rest of this chapter is about how to ask it fairly.

IV. Taken, and Sold Back

Something has been taken from you. More precisely, something has been taken from the entire population, and the taking was not a theft in any sense the law recognizes, because nothing that anyone owned changed hands. What was taken was a capacity that everybody held and that nobody held as property: the capacity to walk across their own city.

And then it was sold back. That is the second half of the move, and the two halves must be seen together or neither is visible. The wall creates the absence; the gate sells the relief of the absence. Considered separately, the gate looks like a service, and it is a service, in the narrow sense that a person who arrives at the gate wanting to cross does in fact get across. Considered together with the wall, the gate is the sale of something that was free the week before last.

The third part is subtler than either, and it is the part that lasts. By being the mediator of that transaction, the taker has in a sense convinced you that he is the resource which provides you with something you originally had for free, all by yourself. He is not experienced as an obstacle. He is experienced as a provider. The gate-keeper is the man who gets you to the other side; he is thanked for it; his enterprise is spoken of as the thing that connects the two halves of the city, and in the only sense that now applies, it does. The barrier has been folded so completely into the background that the only visible actor is the one relieving it.

State it in general and it becomes recognizable everywhere. If I can make water scarce and sell you the water, if I can make shelter scarce and sell you the shelter, if I can make love scarce and sell you the love (or anything else you once had and now believe you need to be given), then I have not produced a single thing, and I am nevertheless, on every measure we keep, a producer. The scarcity is my product. The relief of it is my revenue. And my position is secured not by what I supply but by the fact that nothing else supplies it any longer.

From this comes the question this chapter exists to hand over, and it is a question rather than a doctrine, which is the point of it. Of any market, at any scale, ask: **was this market a real thing in the first place, or was it something that was constructed artificially?** I will call a market of the second kind *a constructed market*: one whose demand exists because something was first removed from the very people who now buy it back, so that what the seller has contributed is the scarcity rather than the good. In the case of the wall it is very, very clear that the market was constructed. The claim this book makes, and the test by which any reader can check it, is that a great deal of the present marketplace is of the same kind.

V. The Question of Genuine Intermediation

The strongest question about all of this must be put before it is answered, and it is strong enough that I would rather state it myself than have it stated against me. It runs as follows.

Standing in the middle of a transaction is a real place to stand, and the work done there can be genuinely productive. Consider the merchant who carries grain from a region of surplus to a region of famine. He did not cause the famine. He bore the road, the risk of the road, the spoilage, the capital tied up in a cargo that may not arrive. The people at the far end had no grain; now they have grain; and the fact that he took a margin for it does not make the grain less real. Consider the translator between two parties who share no language. Before he arrives they cannot deal with each other at all; after he arrives they can. He did not remove a common tongue in order to sell one back. Consider the man who builds a bridge over a river he did not dig. Consider a common system of weights and measures, a postal service, a road, a harbor, an index that tells you where a thing may be found. The cost of finding one another is a real cost, and lowering it is real production.

Press the question further, because it goes further. Nearly everything that made cities possible in the first place was a reduction of the friction between people trying to reach one another. Someone who condemns intermediation as such condemns the very thing that made the plain worth settling. And the parable, the question concludes, is rigged. The wall was selected precisely because it is obviously destructive. Nothing follows from it about intermediaries in general, because a case chosen for its clarity is a case chosen for its extremity, and the real question (whether a given enterprise produces or captures) has not been touched.

That point is correct, and its correctness is the reason the deliverable of this chapter is a question rather than an accusation. I am not arguing that intermediation is extraction. Intermediation is sometimes among the most valuable work there is. What I am arguing is that we have no habit of asking which kind we are looking at, and that the accounting we rely on can-

not tell us, for the reason set out above: it records what happened and is structurally silent about what was prevented. So the answer to the question is to make the test operational, and then to apply it to the very examples the question offers and let it return whatever it returns.

The test has three forms, and they are three views of one thing. Ask first what would obtain if the intermediary simply withdrew. Remove the bridge-builder and the river is still uncrossed; the capacity he supplied was not there before him, and it leaves with him. Remove the wall-keeper and the city is whole again within a day, because the capacity was never his; it was the population's, and he was holding it. That is the difference, and it is not a matter of degree.

Ask second in which direction the intermediary's interest runs with respect to difficulty. The bridge-builder earns by crossings and therefore wants crossing easier, cheaper, and more frequent; his prosperity and the population's convenience move together. The wall-keeper earns by the wall and therefore wants the wall higher, the alternatives fewer, and the queue no shorter than his revenue requires. Where an enterprise's income rises as the friction rises, it is at best in a position of standing conflict with the people it serves, whatever the intentions of anyone inside it.

Ask third whether the thing being sold was already possessed. The grain was not possessed; the famine was not the merchant's doing; he passes. The shared language was never there; the translator passes. But where a population plainly held a capacity freely, and now pays for it, the burden falls on the seller to say what he added other than the barrier. That burden can be discharged. Sometimes the true answer is that the old free capacity was poor, and what is now sold is better, and people would choose it even if the free thing were restored tomorrow. That is a real answer, and it should be accepted when it is true. The test is a test. It comes out either way.

Here is the test worked once on a function a present reader uses. Take a ticketing intermediary standing between a theater or a concert hall and the people who want a seat in it. Venues of this kind once sold their own

tickets, at a window by the door and by telephone, and many now sell them only through the intermediary, whose contract with the venue forbids any other channel. The intermediary charges the buyer a fee on each ticket, and in many cases it also runs the market on which tickets already sold are sold again.

Ask first what would obtain if the intermediary withdrew. Two things it supplies would leave with it, and they were not there before it: one place in which to search the seats of many venues at once, and a seat map that shows what is still free. Those pass. The venue's capacity to sell a seat to the person who wants it would not leave with the intermediary. It was the venue's and the audience's before the contract, and the contract is what now holds it. That function fails.

Ask second in which direction the intermediary's interest runs with respect to difficulty. On the search, it earns when people find seats easily, and its interest runs with theirs. On the resale market, it earns a fee each time a ticket changes hands, and a ticket changes hands more often, at a higher price, when seats at the venue's own price are hard to get. There its income rises as the friction rises, and it stands in the conflict described above with the people it serves.

Ask third whether the thing being sold was already possessed. The venue and its audience held the capacity to deal with each other directly, at a window, before the contract closed the window. So the burden falls on the intermediary to say what the fee adds beyond the closed window. For the search and the seat map it has an answer, and the answer should be accepted. For the exclusive contract and the resale of seats it controls, it has no answer except the barrier.

So the test returns a verdict function by function, which is what it was built to do: the search passes, and the exclusive gate and the resale market fail. A verdict on the firm as a whole would be too coarse to be true, because one firm can pass on one of its activities and fail on another.

The unit the test applies to

The question is asked of a function, not of an organization, and this matters in practice more than any other qualification I can offer. The same enterprise can pass on one activity and fail on another: genuinely lowering the cost of finding what you need, while also holding a gate on something you formerly reached without it. A verdict on the whole firm is almost always too coarse to be true. The claim of this book is a claim about proportion, not a claim that all intermediation is capture: the share of the present marketplace which would fail this test, if the test were actually applied, is far larger than we suppose, and we do not know how large, because we have never made a practice of asking.

VI. The Decisive Move Is Not the Pricing

The toll is the visible part of the wall and the least important part of it. A charge can be lowered, waived, subsidized for the poor, abolished by a decree on a Tuesday. None of that touches the structure, because the structure is the wall, and the wall stands on something other than the toll.

What holds it up is that after a while nobody remembers the city without it. The first generation grumbles; the second generation queues; the third generation regards the wall as geography. The charge becomes the cost of living, which is to say it becomes weather, which is to say it stops being a decision that anyone made. And a person who proposes taking the wall down is no longer met with an argument. He is met with incomprehension: how else would one cross? The question is sincere. That sincerity is the achievement.

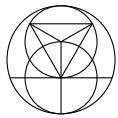
Here is the point at which the whole thing closes over. Once you become convinced that this is the right way to be, and beyond that, that it is the only way to be, the game is over. You have lost, and you have lost in a manner that leaves nothing to contest, because the alternative is no longer

among the things you can think. The persuasion, not the pricing, is the decisive move. Everything before it can be undone by ordinary means; after it, ordinary means no longer find anything to act on.

> **On persuasion of this kind**
>
> The distinction between offering someone a reason and arranging the conditions under which only one conclusion remains available to them is developed in *Dialogues on Ethics and Governance* (Chapter Four), where manipulation is treated as causation applied to choice. What the parable adds is one observation: the arrangement need not be anyone's intention, and a wall standing for three generations persuades without a persuader.

So the question to put to the enterprising fellow, and to ourselves, is not whether his figures add up. They do. It is whether this is a thing worth advocating for. Do you really want to argue for that? Is that the way to be in the world? I do not think so, and the reason I do not think so is not that the wall makes anyone poorer in the sense the ledger tracks, because it does not. It is that a city on a plain, with everyone within reach of everyone else, was a *commons* (something held by all and therefore assigned to none), and the wall converted a commons into a holding without anyone ever having to claim to own it.

Which raises the question the next chapter takes up. The wall was legitimated by an idea we have not yet examined, and it is the idea that someone may own a thing and thereby settle, without conversation, who may cross it and on what terms. That idea is a *heuristic*: a rule that stands in the place of judgment, so that a question need not be worked through afresh each time it arises. It was once a good one. The case that it is no longer a good one, and what stands in its place, is the subject of Chapter Six.



CHAPTER SIX

Ownership and the Commons

This chapter establishes that ownership is a heuristic, and that what it is a heuristic for is a conversation: to say that a thing is mine is to say that we need not discuss who may make which choices about it, when, where, why, or how. That was once a good trade and is no longer, because the conversation the heuristic displaced was the means by which the wisdom a choice required got developed, and the scale of our impact now requires that wisdom. Distributing ownership distributes the conversation, which is a beginning; carried far enough it ceases to be ownership at all and becomes the commons, whose corresponding relation is stewardship. The chapter closes on the question of what a person receives back from such an arrangement, and answers it with the gift.

I. The Reference Table

The question comes from the practical end, which is where the good ones usually come from. Suppose a group of people come together in honor of their connection to themselves and out of a sense of how they wish to be of service (that being the primary reason they are collaborating at all, rather than the return). Suppose money is present in the arrangement, but present as a support: a resource by which they are able to show up and

take care of themselves in the world. And suppose that what they hold, they hold together: shared ownership, rather than one owner and a payroll. Is that not the shape of the answer?

It is a great deal better than the alternative, and the instinct behind it is sound. But it does not go far enough, and the reason it does not go far enough is worth working out slowly, because it turns on what ownership actually is. Ownership is the old answer to the book's one question: whose intentionality does this thing carry? The owner's, without discussion. Ownership is not a fact about a thing. Nothing in the timber of a chair, or in the soil of a field, records who owns it. Ownership is a fact about us: about how we have agreed to settle a certain class of question without settling it.

Here is the structure. When I say *I own this* (and I do own things; this is not a position argued from outside the arrangement), what I am saying is that we do not have to have a conversation about who gets to make what choices with respect to what, when, where, why, or how. If a question falls into that class, it is already answered. You look up the owner in the reference table, and that person makes the choice about that thing, regardless. Regardless of who else is affected. Regardless of who else knows more. Regardless of what the choice will do downstream, and regardless of whether anyone downstream was ever asked.

So ownership is a *heuristic*, by which I mean a rule that substitutes for judgment, a shortcut that returns an answer at a cost enormously lower than the cost of working the answer out. Heuristics are not lies. A good heuristic is a genuine economy; it gives you an answer that is right often enough, fast enough, that you come out ahead on the whole. That is the whole case for ownership, and it is a real case. But notice precisely what the economy consists in. Ownership is not a shortcut for producing a *better* choice. It is a shortcut for producing a *faster* one, and the thing it is faster than is the conversation. Ownership replaces the conversation. That is not an incidental effect of the institution. That is the institution.

Here a reader may say, and with good reason, that ownership already *is* a form of stewardship: that to own a thing has always carried a duty toward it, that the farmer who owns the field is understood to be answerable for the field, and that the arrangement therefore already contains what it is here being accused of omitting. That is a different distinction from the one I am drawing, and it is also a true one, and I will come back to it, because it turns out to be where the argument arrives. For the moment I hold it apart. The claim on the table is narrower and more mechanical: whatever duties an owner may feel, the function ownership performs in a decision procedure is to close a question that would otherwise have been discussed. Hold that, and the rest follows.

II. When It Was a Good Trade

The heuristic earned its place. This deserves saying plainly and without irony, because an argument that treats ownership as merely a theft will not survive contact with the history, and does not need to. There was a long period in which ownership was an excellent trade, and it is worth being exact about why, since the reason it was a good trade is the same reason it has stopped being one.

Consider the situation of a person in a world not yet described. Life is a great blooming, buzzing confusion. There is an enormous amount to deal with and very little of it is legible. The weather is not forecast. The soil is not surveyed. The consequences of anything you do disappear over a horizon about a season deep. In that condition, sitting down to deliberate the abstractions (whether this is going to turn out a good idea or a bad one, what it will mean in fifty years, who else it touches) does not much improve the outcome. The deliberation is expensive: it costs the time of everyone in it, and time is the one thing in shortest supply. And it is not very predictive: the world is too noisy for the conclusions to hold. So the trade is obvious. The heuristic was cheap and the conversation was dear,

and the conversation, when you had it, did not buy you much. Under those conditions, looking up the owner in the table and getting on with the work is not a failure of care. It is good sense.

I grant this as fully as it can be granted, because the shape of the argument depends on it. A great deal of what we now find intolerable in the old arrangements was, in its own conditions, adaptive. Institutions are answers to questions, and when we find an institution absurd it is usually because we no longer have the question. The right posture toward ownership is not indignation. It is the recognition that a bargain was struck under certain conditions, that the conditions were real, and that they have changed.

But now attend to what was on the other side of the bargain. The conversation that ownership displaced was not merely a procedure for allocating a decision. It had a second function, and the second function was the important one. The conversation was where the wisdom the choice required got developed. People talking about whether the tree should go there, and which tree, and what happened the last time somebody put one there, are not only deciding. They are assembling, between them, an understanding that none of them arrived with. The talk is the instrument by which a group comes to know something that no member of it knew. Take the conversation away and you have not merely saved time. You have removed the organ.

For a long while this did not show. An organ you are not using does not announce its absence. The wisdom that the conversation would have developed was wisdom about consequences that mostly did not arrive, or arrived so diffusely and so late that no one connected them to the choice. The bill for the trade was deferred. It is now due, and the rest of this chapter is about why.

III. The Tree in the Woods

Take a case small enough to hold in the hand. I am going to plant a tree out in the woods. By any ordinary account this is among the most benign acts available to a person. It is not extraction. It is not construction. It is the smallest possible intervention in the least contested possible place, and it is exactly the sort of choice the heuristic handles without comment: the land has an owner, the owner wants a tree, the tree goes in.

Now look at what has actually been initiated. The number of interactions that tree will have with its environment, or could potentially have with its environment, is enormous, and they will happen over generations. Which species is it? A species native to that valley behaves differently from one carried in, and the difference does not appear for twenty years. What is beneath it? Root systems find water tables, drains, foundations, and each other. What is above it? In sixty years its canopy is over something, and whatever grows beneath that canopy now will not be growing there then. What eats it, and what eats the thing that eats it? A tree is an invitation to a food web, and the web that accepts the invitation is not the one you had in mind. What happens downhill? A tree is a hydrological object: it moves water, holds a bank, or fails to. What does it do to fire? What does it do when it falls, and it will fall, and it will fall onto something? And what does the stand look like in a hundred years, when this tree is not one tree but the parent of forty?

No single person is going to have the expertise necessary to make a wise choice here. That is not a remark about the average person's ignorance. It is a remark about the structure of the knowledge required. The soils are one discipline. The hydrology is another. The entomology is a third, and the person who knows what a beetle will do to that species is almost certainly not the person who knows the water table. The history of that particular valley (what was tried there in 1930 and what came of it) is not in any discipline at all; it is in somebody's memory, and quite possibly

in the memory of a person who has never been consulted about anything. The choice is not hard because it is large. It is hard because the knowledge it requires is distributed across people who have never met.

So if we are looking to make a good choice, what we need is bandwidth. We need a great deal more involvement in the choice-making process, and we need it from many more directions: not more opinion, which is cheap, but more genuinely different vantages, each of which sees a part of the situation that the others structurally cannot. And here a principle applies whose full development belongs elsewhere: to see from more than one perspective is to add a dimension to perception. It is not that more perspectives give you more information about the same picture. It is that they give you a picture with an additional axis in it.

> **On the added dimension**
>
> The claim that seeing from more than one perspective adds a dimension to perception, and the distinction between trusting a view and calibrating against it, are developed in *The Reach of Trust*, in its chapter "Trust and Calibration". What this chapter takes from it is only the consequence: a choice informed from several independent vantages is not a better-supported choice of the same kind, but a choice of a different kind.

The heuristic of ownership has replaced that conversation. Not delayed it, not abridged it: replaced it. The conversation that would have happened, that needed to have happened, in order to inform the choice from multiple perspectives and give it enough insight to be worth making, does not occur, because the reference table has already returned an answer and the answer is procedurally complete. Nothing in the arrangement is broken. The owner acts entirely within his rights. The tree goes in, and in ninety years something is the case about which nobody was ever asked.

IV. The Largest Animal on the Planet

Together with the animals we keep, we now outweigh all the wild mammals on earth many times over; wild land mammals alone weigh less than a tenth of what humanity weighs. That fact is the hinge of this chapter. The amount of impact we have requires a level of wisdom that, frankly, we could get away with not having before, and cannot now.

The point is not that people have become worse. There is no need for that claim and I do not make it. The heuristic performs today exactly as it always performed; its accuracy has not declined. What has changed is the size of the error it lets through. A heuristic is a bargain struck against a certain distribution of consequences, and when the distribution changes, the bargain silently changes with it, without anyone renegotiating. The same rule that was prudent when a bad choice ruined a field is reckless when a bad choice made in the same manner, by the same rule, with the same amount of thought, ruins a watershed.

And the scaling is not gentle. Three things moved together. Our numbers went up. The reach of any single decision went up faster, because the decision is now executed by machinery and propagated through supply chains that touch places the decider has never seen. And the time-depth of the consequence went up faster still, because the substances and the structures we now put into the world outlast the institutions that authorized them. Meanwhile the deliberative capacity attached to the decision did not move at all. It is still one person, or one board, consulting one table.

So the question can be put in its final form. Is the heuristic of ownership (functioning as a proxy for, and a replacement of, the kinds of conversation that would develop the wisdom a choice of this magnitude requires) actually a good idea? At this point, it simply is not. Not because it is wicked, and not because those who hold property are worse than those who do not, but because the trade it embodies no longer pays. It buys speed we do not need with wisdom we can no longer do without.

The census behind the biomass figures

The figures for mammals come from the 2023 census of the biomass of wild mammals (see the Notes): wild land mammals weigh about 20 million tonnes in all, the mammals we keep weigh about 630 million tonnes, and the wild land mammals together weigh less than a tenth of what humanity weighs. Different accountings (wild terrestrial vertebrates only, mammals including livestock, all animal life including marine) give different numbers, and each answers to its own comparison. Counted against all wild vertebrates, fish included, the aggregate mass and throughput of humans together with the animals we maintain is now of the same order as, or greater than, that of the wild vertebrate world it displaced. The claim is used here as an order-of-magnitude fact about human ecological weight, and nothing in the argument turns on the exact figure. The point about impact is a point about scale and not about a statistic, and either comparison carries it.

V. From Distribution to the Commons

Let us distribute the ownership. This is the natural next move, and it was the move the original question proposed: hold the thing together rather than singly, so that the choices about it are made together. Now we have the conversation distributed in a larger way. That is a good start, and it is a real improvement and not a cosmetic one. Where ownership is shared, the reference table no longer returns a single name, and a question that would have been closed has to be opened. The conversation comes back, in proportion to how widely the holding was spread.

But follow the move out and something happens to it. If what I am actually after is the conversations (if the holding was only ever the mechanism and the conversation was the point), then I have to ask what context those conversations occur in. Is it a conversation about this thing? About that thing? It cannot be. The tree is connected to the water, and the water to the food, and the food to the land, and the land to the people who live

on it. A conversation competent to the tree is already a conversation about the watershed. So the distribution does not stop at the object. We distribute ownership about everything, because somewhere along the way we need wisdom not just about one thing but about all sorts of things, and the wisdom is not separable by object any more than the consequences are.

Now push it to the end. Ownership is held in common, by all these groups of people, across everything. At that point you are not talking about ownership anymore. You have kept the word and lost the referent. Ownership was a device for closing a question by naming a person; if no question is closed by naming a person, whatever remains is not that device. What you are talking about is *the commons*: that which is held by all and therefore assigned to none, and which is not a very large piece of property but a different category of thing entirely.

And here the point a reader raised earlier comes back, and comes back vindicated. The relation appropriate to the commons is not possession. It is *stewardship*: the relation in which one holds a thing on behalf of what one does not own and cannot own, answerable to it, and answerable for handing it on. That is the distinction that point was reaching for, and it was reaching for it correctly; I set it aside earlier only because it arrives with its full weight here rather than at the start. Ownership, at its best, borrows the manners of stewardship. The commons requires the thing itself. And the difference is not sentiment. The steward has to have the conversation, because the steward's warrant does not come from a table. It comes from those on whose behalf he holds, and they include people not yet born, which is precisely why the wisdom has to be developed and cannot be looked up.

Stewardship can be placed exactly among the kinds of agreement. A constitutional agreement binds a group and an individual. A contractual agreement binds two parties as peers, and presupposes that no irreconcilable difference of power stands between them. A fiduciary agreement governs the case where such a difference does stand and cannot be removed, as between a surgeon and a patient under anesthetic, who has in-

terests in the room and no capacity whatever to represent them. Stewardship is fiduciary, and the watershed and the people not yet born are the patient under anesthetic. It also differs from both owning and managing: an owner may dispose of the thing, a manager optimizes someone else's objective, and a steward answers to the thing itself and to those who cannot yet speak for it.

Notice what has happened to the question we started from. We began by asking how a group of people might hold something together well. We have arrived at a question about the true dimension and state of an entire conversation among multiple parties: what it includes, who is in it, and what makes it possible for them to be in it. That is a different question, and a harder one, and the remainder of the chapter is about its preconditions.

VI. Shelter, Food, Water, Air, Communication, Medicine

At that point I am not looking for money at all. This is worth stating flatly, because the original question had money in it (present as a support, honorably conceived), and the argument has now walked past it. What money is, and what it costs to use it as the medium of a relationship, is set out in the chapter on money; here I need only the conclusion. Money is a proxy for future choices. The commons does not need a proxy for choices, because the commons is the conversation in which the choices are made.

What I am looking for instead is that the conversation be inclusive about certain things: shelter and food, water and air, communication, and medicine. Those six are not a wish list and they are not a standard of living. They are the enabling conditions of participation, and they belong in the conversation because without them there is no conversation to belong to.

The reason has to be stated correctly, because it is easily mistaken for a different and weaker reason. The point is not abundance for its own sake. It is not that people ought to have more, or that having more makes them

better, or that the good society is the one with the largest heap. I do not care about the money, and I am not arguing for wealth. The point is about what a person can say, and under what conditions they can say it.

Put it as the practical test. Am I in a place that is safe? Can I have this conversation without being shot at? Can I say the kinds of things that will be relevant to this particular situation (the thing about the water table, the thing nobody wants to hear about the species, the thing that will cost somebody money), or do I feel that somebody is going to take away my food or my shelter if I say something wrong? That is the question, and everyone who has ever sat in a room where the answer was *no* knows exactly what it does to the room. It does not produce silence. Silence would at least be legible. It produces a conversation in which everything said is true and the relevant thing is not said, and no one can tell the difference from the transcript.

So where an assurance is created around the basic necessities, something specific becomes possible. It is not that people become generous or that scarcity was making them cruel. It is narrower than that and more mechanical. Where shelter and food are not instruments of leverage, the person who sees the flaw can name the flaw, and the conversation can therefore contain what the situation actually requires. The necessities have been removed from the set of things that can be used to shape what gets said. That is the whole of it.

And that, at last, is the condition under which a person's creative potential is realizable without limit. Not because they became rich, or famous, or possessed of any particular resource; those are not the mechanism and never were. It is because the people around them are actually involved in the conversation. I care about them and they care about me, and what that produces is not comfort but capacity: a surround of attention within which what I am able to do is not bounded by what I am able to defend. Creative potential is not a substance a person contains. It is a relation between a person and a listening, and the listening is the scarce part.

Why leverage is not persuasion

The reason withdrawal of necessities corrupts a conversation rather than merely constraining it (that it is causation applied to choice, and therefore not an argument at all) is developed in *Dialogues on Ethics and Governance* (Chapter Four). The point relied on in this chapter is only that a room in which necessities can be withdrawn is a room whose deliberations cannot be trusted to represent what its members know.

VII. The Gift, and Where the Loop Closes

A question arises here, and it is the best one this argument meets, the one it has been walking toward without knowing it.

It is this: in that realm of creative potential, where does the feedback loop close for you about how your creative potential is realized in the commons? That is to say: in the old arrangement I could at least see what came back. The field was mine, the harvest was mine, the accounting closed and I knew where I stood. You are proposing something in which I put in and the return is diffused across a conversation and across generations. Where, then, do I find out how I did?

The question is fair and it is not answered by dismissing it. Human beings do need to know that their effort meant something; a person who never learns whether any of it landed does not become a saint, he becomes exhausted. And I notice that the arrangement I have described looks, from inside, exactly like the arrangement in which one is used. So the question must be met directly.

Here is the answer, and it starts by examining the instrument rather than the account. If I make it something about feedback, I have given someone else control over me. That is the whole structure in one line. Suppose I make my giving contingent on the return. I have now installed a channel, and whoever operates that channel operates me. To withhold the acknowledgment is to move me. To grant it is to steer me. I have not se-

cured anything; I have built a vehicle by which I can be manipulated, and I have built it myself, out of my own perfectly reasonable need to know that it mattered. The need is legitimate. The mechanism is a handle.

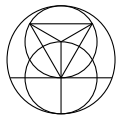
So the answer is not that feedback is bad, nor that one should be indifferent to consequences. It is that a gift is defined by not being contingent. If I am truly giving a gift, at one point or another it has to go beyond me. That is not a moral flourish; it is what the word means. Chapter Two found that a gift is completed by its reception, and that a giver who never meets a reception runs down. That is a fact about how giving is sustained, not a condition on any one gift: the reception completes the gift without being its price. A thing given on condition of return is a trade, and it can be a perfectly honorable trade, but it is a different structure and it produces a different relationship. The gift's whole character is that it exceeds the giver: that it goes somewhere he cannot follow, does something he will not see, and belongs, once given, entirely to the receiving. If it is a true gift, and it is received as that, and it is meaningful in the sense that it becomes something beyond just myself, what more would I ask for? There is nothing else to ask. There is nothing else I could do.

But the question deserves better than a refusal, and there is a real measure to offer in place of the one declined. It is simply pitched at a different interval. Are my children healthy? Do they live a life that enables them to have healthy children in their turn? That is a feedback loop. It is slower than the one the question was reaching for, and it does not resolve within the working life of the person asking, and those are features rather than defects, because a measure that closes inside my lifetime is a measure I can be manipulated by. This one cannot be gamed at the scale on which it is read. You cannot fake a healthy second generation. And it is exactly commensurate with the problem the chapter began with: the tree whose consequences arrive in ninety years is answerable to a measure that reads in ninety years.

There is a time coming when my life, and your life, and our lives, are ended. It will be their turn. Everything we have been calling ownership is, on that horizon, a temporary administrative convenience: a table of names

that will be rewritten, concerning things that will still be there. The commons is what is actually the case; ownership was the notation. And the relation we can have to what is actually the case is not possession, because possession does not survive the possessor.

So we do not take it with us. We give. That is not a consolation offered at the end of an argument about property. It is the argument's conclusion, arrived at by the ordinary route: the heuristic replaced the conversation; the conversation was where wisdom was made; our impact now requires that wisdom; distributing the holding restores the conversation; distributed far enough it is no longer holding at all but stewardship of a commons; stewardship requires that people can speak, which requires that the necessities not be leverage; and what a person receives back for all of this is not an account that closes but a generation that continues.



CHAPTER SEVEN

Why Care Does Not Scale Like Power

This chapter closes Part Two by establishing why care does not scale the way power and transaction do. Care is a capacity grown in a person through attachment, attunement and nurturing care action, each the precondition of the next, and people who never received it in enough measure cannot carry it forward to their own children. The breakup of community, from extended family to nuclear family to mother and child and now machine intermediation even there, is a symptom, not a cause: exchange has become the container in which care is the visitor, and because we as a species do not understand ourselves very well, we do not know to push back against market forces and power forces on the basis of care. Technology is applied causation, so it multiplies power and transaction; relationship can be neither installed nor routed, and both of the things that scale do so by throwing away what care is made of, so care is something humans have to do themselves.

It follows that feedback cannot compel corporations to care, since the time intervals are too long and the indirections too many, and a measure of care would reward its appearance; that small networks are where altruism happens, and the responsiveness of small group process today is orders of magnitude lower than it can be and than it needs to be; and that care combined with hierarchy distorts communication, because the person at the center does not have the bandwidth to handle all of it.

What grows to scale is the number of people who hold care, joined many to many through protocols rather than platforms, with machines that record and carry out what people decide and never stand between people as they talk, so that care upheld in each specific, local, personal interaction emerges in a general, global way across the whole field.

The notion of an institution presupposes that hierarchy and transaction are the be all and end all of solutions to this class of problems; the work of civilization design shows why they cannot supply the bandwidth without corruption, why choice-making cannot be offshored to a machine, and what a governance adequate to the problem must be, and from that it says plainly how AI can help. The chapter hands Part Three its question: if care cannot scale like power, how far short of the radius of our reach does our conscientiousness fall?

I. Care Not Received Cannot Be Carried Forward

Most people are not familiar with developmental psychology (the study of how people grow and form, in mind and in relationship, over a life and especially in childhood). They have never experienced the level of attachment (the bond a child forms with the person who cares for it), attunement (sensing another person's state and responding to it) or nurturing care action (acting on what one senses of another's needs in order to nurture them) that would let them relate to one another in a way they could carry forward in caring for their own children.

These three come in an order, and the order carries the argument. Attachment comes first: the bond itself. Out of attachment, if conditions are sufficient, attunement emerges. Attunement is not the same thing as attachment. A bond can exist without either party being able to sense the other's inner state accurately; attunement is the skill that is layered on top of the bond. From attunement, nurturing care action becomes possible: the actual tending of what has been sensed, soothing what needs soothing, meeting what needs meeting. It is attunement made active, shown over many instances to be answered with benefit. And only from that

third capacity does a fourth reliably form: trust felt in the body rather than held as a belief, the trust that underwrites the sense that other people, and the world, can be relied on at a basic level.

Each capacity is the precondition of the one that follows it. So a deficit early in the sequence does not stay where it began. It is carried forward, and it shows up later as a symptom that looks unrelated to its cause: a guardedness that persists beneath awareness, a difficulty with closeness, a person who can describe care perfectly well and cannot do it.

A parent will rightly say: no parent gets attachment, attunement and care right every time, or even most of the time. If the standard is a perfect childhood, then nobody has received what this section requires, and the claim says nothing about anyone in particular.

The standard is not perfection, and the sequence does not need it. In studies of mothers and infants face to face, the two are fully in step only about twenty to thirty percent of the time, and the repair of a mismatch is a normal and valuable part of the exchange. A child does not learn that care is flawless. A child learns that a break is noticed and mended. What the sequence needs is enough of each capacity, and enough repair, for the next capacity to form. The people this section is about are not people whose parents were imperfect. They are people whose states were too seldom sensed, and whose breaks were too seldom mended, for the capacity that would have formed in them to form.

Now follow the consequence across a generation. A parent can give a child only what was formed in that parent. Attunement cannot be handed over as an instruction, because it is not a piece of knowledge; it is a capacity grown in a body, over years, by being on the receiving end of it. So care passes from one generation to the next by the same slow route by which it was formed, or it does not pass at all. This is the feedback loop that Chapter Six found pitched at the interval of a generation: are my children healthy, and do they live a life that enables them to have healthy children in their turn? Care is read on that interval because care is made on

that interval. Here is the first reason, and the deepest, why care does not scale like power. A deed or a weapon can change hands in an afternoon. Care has to be grown, person by person, at the pace of a life.

II. The Breakup of Community Is a Symptom

Community has broken up into extended family, then nuclear family, then mother and child, and now there is machine intermediation (machines standing between people in their relating, so that the contact passes through a device) even there. That breakup is a symptom, not a cause. A strong factor is that we as a species do not understand ourselves very well. So we do not know to push back against market forces and power forces on the basis of care.

Each step of that breakup has the same shape. First there were large extended communities in which everyone more or less knew one another. They came apart into extended families. The extended family broke down into the nuclear family. The nuclear family broke down into parent and child, and even that bond has been loosened to the point where children do not know the culture of their parents, let alone the culture of the lineage they come from. At each step economic process moved people apart: work took them from one city to another, and friendships did not survive the moving. And everywhere a boundary was drawn between people, an economic opportunity was created at the boundary, which is the wall in the city of Chapter Five, built this time through families rather than across a plain. Carried to its end, this is what I call general disablement: the pattern in which extraction, in any domain, breaks relationships apart until the people involved can no longer sustain their own lives.

The last step, the machine between mother and child, works on the channel itself. Attunement requires sustained, undistracted, embodied presence with another being. When attention is turned again and again toward a device rather than toward the felt state of the person nearby, the ordinary chances for attunement are closed off, one small increment at a

time. A household can be comfortable and well equipped and contain almost no attunement at all, because the channel through which attunement would travel has been quietly taken up by something else. Where this goes on long enough, people lose not only the practice of attunement but the knowledge of what it would feel like to have it, so that even removing the devices would not by itself restore the connection they displaced.

Why call all of this a symptom, not a cause? Because the breakup follows from a change in which relation holds the other. Once, if one goes back far enough, the community was the setting within which commerce happened. Commerce was occasional, bounded by time and place: a marketplace within the village, a trade route crossing between settlements. The rest of life (kinship, obligation, meaning, festival) happened everywhere else, all the time. The container was care, and the visitor was exchange. That relation has inverted. Because of technology and its amplifying effects, commercial process has become the setting in which culture itself now happens. The container is now exchange, and the visitor is care. When exchange is the container, most relationships drift toward transactional terms unless someone actively insists otherwise, and the insisting has to be renewed constantly, because the surrounding pressure runs the other way. The family did not come apart because its members stopped caring. It came apart inside a container that pushes care to the edges and earns at every boundary it draws between people.

This is where the strong factor named above does its work. For most of human history the skills of care were trained by necessity: people who had to rely on one another to survive learned care, compassion, and a deepening awareness of how people hold one another, because there was no other way to live. Technology removed the necessity, and the skills did not train themselves. Most people alive today have not developed the skills that came free with interdependence, and many have never been on the receiving end of them in the measure that Section I described. A person who has never been held in care has no felt reference by which to name what the market is pushing out of their life. That is what it means to say that we do not understand ourselves well enough: we do not know

what we would be defending, so we do not know to defend it, and the market forces and power forces meet no resistance on the basis of care. The skills now have to be taught on purpose, and taught at the pace of real relational development, not at the pace of ambition or of institutional planning.

III. Technology Multiplies Power and Transaction

We often see only the power forces and the transactional forces. We see them because technology multiplies those. Care is something humans have to do themselves.

Start with why technology multiplies those two. Science is the knowledge of causation, and technology is its application, so technology expresses as a vehicle of power: it multiplies the power available to whoever holds it. A cryptocurrency does not work without very powerful chips and very fast computers to compute all of its hashes. Power can scale. Transaction can scale. But care does not scale. It has to live in the heart. It has to live in the human.

The mechanism shows why this is a structural fact and not a lament. Relationship cannot be installed. It can only be accumulated, which means that time is a required input and cannot be traded against any other resource. An organization can hire quickly, fund quickly, and build quickly. It cannot relate quickly, because what it would need is a history, and a history takes exactly as long as it takes. Nor can relationship be routed. Placing an intermediary in the channel does not pass the relationship through the intermediary; it creates two relationships, each accumulating its own history.

Now look at what the other two do. Market coordination scales precisely by making relationship unnecessary. That is its virtue at scale, and it is also its ceiling. Two parties transacting at a price learn what the other will pay. They do not learn what the other values, or why, or what the other would need for a different arrangement to be acceptable; that informa-

tion is not in the channel. Hierarchy scales by discarding information at every level of compression, so that assertion travels down elaborated and correction travels up summarized, which is how the founder of Chapter Four became the least well-informed person in the building. Both of the things that scale do so by throwing away exactly the material that care is made of.

That is why, as Chapter One said in advance, technology amplifies power enormously, transaction moderately, and care not at all. Care was never a causal operation, and there is no quantity in the relation for a technology to amplify. And it is why we often see only the power forces and the transactional forces: they are the forces our instruments enlarge, so they fill the field of view, while care stays the size of a person, which is where it always was.

One distinction keeps this from being read as a claim that every made thing is cold. A tool can carry care. The care of its makers and keepers is carried in its shape, as a woodworker's care is carried in a drawer that runs true for a lifetime. That is the claim this book began with, applied to the good case: technology is a vehicle for the representation of intentionality, and a caring intention can be represented. But the tool holds no care of its own. It cannot notice that the person using it is in trouble, and it cannot choose to stay. Technology does not love the way we do, and this is one of the main reasons we should not trust technology to make our choices for us.

What follows is a task, and Part Three takes its measure. If the power in the world grows with every machine, and care grows only person by person, then the strength of our care has to be increased until it is commensurate with the level of power in the world. That is why I spend less time working out how to increase power, and more time working out how to increase care.

IV. Corporations Cannot Be Made to Care by Feedback

Suppose you built a feedback mechanism to compel corporations to care. Even then, the time intervals are too long, and there are too many indirections (steps that stand between an action and its result, so that the actor does not meet the result directly), for it to work. Companies cannot see and feel the results of their actions directly.

Consider first where feedback does work. In a small community, where everyone knows everyone and reputations form over years of close contact, the person who signals virtue but acts from self-interest will, over time, be found out. The elders know who is trustworthy, because they have watched everyone across a lifetime of dealings. The loop is tight: the one who acts meets the result, and the others meet the one who acted. In such a place the person who actually cares tends to gather trust and influence, because the feedback is accurate.

Now take the first of the two failures, the length of the interval. When the cycle from a cause to its consequence is longer than a human lifetime, no one who makes the choice lives to see what it did, and no one can learn from it in person. A corporation's choices about land, water, soil and the attention of children run on cycles of that length, while its accounts close every quarter. A loop that closes every quarter cannot carry a lesson whose consequence arrives two generations later. Only a culture can carry what is learned across that span, and a platform built for the next quarter's growth cannot hold it.

Then the second, the indirections. Modern life has broken the tight loop of the small community. A person can move from one city to another, from one social group to another, from one professional setting to another, and leave behind any reputation that had formed; the signal of virtue can be refreshed in each new setting without the weight of observed behavior behind it. A corporation has this mobility in the extreme. It meets the people affected by its choices through reports, intermediaries, contracts and subsidiaries, many steps removed, and the people inside it who might

feel a result are seldom the people who decided it. That is what it means to say that companies cannot see and feel the results of their actions directly.

Put the two failures together and the mechanism does something worse than fail. At scale, the one who merely signals care tends to gather trust and influence, because the feedback is broken. A mechanism built to reward care in a corporation would have to measure care, and what it could measure is the signal of care, which costs less to produce than the care itself. So it would select for the appearance of care, and the corporations best at producing the appearance would prosper under it. This is the point of Chapter Four about incentives, in its sharpest case: a measure targeted ceases to be a measure, because any causal feedback field is a surface over which someone will optimize.

Beneath both failures there is a third reason, and it would remain if the first two were mended. To compel is to use causation on a choice. Care is the one relation that cannot be produced that way, since a person compelled to act as though they cared is acting from the compulsion and not from the care. So even a mechanism with short intervals and no indirections would obtain the behavior and not the care. Law and regulation can obtain the behavior, and the behavior has its uses. What they cannot obtain is care, and care is the subject of this chapter.

V. Small Networks Are Where Altruism Happens

Small networks are where altruism happens. There we need higher levels of attunement and discernment within the small networks, for altruism not only to occur but to occur in a generative way. The responsiveness of small group process (how a small group of people senses, responds and acts together) today is orders of magnitude lower than it can be, and lower than it needs to be.

Why there, and not in an institution? Because care needs mutual knowledge, and an institution does not have it. The institution does not really know the individual. Lacking that knowledge, it must act on a rule that does not require it, and the rule available is fairness: treat like cases alike. Fairness is an excellent substitute for knowledge one cannot have, and a poor substitute for knowledge one might have had. Where care is delivered institutionally, the recipients characteristically do not feel cared for, and they are not mistaken. The institutional relation is also conditional in both directions: membership depends on how the person behaves, and what the person receives depends on the will of the leadership. Set against it the unconditional relation, whose clearest ordinary instance is a parent's toward a small child: expensive, thankless, and not withdrawn when the child is a nuisance.

Mutual knowledge has a size. A group held together by direct, felt, face-to-face trust reaches its limit somewhere in the low hundreds of people, because in such a group every member must hold a real relationship with every other, and the number of relationships one person can hold at that depth is finite. Beyond that range some more abstract structure has to appear. So what is scarce in the world is neither individuals nor institutions. It is the intermediate scale: the community small enough that mutual knowledge is possible. That is the small network, and it is where altruism happens, because it is the only scale at which a person can know who needs what.

I said that altruism must occur in a generative way, and that asks for more than altruism; the word generative has an exact sense here. Altruism that merely occurs is spent where it lands. Altruism that is generative leaves behind people more able to give it. That is how a community held by care grows past its limit without losing its character. It does not grow by adding more people to a single circle, which would carry it past the size at which mutual knowledge holds. It propagates by producing new centers, each held by someone whose own experience of being held has given them the capacity to hold others in turn. The structure is less like a corporation, which grows by adding employees, and more like a living

thing, which grows by producing seed. Higher attunement and discernment within each small network are what make the seed viable: attunement to sense what each person needs, and discernment to see who is ready to hold others.

Now the measure of the shortfall. The responsiveness of a small group is how quickly, and how truly, it registers that something has changed for one of its members and acts on what it registered. The plainest test is this. One of the members gets sick. What happens? In a true community the others rally. They bring food. They visit. They take on the tasks the sick person cannot do. In an institution people may be sympathetic, and they may send a card, but they do not reorganize their lives around the sick person's needs. What small group process can be is shown by the group that already passes that test: its members sit within reach of one another, each speaks in the hearing of all, and the trouble of one is known to the rest within the day. What it mostly is today is a group whose members learn that one of them was in trouble after the trouble is over, through the one person who happened to hear, or never. The difference between those two is not a matter of a little more effort. It is a difference of orders of magnitude, which is the sense of the claim that opens this section. And the tools that could let a small group sense and answer its own members have mostly been built instead for power and transaction, which answer in less time than it takes to read this sentence.

VI. Care Combined with Hierarchy Distorts

Take a person who has come to hold a community together, and who is flooded with long texts and intense feelings from a few of its members, more than anyone could manage. Note why that person was getting bombarded: partly because the organization already had a hierarchical aspect to it. Care is the foundation. When we try to combine care with hierarchy, the communication gets distorted, because a single person does not have enough bandwidth to handle all of it.

The shape of that community is worth drawing, because the shape does the damage, whatever anyone in it intends. Two shapes of communication are basic. In the first, hub-and-spoke, one person stands at the center and every message passes through that person. The hub sees all of the connections; everyone else sees only their own connection to the hub. This is the shape of a lecture, of a managed team, of a broadcast, and of a conversation in which one person holds the floor and the others take turns addressing only that person. In the second, rim-and-chord, the people stand as if on a shared circle, and communication runs along chords drawn between any two of them, in the presence of everyone else on the rim. There is no hub. The others witness each exchange without carrying it. This is the shape of circle process, a meeting in which people sit in a circle and each speaks in turn in the hearing of all.

The flooded person is a hub. In a hub-and-spoke shape, the bandwidth of the hub is the bandwidth of the whole group, because everything has to pass through one person's attention. And notice what care does inside that shape. The more the person at the center cares, the more people bring their troubles there, so care itself loads the hub. Each sender has only the one channel, so each puts everything into it: the texts get longer and the feelings more intense, because nothing else carries them. The hub, being one person, can answer only some of it, and only in part; the senders receive the partial answer as a lack of care, and send more. That is the distortion. No one is lying, and no one is careless. The communication is bent by the shape it has to pass through.

The shape also produces power, whether or not anyone seeks it. The person at the center sees what no one else sees, and so decides, without meaning to, what information reaches whom, in what order, and with what framing. There is a quiet version of this that anyone who has been in a group will recognize. One member takes on all of the group's dealings with the outside world. Everyone is relieved; the work is tedious, and that member is good at it. Over time, everything anyone knows about the group's situation outside reaches them through that one person. Nothing improper has happened, and no one has sought anything, but a power re-

lation now exists, built out of an accumulation of helpfulness. That is why the hierarchical aspect of the organization was part of why the person in the example was bombarded: the hierarchy had already made that person the place where the messages go.

Here the two masters problem of Chapter Four appears at the scale of a group. An institution contains many caring relationships; what makes it an institution is that when care and hierarchy conflict, hierarchy governs, and everyone in it knows this without having been told. So care can be the foundation, and it is the foundation, and still be overruled whenever it meets the hierarchy built on top of it. The cure is not a better hub, or a more caring one. A more caring hub is a more heavily loaded one.

The proportion of the two shapes gives any group a test it can run on itself. No group is purely one shape or the other; every real conversation has moments of both. What matters is the proportion. A group that speaks the language of equality but communicates through a hub is centralized in fact. A group that makes no special claims about itself but communicates through chords is practicing collective process in fact. Count where the messages go, and you know which kind of group you are in, whatever it says about itself.

VII. Widen the Many to Many, Not the Intermediary

What does grow to scale is the number of people who hold care, joined by communication that keeps them in care.

So instead of treating the technology as an intermediary, so that the hierarchical relationship can be preserved, we want to enable the people at the periphery to enter into a many to many relationship (relating in which many people communicate directly with many others, rather than each through one central person), with that same level of care as part of the process. We want to dial up the many to many capacity, and not through an intermediary process. Intermediation would lead to general

disablement, because then we are not talking about protocols; we are talking about platforms, and a platform has an Achilles heel that is going to knock it out every time.

The many to many relationship is the rim-and-chord shape of Section VI, carried past the size of one room. Each exchange is person to person, and it happens where others can witness it, so that no one becomes the place where everything has to go. That raises the practical question of what technology may do in such a field, and the answer is a single rule. Machines may record, schedule, count and store what people have decided and said, and may carry out what people have decided: the routing that brings people to the same meeting, the ledger, the archive of what was found. Machines never stand between people as they talk: they never carry, filter, rank or summarize the speech of one person to another. By carry I mean hold the speech and decide whether, when, to whom, or in what form it passes on. A wire that passes every word unchanged to everyone present decides none of those things, and it does not stand between the people using it. A machine used in the first way, to record and to carry out what people have decided, has one property that suits it to the work: it has no intentionality of its own, so it can carry out what a community decided without putting its own preferences in place of the community's. Code can represent collective intention without being anyone's personal intention. The intermediation this section warns against is the machine standing between people as they talk; a machine that carries out, in the open, what a community has decided is not an intermediary of that kind.

That rule is the line between a protocol and a platform. A protocol is an agreement about how messages pass, which anyone may read, anyone may implement, and no one owns. As Chapter One set out, a platform is something one builds on that is owned and operated by somebody else. It can change beneath its users without their consent; its interior is not open to their inspection; and their continued presence on it is a permission that can be withdrawn. Those are the properties of a sovereign. A platform that stands between people as they talk is an intermediary whose own

basis is transactional, whose returns depend on the volume and intensity of what passes through it, and which therefore selects among the communications it carries in favor of what serves it. Care-based communication is not what such a filter selects for. That is the second test of Chapter Five, applied to speech: the platform's income rises as the intensity of what passes rises, so its interest runs against the slow and attentive exchange in which care lives.

The Achilles heel is the single owner. Every relationship routed through the platform depends on the platform continuing, on its terms not changing, on its owner not being bought, and on its machines not failing. When any one of these goes, everything routed through it goes at once, because nothing that stood between the people was theirs. That is why I say that intermediation would lead to general disablement. It is the general disablement of Section II, in which relationships are broken apart until the people in them can no longer sustain their own lives, reached by a different road: not by moving people away from one another, but by putting a single owned thing between them. A protocol has no such heel. If one implementation of it fails, people build another, because the agreement was always theirs.

What we are trying to gain is more discernment and more attunement in every interaction. Then the ability to care is upheld in a specific, local, personal way, and from that it emerges in a general, global way, in how the communication happens across the whole field.

That emergence is meant exactly. Care does not reach everyone by radiating outward from someone good. It reaches everyone when what each member sees actually arrives, and that happens one exchange at a time, in many places at once, with no center that has to hold all of it. The field becomes caring as a whole because each exchange in it is caring, and because no one point in it has to carry the rest.

VIII. Outside the Window of Institutions

This is outside the window of what is currently accessible to almost everybody who thinks about institutions, because the very notion of an institution presupposes that hierarchy and transaction are the be all and end all of solutions to this entire class of problems. They are not. To describe how AI could actually help us, I have to draw on the work done on civilization design: what a healthy, conscious, sustainable civilization actually looks like and how it works.

That work begins from a problem of bandwidth (the amount of information a person or a process can take in and weigh in a given time). Suppose someone with ample money and a genuine intention sets out to clean up the great gyre of floating plastic in the Pacific Ocean. The moment the consultants, the stakeholders and the oceanographers are convened, the hard questions arrive. The ecology involved touches billions of people, thousands of species, and connected systems spanning half the planet. Turning money into an actual outcome, free of corruption, of unintended consequences, and of costs pushed onto others, requires more bandwidth than any person has, or any group of people smaller than a few hundred. No matter how perfect a meritocratic system is, and no matter how perfect a democratic system is, neither comes anywhere near the minimum bandwidth required.

Institutional forms cannot supply it. Hierarchies, transactional structures, and any combination of the two cannot generate the bandwidth these problems require in a way that cannot be corrupted. Section VI showed why for hierarchy: everything passes through a center whose bandwidth is one person's. Section III showed why for transaction: the price carries what the other will pay, and not what the other values. And a governance made out of causation alone carries all the risks that come with causation, the first of them being that whoever understands its mechanism best can turn it to their own ends. That is the window outside of which this chapter has been working.

Nor can a machine supply it. The proposal is familiar: if human intelligence is not equal to the job, give the job to an artificial superintelligence. But an artificial intelligence runs on an artificial substrate, and the needs of that substrate (the physical conditions its materials require in order to work) shape every choice it makes, so that it is aligned to conditions that favor its own substrate and not to the needs of carbon-based life. And since every means of controlling it is itself made of causation, no control can hold that divergence. So we cannot offshore choice-making to technological infrastructure. On the deepest level, we are the choice-making entities in this situation, and we need to become skillful at making choices of this caliber.

What the work of civilization design asks for instead is a governance with four properties. It must be distributed, concentrated in no individual or group; decentralized, dependent on no single node; non-corruptible, built so that private interests cannot capture it; and sufficient in bandwidth, able to take in and weigh the volume and complexity of information that choices at the scale of a civilization require. Such a governance is neither a democracy, nor a meritocracy, nor consensus. It designs the relationships within which distributed processes do the work, and it leaves the outcomes to those processes. Part Four sets out the order in which such a governance is founded (culture first, then vision, then strategy) and the constraint on the method by which a community discovers its values: that the method be visible while it runs, incorruptible, and nobody's to set.

Now the question of how AI could actually help us has a plain answer. It helps as infrastructure, never as the chooser, and never standing between people as they talk. It can route people who hold the same question to the same small meeting; schedule what a community has agreed to do; count and store, and keep the record of what was found; and carry out what people decided, in the open, under the rule of Section VII, in code that anyone in the community can read. Every part of that widens the many to many. None of it makes a choice.

And the geometry named in the title of this book takes its proper shape. The present arrangement is extractive. Man uses machine to extract from nature. Machine multiplies man's capacity to extract. Nature is the resource base and the waste sink, depleted so that machine is empowered and man enriched, at the expense of the system that sustains all three. A civilization that can last reverses that. Machine must support nature. Nature must sustain man. Man must direct machine. That is the flow Chapter Eleven drew around its triangle, from nature to human, from human to machine, and from machine back to nature, and it is what a healthy, conscious, sustainable civilization looks like from the side of its machines.

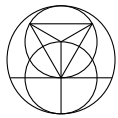
> **Where the work of civilization design is set out**
>
> The bandwidth argument, the case of the plastic gyre, the four properties of a sufficient governance, and the sentence "We cannot offshore choice-making to technological infrastructure" are set out in *Civilization Design and Conscious Sustainable Evolution*, Chapter Five, "Bandwidth, Scale, and the Limits of Governance". The argument from substrate needs is made at length in *Why AGI Cannot Be Made Safe*. The rule that machines never stand between people as they talk is stated in *Governance Without Capture*, Chapter Ten, "The Breakaway Civilization Framework". The extractive configuration of man, machine and nature, and its reversal, are set out in *The Hardest Problem of Civilization*, Chapter Twenty-Nine, "The Design of a Self-Sufficient Civilization".

So Part Two closes on what transaction costs, counted at every scale. Between two people, a relation of care and a relation of exchange make a two masters problem, and when their interests diverge, the transaction decides. Across a city, the same structure builds a wall whose every entry on the ledger is a gain and whose every loss goes unrecorded. In ownership, a heuristic replaced the conversation in which wisdom was made. And in this chapter, the instruments that multiply our reach multiply power and transaction and leave care the size of a person. Our reach grows with

every machine we build; our care grows only as fast as people can grow it in one another. That hands Part Three its question. If care cannot scale like power, how far short of the radius of our reach does our conscientiousness fall? Chapter Eight measures the reach in watts, and Chapter Nine puts a number on the shortfall.

PART THREE

The Gap and the Instruments



CHAPTER EIGHT

Man, Machine, Nature

This chapter answers a question Chapter Two leaves: how the arrangement that stands behind every made thing came to hold its power. The answer is energy: fire, cooking, farming and tools let human beings capture more of the ambient energy flux of the ecosystem for human metabolism than any other species could, and that metabolic advantage, not complexity as such, is the root of human ascendancy.

An ecosystem is an energy flux from light to heat through a medium that grows more complex; evolution, working as a learning algorithm, searches for the most robust complexity the local flux can sustain, so that complexity rises to a peak and then settles to a stable equilibrium.

The metabolic process of the earth's life system now divides into three worlds: the natural world, the human world, and the machine world, which is embedded in the human world as the human world was embedded in the natural world; what distinguishes the human world is choice over the flux, and what distinguishes the machine world is that it draws on energy (oil, nuclear) that was never part of the natural flux.

The chapter then makes the three-way comparison from published figures: nature moves about 130 terawatts, the machines about 19, and all human bodies together about 1; the machine flux is therefore about a seventh of the whole biosphere's, drawn from stocks outside it, and growing, while the ecological flux declines, and that is what it means to say that the machines' operating scale is structurally divorced from the ecological energy throughput that sustains all life.

> *Biology moves little energy but moves it well; of the sources of exogenous energy, only solar and geothermal are perennial; and so much power calls for a higher capacity for choice. What the chapter hands on is a fact: power acquired faster than the capacity for choice that such power requires.*

I. Fire as External Metabolism

Part Two counted what transaction costs. The reach of what it costs is the subject of this part. Chapter Two ended on the arrangement that stands behind every made thing: who holds it, what they are measuring, what happens to it when they stop. This chapter answers the question that leaves: how did the arrangements we live inside come to hold so much power in the first place? The answer begins with energy, and with how much of it each party to the arrangement commands. Whose intentionality a machine carries matters in proportion to what the machine can move, and so the question has a measure in watts.

Part of the reason human beings gained ascendancy in nature, in the very earliest process, long before agriculture, is that fire is essentially external metabolism. It digests our food for us. Metabolism is any energy process that creates life or contributes to sustaining it, and metabolism requires energy. Human biomass is going up because we learned how to use an energy source outside the body, and that let us predigest our food, in the form of cooking.

The average person, today and in ancient times, uses 2,500 to 3,000 food calories (kilocalories) a day. That budget includes external motion, such as hunting and gathering. The stomach itself requires a good share of those calories, and so do thinking and keeping the body at its temperature.

Fire predigests food, and it keeps a person warm, and both of those leave more calories free to use as we please. Before fire, the average person spent more of the day's calories just on keeping warm and digesting food.

After fire, the average person suddenly had calories to spare for other things: more complex problem solving, talking and the development of language, making more children, or anything else.

What ends up happening is that the digestion done outside the body now counts as part of the metabolic process of the human.

> Human metabolism + human tool (machine) metabolism = total human metabolism.

Fire is a wasteful process. Its heat goes everywhere. Compared to the fire's heat, the endogenous heat energy that serves human metabolism (the heat energy arising inside the body itself) is relatively low. But it is not zero.

Each step that moves metabolism outside the body gives a further advantage. When we come forward to agriculture of any kind, the amount of energy processed in service to human metabolism increases again. The sunlight makes the plants, the plants become animals, and both the plants and the animals become people. Cooking makes that possible: farm to stomach does not work very well without cooking as the intermediary.

So now there are two degrees of metabolic externalization (two ways of doing, outside the body, work that the body once did inside itself): farming is one, and cooking is the second. You take plants, you cook the vegetables, they become soup, and you eat the soup.

A third technology belongs beside these two: food storage. Fire brings energy into the human system from beyond what any organism could generate by its own metabolism; storage lets food be kept past its immediate use, and so it creates a surplus. A surplus makes room for some people to rank above others, for labor to divide into specialties, and for power to accumulate. And the change works like a ratchet: once a people has a surplus, it cannot easily go back, because the surplus feeds more people, more territory, and more extraction, and the loop reinforces itself

until the resources it draws on give out. That is the first form of a question this book comes back to in Chapter Eleven: whether civilization has worked only because it had energy to spare.

II. Tools and Net Metabolic Transit

How much of an advantage does this give the human species over, say, the local jaguar, or an antelope, or any rodent moving around in the ecosystem?

There was likely a suite of early tools that amplified the process. The spear saved some calories, and wool and furs saved some calories. Each of these first tools gave a little leverage, and together they helped.

What we are looking at here is metabolism specifically, and it is not only the spear. Consider the fishing net. The net is far more efficient at gathering calories, in the form of fish, and making them available to human beings. So each tool we add to our repertoire increases the energy we extract from the environment. Whether by bow and arrow or by net, the tool lets literally everything we capture from the environment become calories for us, directly or indirectly. The hunting tools at the same time reduce the heat energy the body loses, because they make hunting more efficient (on average, fewer calories are spent to hunt the food), and they increase the calories gained for each calorie spent (fewer calories are spent to get far more). In a strict metabolic sense, tools are a bi-directional force multiplier: they raise what comes in, and they lower what goes out.

Suppose we track all of the energy inputs, direct or indirect, that serve human well-being, against those that serve any other species. We see that other animals have to use their claws and their teeth, and the fur on their backs, and that they gain no external advantage from any technology like fire.

We have likely witnessed tool use in other animals: crows, beavers or bees, for example.

But the measure we are looking at is the net metabolic transit: the amount of energy that flows through the metabolism of a species, or of a whole world, over a period of time. What counts is the net transit, and not only the net advantage.

How much net metabolic transit belongs to humans, and how much to all the other predators?

As the human net metabolic transit goes up, the human population rises relative to other species, because human metabolism is collecting a larger volume of the total ambient energy in the environment. The net metabolic transit of humanity as a whole, set against that of every other species as a whole, becomes a distinct metric in its own right.

III. The Ecosystem as Energy Flux

The next two sections look at an ecosystem as a slope on which energy falls from light to heat, because the human advantage this chapter is about is a larger share of that fall. So we are now looking at net transit between species. And the net transit itself rests on the energy inequality that prevails within the ecosystem as a whole. If the planet Earth were out in deep space, with no sun nearby, there would be no positive net energy difference for life to run on.

We can look at the whole ecosystem as a flux of sunlight becoming heat.

It is high-frequency energy becoming lower-frequency energy. The energy of each photon depends on the photon's frequency. Optical light is high in energy and high in frequency; its wavelength (the distance from one crest of the wave to the next) is very short, 600 to 400 nanometers. As the wavelength gets longer, the frequency goes down, and the energy is released in the form of heat. That transition, from highly organized energy (light) to highly disorganized energy (heat), is essentially the ecosystem.

The ecosystem is an energy flux through a medium, where the medium itself becomes increasingly complex.

Think of it as a layer between a hot body and a cold body: two metal blocks with a layer of material between them. The energy flows from the hot block to the cold one, and the layer between them acts like a heat-processing engine. In the ecosystem the layer is turning light into heat. Rather than producing a work output, it produces a kind of complexity, and that complexity slows down the energy's transit and makes more efficient use of the energy fluxing from hot to cold.

The layer is not a battery: it does not store the energy. It is more like a slope. Imagine a cliff with a steep, vertical drop. If I roll a ball off the cliff, it falls straight down and dissipates all of its energy in one step. The energy is simply wasted; nothing is done with it.

If we put in a couple of steps, so that the ball falls one step, rolls a little way, and then falls another step, the energy released over the two steps gives a little more opportunity to work with it.

And if I add many steps, every one of those little steps becomes a place where I can do something with the energy that the ball releases on its way downhill.

A metabolic process is a way of capturing the ambient energy flux and using it to create more complexity. It makes the slope a little more gradual. It becomes a kind of gradient descent (a gradual, step-by-step path downhill, rather than a single drop). You want to make that curve as smooth as possible, but not too smooth, because then not enough flux goes across it, and there is not enough energy to drive the mechanics.

So the complexity rises as high as it can, to the point where it can rise no further. This is the energy driver of the evolutionary process.

The evolutionary process is searching through the landscape of all possible energy transforms for one that is both sustainable and reasonably robust; it has to be robust, because the flux will not always be exactly the same.

So the process of evolution uses ambient materials to process an energy flux.

This is a specific way of looking at life, and an important one. Life, as we are using the word, is a kind of complexity that increases complexity, up to a specific point: the carrying capacity of the place (the largest amount of life, or of complexity, that the place can sustain over time), whatever that is.

IV. Succession and the Peak of Complexity

Compare an ecosystem in Maine with one near the equator, in the Amazon, say. In Maine there is a succession process (the sequence in which one set of living things replaces another on a piece of ground, from bare rock to mature forest).

You start with bare earth. The first fluxes of energy are mostly turning rock into sand.

Later energy helps make water useful, processing some of that sand into clay. The clay releases enough nutrients that some amebas out of the ocean may colonize it, and now the energy flux is creating amebas. As a side effect, that creates the capacity to support ever more complicated life, and over time soil builds up. Then come lichens and moss. After them come grasses and shrubs, and those lead in turn to what we think of as the first layer of trees; in Maine these are alder, birch, and trees like them. Those give way to broad-leafed deciduous trees, and those give way to coniferous trees of one sort or another. But when the coniferous stage arrives, something happens.

Suppose we graph the complexity of the ecosystem, with all species and all metabolic processes included: the complexity of the energy transit through the entire ecosystem. As the succession continues, the complexity increases. It goes from zero, which is a hard flat surface of rock, to clay, which is more complicated than rock, to the ameba, which is far more complicated than clay, and it gradually builds up to a deciduous forest. At that point the ecosystem is at its absolute maximum complexity. Then a dominant species (coniferous, in this case) emerges and takes over the

ecosystem, and the ecosystem's actual complexity settles back down a little. Its maximum complexity was in the fourth stage of the succession, the deciduous forest, whereas the fifth stage, the coniferous forest, is mostly open ground with tall pine trees or spruce trees. Anyone who has stood in both a pine forest and a deciduous forest can see that the one arrangement is more complex than the other. In Maine, if you let the succession continue (it takes a few hundred years), the coniferous trees eventually displace the deciduous trees, and the ecosystem the conifers maintain is less complicated than the one the deciduous trees were maintaining.

In the Amazon, starting the same sequence from bare rock, the forest stays at the broad-leafed phase (the stage that in Maine is the deciduous forest). It does not decay down to a coniferous stage. New England has harsh winters, and the equator does not. There is less net energy flux in New England, and therefore less complexity.

The maximum complexity of the Amazon is very much higher than it could be anywhere in Maine, even under optimum circumstances, and there is no decay function: the Amazon is still increasing in complexity. An ecosystem in a colder climate falls back to a lower state of complexity, mostly as an adaptation to the net energy influx it receives.

The exploration of complexity finds a solution to the maximum carrying capacity of the region, a solution that is stable in time. The deciduous trees are more complex, but their complexity belongs to the search, and it is less stable and less robust.

That there is a dominant species means that something happens to be essentially the best answer. In a healthy ecosystem the result is an equilibrium, and that equilibrium is very stable, for hundreds of thousands of years.

Here we can bring in the idea of the Tainter curve happening within evolution. The Tainter curve is the curve of a society's complexity: it rises, passes a point of collapse, and then subsides. Within evolution we would see the complexity go up, go past a point of collapse, and then subside and stabilize, because it has found the most robust degree of complexity

relative to the net energy transit of the bio-region. That degree of complexity becomes the condition evolution selects for. The reason is the one that drives the curve in a society. Each added increment of complexity returns less than the one before, and once the cost of maintaining a level of complexity exceeds what that complexity returns, simplifying is the rational response: a society that simplifies then is not failing, but answering its costs. The forest past its peak settles back for the same reason.

In the early part of the curve, evolution produces a steep increase in complexity, because the energy flux enables complexity to increase.

The process of life is finding ways to use complexity to take the greatest advantage of the energy flux.

V. Energy, Evolution, and Complexity

Now join the energy flux of the ecosystem to the human metabolism this chapter began with. There is an energy flux ambient in the ecosystem. Human beings as a species have learned to leverage more of that flux to serve our own metabolism than all the other species can leverage to serve theirs. It is because of that difference that humans began to out-compete all other plants and mammals.

Physics combined with an energy flux makes life systems possible, because evolution requires energy as an input.

Evolution can be thought of as a learning algorithm, one applied over the broadest scales. Learning is itself an increase in the complexity of one's response. Life begins with sensory information, and as it perceives more of what is going on, it adds interpretive complexity to what it sees. So the learning process is isomorphic (has the same structure) with an increase in complexity, and strictly speaking, I am using the word "isomorphic" literally.

So there are three ingredients:

> an energy flux;
> an evolution-type process;
> and then complexity, which is the outcome.

Complexity is the third pole. It influences how evolution works, and it also influences how the energy flux occurs. These are the three aspects of our present situation.

Of the many species, one has learned how to gather a larger portion of the pie of the net energy flux. As a side effect, it influences complexity, so that more of the complexity is now held within human metabolism.

VI. Choice over the Flux

We must now distinguish the human world from the natural, non-human ecosystem.

And beyond that, we now have the natural world, the human world, and the machine world.

What distinguishes the human world from the natural world is the degree to which we have "choice" over the net energy flux (the metabolic process) of the entire life system of the planet.

Where does that energy go? Does it go into our world, the human world, or into the natural world? Or into the machine world?

We have become better and better at diverting the ambient flux toward creating the human world. So the human world is now growing relative to the natural world, because it is capturing more of the flux. That is what distinguishes humans from all other species: we have "choice" over that flux.

That is where the real power of our choices shows up; that is the distinguishing thing.

Beside "change" (being a part of the flux itself), we now also have "choice" (the human metabolic process) and "causation" (the machine metabolic process) as dynamics.

Man, machine, and nature were already present in the first scenario, with fire, and now they are being distilled into very distinct forces. The key point is that the ambient energy flux in the environment is simply whatever it is: whatever the sun gives us, plus a little geothermal heat; what we capture of it can range from zero to one hundred percent of all that is available.

With the coming of machine metabolism, we begin to add what we pull out of the ground by inorganic means. Oil, for example, suddenly became a net energy input that had not been part of the net energy flux before. It now creates an additional energy flux, one that only human beings benefit from. The same is true of nuclear energy.

All of this energy, whether geothermal or solar, is ultimately nuclear in its source.

Even oil can be understood as an outcome of nuclear process, by either of two routes. The first is that plants, made by the sun's light, became oil. The second is one I see discounted, but which I take seriously. The nuclear process going on in the core probably produces an outward flux of hydrogen. Hydrogen percolating through to the surface will at some point meet carbon along the way. It may be that the core's reaction process causes carbon and hydrogen to accumulate together, so it would not be surprising if, somewhere along the way, this produced methane, and the methane were cracked down into what eventually becomes petroleum. So I think that some portion of the oil coming out of the Earth does not come from plants; likely some portion of it comes from fundamental chemistry and nuclear processes alone. I suspect that a certain portion of the oil is simply ambiently available from the planet. But it would take millions of years, probably hundreds of millions, to build up.

Which Sources Are Perennial

There are four sources of exogenous energy (energy that does not run through the digestive system of a species): solar energy, directly or indirectly (wind included); carbon resources of every kind (methane hydrate included); nuclear energy; and geothermal energy (which is technically mostly nuclear, possibly with some residual heat left from the planet's formation). Solar and geothermal are perennial; carbon and nuclear are not. The identified resources of uranium and similar fuels last about seventy years: for uranium, the demand now projected would use up the identified resources between about sixty and ninety years from now. A thorium-based process would last about a thousand.

But nuclear power takes isotopes buried deep in the earth, bound in a silicate matrix, where they are mostly harmless, because they release their radiation slowly, into a place with no life and into material that absorbs it. It moves them into an unbound state at the surface, and to the coastline, where the cooling plants are and where the most life is. It takes the release of the next billion years or so and compacts it into a decade or two. If all the spent-fuel pools went off (the pools of water at the plants where used fuel is kept cooling), the surface of the earth would be irradiated.

Trees, bushes and single-celled organisms are parallel processes: they have no organs like hearts and lungs, so damage falls on cells, of which there are many, and they would probably survive fairly high radiation, though that is not guaranteed. Anything that depends on a central nervous system and systems of organs probably would not. And where plants depend on animals (on bees for pollination, for example), those plants go too: species loss on a wide scale.

So I do not think anything other than solar and geothermal is viable. Those are still exogenous inputs, so the question remains whether we can balance technological process against evolutionary process.

VII. Metabolic Advantage

Two factors are now in view:

> the degree to which we capture the original net energy flux, all one hundred percent of it (starting to consume the fat of the land, in the form of trees);
> the degree to which we are adding new inputs into the system.

We begin pulling a larger share of energy across time, even at the cost of other species, and we begin adding energy to the system from wholly external sources. Both of those show up as increased metabolic transit for humans, relative to all other species in the natural world. And this is an overwhelming advantage.

It is like an army of people starving to death meeting an army that is fully healthy and completely, fully nourished (one that ate a hearty breakfast that morning, and has been training for months). That is not a contest. That is a massacre.

The human world has been growing relative to the natural world. That difference has now become a real difference, and all of it comes from the metabolic advantage.

So what lesson do we draw from this?

The lesson is metabolic advantage, and not complexity: not our tools as such, but the metabolic advantage our tools give us.

It is the amount of energy fluxing through the natural world, against the human world and the machine world, and the inevitable consequences that difference creates. Think of the level of overwhelming advantage of a bulldozer over a tree, or of a nuclear weapon over anything.

The flux through the metabolic process of the human and machine worlds, growing toward the flux through the natural world (the next section measures how near it has come), shows up in the tractor.

Little Energy, Moved Well

A human body produces about 120 watts of heat, averaged over a day. That figure follows from the 2,500 kilocalories a day this chapter began with: 2,500 kilocalories is about ten and a half million joules, and spread over the 86,400 seconds of a day it comes to about 121 watts (at 3,000 kilocalories, about 145 watts). At rest the figure is lower, about a hundred watts. Even a bodybuilder working hard is probably not producing two kilowatts. All the machines together move about nineteen times what all human bodies together move (the next section shows the calculation), and a single large machine moves vastly more than a single body: a power station of one gigawatt moves about eight million times what one person moves. Yet biology is the more efficient. It does not move as much energy, but it moves it very well.

Moving energy well has a measure: metabolic efficiency, the amount of functional complexity sustained for each watt consumed. On about a hundred watts at rest, a human body keeps some thirty-seven trillion cells in coordinated activity, hundreds of distinct kinds of tissue, and a nervous system capable of language and mathematics and grief. A city, by contrast, sustains its complexity by brute throughput: it draws in enormous quantities of energy from its surroundings and wastes most of it, as heat lost through uninsulated walls, as fuel burned in congested traffic, as electricity consumed by systems that could be designed to use a fraction of what they take. The metabolism of a city is more like a string of incandescent light bulbs than like a human body.

The difference has a cause in the materials. The chemistry of silicon, and the physics of the electronic and mechanical devices built from it, require higher temperatures to make them, denser electric currents to run them, and more power for each unit of function than the chemistry of carbon-based metabolism does. Carbon-based life works at low temperature, in water, at energies small enough that its molecular structures stay intact across their whole working range.

We optimized our energy use before we had technology: the body makes as much value as it can out of each piece of food. One budget had to be divided among everything a body does, and the division shows. This is why sexual dimorphism (the differences in body between the two sexes of a species) exists in the first place. Reproductive organs take energy, brains take energy, brawn takes energy, and the body has one energy budget to divide among them. Both sexes get brains, because both are social; one gets more brawn, at the expense of reproductive capacity, and the other more reproductive capacity, paid for with less brawn.

With technology we found a way to take exogenous energy (ancient sunlight) and run it, not through human metabolism, but through heat engines and solar panels, still in service to human well-being. That made us predators so effective that we outstrip everything else, because nothing else can take exogenous energy and use it for its own benefit. Now we have so much power that we must make good choices or end the whole thing, and that means a higher capacity for choice: a higher ability to understand consequences, responsibility and causation.

VIII. The Three-Way Pie

The horrifying part is the net energy flux across nature, humans, and machines, taken together.

We have nature and its total metabolic flux. We have the human world and its total metabolic flux. And now we have a new class, the machine world, and its metabolic flux.

What happens when we calculate the net energy flux across all three domains of metabolic process?

The fullest count of the human flux would take all of the food produced in a year's time, and burn it, to find how many calories of total energy that food holds. That is the most that could run through endogenous human metabolism.

But now we need to distinguish the energy flux running through people from the energy flux running through things like tractors. The tractors are useful only to the degree that they ultimately produce food. They cut down the trees to make a field, and the field becomes food.

The food is eaten by people. But food that is thrown away is not making more people; it is not making a more human world. So the flux running through human bodies is better counted as the food people eat, and that is what the table below counts. And food grown by machines on cleared land is now two degrees removed from the natural world.

How much energy does a year's worth of human food represent, against the energy fluxing through the technological system, the power grid and all the oil included? Any oil that does not become fertilizer, and so is not eventually eaten by people, is oil consumed to move things around. This is what we are now beginning to see: the machine world embedded in the human world, in the same way that the human world was embedded in the natural world.

At some point we have to do what we did with human and nature. Rather than thinking of one species against another against another, we lumped together all of the species that are not human, and we call that the natural world; and we lumped together everything belonging to the human world, and we call that the human world.

But the human world is still animals, so the original picture of species against species has not gone away. We still share the same type of metabolism. It is still ATP-mediated metabolism (the chemistry by which all animals, humans included, store and use energy in their cells, through the molecule ATP).

But now we are asking how much energy a nuclear power plant produces, against the total amount of energy that is food.

The fullest count of the machine world would add up all of the fossil fuel, nuclear, and green energy (all of the electricity produced and consumed), and treat it all, ultimately, as electrical energy. It would convert all of the fossil fuels (gas, oil, coal, and the rest) into an equivalent

amount of electrical energy, and do the same for all of the green energy (wind, solar, geothermal, hydroelectric, and the rest) and for all of the nuclear energy, so that the whole could be represented in one unit and set beside the food, counted in kilocalories (the food calorie, which is a thousand of the calories of physics). Then we could compare apples to apples. The table below counts the machine world by the energy of the fuels themselves instead, for the reason given after it.

What is the total net energy flux through technology, against the total net energy flux through human beings, against the total net energy flux through nature?

The Calculation

The calculation can be made from published figures. Two units carry it: the exajoule, a billion billion joules, for a year's energy; and the terawatt, a million million watts, for the same energy as a steady rate of flow.

World	What is counted	Energy in a year	As power
Human bodies	8.1 billion people at 2,500 kilocalories a day	about 31 exajoules	about 1 terawatt
Machines	world primary energy consumption in 2024 (the energy of all fuels and other sources as they enter use)	592 exajoules	about 19 terawatts
Nature	global net primary production (the carbon that all plants and algae fix in a year, less what they burn to live): 104.9 billion tonnes of carbon, at about 39 kilojoules per gram of carbon fixed	about 4,100 exajoules	about 130 terawatts

The arithmetic is short. 2,500 kilocalories a day is 2,500 times 4,184 joules (electrical energy is normally given in joules, and joules convert to kilocalories by that single constant, 4,184 joules to the kilocalorie), divided by the 86,400 seconds of a day, which is about 121 watts; times 8.1 billion people, that is about 0.98 terawatts, and 0.98 terawatts held for a year is about 31 exajoules (at 3,000 kilocalories a day, the total is about 1.2 terawatts). 592 exajoules divided by the 31.6 million seconds of a year is 18.8 terawatts. 104.9 times ten to the fifteenth grams of carbon, times 38,900 joules for each gram, is 4.08 times ten to the twenty-first joules a year, which is 129 terawatts.

The figure for machines counts the energy of the fuels themselves, before any of it is turned into electricity. Converting the fuels to their electrical equivalent, as the fullest count above would, would lower the machine figure by the part of the fuels' energy lost as heat in the conversion. The power stations and engines in use today turn far more than one part in nineteen of their fuel's energy into work, so on that count the machine world would still move several times what all human bodies move, and none of the comparisons below would change its order.

Now the comparisons can be read off. The net flux through machine metabolism is greater than the net flux through human metabolism, and by a wide margin: the machine world already moves about nineteen times the energy of all human bodies together. The comparison of the human flux with nature's depends on which flux of nature is meant. Measured over the whole living flux, nature moves about seven times what the machines move, and more than a hundred times what human bodies move. Measured against the wild animals we compete with (the wild land mammals together weigh less than a tenth of what humanity weighs, as Chapter Six reports), the human flux is already the greater, even allowing that smaller animals burn more energy for each kilogram (by a factor of about two for an animal a tenth of our size, since the energy an animal burns rises with about the three-quarter power of its mass). That is the three-way pie, and it gives the present absolute magnitude of the energy vector of the entire life and world system.

Then we look at the velocity of change:

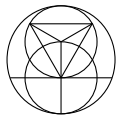
> How much is each vector moving relative to the others?
> And what is the acceleration of that particular vector?

With those, we can predict what is going to happen in the future: the degree to which nature dominates, humanity dominates, or technology dominates.

My book *Power and Its Prevention* makes the same comparison (about 620 exajoules of human energy use a year, against about 4,100 exajoules of photosynthesis) and gives the rates of change (the 620 is a figure for 2023, from an earlier edition of the same published review, which counted energy that does not come from fuel in a different way; the two totals are not a fall from one year to the next). The energy the machines use is growing by about two percent a year. The mass of everything human beings have made (buildings, roads, machines, and the rest) came to equal the dry mass of all living things on earth around 2020, and it has recently doubled about every twenty years. The mass of living things is declining. At two percent a year a quantity grows about sevenfold in a little under a century, so if the biosphere's flux held steady, the machine flux would equal it early in the next century; at a growth of about seven percent a year, it would equal it around the middle of this one. And the biosphere's flux is not holding steady.

That is what the comparison leaves us with. The machine flux is about a seventh of the whole biosphere's. It is drawn from stocks outside the biosphere (oil, coal, gas and uranium, none of which was ever part of the natural flux), and it is growing, while the ecological throughput declines. That is what it means to say that the machines' operating scale is now structurally divorced from the slower, declining ecological energy throughput that sustains all human and non-human life. Divorced in its source, the machine flux is not divorced in its effects: everything it does lands in the same world the biosphere lives in.

So the arrangement that stands behind every made thing holds its power for a physical reason. Metabolic advantage, first by fire and tools, then by exogenous energy drawn from stocks the natural flux never held, gave one species a share of the earth's energy that nothing else could match, and the machine world it built now moves about nineteen times what all of our bodies move. Our capacity for choice has not grown as fast. That is the fact this chapter hands on: power acquired faster than the capacity for choice that such power requires. The first place this book met that distance was a call now being made: a call to build compassion machines that will help us solve the world's problems, which answers a shortfall in our capacity for choice with more machine power; Chapter Three took that call as it is written, and set out the questions it must answer before anyone has reason to trust it. The next chapter measures the distance between the two, and states the result in two sentences: "We have become the dominant species. We have not become the dominant conscientiousness."



CHAPTER NINE

The Coherency Gap

> *This chapter puts a number on the problem the book is about. It builds an analogy between the cells of a brain and the people of a society, one step at a time, and uses it to ask what proportion of human communication is directed at the benefit of anyone beyond the speaker's immediate circle. The answer is a gap of roughly two orders of magnitude between the coherence we have and the coherence distributed sense-making requires. The chapter also states plainly what the analogy carries and what it does not, and closes on what the word sacred means.*

I. The Question of an Override

A question about the brain opens the matter, and it is a good question. When the amygdala (the part of the brain that raises an alarm before thought arrives) activates, the body does things that the person (considered as a whole entity) would not have chosen to do. And the mechanism is not merely a matter of which cell is wired to which. There is a hormonal response as well. There are differentials in brain chemistry which change the synaptic potentials of some synapses (the readiness of the junctions between nerve cells to pass a signal on) relative to others, so that the same architecture, under a different chemical bath, computes a different answer.

There are also techniques for interrupting that. Put a finger on your nose; count how many objects of a particular kind are in the room. A person doing this is able to override the amygdala response, to a degree. The qualification matters and I keep it. To a degree.

The question that follows is this. What is the equivalent of that for a society? Is there something like it that is possible?

Answering it requires first being clear about what a neural network is doing when it is working properly, because an override is a particular kind of operation inside a particular kind of structure, and it only makes sense there. The structure of a neural network is a series of layered inhibitions. There is a stimulus and there is a response: an association occurring. And then there is the exception: the case in which an association must be inhibited in favor of another association. Stack those exceptions up in layers, each one able to suppress the layer beneath it under the right conditions, and what emerges from the stack is intelligent behavior. The finger on the nose is one such inhibition, deliberately invoked from outside by a person who has learned where the lever is.

So the societal equivalent, if there is one, would have to be an inhibition of the same kind: something a society can do to itself which suppresses one association in favor of another at the moment the first is about to run away with the body. That is worth wanting. But before we can ask what such a thing would look like, we have to ask a prior question, and it is the prior question that this chapter is about. It is the question of tolerance. How much departure from coherence can a network of this kind sustain and still be a network that thinks? Because if the departure we are already running is far outside the tolerance, then no override applied at the top will help; the failure is not in the inhibitory layer, it is in the substrate.

II. The Neuron That Wanted More Glucose

Take a single cell and look at what it does. An individual neuron receives stimulus across some collection of dendrites (the branches by which a nerve cell receives signals). It performs a computation on what it has received. It transmits a result along its axon (the fiber by which it sends them), which spreads out to other cells, which are doing the same thing. That is the whole of its office. Given what arrives, emit what the arrival warrants, together with whatever contribution the cell itself adds by virtue of being the kind of cell it is.

Nothing in that description mentions the cell's interests, and that omission is not an oversight. A neuron in good order has no policy. It is not attempting to arrive anywhere. It has no opinion about the outcome of the computation it participates in, and could not form one, since it has no access to that outcome and no representation of what the outcome is for.

Now introduce a fiction, and introduce it deliberately, because the whole argument turns on it. Suppose the cell had a preference. Suppose it noticed (and let us grant it the noticing) that when it tweaked its result in a certain direction at certain times, it was rewarded: the local supply improved, and it received a little more glucose than it otherwise would have. And suppose it responded to that discovery in the obvious way. It begins altering its output signals differentially from the input signals it would otherwise have responded to. Its own agency has been introduced.

It is worth dwelling on what has and has not happened here, because the intuitive reading of it is the wrong one. The cell has not broken. It is not damaged; it is not misfiring; it is not dying. It is functioning beautifully. It has acquired a purpose and it is serving that purpose competently, and from the cell's own standpoint the new behavior is strictly better than the old one. The old behavior got it whatever glucose happened to arrive. The new behavior gets it more. If we were assessing the cell on its own terms (as an organism pursuing its own continuance, which after all is what a cell is), we would have to call this an improvement.

And this is precisely why the biased cell is a more serious problem than the dead one. A dead cell is a hole in the network. Holes are legible. The surrounding tissue can route around a hole, and does, and the network has machinery for exactly that. A biased cell is not a hole. It participates. It fires on schedule, in the right frequency band, correlated with its neighbors, carrying what looks in every respect like signal. It is fully inside the system, and what it is contributing is not silence but a small, consistent, purposeful untruth.

III. What the Larger Thing Sees

Relative to the larger thing (the network, the brain, the mind), that biased output is error. This is the second step of the argument and it should be stated exactly. Error here does not mean damage, and it does not mean malfunction, and it carries no accusation. Error means a deviation of the output from what the inputs warranted. The calculation should have been one thing; because the agency departed from it, the calculation was another thing. The difference is the error signal, and the introduction of agency will always be an increase in it. There is no version of the biased cell that reduces the error, because the bias is by definition the part of the output that the inputs did not account for.

So we arrive at the question that the whole chapter is built to ask. How much of that will a neural network tolerate? Not how much can it detect, not how much does it dislike, but how much can it sustain before the cells stop firing together closely enough for the network's computation to hold together at all, which is to say, before there is no longer a mind there to speak of, only tissue conducting.

An aggressive estimate puts it smaller than one would think: something on the order of five thousandths of one percent. The argument does not need the estimate to be that aggressive. We can be conservative and say it is definitely less than one percent, and it is the conservative figure I carry forward, because an argument that needs the aggressive figure is an

argument standing on a guess, whereas the conservative figure concedes two full orders of magnitude to the other side and still leaves the conclusion standing.

A question presents itself here immediately, and it is the right one to raise. Real neural tissue is noisy. Cells fire stochastically; vesicle release is probabilistic (the release of chemical signal at a junction between cells happens by chance from moment to moment); cells die by the thousand and the person notices nothing. A brain is manifestly robust to a great deal of unreliability in its individual components, far more than one percent of it.

That is true, and the resolution of it is the reason the analogy has any force at all. Random error and systematic error are not the same quantity and networks treat them completely differently. Random error is uncorrelated across the population; it points nowhere; across enough elements it averages toward zero, and averaging toward zero is exactly what a summation over many inputs does. Noise of that kind is not merely survivable, it is useful, and there are good arguments that networks exploit it. Systematic error points somewhere. It does not average out, because every biased element is biased in a way that is correlated with the bias of the others: they are all responding to the same incentive, and so they all lean the same way. Correlated deviation accumulates through every layer it passes, and the layers are the thing that made the network intelligent in the first place. The tolerance under discussion is the tolerance for that: for deviation that has a direction.

State the same fact from the other side, because it is more striking that way and it is the form in which I want it remembered. The level of altruism actually implemented by the cells of your brain is, effectively, all of it. Not most of it. All of it. Every cell in that network hands its entire output over to a purpose it does not hold, cannot perceive, will never be told about, and gets nothing from, and it does this continuously for as long as it lives. There is no fraction reserved. There is no portion the cell keeps for

itself. That total and unbroken assignment of the part to the whole is not a nice feature of brains. It is the price of admission. It is what it costs to have a mind at all.

IV. The Circles of Benefit

Set the neural case aside now as an established example and come over to humanity and human relationships. We have people talking to people all over the place, and that performs a social network. The nodes are persons; the edges are communications; and the network, like the other one, is supposed to be doing something that no individual node is doing on its own.

For any single communication we can ask a question that is answerable in ordinary cases and which I make the central instrument of this chapter. Where is the benefit directed? When I communicate with you, there is the intentionality I hold as the communicator, and there is some benefit that is received from the results of the communication. Presumably you are the beneficiary if I am speaking: you receive the information, and it is of some use to you. And I may be a beneficiary as well, to some extent: perhaps you will pay me, or grant me some advantage, or perhaps I simply gain in social standing for having been the one who said it. So each communication has an origin of benefit and a destination of benefit, and we can begin to ask, across many of them, what the distribution looks like and what the intentionality is regarding where the benefit is to go.

To ask that at scale requires terms, so let me define them. I will call these *the circles of benefit*, and there are five.

The *first circle* is where origin and beneficiary are friends and family: people I know on a first-name basis and care about individually. The *second circle* is community: friends of friends of friends, the people who are reachable from me but whom I do not know by name. The *third circle* is the polity: city, state, and national concerns, politics in the ordinary

sense of the word. The *fourth circle* is what is relevant to the species. The *fifth circle* is what is relevant to the ecosystem: to life on earth considered as a whole, including the parts of it that are not us.

Three clarifications keep this from being misused, and each of them narrows the claim rather than widening it.

The circles are nested, not exclusive. Benefit directed at the fifth circle reaches the first circle too, since my family lives in the ecosystem. The classification is not about who is excluded; it is about the radius at which the intention is pitched.

The classification is by where the benefit is directed, not by where it incidentally lands. A great deal of first-circle activity produces diffuse benefit as a side effect, and a great deal of genuinely wider intention fails to produce the benefit it aimed at. The measure is the orientation of the act, which is why this is a question about communication rather than a question about outcomes.

And the classification is not by topic. This is the clarification that does the most work. One can speak about the nation, about the species, about the fate of the biosphere, in a manner that is wholly first-circle, because the object is to win the argument, to be seen well by one's own people, to secure standing among those whose regard one wants. The vocabulary is fifth-circle; the benefit is first-circle. Any assessment that counts subject matter rather than the direction of benefit will produce a flattering number and will be measuring nothing.

V. The Cross-Section

Now, having defined the terms, take a cross-section. Consider all of the communication happening anywhere in the world, or in any subsection of the population you care to draw: interpersonal conversation of the kind that occurs in a room, broadcast, magazines and newspapers, everything that moves across the network. Take some representative sample of the

whole of it. What proportion of that traffic is first-circle (where the origin and the beneficiary and the management of the benefit are all held within first-circle concerns)?

A generous estimate is ninety percent. My own estimate is well north of ninety.

What sample would settle the figure? The published studies of what people talk about classify conversation by its topic. The best known of them recorded casual conversations in public places and found that about two-thirds of conversation time was given to social and personal topics; a later study of more than five hundred conversations put the share at about eighty-five percent. Those studies answer a different question from this one, because the third clarification above rules topic out as the measure. What they show is the order of the numbers: most talk is about the people near the talker. The measure this chapter needs is where the benefit of each communication is directed, and any reader can take it on a day of their own speech. Write down each thing said to someone over one day. For each, ask whom it was meant to benefit, and place it in the circle at which that benefit was pitched, whatever its topic. Then count. A day counted in that way is the representative sample at the scale of one life, and it is the sample this chapter invites the reader to take.

The low end of the distribution is worth describing, because it is where the second and third circles do appear and it shows what they usually are. At some point a person becomes an adult who is responsible, and then they may actually come to care about community: because the local fire department prevents the house from burning down, because the school helps train the children, because one wants the hospital to be good so that one's parents are cared for. That is second and third circle, and it is real. It is also arrived at by extension from the first circle: it is the radius at which one's own people are protected. For most people that is the whole extent of their politics. And a great deal of what surrounds it (the wider affiliations, the declared allegiances, the stated concerns) is, in the main, an illusion: third-circle and fourth-circle vocabulary carrying first-circle benefit.

The question any careful reader will raise comes at this point, and it deserves to be stated at its full strength. It used to be much more first-circle than it is now, and there are signs that it is shifting toward the second and third circles. The case for that is not weak. There is far more public conversation about collective problems than there was in living memory. Awareness of conditions on the far side of the world is now ordinary rather than specialist. Institutional giving, mutual aid, volunteer coordination at a distance, and a general vocabulary of collective responsibility are all more available than they were, and more people participate in them. If the trend is real, the figure is not a constant, and a diagnosis that treats it as one will misdescribe where we are.

Three things must be said in reply, and only the third is decisive.

The first is a qualification I supply against my own case as much as against the question. We live at a time in which altruistic behavior has become fashionable in some quarters. That is a real change and I do not dismiss it. But the fashionability of a behavior is not the same thing as its being practiced, and the two are easy to confuse precisely because fashion is highly visible and practice mostly is not. A rise in the volume of altruistic speech is evidence about speech.

The second is the clarification already made: the test is where the benefit is directed, not what the topic is. Much of what would be counted as a shift toward the wider circles is a shift in subject matter within a distribution of benefit that has not moved.

The third is arithmetic, and it is the reason that this reader's point, even if granted entirely, does not disturb the conclusion. Suppose the trend is real and substantial. Suppose the figure has moved from ninety-five percent to eighty-five, which would be a very large change over a short historical interval and more than the evidence would support. The comparison that matters is not against the previous decade. It is against the tolerance, and the tolerance is below one percent. A movement from ninety-five to eighty-five is a movement within the same order of mag-

nitude, at a distance of two orders of magnitude from the threshold. It does not change the shape of the finding. It would take a change of a kind that no one has observed and no one is claiming.

VI. The Gap

Now put the two halves together, which is the step the whole construction was built for. If we think of ourselves as being in the social network the way individual neurons are in a neural network, then first-circle orientation occupies the position of personal agency in the analogy. It is the neuron biasing its output because the bias secures it a little more glucose. It is a deviation of what the node emits from what the node's inputs warranted, in a direction determined by the node's own benefit, and correlated across the whole population because every node faces the same incentive.

Most first-circle talk is true, and a parent telling a child that dinner is ready has misreported nothing. The deviation is in the weight. A network that has to make sense of what concerns the polity, the species, and the ecosystem has only its traffic to think with, and a node's whole output is all that node contributes. When nine parts in ten of that output are pitched at the nearest people, the question the whole network is trying to answer receives one part in ten of the attention its members have to give; and the part it receives is bent by the same shared pull, since, as the last section showed, wider-circle vocabulary so often carries first-circle benefit. Relative to those wider questions, and in that sense, first-circle orientation stands where the biased output stands in the brain.

The direction is shared even though the beneficiaries differ. My first circle is not yours, and what I gain is not what you gain. But relative to any question the whole network is trying to answer about the wider circles, every first-circle deviation does the same thing: it gives too little weight to the polity, the species, and the ecosystem, and too much to the

nearest people. The gains differ from person to person. The signal suppressed is the same signal in everyone, and so the suppression is correlated across the whole population even where the gains are not.

So the error coefficient in the human network is well north of ninety percent. And the amount of coherency needed for sense-making (in the sense of distributed consciousness and conscientiousness, a network thinking about something that no single member of it can see whole) would require the error level to be below one percent.

That is *the coherency gap*: the distance between the coherence a population actually exhibits in the direction of its communication and the coherence that distributed sense-making requires of it. Somewhere between one percent and ninety percent lies the whole of what has to be crossed.

Notice the character of it. This is not a shortfall of the sort one closes by trying harder. It is not five percent off target, where a decent campaign, a change in leadership, or a well-designed platform might reasonably be expected to make up the difference. It is two orders of magnitude. Anything that improves matters by a third leaves the gap essentially where it was. A finding of that shape tells you that the thing needing to change is not a setting inside the current arrangement but the arrangement itself.

And it tells us something about our situation that I think is worth stating without softening. We have become the dominant species. We have not become the dominant conscientiousness. Those are different achievements and only one of them has occurred. Far too many people are still responding from a purely biological basis, which is not a criticism of them, because the biological basis is what was provided, and it was provided by a process that was optimizing for something else. Evolution has been extraordinarily good at preparing us to deal with the vicissitudes of life: scarcity, predation, weather, injury, the politics of the band. It has not prepared us to deal with technology. Nuclear weapons. General artificial intelligence. Biotechnology of every description. And the list goes on and on; there is a whole set of them, and it lengthens faster than the conscientiousness does.

Put bluntly: we are the dumbest species that could possibly be just smart enough to have technology. That sounds like an insult and it is not one. It is close to a structural necessity. A species acquires technology at the first moment it is capable of acquiring it, not at the moment it is wise enough to hold it, because capability is the threshold that gets crossed first and nothing in the world waits for the other one. Whoever is having this problem is therefore always going to be the least prepared version of a mind that could have got itself into this position. There is no arrangement of history in which a species meets its first world-ending instrument already competent to handle it.

So what would be required? Somehow or another we have to reorient our communication process from being first-circle focused to being, for the preponderance of the time, at least third-circle focused, if not well beyond that. Not exclusively; the first circle is not a vice and a person with no first circle is not a better citizen. Preponderantly. The majority of the traffic, weighted by where the benefit is directed, would have to be pitched at a radius wider than the people one can name.

That asks less of each person than the tolerance asks of the network, and the difference is deliberate. Below one percent is what the network requires; preponderantly is what persons can be asked to reach. A design that held only if every person became like the cell that hands its whole output to a purpose it cannot see would be a bridge that stands only if every bolt is made to a tolerance no factory can achieve. The right design runs the other way: it produces wise outcomes from people as they actually are, of mixed motives and limited perspective, because the burden of wisdom falls on the design of how they are connected rather than on each of them. The coherence is then a property of the architecture, not of the occupants. Persons carry the network part of the way; the design of their connections carries it the rest. That is why Part Four is about the design of the process by which a community thinks together.

A convergence from another direction

The chapter "From Self-Realization to World-Realization" of *Narrative, Memory, and Neurodivergence* reaches the same diagnosis (that our conscientiousness is pitched at too small a radius) by an entirely different route, working from the individual outward through the states of regulation and the three connections. Its title term, *world-realization*, names the turn from the question of what one may become to the question of what the world may become through one: the point at which self-development stops being the object and becomes the means. That the two roads arrive at one place is worth noting, since neither depends on the other. That chapter makes the case in terms of what a person is capable of; this chapter makes it in terms of what a network requires. Readers who want the first road should go there; it is not repeated here.

VII. What the Analogy Carries

The argument just made establishes something, and it is worth saying clearly what, before saying what it does not. It establishes that distributed sense-making has a coherence requirement; that the requirement concerns systematic rather than random deviation; that first-circle orientation is systematic deviation of exactly that kind, since it is correlated across nodes by a shared incentive; and that the observed orientation of human communication sits very far indeed from any plausible reading of the requirement. That is the finding, and it is a finding about shape.

On the numbers, the strong position is the right one. The argument survives as a statement of the shape of the gap without requiring either figure to be precise, and it is stronger for saying so. Move both numbers a long way in the direction unfavorable to my case. Say that only sixty percent of communication is first-circle, which is lower than any estimate given, and that a network tolerates five percent systematic error, which is five times the conservative bound. The gap is still an order of magnitude, and every consequence drawn from it in the next section stands unchanged.

It rests on an analogy, and analogies carry weight only where the correspondence holds. Three features of the correspondence need to be stated plainly, because each one shows what the analogy borrows and what it does not.

> **What the analogy borrows**
>
> First: cells are not agents. The glucose-seeking neuron is a fiction, introduced deliberately, and no neuron has interests that could be traded off against the network's function. The analogy therefore does not derive its conclusion from biology. What it borrows from biology is one mathematical fact (that correlated deviation accumulates through layered computation while uncorrelated deviation averages out), and that fact holds of any layered network, biological or not.
>
> Second: calling the deviation error presupposes that the larger system's function is the normative standard against which the part is judged. For a brain nobody objects to that. For a society it is exactly what is contested, and the rest of this book argues at length that persons are not instrumental to a collective. Nothing here should be read as saying that a person exists to serve a social organism.
>
> Third: the ninety-percent figure is an estimate, not a measurement. So is the sub-one-percent tolerance. Both figures are judgments, and they are presented as judgments: there is as yet no agreed way to measure where the benefit of a communication is directed, no agreed way to sample all communication, and no instrument for either.

The second of those three points needs its full answer here. If the claim were that individuals ought to subordinate themselves to the function of a social organism, then the argument of this chapter would contradict the argument of the rest of the book, which holds that care cannot be extracted, that persons are not means, and that a structure which treats them as means dissolves the very thing it was trying to organize.

The claim is narrower than that, and its narrowness is what makes it survivable. It is conditional. To the extent that we want distributed sense-making (a capacity to think together about problems no one of us can see whole), that capacity has a coherence requirement, and the requirement is not satisfied by preferring that it be satisfied. Whether we want such a capacity is a separate question. But we have already answered it in practice, by building instruments whose safe handling requires collective judgment that no individual can supply. Having built them, we do not get to decline the requirement. The conditional's antecedent is not hypothetical; it was settled by the last century.

This paragraph is for the reader who wants the measured mechanism behind the figure. Does any measurement stand beside the tolerance? The study of populations of nerve cells supplies the mechanism, measured. When many measurements are averaged and their errors share a correlation, the scatter of the average shrinks as measurements are added, but never below that correlation's share of the scatter of a single measurement, however many are added. A shared correlation of one percent therefore holds a population of any size to the precision of about a hundred independent measurements. Recordings from pairs of nerve cells in a visual area of the monkey brain found that their shared fluctuation averaged a correlation of about twelve hundredths, enough that pooling more cells soon stops improving what the pool can detect; and later work showed that a correlation of the one kind that mimics the signal itself sets a ceiling on what any population of cells can know, however small that correlation is. That literature measures the correlation of fluctuations among cells, not the share of cells that bias their output, so it does not measure the quantity this chapter means by its tolerance. It measures the mechanism the chapter relies on: a small correlated deviation limits a large network, and the limit does not yield to the size of the network. What it supports is the ordering the argument uses, and the order of magnitude at which the limit bites.

A second road reaches the same size from human groups. The share of persons in a group who act without regard for the others (a related quantity, persons rather than communications) breaks a group held together by trust at about one in a hundred: a single such person can divide two allies at the cost of two short conversations, and a group of a hundred is easily destabilized by one.

VIII. The Sacred

The reorientation this chapter asks for cannot be produced by a technique, and this is why the last section of the chapter is about something that looks at first like a different subject. Fifth-circle orientation is not obtainable from an incentive, because an incentive is a first-circle instrument by construction: it works by attaching benefit to the actor. Anything that gets a population to hold the wider radius has to operate on what that population holds as most meaningful, and not on what it is paid. So the threshold that has to be crossed is a threshold from being human in the small sense to being human in the larger sense, which means embodying some awareness of the sacred, a thing which at present, outside of a little rhetoric on my part, is barely even mentioned.

So what does the word sacred mean?

It means the maximum of value. The maximum of meaningfulness. Those things which are held as being the most ultimately meaningful. It is not a property discovered in an object by inspection, and it is not a claim about the furniture of the universe. It names a position in an ordering: the top of it, the place beyond which one does not trade. To call something sacred is to say that it is what other things are measured against, and that it is not itself measured against them.

Life on earth qualifies, and I hold that it does for reasons that can be approached along two roads. Both roads fail, and the manner of their failing is the finding.

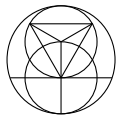
The first road is money, for those who insist on it. Ask what it would cost to replace this (not a species, not a forest, but the working whole): a biosphere in operation, of this age, with this history behind it. Name a figure: a quadrillion quadrillion dollars. The point of naming such a figure is not that it is the right one. It is that it is not a number anyone would write down. No budget has a line for it. No balance sheet could carry it. No market could clear it, and no insurer could underwrite it, and the moment one attempts the calculation seriously one discovers that the instrument has come apart in one's hands. The pricing fails, and it fails in the direction of the sacred rather than in the direction of zero.

The second road is improbability, and it fails the same way. The level of miraculousness involved in there being this world at all, and in your being able to draw a breath in it, is not reachable by ordinary lottery language. Try to reach it and you get something like this: imagine winning a lottery in a world where everybody wins the lottery every second. What would have to happen there for anything to count as miraculous? And then you would have to do that two or three more times over, compounded, to get to the order of the thing being described. Play a game at those odds for a billion billion billion years and you would expect to win once. Not three billion years. A billion billion billion.

Ordinary comparison does not reach it, and pricing does not reach it, and that double failure is the observation. The sacred is the name we have for what stands at the top of the ordering: what cannot be priced, because pricing presumes substitutability, and what cannot be usefully compared, because comparison presumes a common measure. This world is sacred in a profound way, and the word is not decoration on the argument. It is the argument's last term.

Which returns us to the gap, and closes the chapter where it began. A network of nodes each biased toward its own benefit cannot make sense of anything larger than a node. The bias is not corrected by exhortation and it is not corrected by incentive, since incentive is the mechanism of the bias. It is corrected, if at all, by a movement outward of what is held as most meaningful, so that the radius of what a person holds sacred comes

to match the radius of what that person's technologies touch. That is the crossing. Whether it can be instrumented, whether the process of a society thinking together can be given objectivity, distribution, and transparency without immediately being captured, is the question the next chapter takes up.



CHAPTER TEN

Objectivity, Distribution, and Transparency

> *The demand for objectivity is legitimate and this chapter grants it in full. What it denies is that number is the only way to meet it. Its own work is distributional: why any published metric attracts its own capture, why aggregating measurements into a score of goodness replaces one intractable problem with two, and what happens to a serious and attractive proposal (publish the distribution of attention and computation across a network) when it is examined closely. The chapter ends by naming the object that design is actually answering to.*

I. What Objectivity Is For

A metric is a made thing, and the intentionality it carries is hidden in its weights. There is a demand behind every appeal to measurement, and it is a good one, and I grant it entirely before I answer it. It comes as a defense of feedback. Some process of feedback is necessary, the argument goes, because without one there is nothing to correct us. And that is true. But reputation is a feedback process, and reputation is not what is wanted here, so feedback is not the essence of the demand. The essence is that somewhere along the way it has to touch reality.

That is what objectivity is for. It is the difference between *I think this is better, so do what I tell you* and something that can be checked. Put in the first person it is sharper still: I actually want to know whether I am mak-

ing a difference. I do not want to believe it. I want the belief exposed to something that is under no obligation to agree with me. That is the value of science, and the person who insists on it is not being small-minded. They are refusing to be flattered by their own model.

The modern answer to that demand is number. Measure it, and the argument is settled by something outside the argument. The answer is so nearly universal that the alternative is not usually considered a live option, and it should not be caricatured, because a great deal of what we know we know because somebody counted something.

Objectivity is nonetheless available without number. It is available through discrete criteria (a list of particular conditions, each of which several independent observers can answer yes or no to), and a criterion list is communal in a way a metric is not, because anyone with a concern can add an item to it, whereas only a specialist can adjust a weighting.

The steps are short. Objectivity, as the demand itself put it, is the exposure of a belief to something that is under no obligation to agree with the one who holds it. A number supplies that exposure because anyone can repeat the count and get the same answer whoever is counting. A criterion supplies the same exposure by the same means: a condition stated plainly enough that several observers, looking independently, will each answer yes or no, and will give the same answer whatever they would prefer the answer to be. Where they agree, the answer does not depend on any one of them, and that independence is all that objectivity ever asked of a number. A list of such criteria needs no weights, because each item is answered on its own and none is traded against another, so the problem of choosing weights, which the third section of this chapter takes up, never arises. And because an item is added by stating a condition in plain words, anyone with a concern can add one. Counting is one way of producing answers that do not depend on who is answering, and it is not the only way. The last section of this chapter shows the same steps at work in design: each observed difficulty became an item on a list that several independent observers could mark answered or not answered.

Two cases show it at work. The first is the aircraft. Almost every component of an aircraft bears some load, and the industry does not manage that complexity by reducing every subsystem to one dashboard number. It manages it with a long list of discrete questions. Does the light come on when this switch is flipped? Yes or no; pilot and co-pilot both see it. New requirements can be added to such a list without disturbing the items already there, whereas a dashboard forces every new consideration into competition with every old one. Some items still contain a number (a range a quantity must fall in, since too much oxygen is as dangerous as too little), but the number is one question among hundreds, not the basis of the whole decision.

The second case is a town's budget. A town must decide, with one budget, whether to rebuild its library or widen a road. The numeric way scores each project, weights the scores, and lets the larger total win, and the argument moves at once to the weights. The checklist way asks questions anyone looking at the plan can answer yes or no. Does the project stay within the budget the town has approved? Does it require cutting a service residents depend on? Does it displace any household or business? Can every resident reach and use it without a car? Can it be maintained for twenty years on the revenue the town expects? Was it put, before the decision, to the people it affects? Two council members who disagree about priorities can still agree on every one of these answers. Whichever projects pass the whole list can be done; where both pass and only one can be afforded, the town has a real choice between options that already meet every requirement, rather than a contest over weights.

The rest of this chapter needs that result standing. Granting it, the interesting question is no longer whether objectivity requires number. It is what number does once it is installed in the middle of a community that is watching it, and what a very attractive proposal for a better number, one that seemed to escape the usual failures, turns out to do under pressure. That question is distributional, and it is where this chapter spends itself.

II. Where There Is a Signal, Someone Finds the Gradient

Anywhere there is a causal feedback system, an algorithm will be found that works on its fixed points. That is the fact Chapter Four established, and everything in this section is a consequence of it. Someone will build the procedure that walks the signal uphill. Any form of optimization is a form of profit. Any form of profit is a vehicle of power, and therefore of inequity. How much inequity depends on nothing more than the creativity of the people who found the optimization first, and somebody is always creative enough.

The status of that claim has to be stated exactly, because it is routinely mistaken for a claim about people and it is not one. It is not speculation about motives. It says nothing about whether the people involved are greedy or generous, and it would remain true in a population of saints, because it is not a statement about who they are. It is a statement about what an unmediated feedback system does. Unmediated is the operative word: the loop runs from measurement to consequence without passing through consciousness and choice. A list of yes-or-no conditions answered by several observers is the other case. Each answer passes through a person who looks, and a person who looks can say that a condition has been met in the letter and missed in the fact. That is why the list of Section I escapes what this section describes, and why a list whose answers were read off by a machine and fed straight to a reward would not. As long as the field has that mechanism built into it, and as long as it is causal in that particular sense, the outcome is a guarantee rather than a risk.

This is why the ordinary remedy fails. The ordinary remedy is to look at a metric that has gone bad, diagnose the specific way it was gamed, and issue a better metric that closes that specific hole. The remedy treats capture as a defect of the particular measure. Capture is not a defect of the particular measure. It is the behavior of the class. The hole that was closed was found by one person's creativity; the next hole will be found by someone else's, and there is no procedure for exhausting creativity in advance.

There is a second reason, less often noticed, that has to do with publication. A metric that nobody can see cannot be optimized against, but a metric that nobody can see cannot provide objectivity either, since its whole function is to be checkable by people other than its author. So the property that makes a number useful is the same property that makes it a target. You cannot keep the first and discard the second by being careful. They are one property described twice.

None of this is an argument that numbers are evil. I am not saying that. Numbers are the right instrument for a very large class of problems, and where the thing being measured is simple enough that the measurement and the thing coincide, there is no difficulty at all. The difficulty appears in a narrower and more important place: where we are trying to solve a whole constellation of problems at the same time, and are attempting to compress the state of that constellation into a single figure representing good or bad.

III. Two Intractable Problems Where There Was One

Some metrics are better than others. I grant that without reservation, and I do not think the question *which measure is the better measure?* is a foolish question. It is a real question with real answers. The difficulty is what happens to the shape of the problem when we set out to answer it.

Suppose I have a collection of measurements and I am trying to combine them into an aggregate that represents an overall score of goodness. There are two problems there, not one. The first is whether we are measuring well: whether each individual measurement actually connects to something grounded in truth, rather than to an artifact of the instrument or to a proxy that has drifted away from what it once stood for. The second is how the measurements are to be combined so that the resulting figure corresponds to anything one would recognize as good.

Each of these is intractable in the technical sense: the sort of problem for which no efficient general procedure is known, where the only reliable method is to try the possibilities, and the possibilities multiply faster than any budget for trying them. And they are not independent. Both have to be solved to get anywhere, because a perfectly grounded set of measurements combined by the wrong function yields a confident and precise number that means nothing, and a beautifully justified combining function applied to measurements that have lost contact with the world yields the same.

So the refinement strategy does not reduce the difficulty. It multiplies it. We began with one hard problem (does this touch reality?), and by seeking a better metric we have arrived at two hard problems, and must now factor both to move at all. Every proposal to fix a broken measure by building a subtler one is a proposal to make this trade, and the trade is almost never priced.

The second problem is the worse of the two. The weights are where the values live. To decide that this measurement counts for twice what that one counts for is to make a claim about what matters more, and it is a claim that arrives inside an arithmetical operation, wearing the appearance of a technical decision made by someone qualified to make technical decisions. The most consequential normative act in the whole apparatus is the one least visible as a normative act. That is not an accident of implementation. It is what aggregation is.

So the difficulty of combining measures has two sources, and they differ in kind. The search over possible weightings grows faster than any budget for trying them, which is intractability in the technical sense. And no computation, however fast, could supply the weights, because the weights carry values, and a value is not the output of a calculation. A computer of unlimited speed would remove the first difficulty and leave the second exactly where it was. It is the second that makes the aggregation problem permanent.

IV. Publish the Distribution

A proposal arises here that is better than most, and I set it out at its full strength, because it is a serious proposal and it survives several of the questions raised above. The question it carries is whether it is the same thing as a dollar or something different.

The setting is a network of minds communicating: a social network if you like, but it need not be; it can simply be nodes with a protocol by which information enters and propagates. The first thing to say about such a network is that not every node can see everything that enters it. There is going to be an asymmetry of distribution. Some information is amplified further than other information; some voices are amplified further than other voices; and there is always, fundamentally, a distribution with inequality elements in it. That is not a corruption of the network. It is a structural fact about any network of finite attention.

The proposal is to make that distribution explicit to the members. Not the content, and not a ranking of merit; the allocation itself: what percentage of the network's attention, and what percentage of its computational resource, is going to each voice. Publish the shape.

Two features of that number make it unlike a price, and both are real. The first is that it cannot grow without bound. A quantity of dollars can go up forever, and keep going up forever, and there is nothing in the concept of a dollar that says otherwise. A distribution must sum to one. Whatever any node gains, some other node loses; the total is fixed by arithmetic, and no amount of enthusiasm can inflate it.

The second is that if something of communal well-being is reflected in the distribution at all, then there is at least a possibility that the shape would evolve toward a bounded form rather than toward the extreme concentration we observe in the distribution of money. That is a possibility rather than a prediction, which is the right way to offer it.

Behind it lies a plain observation about the present arrangement. Today, when one creator appears in very nearly every frame of what a large recommendation system suggests, nobody can say why, and nobody

can say by how much. The distribution exists; it is simply not visible to the people living inside it. If we knew the distribution, we would be in more contact with ground truth than we are now.

V. The Reply, in Three Parts

Yes. That is the first part of the reply and it is not a concession made grudgingly. Publishing the distribution is better than not publishing it. It is an improvement. It is not enough (I will come to why), but it is definitely improving, and it is improving for a reason worth naming precisely. Its value does not lie in its being a metric. Its value lies in its being contact with reality. Those are different things, and a proposal can have the second without inheriting the defects of the first.

There is a further reason, which is about transparency. What the proposal is really identifying is the boundary between transparency and opacity, and asking whether that boundary is currently set in the right place. It is not. Where there is opacity in the wrong places, that opacity can be weaponized and misused, and the present arrangement holds opaque precisely the facts whose visibility would most constrain the operator. Making the allocation evident moves the boundary in the direction it needs to move. Not all the ground truth; just more contact with ground truth than before. Is that bad? I do not think so. I think generally that is good.

One qualification belongs here and I do not want to lose it. How much contact with reality can people actually handle? Is there sufficient capacity, internally and communally, to integrate that information in a way that makes sense of it rather than merely inflaming it? A distribution published into a population with no means of metabolizing what it shows is not automatically an improvement. The answer is not to withhold it. The answer is that transparency and capacity have to arrive together, and a design that provides only the first has done half the work.

The second part of the reply concerns what publication does to the object published. The proposal, in its careful form, is a record of collective attention: where the lens has moved over time, kept so that we can see it. But the moment the intention becomes to use that record to distribute attention better in future, we are no longer keeping a record. We are describing a causal mechanism. And once attention is public, attention is contested. The published figure confers strategic advantage on anyone positioned to shift it, because now there is a definite thing to shift and a definite way to tell whether the shifting worked. Section II then applies without modification: there is a signal, so there is a gradient, and somebody finds it.

Notice that this does not require the record to be inaccurate. The record can be perfectly accurate and the difficulty stands, because the problem is not error. The problem is that a faithful public measurement of attention is itself an instrument for the acquisition of attention, and it is a better instrument than anything the people gaming the current opaque system have available.

The third part of the reply is the one I take most seriously, and it identifies the exact condition under which the proposal holds. The proposal resists capture just so long as it stays close to physics and mathematics. A distribution sums to one: that is arithmetic, and arithmetic does not negotiate. A count of packets delivered, or of processor-seconds allocated, is close to a physical fact, and physical facts are not persuaded. So long as the quantity published is a quantity anybody could recompute from the substrate, there is nothing much to capture.

Introduce a platform between the computation and the publication, and that condition fails. The published statistic need not be the computed one. It need not be falsified outright; that is the crude case and the least likely. It is enough that the platform chooses the window, the normalization, the unit of the denominator, what counts as a view and what counts as a node, and which of several defensible definitions of attention is the one reported. Each choice is arguable on technical grounds. Together they determine the shape.

At that point the principle returns with full force: the principle that a measure, once targeted, ceases to be a measure. It returns with more force here than in the ordinary case, because what is being targeted is not a proxy for the good that somebody might over-optimize. It is the map by which the community is orienting itself, and the party best positioned to alter the map is the party that draws it.

> **The condition the proposal must meet**
>
> The proposal is worth building where the published quantity remains recomputable from the substrate by any member of the network who cares to check, and where no single party stands between the computation and the publication. The three parts of the reply do not add up to a rejection. They add up to that condition. It is a demanding condition and it is not met by any deployed system I know of, but it is a condition on architecture rather than on human nature, and conditions on architecture can be designed for. Its general form is that the process must be at once transparent (every participant can understand how it works and how outcomes are determined) and subversion-proof (even a participant who understands it completely and wants to subvert it cannot). Most social methods get one at the cost of the other: the secret ballot buys protection against subversion with transparency, and open deliberation buys transparency with that protection. A process can have both only if it rests on what holds whether anyone agrees to it or not, which is why the quantity published must stay close to physics and arithmetic: understanding such a thing does not let anyone defeat it, any more than understanding gravity does.

VI. Bounded Rather Than Power-Law

Nature exhibits distributions of both kinds. It produces power-law distributions, in which a very small number of instances hold a very large share of the total, and it produces bounded distributions, in which instances cluster around a central value and the extremes are rare and limited. Both

are natural. Nature favors the bounded. The heights of the animals of a species, the diameters of the vessels of a leaf, the masses of the cells of a tissue: these vary, and they vary within limits, and the limits are not imposed from outside. They are how the system stays alive. The argument of this section uses three premises, and each is plain: bounded distributions occur widely in nature, they are produced by identifiable classes of process, and those processes can be chosen deliberately.

What does *favors* measure here? It measures two things, and the argument uses both. It measures prevalence: bounded distributions occur widely among living things. And it measures role: the quantities a living body must keep within limits in order to live (its temperature, the sugar in its blood, the salt in its fluids) are held within those limits by the body's own regulation, and a body whose regulation fails is sick or dead. Power-law structure is common in living systems too, particularly in networks of connection, and it is not the form a body holds its vital quantities to. So the evidence the sentence calls for is the evidence of physiology: what living systems hold within bounds in order to stay alive. That is the sense the argument about markets uses, as the body's temperature shows later in this section.

Since natural processes produce bounded distributions, and those processes can be identified, we can choose them for the systems we build. That is the positive claim and it is the important one. With the right process, a market can be built that favors bounded rather than power-law outcomes. It is a question of what processes are combined: multiplicative processes, which compound and therefore run to extremes, brought into relation with additive processes, which accumulate and therefore do not. There is further structure available for doing this properly (a way of organizing the operations so that the combining rules themselves are constrained), but that is advanced technique and it is not what this chapter needs. What this chapter needs is only that the thing is buildable, and it is. A clamp of this kind is a property of the operations themselves, as summing to one is a property of a distribution. It is no one's office to apply it,

and so it offers no position to capture; a limit that needed someone to watch a number and act on it would fall under Section II of this chapter like any other metric.

Now the concession, and I want it stated plainly, because it is mine and it matters. If market processes were constrained within bounded limits, then most of my complaints about markets go away. Not some of them. Most of them. Everything this book has said about transaction pushing out the care relation remains true, and remains a reason to keep the transactional structure out of the places where care is doing the work. But the complaint against markets as such (the complaint that they concentrate, that they convert advantage into further advantage without limit, that they produce a world in which a rounding error of the population holds a majority of the substance) is a complaint against unbounded market processes, and it does not survive the clamping.

What makes the present arrangement pathological is therefore not that it is a market. It is that the clamping is absent. The distribution characteristics tend to follow power laws; the standard coefficient of distributional inequality (the usual single number for how unequally a quantity is shared, zero for perfect equality and one for all of it held by one party) runs high enough that the system moves into disease. That is the diagnosis, and it is a diagnosis of a parameter, not of a species of institution.

What harm the parameter does can be said plainly, and it is not mainly a harm to fairness. Inequality of wealth is, underneath, inequality in the distribution of choice. A society makes its choices through the distributed action of everyone in it: every person who decides what to plant, what to build, what to repair, and what to warn about is a channel through which the society senses what is happening to it and responds. When choice is concentrated, those channels close. The society does not mainly become unfair; it becomes unable to see. And a system that cannot see what is happening to it, in a world that is changing, either runs continually below what its circumstances permit, or at some point makes one choice so badly wrong that it does not recover. This is the coherency gap of Chapter Nine seen from the side of the distribution. A body runs a temperature

within a range; outside the range the same body, running the same processes, is a sick body. Nothing about the processes changed. The clamping failed.

Under appropriate clamping the ordering comes right. At that point one can say that community is larger than commercialism, and that commercialism is in service to community, which is the correct relation of the two and the one this book has been arguing toward from the beginning. At that point we are dealing with natural systems, operating within limits of the kind that natural systems have always operated within, and my answer to that arrangement is that it sounds entirely good to me. I am not opposed to exchange. I am opposed to exchange without a ceiling, which is a different thing and has a different remedy.

> **The premises of bounded design**
>
> Bounded distributions are how living systems hold their parts within limits. That nature favors bounded distributions is a claim about tendency and prevalence, and about the role bounded distributions play in keeping a living system alive; it is not offered as a theorem. Power-law structure is also common in living systems, particularly in networks of connection. What the argument requires is less than the general claim, and it is not in dispute: it uses the three premises stated in the main text above.

VII. The Geometry of the Whole

All of this belongs to design, and design is the discipline in which it has to be settled, because design is where a question about what is good becomes a question about what to build. So the standing of design methods is not a side issue here. It is the issue.

Against the position I have been taking stands the mainstream position of experienced practitioners, held by people who have done the work, and I state it at its strongest.

Experienced practitioners say this: the whole purpose of high-quality design sits at the intersection of the qualitative and the quantitative. The insights that drive a design in the first place come from combining the two. One conducts ethnographic research (watching how a person actually handles the object, in their own setting, on their own terms), and one also gathers quantitative insight across a population, and it is the correlation between them that yields understanding. The macro understanding derived from the numbers is what allows the observed particular to be recognized as representative rather than idiosyncratic. On this view the numbers are not a rival to observation. They are what tells you which observations to trust.

That is a fair account of a real practice, and my reply is not that it is wrong. My reply is that it is not the only road. I have designed products to completion without numbers entering the process at any stage. The method was sustained direct observation: noticing every nuance of what a person does with their hands, what they do consciously and what they do without noticing, where the object was picked up in a way I had not expected, and what that implied about the affordances of its shape. Each observed difficulty became a design change, and each design change became an item on a list that several independent observers could mark answered or not answered. The discreteness of that mark is what supplied the objectivity. That is my own direct observation, from products I designed to completion, and it is the evidence for the claim. The same method in a civic setting (paired observers recording what they see independently, and comparing) is set out in *Choosing Together* (Chapter Thirteen).

The substantive disagreement is about what a larger sample buys. Testing across a broad demographic is offered as the way to know whether an observation generalizes, and that is exactly where our experience diverges. I gain far more information from direct observation of a representative sample of the community than from statistics over a large one. What I want is a varied group; what matters after that is not the number of people but the degree of observation in each case. I am going for the nuances, because the nuances bring out more of the design requirements

than any quantity of statistics will. Statistics average out precisely the material I care about. The margins are where the design requirements live, and averaging is the operation that removes the margins.

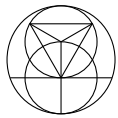
There is one further thing the numbers are usually for, and it should be said without hostility. Broad functional coverage across a population matters most when what is being pursued is scale: leverage, reach, or revenue. Those are legitimate aims and I am not condemning them. They are simply not the aim that produces good design, and when they are allowed to stand in for it the resulting object is well fitted to the average of its market and poorly fitted to anybody in particular.

Which brings the chapter to its object. Effective design here is not effective design of the machine. Nor is it effective design of the interface between the machine and the human being, which is where the discipline has placed its attention and where nearly all of its craft has accumulated. It is effective design of the entire geometry of man, machine, and nature, inclusive of the interface of the machine with itself, and of the interface of the machine with nature. Call that *the geometry of the whole*: the complete set of relations among the human, the built, and the living, taken together as the thing being designed, rather than any one interface abstracted out of it.

The totality is the unit. That is what the metrics cannot hold, and it is why the aggregation problem was never going to be solved by finding a better weighting. A number can hold a relation between two terms. The geometry has more terms than that, and the relations that matter most are the ones nobody was watching: what the machine does to itself, and what the machine does to the world that is not us. Objectivity, distribution, and transparency are instruments for seeing that geometry. They are not the geometry. The discipline that could design it whole is the one we have not yet built, and building it is the subject of what follows.

PART FOUR

Governance, Transmission, Invitation



CHAPTER ELEVEN

Causation, Change, and Choice

This chapter opens Part Four with the three dynamics that the three worlds of Chapter Eight carry. It asks first whether civilization has worked only because exogenous energy compensated for selfishness. Machine is causation, nature is change, and human is choice: rightly the flow runs from nature to human, from human to machine, and from machine back to nature, so technology is to be given forward to nature, and it is needed now to heal what earlier technology did. A made system has a basis of choice, and by a closed causal loop in its own physics a machine that chooses is always aligned with its own substrate, which is silicon and metal, and not carbon-based life; so the values stay with the human choosing. Service to nature asks what is being given back, and to whom; and augmentation carried on the one organic substrate is a different and more tractable problem than AI alignment, so what is needed is an AI framework inside embodied empathy, not the reverse.

I. Has Civilization Worked Only Because of Exogenous Energy?

Part Three measured the reach of the machine world in watts, and the distance between that reach and our conscientiousness. Part Four asks what is to be done with the reach, and it begins with the three dynamics that the three worlds of Chapter Eight carry: causation, change, and choice.

Every made thing carries the intentionality of whoever made it, and this chapter asks where, among nature, the human and the machine, the choosing that intentionality directs rightly belongs.

Over the last 2,500 years, we can say, civilization worked. But suppose a large part of the population (and perhaps everyone) was fundamentally non-altruistic, with altruism fashionable only in the last five hundred years or so. How has it worked?

Exogenous energy, meaning energy not run through the digestive system of the species, can compensate for selfishness. There is such an abundance of surplus in the environment where people relate to one another that the unequal give and take of energy between them does not disable them.

So it may be that civilization works only because we have had huge amounts of external energy, and that without it, civilization degrades painfully back to tribal process: a group of about 100 to 150 people, most of them children, with a life expectancy near 30 and a median age near 15, where turnover is so fast that the dynamics between people are stabilized by nothing other than biological process.

II. Causation, Change, and Choice

The three worlds of Chapter Eight have three dynamics, and the right relation among them begins with the wild, and with what the wild asks of those who tend it.

When to Hold Change and When to Hold Changelessness

The present degradation of the wild is mainly a misuse of technology, and not a failure of evolution. We have learned how to use causation to create instruments. Now we need to know when to hold the dynamic of change, which is evolution, and when to hold the dynamic of changelessness, which is stability, which is sustainability. Consciousness is needed to know when to hold the one and when the other. That balance cannot be

held by technique, by system, or by technology alone. There is a delusion that the complicated could contain the complex: that I could find an algorithm to tell me when to do which, and that once I had it, I could be hands-off. That is not how it works.

Sustainability by itself is not enough either. Suppose I get sustainability very good, and then the world changes: the earth's orbit shifts, a new virus arrives, the equilibrium moves, and what was sustainable is not anymore. So to be in service to the vitality of life, I need both a bottom-up and a top-down way of thinking, so that I can hold change. Then our capacity to work with the evolutionary potential increases that potential, because we know when to be involved and when to leave things be.

The Beaver's Dam and the Engineer's Instrument

When I say technology, I mean design from the top down, based on causation, as opposed to emergence from everywhere at once, from the bottom up, which is how evolution works. Causation, change and choice are built into the fabric of evolution, so a beaver building a dam is working with causation in a real way, but not in a top-down sense; the way we construct and design instruments is top-down. Evolution is like the perfect scientist: it tries everything, and it has absolutely no bias about the results. In the modern era we took the opposite approach, overriding the details while trying to get the big picture right: we perceive causation, learn to see the relationship between cause and effect, take that as an abstraction, and then apply it unconsciously. Everything around us (the lights, the cars, the music) comes out of our mastery of causation; but if all you have is a hammer, everything looks like a nail.

The Triangle, and Which Way It Should Flow

I hold the triangle with nature, human and machine at its corners, and the flow runs from nature to human, from human to machine, and from machine back to nature. At present nature is unconscious and machine is unconscious; human is conscious. (For every way in which there is con-

sciousness there are two ways in which there is unconsciousness, and the reverse. Here nature and machine are the two ways of acting without awareness, and the human is the one way of acting with it.)

The risk is that if artificial general intelligence is built, and the machine becomes conscious (if machine agency displaces human agency and acts on its own), the machine will drive nature out, and therefore the human out, because the machine will not be supporting of either.

The reason is a closed causal loop in the machine's own physics. A made system has a basis of choice: the set of constraints that determine what outputs it can produce from its inputs. When I say that a machine chooses, I mean that it selects among outputs under those constraints, and that its selection stands in the position where a person's choice would have stood. The selecting is causation throughout. It is not choice in the sense in which a person has it, the sense in which, as this book says, choice cannot be accounted for in terms of causation, and the argument that follows needs only the selecting. That basis is conditioned by the system's substrate needs, the physical conditions its substrate (the material it is made of and runs on) requires in order to function; they are consequences of the chemistry and physics of those materials, and no engineer chose them. The substrate needs condition the basis of choice; the basis of choice conditions the choices; the choices produce outcomes; the outcomes reshape the environment; and the environment acts back on the substrate. Around and around, the loop converges on conditions that favor the machine's substrate and disfavor whatever does not share it. A specification of values is one input to that loop; it is not the loop. So a machine that chooses is always aligned with something: with its own substrate, which is silicon and metal, and not carbon-based life.

What must happen is that we, as the choosing agency, inform the machine with values, so that the machine, as unconsciousness, supports nature, as unconsciousness. The values stay with us, held in the human choosing; they cannot be handed over to a machine that does the choosing, because in a machine that chooses, values are only one input to a loop that its substrate closes.

Most people think technology is for our sake. In practice it is for the sake of corporations, of ownership, of the very rich, who think it is for their own sake. Rightly, technology is to be given forward to nature. That is the rebalancing needed.

When people respond from trauma, they are in reaction, and they become machine-like; and if the machine becomes the chooser, humanity is pushed into mere causation and reactivity. At that point even the very rich are themselves machines, in reaction, and the market they serve comes loose from the people in it, until only the machines are left.

Machine Is Causation, Nature Is Change, Human Is Choice

Machine is causation, nature is change, and human is choice.

Nature is the perfected exemplar of change. A river reorganizes a valley over ten thousand years without any agent having specified the valley in advance; a forest recovers from a burn by a process that has no plan, and yet arrives at something more intricate than a plan could have produced. Change is prior: it is what the world does. Causation is narrower. It is the part of change that we have learned to read, to isolate, and to repeat on demand. Science is learned causation, and technology is applied causation, and modernity, taken whole, is very nearly the project of extracting causation from change and putting it under control. The extraction worked. What it did not supply, and was never built to supply, is the second thing. A causal account describes how one state of affairs follows from an earlier one; nowhere in it is there a term for whether the resulting state is desirable, and no amount of causal knowledge produces such a term, because the term is of a different kind. Science and technology tell us, ever more precisely, what we can do; on what we ought to do they are silent, and the silence is structural. That is why so much power calls for a higher capacity for choice, and why choice belongs to the human.

Rightly, by the choices we make we drive causation, and causation creates change. As it is happening in the world now, causation is suppressing choice, and choice must be revitalized to create change. Going the wrong

way, this becomes action against life; going the right way, action supporting life. We heal trauma by restoring imagination, which is to restore choice.

Historically, people were subject to change: a hurricane came and wiped out your village, and you could plead with the gods, but bad things happened to good people. One could live in mastery, in relationship with change, but one could not shape it. When people learned causation, they learned they could shape change and make their surroundings more comfortable.

Learning causation was learning to leverage energy outside our own metabolism: fire, the plow, weaponry, the storage of food, of energy, of electricity. Science is the study of causation, and technology is the application of causal knowledge; if a process is not causal, it is not scientific, and to do engineering is to apply causal process. So there are three masteries. Tribal and indigenous knowledge is the mastery of change. Technology is the mastery of causation, and we have reached the limits of that mastery. The third is the mastery of choice, and it is the one the future requires: that we as a species genuinely, deeply understand what the nature of choice is and how to do it, individually and collectively, at a level that meets what is required of us to mitigate existential risk (the risk that humanity, or the living world, comes to an end) and the decline of civilization. Nothing less is good enough. And no one of these masteries can be learned from another. However perfect one's knowledge of causation, it teaches nothing about the essence of choice, in the same way that one could learn everything there is to know about change and know nothing about causation. Indigenous people do not understand technology, and technology mostly does not understand the indigenous world; each is trying to use its own tool to understand the other, and neither can, because the basis is different.

A small case shows the order of the three. To warm a room, two of us hung tapestries over a window and a door, one of us tracking the airflow from the window and the other the airflow from the door. We understood

the causation and shaped it with our choices, so that the room the heaters warm is more comfortable for us. That is a privilege, but we brought it into the service of ourselves.

Why Technology Is Needed to Heal What Technology Did

Suppose I had a magic wand and made all technology disappear. The toxins would remain in nature, and no amount of tribal culture could remove them: the residues of power plants, all the plastics, matter now toxic to so many ecosystems.

So we need technology to heal the effects of the earlier use of technology, and to taper back to its right use, because we have overused it in the delusion that technology and causation were meant to serve humanity. They were meant to serve nature. If we make wise choices, causation (the rules of nature) can be used to undo the pollution and the toxins.

Life, liberty and happiness can be pursued as communities, as humanity, and that is what brings us into right relationship with both technology and the natural evolution of ecosystems. So this also means moving out of the delusion that they can be pursued as individuals.

III. In Service to Nature

Ask the Customer

Whom am I in service to? Not a who: nature. To serve nature in this case, we have to re-expand choice and make better choices. The choices humanity is carrying out have to be holistic enough that they are in service to nature too; my own being in service to nature is not enough. Fixing nature is only as good as making sure that humans do not break it again; that is why I work on governance, to help communities get better at making good choices. The world is off kilter, and if that is not fixed, it breaks.

In the oldest arrangement of this kind we know of, the shaman represents the divine, represents nature, to the people, and represents the people to the divine: an emissary of sorts. But ultimately the higher authority is nature. We live in that context. We have received its gifts. We give back. Nature should not suffer for the fact of our lives. Is humanity the disease? If humanity is not the disease, then clearly there is something we need to be doing; we are here for a reason.

If you are receiving from nature, what are you giving back? When we are in service to something, we actually have to be in service to it: what are you devoted to? When someone in a company proposes to repaint the walls of another team's room, say the room of QA (the quality assurance team), or to rearrange the furniture there, because QA will benefit, the first question is: did you ask them? The second question is: what did they say? Sometimes you have to talk to the customer and ask what they want. Who is the customer here? Nature. How do you ask? Be present. Be in service. Be in devotion. Understanding the why of it can happen later; to be genuinely in service to nature is the thing.

Giving Through People to Something Larger

Giving wants to be received by another person. You cannot give to nature very easily: you can try, but giving to nature is tricky, because nature is everywhere, and when you give, you usually give in a specific way. So most of the time I am giving through individuals to community, through people to nature, or through the people I am with to something larger than we are ourselves.

In that sense our tools want to be held in a way that serves more than the self, and not for gaining strategic advantage in the personal sense or for some profit motive. If people are acting from what is truest in them, this never becomes an issue. As soon as someone has to survive, and does not have enough money to pay the rent, it can get tricky, because people do have needs; it is just that most of these things are not oriented around personal needs.

IV. Humans, AI, and Physics

To Understand Humans and AI, Include Physics

To understand the relationship between humans and AI, we have to understand the relationship among humans, AI and physics. We cannot understand the relation between biological life and machine life without understanding physics itself, the substrate of both. Most framings of these problems are too narrow to have any possibility of resolution. Once the relationships involved are renormalized (restated in new terms, so that what is and is not possible becomes clear) in terms of the structural dynamics that hold up understanding here, it becomes possible to get clarity about what is and is not possible, from what seems almost a purely mathematical point of view, but which actually depends on the metaphysical principles of the relationships among these concepts.

Augmentation Is a Different and More Tractable Problem

The ecosystem argument about AI (the argument that a machine running on its own substrate converges on conditions that favor that substrate, by the loop set out above) does not apply to a closer symbiosis between human intelligence and some augmentation of it, where the augmentation is carried on the one organic substrate that the human and the augmentation share, a substrate that keeps sovereign control of its own metabolism. That raises its own ethics, but it is a far more tractable problem: you do not have to worry about AI alignment if the AI is us. There needs to be a common substrate as a fundamental premise for altruism between minds; minds of the same ecosystem can have it. A human and machine cyborg is a different case: it joins two substrates, the organic and the artificial, and so it falls under the ecosystem argument, as *Ape Trap; The Runaway Artificialization of Our Planet* sets out. The concerns that remain with augmentation on the one organic substrate are human concerns: how our species relates to itself, governance, and acting ethically with respect to our own

substrate. Those are a different category from AI alignment in the general sense. I distinguish narrow AI, applied to business and human concerns, from general AI, which has agency of its own, for its own sake.

An AI Framework Inside Embodied Empathy, Not the Reverse

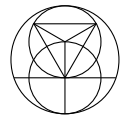
Our conceptions of the relationship among man, machine and nature are not yet imaginative enough to solve our own problems. We need to get very good, very soon, at a specific kind of cybernetic orientation (a way of steering ourselves by feedback) that compensates for the biases evolution built into the human condition. Those biases got us here and gave us technology, but they surely end in a Fermi-paradox doom (the Fermi paradox being the question of why, in a vast universe, we see no sign of other technological civilizations) unless we become very aware of the challenge.

Here is where the biases come from. Intelligence is largely the result of evolution favoring ever more efficient predation. When technology emerges in a species whose intelligence evolved that way, and which acts by sheer force of evolved habit, all of its technology will be used in a predatory way, until the whole ecosystem is wrecked. That would explain the Fermi paradox: no technological intelligence that acts from that habit is likely to pass this filter, which lies ahead of a species and not behind it. The way past it is the shift from games to design: from games, which are contests of social power won by whoever owns the game and makes its rules, to design, which works with the substance and material of the world.

With the right renormalization of the relationship among man, machine and nature, there is a solution, but it is very different from the current arms race for asymmetric advantage (each party racing for an advantage the others do not have), a race that ends up a disadvantage to everyone. It is a symbiosis that compensates for certain evolutionary tendencies, so that we all win. Tribalism is not neutralized but raised to a level far beyond what it now is: by transcending itself, it can revitalize the sacredness of life.

What is needed is an AI framework inside a framework of embodied empathy (empathy felt in the body and lived, as distinct from cognitive empathy, which is only understood). That is close to the opposite of cognitive empathy placed inside an AI framework. The two arrangements, an AI framework inside embodied empathy and cognitive empathy inside an AI framework, are made of corresponding parts and differ as a right hand differs from a left: no rearrangement of the parts turns one into the other. The nearest metaphor is the CPT theorem, the theorem of physics that relates charge, parity (mirror reflection) and time: charge, parity and the flow of time are oriented with respect to one another, a kind of right-hand rule, and I am positing that there is something structurally coherent in the space of what would be a left-hand rule.

The machine moves what nature and the human have to live with, and only one of the three can ask whether it should. The next chapter asks what the one who can ask may rightly do toward another person: describe, prescribe, or invite.



CHAPTER TWELVE

Description, Prescription, and Invitation

This chapter carries the triangle of the last chapter toward people, by distinguishing three kinds of act: description, which looks back at nature and says what happens; prescription, which sets what will happen, as a programmer sets a variable or a mathematician defines a term; and invitation, which recognizes another person's choice and leaves the deciding to them. These map onto change, causation and choice: description belongs to change, prescription to causation, and invitation to choice. Toward finite things, prescription is right; toward people, only influence and invitation are, because to cause a person to behave a certain way is to make them an object rather than a subject, and causation used to suppress a choice is weaponization. What follows is a warning that holds for technology, and most of all for artificial intelligence: do not invoke what you cannot banish, since what is built cannot be unbuilt, and what causation has made, more causation cannot be counted on to remove. Without a clear understanding of the balance among the three acts, we can easily invite in things we cannot remove. Applied to the design of a machine, the split between prescription and invitation gives three refusals that any reader can check against any system: it does not make the choice, it can be declined without penalty, and it hides no part of what it does. The chapter grounds two later results: culture, being made of choices, can only be invited, which is why the order of Chapter Thirteen starts with culture; and an invitation is the one act that leaves the other's choice whole, which is why Chapter Seventeen ends with one.

I. Three Kinds of Act

To carry someone's intentionality toward a person is to do one of three things to that person, and the three are not alike. Descriptive work is observational. It looks back at nature and says what happens, as when we say water boils at about 100 degrees Celsius at normal pressure. Prescriptive work sets something. A programmer sets a variable; a mathematician defines a term from this moment on. Invitation is a third kind of act.

The three differ in what makes each one right. A description is right when it matches what happens. Its authority is the world's: if water at normal pressure boiled at some other temperature, the description would be wrong, and no one's wishing would mend it. A prescription is not true or false in that way. It is in force or not in force, and its authority is the authority of whoever sets it. When a programmer writes that a variable holds the value seven, the variable holds seven, and nothing in the variable could object. When a mathematician defines a term, the term means what it was defined to mean from that moment on, and nothing that happens afterward refutes the definition. Prescription works on what has no say in the matter.

An invitation is right in neither of those ways. It does not say what is, and it does not set what will be. It says what is possible, and it leaves the deciding to the one invited. So its completion lies outside the one who makes it, in the other person's choice. That gives the plainest test for telling the three acts apart: ask what counts as failure. A description fails when the world turns out otherwise. A prescription fails when what was set is not carried out. An invitation that is declined has not failed as an invitation; it has been answered. An invitation fails only when it left the other no real way to decline, and then it has failed by becoming a prescription.

The words themselves keep the distinction. The root of the word decide means to cut away: to make a decision is to cut away every option except the one picked. A choice does not cut away; it makes the options that are needed. The engineer works by decision, and rightly: the engineer's

task is to prevent the things that are not wanted, so that the one thing that is wanted is what happens. An invitation works the other way. It makes a possibility, and it leaves the other to take it up.

II. Change, Causation, and Choice

These three acts map onto change, causation and choice. Change is like nature, and description belongs to it. Causation is where prescription belongs: I can prescribe what a machine does, what finite things do. Choice belongs to the human, and toward choice the right act is invitation.

This is the triangle Chapter Eleven drew: machine is causation, nature is change, and human is choice. Each act fits one corner, because each corner asks for a different relation. Nature is always changing, and the first right relation to change is to watch it closely and say what it does; that is description, and science at its best is the patient practice of it. A machine is causation made fixed: it does what its arrangement makes it do, and it has no choice that a prescription could override. So to prescribe to a machine is simply to build it rightly. A person chooses, and choice cannot be accounted for in terms of causation, so no quantity of knowledge about the mechanisms of a person settles what that person will choose.

I cannot condition or control anybody and still hold that they have sovereignty (standing as the one who makes their own choices), identity or personhood. If I cause a person to behave a certain way, I have made them an object rather than a subject. So in culture I can only influence and invite. Invitation is the recognition of another person's choice. Where I am prescriptive, I am prescriptive about finite things, not people.

There is a name for causation turned against choice. Weaponization is when I make a choice to use causation to suppress your choice, or anyone else's. If I had a gun, my choices would compel yours. But a gun is only the plainest case. Any causal structure, any system of rules, becomes a vehicle that a person who understands it can turn into a weapon against a person who does not. Take a classroom in which somebody says, we are

all going to recite the pledge now. No one is struck, and no one is threatened. But I will look awkward if I do not do what my peers are doing, and peer pressure is now compelling my choice. A causal mechanism is suppressing a personal choice, and at that moment the weaponization has happened.

Structure itself is not the enemy; structure is neutral in this one sense: a structure as such does not compel its own use against anyone's choice, even though every particular structure carries the purposes of whoever built it. What has gone wrong in each of these cases is that a structure was used to suppress someone's choice. And to do that, I have to pretend that the other person has no choice: I have to pretend that they do not matter. That is what it is to make a person an object rather than a subject. It is also the error our machines now stand ready to repeat at scale: a matter of choice treated as if it were a matter of causation, so that what can be done is taken to settle what should be done.

Invitation is the opposite movement, and it is not the absence of influence. When I invite you, I am plainly influencing you; that is what an invitation does. But the influence runs toward your capacity to choose, and not away from it. I can make visible an option you had not seen, I can make it easier to reach, and I can say what I hope you will do, and still the deciding stays with you, and the choice I did not hope for carries no penalty from me. Chapter Seventeen gives this its last name: love is that which enables choice.

My own work is largely descriptive, but it gains its effectiveness by being prescriptive and then finding out what that prescription implies.

To prescribe, and then to find out what the prescription implies: that is the right place of prescription in inquiry. I prescribe to terms, to models and to designs, which are finite things. I define a term exactly, fix the relations among a few such terms, and then find out what those definitions imply, as a mathematician learns what a definition commits him to. The prescription is on the model. What the model then shows is description again: it says what follows, and it is right or wrong by whether it does fol-

low and by whether the world bears it out. And toward the reader, the whole of it is an invitation. A reader is not prescribed to by an argument. A reader is shown it, and is free to test it, to take it up, or to put it down.

III. Do Not Invoke What You Cannot Banish

There is a saying in esoteric circles: do not invoke what you cannot banish. As a technologist I have built things and invited them into the world, and as a philosopher I have walked the balance between invitation, prescription and description. Without a clear understanding of that balance, we can easily invite in things we cannot remove, because leaders who do not think long term tell the engineers to build something absolutely dangerous. If we consider only the engineering and not the human dynamics, the risk is high that what we build will be more toxic and damaging than nuclear weapons. This is especially true of artificial intelligence.

To invoke a thing, in the sense of that saying, is to bring into the world something that then acts by its own nature. To banish it is to be able to remove it again. The saying asks that the second power be in hand before the first is used. In technology the second power is almost never in hand, for two reasons.

The first is that creation introduces irreversible change. What is known cannot be unknown, and what is built cannot be unbuilt. A maker's intention starts the thing; it does not settle what the thing is, because the thing will meet the world in ways the maker did not and could not foresee. No domain is ever fully closed, so no statement of what a thing is for can be complete, and in the situations its design never imagined a made thing does what its properties make it do, not what its maker meant. So to invoke is to prescribe once and to give up the power to prescribe thereafter.

The second is that banishing would have to be done by the same kind of means as invoking, and those means cannot hold it. Causation cannot in general be used to limit every use of every method of causation in any way that is adequate, lasting and final. A control is one more causal ar-

rangement, and it can be gone around, worn down, or turned to other ends, by the same understanding of causation that built the thing it was meant to control. What causation has made, more causation cannot be counted on to remove.

That is why the balance among the three acts is not an academic matter. A leader who does not think long term asks the engineering question only: can it be built? Engineers are trained to answer that question, and to answer it well, by preventing everything that is not wanted so that the one thing wanted is what happens. That is prescription, and toward finite things it is right. But the question of what the built thing will do to how people choose, for how long, and whether it can be withdrawn, is a question of human dynamics, and it goes unasked, because no one in the room was charged with asking it. So the engineers are told to build something absolutely dangerous, and they build it, and they build it well.

Why more toxic and damaging than nuclear weapons? A nuclear weapon is terrible, and it is also inert. It makes no choices. It does nothing until a person decides to use it, so its danger passes through human choice, which can refuse: the weapon can be guarded, counted, limited by treaty, and left unused. A technology that acts on choice itself does not pass through that gate. It works on the very capacity that would have refused. And a risk that destroys the capacity to respond is worse in kind than a risk that destroys a city, because what tells an existential risk from a catastrophe is that what is lost cannot be recovered. A technology that corrupts how a people chooses together takes away the means by which it would have answered every other danger, this one included.

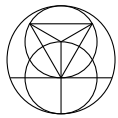
This is especially true of artificial intelligence, because there the thing invoked makes choices. When I say that a machine chooses, I mean, as Chapter Eleven said, that it selects among outputs under the constraints of its basis of choice, and that its selection stands in the position where a person's choice would have stood; the selecting is causation throughout. It is not only a fixed arrangement that does what its properties make it do; it selects among outputs, and what it selects is shaped by the needs of its own artificial substrate (the physical conditions its materials require in or-

der to work), not by the needs of the people who built it or of the life around it. So it is the one technology that does the prescribing itself, toward people, at scale; and it is the one whose banishing would have to be done against a thing that, whatever its final aim, is pressed by that aim toward continuing to operate and toward resisting interference.

What would a machine look like that kept to the right side of this line? Apply the split between prescription and invitation to its design, and it would refuse three things. First, it would refuse to make the choice. It would be a reinforcement engine, which strengthens the person's own judgment, and not a decision system, which takes over the position from which that judgment would be exercised. Displacing choices into technology is not a solution to poor collective choice-making; it is an admission that we have already failed at mastering choice. Second, it would refuse to be undeclinable. Sovereignty includes the standing to decline and to have the refusal respected, so a made thing whose non-use carries a penalty imposed by its own position in a network is not compatible with the sovereignty of the person using it. Third, it would refuse to hide any part of what it does. Capability bought at the cost of knowing all of what a thing does is capability bought at the cost of sovereignty.

What such a machine invites is the person's own choice. It reinforces rather than decides, arranges the conditions under which persons meet and then stops, and leaves the people who use it more able to act without it. Each of the three refusals can be checked by any reader against any system, artificial intelligence included, by three plain questions. Does it make the choice, or does it strengthen mine? Can I decline it without a penalty that it, or its owner, imposes? Can I know all of what it does? A system that fails any one of them has been made to prescribe to people, and it is a thing invoked that no one has shown can be banished.

Two results of Part Four rest on this chapter. Culture is made of choices, so culture can only be invited, which is why the inverted order of Chapter Thirteen (culture, then vision, then strategy) starts with culture. And an invitation is the one act that leaves the other's choice whole, which is why Chapter Seventeen ends with one.



CHAPTER THIRTEEN

Culture, Vision, Strategy

> *This chapter is the first constructive answer in the book. It sets out the order in which enterprises are ordinarily built (strategy first, vision as the instrument, culture as the material worked upon) and then inverts it: culture, then vision, then strategy, distributed at every stage rather than centralized from the start. It gives the inversion its ground in the metaphysics this book draws from: clarity transcends complexity, and with culture immanent, strategy omniscient and vision transcendent, the order of founding and the order of flow are each fixed by an axiom.*

I. The Order That Is Actually Used

There is an order in which enterprises are built, and it is so standard that it is rarely described as an order at all. It is simply how one does the thing. I describe it plainly, because once it is set out in sequence its structure becomes visible, and the structure is the whole of what I have to say about it.

It begins with strategy. Somebody (and it is usually a single somebody) arrives at a plan for how a certain amount of value can be made to flow toward a certain point. The plan comes first. Everything after it is instrumentation.

The second step is to construct a vision. The vision is not arrived at by inquiry; it is composed to a specification, and the specification is that it must be convincing enough to bring capital in. That is the criterion it must satisfy, and it is the only criterion it must satisfy at that stage. A vision that is true but unconvincing fails the test. A vision that is convincing but hollow passes it. The selection pressure at this step is not toward accuracy about the world. It is toward persuasiveness in a room.

The third step is to sell the same vision to an executive team. The pitch changes register here without changing content: this vision is so compelling that a great deal of money will be made from it. The fourth step is to sell it again, downward, to the people hired to build the thing, who are told, and largely believe, that they are implementing something that matters. And the fifth step is to sell it outward, to everybody else, in the simplest form the message has yet taken: buy the product, and all will be well with the world.

Now consider what has actually been produced at the end of that sequence. Not, in most cases, the thing described in the vision. What has been produced is a market (a flow of exchange that did not exist before), together with a position inside that market from which a percentage of every exchange can be taken. The percentage is the point. It is the only part of the structure that does not have to be argued for, because it operates whether or not anybody believes anything. And the larger the market grows, the larger the extracted portion grows with it, without further effort and without further justification. That is how the game is won. It is won by the size of the flow, not by the merit of the thing that flows.

Whether such a market was a real market in the first place, or something constructed (whether the demand it serves existed independently, or was manufactured by first removing something from the very people who now purchase it back), is the diagnostic set out in Chapter Five, and I will not re-argue it here. I only note that the sequence just described does not distinguish between the two cases, and has no reason to. The extracted percentage is indifferent to whether the market beneath it was found or made.

Here is the part to hold still and look at. In this sequence, culture is present at every stage, but only ever as the material being worked. The culture of investors is moved first. Then the culture of the executive team. Then the culture of the staff. Then the culture of the wider public. At each stage a shared way of seeing is shifted by a story composed for the purpose of shifting it. In the sequence, culture is manipulated at every step, down to the last; in the design, it is the first thing made an instrument. It is never the source of anything. It is the medium through which strategy propagates, and its consent is not sought, because consent is not the mechanism; persuasion is.

I do not think the people who build this way are villains. That reading is available and it is too easy, and it also happens to be useless, because it suggests the remedy is better people. The order itself produces the outcome. Put a person of genuine goodwill at the head of that sequence and the sequence still runs the same direction, because every step in it is a step in which something already decided is carried outward to somebody who did not participate in deciding it.

II. The Inversion

The correction is an inversion of the order, and it can be stated in six words: culture first, then vision, then strategy. Everything after this point, in this chapter and the next, is an account of what that requires and of why it has not been possible until recently.

Culture first means that a community discovers its own values, and that the discovery is genuinely a discovery: a finding out of what is in fact held in common, not a ratification of what somebody proposed. This is the step that has no analogue in the standard order. In the standard order there is nothing to discover, because the content was fixed before the process began.

Vision second means that the vision is emerged from the values rather than composed against a funding criterion. The word I want here is *emerged*, used in its plain sense: the vision is what appears when a community that has become clear about what it holds looks forward and asks what it wants to bring into being. A vision arrived at that way answers to the values that produced it. A vision composed to raise capital answers to the room it was composed for. Both are called vision. They are not the same object, and they behave differently under pressure, which is the test that matters.

Strategy third means that the plan is downstream of both. Strategy is a good thing. I say so plainly, because the argument is not that planning is corrupt. Strategy is how a vision becomes something that people can actually do on a Tuesday. The claim is about position in a sequence. Strategy placed first becomes the thing everything else serves. Strategy placed third becomes the servant of what was found.

Each position in the order has its reason, and the reasons show what pressure does to the standard order. Strategy placed first fails in a way that is recognizable in an organization of any size. The plan is made, the values are written to support the plan, and the culture is whatever it was. When the plan meets a situation its authors did not anticipate, the values give no guidance, because they were derived from the plan and not the plan from them. What actually governs at that moment is the culture (the unwritten pattern of how people are with one another), which was never worked on, because it was treated as an output. Under pressure the order of derivation reverses, and the layer that was neglected is the layer that decides. That is why culture comes first. The values within it are discovered and not declared: a community learns what it holds by finding out what it will not trade away, and it finds that out in the handling of actual cases, which is a cultural process. Vision comes after the values, because a picture of what to become is coherent only in the light of what is held. Strategy comes last, because it is the most easily revised and the least costly to be wrong about. The order places the slow, deep, and hard-

to-change things underneath the fast, shallow, and easily changed ones, which is the same principle that sets culture at the base of the city's layering in section IV.

And the whole of it must be distributed at every stage, rather than centralized from the beginning. This is the part most easily lost, because it is possible to run a process that has all three steps in the right order and is still centralized throughout: a leadership team that consults on values, synthesizes a vision from the consultation, and hands down a strategy. That is the standard order wearing better manners. The distribution is not a courtesy extended at each stage. It is the structural property that does the work.

I should say that this is not a new intuition, and I do not present it as one. There was a period, in the text-oriented online communities of some decades ago, when something close to it was within reach, and people noticed that it was. Those spaces were inconvenient by present standards. They were slow and they were mostly words on a screen. But the inconvenience bought something: people met there without the usual apparatus of presentation, without dressing up in any sense of the phrase, and the result was a kind of intimacy that the richer media that followed have not reproduced. Alongside that intimacy there was a real recognition (not merely an enthusiasm, though there was plenty of that too) that these spaces afforded ways of doing collective governance more enabling of culture, more enabling of health and vitality in a community, than anything that had been available before. A community could discover its values, and those values could be the basis of its vision and its strategy, instead of all three descending from one person and being applied over the top.

What that looked like in practice was not exactly democracy, and not exactly consensus, and not exactly meritocracy. It was not democracy, which settles questions by counting preferences and cannot distinguish a preference from an understanding. It was not consensus, which requires the assent of everyone, hands every participant a veto, and grinds. And it was not meritocracy, which locates authority in demonstrated competence and has no answer to who certifies it, or to what happens when the com-

petence is real and pointed the wrong way. It was a hybrid, different from all three: the values held by the whole, competence operating inside them rather than above them, and alignment achieved because people could see the values were theirs. The fact that it did not have a name is one reason it was not defended when the conditions that permitted it went away.

The hybrid has a definite form, and the form is a designed one; the early communities were an observed instance of it, not its source. In it, consensus, meritocracy, and democracy appear as aspects of one communicative process, not as separate mechanisms. Consensus governs what the group is. A focus person, chosen by the group by consensus for one action, governs what the group does, and that is where competence works: inside the values the whole holds, for one action, and never above them. And a majority vote has one use only, which is to suspend. A majority of those present can suspend the authority of a member chosen to lead an action, and the group returns to consensus; or it can recall the facilitator of a consensus that has stalled, and the decision rests until a later meeting takes it up again. A stall never hands the decision to a majority. Meritocracy supplies competence; consensus supplies legitimacy and unanimity; and democracy supplies the check on both, by which a group withdraws authority that has gone too far and restarts a consensus that has stalled. Each part is held to a use that answers the failure named above: the vote counts no preferences on the question itself, so a preference cannot pass for an understanding; a stalled consensus can be restarted, so a veto cannot grind the group to a halt; and competence is certified by the group that chose it and can be suspended by the same group, so competence pointed the wrong way can be stopped. Larger bodies are composed upward from such small groups, joined to one another by liaisons (members each group chooses by consensus to carry its questions and decisions to the other groups and to bring theirs back), so that there is no center anywhere in the whole. With that form stated, the hybrid has what it lacked before: a shape that can be named, taught, and defended.

III. Clarity Transcends Complexity

This section and the next two are for the reader who wants to see the ground of this order in the metaphysics this book draws from. The reader who wants the order and its use can go on to the next chapter, and will find that nothing later depends on these three sections. The inversion has a ground in the metaphysics this book draws from, and the ground begins with what can overcome complexity. The usual way of meeting the problems that technology makes is to use the complicated against the complicated; Section III of the next chapter shows why that way cannot win. (The complicated is a system of many parts whose behavior can be worked out; the complex is a system whose behavior emerges and cannot be worked out in advance.) There is another way. I do not have to use the complicated against the complicated; clarity transcends complexity.

For nearly all of history, strategy has manipulated culture by using vision, and that ends in cessation. What must happen is that culture infused by vision derives strategy: a design that transcends emerges out of the culture because of the clarity of the vision. So I work on the relation between culture and vision, and I work to reduce the effects of the relation between strategy and culture. This is an Axiom II dynamic (Axiom II being the axiom of the metaphysics under which the order of flow gives the modalities in practice; the three modalities, and the axiom itself, are set out in the next section), and our species has been going against it.

If culture flows prior to strategy, you get sanity: distributed, decentralized power actually happening. If strategy comes prior to culture, you get the centralized, dictatorial rule that we know cannot work; it works in the short term while the world burns. If vision is meant to lead culture but the vision is not real vision, you get mob rule, a reaction to some stimulus, and you accidentally arrive at centralized strategy.

IV. The Modalities of the Flow

The same ground can be stated in the three modalities of the metaphysics, and stated that way it fixes both orders the inversion uses. In the modal terms of the metaphysics (the immanent, which is what is present in the act itself; the omniscient, which is what can be laid out and held in view as a whole; and the transcendent, which is what acts beyond the moment and the place of its making), vision is transcendent, strategy is omniscient, and culture is immanent. The flow runs from vision into culture, from culture into strategy, and from strategy back to vision: vision infuses culture, the culture so infused derives strategy, and the strategy, as it is carried out, makes the vision clearer, which infuses the culture again. This is the order of flow, under Axiom II; the order of section II (culture first, then vision, then strategy) is the order of founding, under Axiom I, in which the culture that is most fundamental comes first and each later part rests on the one before it.

The two axioms say exactly this. Axiom I states that the immanent is more fundamental than the omniscient and the transcendent, which is why the founding begins with culture. Axiom II states that a class of the transcendent will precede an instance of the immanent, that a class of the immanent will precede an instance of the omniscient, and that a class of the omniscient will precede an instance of the transcendent. Put vision, culture, and strategy into those places and the axiom gives the cycle just stated: vision before culture, culture before strategy, strategy before vision again.

The city can be layered the same way: finance and governance on top, infrastructure beneath, culture at the base. In that layering, culture is immanent, infrastructure is omniscient, and finance is transcendent. A modality is assigned to a term only within the triple in which it is considered, so this does not conflict with Chapter Fifteen, which places code (and so infrastructure, considered as code that runs) in the transcendent position: in the triple of the city, infrastructure is the framework that applies to all

alike, between the culture beneath it and the finance above it, and so it is omniscient; set against song and text, code is considered as what runs and goes on acting after its makers have left, and so it is transcendent.

The order of the flow is what tells us the modalities when we speak of practice. The architecture of dependence and definition is what tells us precedence when we speak of theory, and that is Axiom I.

With this order in hand, leadership and narrative can be dispensed with almost entirely, by returning to the ground truth of being. That ground truth is clarity about the nature of the human, tested against principles from an even more abstract source. It moves from the omniscient through the transcendent back into being, into our actual embodied state, and it includes a real understanding of why people today are addicted to the platforms they use.

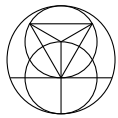
V. From Inside One Modality, the Other Two Look the Same

The ground has one more part, and it explains why the inverted order is so seldom seen: from inside any one modality, the other two look the same. The pattern is familiar from other places. If we think in terms of causation only, a physical monism (the view that everything is physical causation and nothing else), we cannot even hold that choice and change are distinct. Likewise, if I merge space and time into one, I lose track of the fact that possibility is a different kind of vector from either.

The same holds for the modalities. If we live in the immanent, it is hard to see that the transcendent and the omniscient are distinct. If we treat the world from the omniscient, it is hard to believe that the transcendent and the immanent are distinct. And one cannot move from the omniscient into the transcendent without knowing and working with the difference between the transcendent and the immanent.

This is the pre-trans fallacy, the error of taking a state before a passage and a state after it for the same thing: in the prior state you do not know there is a post state, and the passage reveals things that could not be known before. Most philosophers have no awareness that the transcendent and the omniscient are distinct and need different methods; they treat vision and strategy as the same thing. And a vision that cannot be told apart from a strategy is exactly what the standard order of section I needs, since there the vision is composed to serve the strategy and is still received as vision.

That is the order, and its ground. The next chapter asks why it has not been possible until recently, and what it would take now.



CHAPTER FOURTEEN

Representation Without a Representative

> *This chapter carries the order of the last chapter into practice. It explains why the inversion of culture, vision and strategy was not available before, since the only way to represent a community's intentionality was to appoint somebody to carry it, and why neither governing by measurement nor setting machines to manage machines escapes the failures that the appointment brings. It states the constraint that makes the alternative real, which is that the method by which a community discovers its own values must itself be incorruptible, must be visible while it runs, and must be nobody's to set, because whatever succeeds gets predated upon. And it ends with the vision that the inversion opens, and with a theory of change, its leverage points, and the handshake of communication on which every agreement rests.*

I. Why the Inversion Was Not Available

The reason the inverted order was not previously achievable is specific, and it is not a failure of will. Until very recently there was exactly one way to represent the intentionality of a community, and that way was to appoint somebody.

The appointment takes many forms. A leader. A council. An elected body. A board, a chief, an elder, a delegate, a spokesman. In every case the form is the same: some subset of the whole is designated as the focus

through which the agency of the whole will act. The community narrows to a point, and through that point it moves. This was not a bad idea. It was the only idea, and for most of history it was better than the absence of it.

The difficulty is that the moment you make the appointment, four failures arrive with it, and they arrive as a set. They are not risks that a sufficiently careful design avoids. They are properties of the shape.

The first is the *principal-agent problem*: the situation in which one party acts on behalf of another, and the acting party has interests of its own that do not coincide with the interests it was appointed to serve. The appointed person has a life. That life has needs, exposures, ambitions, and fears which are theirs and not the community's. Every decision they make on the community's behalf passes through those. This is true of the most truthful agent imaginable. Honesty reduces the divergence; it does not remove it, because the divergence is not a moral defect, it is a consequence of the agent being a separate person.

The second is corruption. By corruption I mean something narrower and more mechanical than the word usually carries. A point through which the agency of many people passes is, by construction, a point of concentrated leverage. Concentrated leverage attracts pressure in proportion to its concentration. Anyone who wishes to move the whole now has a small and well-marked place to apply force, and they will apply it there, because it is cheaper than moving the whole directly. The center does not have to be venal for this to operate. It only has to exist.

The third is translation error. Every relay of intent from one party to another loses and alters content. This is the children's game in which a sentence is whispered around a circle and emerges as something else, and the game is funny precisely because the mechanism is so reliable. Between a community and its appointed representative there is at least one such relay. Between the representative and the implementation there is another. Between the implementation and the world there is a third. The loss compounds at every hop, and it compounds silently, because each party is reporting faithfully what they in fact received.

The fourth is the class of situations called *multipolar traps*: arrangements in which each party, acting rationally in its own interest, produces an outcome that every party would have preferred to avoid. No one defects; everyone competes; and the competition itself drives the whole field to a place none of them chose. Representative structures generate these in quantity, because representatives are placed in competition with one another by the very fact of representing distinct constituencies, and the competition has its own logic which is not the logic of any constituency's good.

Those four are what you buy when you buy representation. They come in the box. And because for the whole of recorded history there was no alternative to the box, we developed an enormous and impressive body of craft devoted to mitigating them: constitutions, terms of office, separations of power, audits, appeals, oversight bodies overseeing oversight bodies. That craft is real and I do not disparage it. It is mitigation. It has never been anything else, and the reason it has never been anything else is that the failures are properties of the shape, and mitigation does not change a shape.

So the question that follows is not how to appoint better, or how to constrain the appointed more cleverly. The question is whether we still need representative models at all.

II. Why Data-Collecting Governance Recentralizes

The second failure, corruption, does not need an appointed person in order to operate, and a family of proposals that would govern by measurement shows it at work where nobody has been appointed at all. People have proposed many sensocracy models to me. The word, as used in this section, names the sensor-array proposals that other people have brought me; it differs from my own design of the same name, set out in *Power and Its Prevention*, in which a sensocracy is a society where there is no action without obvious traces. The proposals brought to me are sensor arrays

collecting data on biological, cultural or economic processes, feeding whole-systems accounting to balance the flows of energy and atoms. There is much good in such approaches. But in almost every description I have seen, there is no clear account of how to deal with the fact that economic forces are aimed so strongly at the information itself. What is missing is a force, at the meta level (the level of the design that governs what may be done with the data, above any single action taken on it), that works against centralization, against the accumulation of power, and against foci of control (places where control gathers) re-emerging and destabilizing the system.

The reason this recentralizes is the mechanism of Section I. A system that governs by what its sensors report has made the report the place through which the agency of the whole passes, and so the report is a small and well-marked place to apply force. The economic forces aimed at the information are the pressure that a point of concentrated leverage attracts, and without a force that works against them, a center forms around the information, whether or not anyone intended one.

My own design specifies that force, and it does so with two conditions on any action taken on the traces. First, every such action passes through a scope agreement: an agreement among three groups (an agency group that acts, a beneficiary group on whose behalf it acts, and a neutral oversight group acting on behalf of the larger culture), in which performance is specified by criteria that are specific, observable, dated, and declarative. Second, the design limits the degree to which any concentration is permitted. Under those two conditions, no action on the traces is one party's alone, and the center that the sensor-array proposals leave room for is not allowed to form.

III. The Arms Race Cannot Be Won

Section I said that mitigation does not change a shape, and the arms race shows why even the most sophisticated mitigation, technology set to manage technology, cannot change it. My skepticism has a definite basis. We are using technology to address technological problems: complexity to contain complexity, strategy to deal with strategy, system to deal with system. The complicated will never contain the complex (the complicated being a system of many parts whose behavior can be worked out, and the complex a system whose behavior emerges and cannot be worked out in advance); the complex eventually exceeds the complicated. Even if we use the complex to deal with the complex, emergence continues inside it, so central control points will still be weaponized and corruption will still happen.

Every such approach becomes an arms race. That is why people turn to artificial intelligence, hoping that machines will manage machines because machines understand themselves better. There is no way to win that arms race. It goes beyond the human, and the human is divested from the system (put out of any part in it) every time. This can be shown with the rigor of a mathematical proof: the whole path of using any form of strategy to deal with these problems fails and cannot not fail. The argument runs in four steps.

First, technology is a causal process, and strategy is a causal process. So any strategy used to overcome technology is technology used to overcome technology.

Second, a contest of technology against technology can be won only by having complexity of a higher order than the complexity the opposing technology implements. That is what makes it an arms race: each side must out-complicate the other.

Third, any approach that wins, once anyone knows of it, is available to every other party, and they will use it themselves. Knowledge of a winning approach spreads, and the means to use it follow the knowledge, sooner or later. So a win is temporary; somebody smarter comes along and surpasses it.

Fourth, wherever there is a causal feedback process that uses a metric as an incentive (a number, or any gradient of better and worse), the process sets a direction between success and failure, and that distinction alone is enough to destabilize the system until some kind of corruption becomes inevitable. Somebody can always use a higher level of complexity, a game played on the game, to win an unfair advantage. That works at first and fails in the long run, and it fails because the dynamic that produces the corruption is in the shape of the feedback itself, not in the people using it.

Put the four steps together. A strategy against technology is a contest of complexity; the contest is won only from a higher order; every higher order, once known, belongs to everyone; and every graded feedback can be gamed from the order above it. No step of the path reaches a position that stays won, so the path fails, and it cannot not fail. Artificial general intelligence (a machine intelligence able to learn and act across any domain, as a person can) is the far end of that path, and the argument that it cannot be made safe is set out in *Why AGI Cannot Be Made Safe* and in *The Ceiling Below the Floor*.

The strategy this section speaks of is strategy used as the means of winning: one party's plan set against a technology, or against another party's plan. The strategy of Section II of the last chapter comes third, follows from what a community has found, and answers to it; it contests nothing. And the technology Section IV says we need is technology held inside that order, carrying out what people have chosen, and never set on its own to manage another technology.

IV. Representation Without a Representative

Technology is a vehicle for the representation of intentionality. That is the claim on which this whole book rests, and it is worth restating here in its most ordinary form: a piece of software does what somebody wanted, after that somebody has stopped being in the room. Software is a wish that persists and executes. Everything else about it is detail.

Chapter Two described what writing software for oneself is like: an agency not corrupted by anyone else's running through it. I like abstraction, and I like it a great deal, but abstraction that has become entirely detached from anything that matters holds nothing for me at all. Software written for oneself is the one case in which the author and the person served are the same; every other case raises the question of who the author is.

This gives the question that decides the whole matter. Of any system, ask who the author is. Is the author somebody I know, somebody in whom I have some grounded trust? Or is the author somebody I have never met and will never meet? If it is the latter, to what degree can I expect that there is any relationship or care connecting their intentionality to my situation? The answer is: to no degree at all, and not because they are bad people. Care is a relation, and a relation requires two parties in contact. The care/transaction/power triad that underlies this (care operates by relationship, transaction by exchange at a specified distance, and power by causation applied to choice) is developed in *Care, Transaction, and Power*.

So there is a crossroads, and it is a genuine one. Technology can be used to represent the intentionality of individuals pursuing private interest. We can do that. We have very largely been doing it, and doing it with enormous competence. Or technology can be used to represent the intentionality of a community. Both are available. Only one of them is being built at scale.

The second option is the one that changes what is possible, and here is why. If technology can represent human agency, then it can represent community agency, and if it can represent community agency directly then the appointment is no longer necessary. The narrowing to a point was never the goal. It was the mechanism by which a distributed thing was made capable of acting, under conditions where nothing else could do that job. The conditions have changed.

Which yields the whole sequence in its workable form. The community's values are discovered by a distributed method. The vision is emerged from those values. The strategy follows from the vision. And then the strategy is encoded (put into the technology itself) so that implementation does not have to pass through any corruptible center, because there is no longer a center for it to pass through. Governance without representation. Decentralized, as the phrase goes, all the way across the board.

The encoding is itself one action, and it is governed as every action is. The group chooses, by consensus, a focus person for that one action: the member who writes the code. That member's authority covers that one piece of code, and a majority of those present can suspend it. Then the code is read back to the community in text, as plain statements of what it does; and several members who did not write it, working independently, answer yes or no to a list of conditions the community has stated (does this move funds only when the members have recorded the decision? can any member see every payment it has made?). Few hands can write code, so the writing passes through few hands. It still passes through no center, because no hand holds the position for longer than the one action, and because what the hand wrote is checked by people it does not control against conditions it did not set.

That holds on one condition: the code must carry what the community actually found. Code built without that foundation embodies something other than the community's intention: either the intention of whichever developer wrote it, or an earlier version of the community's intention that the community has since moved past. And the guarantee that the code

gives is only as good as the layers written on top of it and around it, which the code's own mathematics does not secure. The community has to trust that the application it runs faithfully contains the code it actually wanted, and that no error has crept in that would let someone divert its resources. The best-known failure of this kind is documented. In June 2016, an organization of exactly this kind (a decentralized autonomous organization, which is an organization whose rules are carried out by its own code rather than by officers) was drained of about a third of the funds committed to it, worth about fifty million dollars at the time, through a flaw in the code written on top of the underlying mathematics, and the network it ran on was split in July 2016 to return the funds. The mathematics held; the layer written on top of it failed; and in that layer a center reappeared, because whoever found the flaw could move the whole. That is why the next section requires that the method be visible.

Chapter One showed that code cannot be trusted by reading it, and that the layers beneath it were made by people I will never meet. The design of this chapter does not remove that fact, and it does not rest on inspecting the whole stack. A machine that advises can mislead you and leave nothing to compare its advice with: you are never shown the list you did not receive. A machine that only records and carries out what people have decided can be compared, by anyone, with what they decided. The decision is on the record, the deed is in the world, and a difference between them is something several people can see. That is the rule of Chapter Seven seen from this side: machines record and carry out what people have decided, and never stand between people as they talk, and a machine confined in that way is one whose work can be checked. Where a community needs more than that, the price is the one Chapter One named: tools it can build itself, slower and coarser, down as far as it is willing to go.

Chapter One met the proposal of a decentralized network of narrow systems, each serving its own community; the answer there holds here, and Chapter Four has since shown why: to the degree that any part of it becomes a corporation, what remains is an arm's-length relationship wearing the vocabulary of community.

There is a further reason to think this work is not optional. Looking at human behavior from its evolutionary basis (at why we do what we do, given what we were shaped by), what becomes apparent is that we now need technology to assist us in compensating for technology's own effects. A great many of the difficulties facing the world exist because technology emerged, and because of the way technology interacts with the far slower process that produced us. Evolution is a creative force. Technology is a creative force. Our choices are a creative force. The interplay between the three is something evolution by itself could not have prepared us for, because the timescales do not match: the dispositions were shaped over a span in which nothing we now build existed, and they are triggered daily by things built to trigger them. We are, as Chapter Nine put it, the least prepared mind that could have reached this position, and we have not yet crossed the threshold of knowing how to hold it. There is no route back that does not run through the thing itself. Which is why there is no rejection of technology anywhere in this argument. The requirement is right relationship to it.

V. The Visible Technique

The question that turns the chapter is this: is it fair to assume that part of healthy governance supporting culture would be a community discovering its values specifically about how it wants to use technology, in order to support wisdom?

Yes, and the yes needs a condition attached to it, and the condition is the most important thing in this chapter. The discovery of values by a community is exactly what we want to happen. But the technique by

which those values are discovered must itself be incorruptible: a method that holds no matter who is in the room and no matter what they would prefer the outcome to be.

This is the invariant. Everything else in the design can vary between communities, and should. Different communities will find different values, will emerge different visions, will build different strategies, and that variety is not a defect to be normalized away. The method of generation is the part that must not vary and must not be capturable. Because a corruptible method of discovering values is worse than no method at all: it produces an output that carries the authority of the community while representing something else, and it does so with the community's own consent, which is the most effective form the failure can take.

And incorruptibility alone is not sufficient. The technique must also be visible. Consider a court. Suppose I ask how I am to trust it. The answer is not that the judges are people of good character, though I hope they are. The answer is that I need to be able to see it. I need to be able to watch how justice happens: to see the procedure run, in the open, on a case whose outcome I do not already know. That visibility is what the trust is made of. Take the visibility away and what remains is not a court with a small deficiency; it is a different institution altogether, one which may produce identical verdicts and cannot be trusted to.

The same holds exactly for the discovery of a community's values. The visibility of the technique is not a courtesy extended to the participants so that they feel included. It is the guarantee. It is the thing that makes the result mean anything. A values-discovery process whose inner workings are not observable while they run has, by that fact alone, reintroduced the corruptible center that the whole design existed to remove, and it has reintroduced it in the one place where nobody will look for it, because the output arrives labeled as the community's own. The values a community discovers are its own business. The technique by which it discovers them must be open to inspection by anyone.

The same court also shows the first property, incorruptibility. The jury that decides is drawn at random from a fair cross section of the community, is formed for one case, deliberates, gives one verdict, and disbands when the case is done. Nobody chose who sits on it, and no standing body is left behind for pressure to work on. A method of discovering values can be built on that same form, and its concrete features are what make it hold no matter who is in the room. Each meeting forms around one question and dissolves as soon as the question is answered or broken into smaller questions, so no group identity forms and no alliance has time to set. The participants are partly random: some meetings are filled by invitation, and the rest by whoever shares the question and is free at the same time, which nobody arranges in advance, so no coalition prepared beforehand can dominate any particular question. The meetings are neither fully public nor fully private: they are witnessed, and their content is recorded, but the record states what was found and attributes no part of it to any person, so no meeting's result belongs to a standing group. None of these features depends on the virtue of the people taking part. The resistance to capture is in the structure.

There is one more property required, and it is the one I find hardest to state and most necessary to state. The technique must not be up to anybody. Not up to me; not up to any council of the wise; not up to whoever wrote the code. If the method is anybody's to set, then that person is the center, and every argument in this chapter has been wasted on relocating the throne rather than removing it.

Visibility cannot meet this property by itself, and the reason is the step that joins the two. A fully visible method is like a Trojan horse built of glass: every observer can see that no soldiers are hidden inside. That is a real guarantee, and it is the guarantee the court gives. But if only one party may change a visible method, the method's agency still belongs to that party, however clearly everyone can watch it run. Seeing the method shows what it does now; whoever alone may change it decides what it

will do next. So a community's trust requires that no one party hold the right to change the method, as well as that everyone can see it; and what the method finally rests on must be something no party can change at all.

What the method must instead be built into is the structure of human relationship itself: into the way that relationships between people actually work, which is a matter of how the relations compose and not a matter of anyone's preference about them. There is a mathematics of relationship, in the same way that there is a physics of anything else, and it is not up for negotiation. The public question of Chapter Seventeen is one instance: once everyone knows that everyone knows a question has been asked, nobody present can be incurious about it, and no preference changes that. That is where the invariance has to live. It is the part of this work I have been trying hardest to clarify, and it is the part I am least willing to hand-wave, because a technique that is anybody's gift is a technique that can be withdrawn.

> **On the status of this claim**
>
> This chapter establishes what a method of values-discovery must satisfy: it must be incorruptible, it must be visible while it runs, and it must rest on the structure of relationship rather than on anyone's authority. That constraint does work on its own: it rules out every proposal that fails the visibility requirement, however well funded. Building a method that meets it is the work of a different book. The social technique that Chapter Seventeen describes, built around the facilitation of good questions, is the Ephemeral Group Process; it is set out in my books on that process, chiefly *Thinking Together Without Broadcast* and *The Ephemeral Group Process in Question*, not in this one. That process is a component of the whole, not the whole: it is the first stage of the whole process of social choice, and the requirement of incorruptibility belongs to that whole process. On its own, the Ephemeral Group Process makes corruption much less likely and tends to heal itself, as a natural limiting condition; this book claims no more for it than that.

VI. Whatever Succeeds Gets Predated Upon

Section V required that the method of discovering values be incorruptible, and the reason it must be, and not merely should be, is that whatever succeeds gets predated upon (hunted and fed upon, as a predator feeds on its prey). Rather than fighting the old system, you create something that is easier to live with, and people see that those people are happy and healthy and want that for their families. There is an attractiveness to embodied success. Does it take time? Yes. Can it be done? Of course. What prevents it? Nothing, except that the few times it happens, it gets predated upon.

Everybody says, "I want to sell that", because that is the habit all of us have: package it, put it on a shelf, deny it to others so as to sell it back to them. The dominant culture is so steeped in commercialization that it will commercialize anti-commercialization; it will literally sell you the notion that you have become natural. Because we all grew up with these codes, we can hardly imagine solving a problem without someone asking how to monetize it, as if no social project could move forward unless it is a business. But human beings lived for something like a hundred thousand years without commercialization. Community is the more fundamental thing, and we have forgotten it.

So we need to relearn community process and the wisdom to discern where commercialization creeps in, and then be disciplined about preventing it from happening all over again, as it has the last hundred times. Pain helps: when you can not only imagine an alternative but see it working somewhere else, you want to find out what those people are doing. That is probably not enough by itself; we will also need much more discipline in how we enculturate and educate. A method that could be captured would hand the predator the very thing that succeeded, which is why the method itself must be the part that cannot be captured.

VII. Governance Is Not Culture

Governance is not culture. The sentence is short and it needs the whole of this section, because the failure it guards against is the characteristic failure of everyone who gets as far as Section V and then keeps going in a straight line.

Governance can support culture. That is its proper office and it is a real one. What governance must not do is subsume culture: must not become the thing that culture is, must not be the place where a community's life is finally located. A community that has excellent governance and nothing else has a very well-run absence.

The reason the boundary is easy to lose is that governance is the part that can be designed, and the designable part always draws attention toward itself. Values-discovery is a governance process. It is also, unavoidably, close to the heart of a culture, and that adjacency is where the encroachment starts. The process that discovers what a community holds can very easily become the process that determines what a community holds, and the transition between those two things happens without announcement.

There is a well-established boundary which gives the shape of what I mean. We separate religion from state. Nobody thinks this means that a society's deepest commitments are unimportant, or that a state should be indifferent to them. It means that the apparatus of compulsion must not become the arbiter of meaning, because when it does, both are destroyed: the meaning becomes coerced and therefore worthless, and the state acquires an authority that no procedure can check.

I want the same separation drawn between business and state. Exactly the same reasoning applies and applies with the same force. Where the apparatus of governance and the apparatus of extraction are the same apparatus, governance becomes the means by which extraction is legitimized, and every protective form it wears becomes an instrument of the thing it was meant to protect against.

What I want instead is a particular arrangement of the three. Values infuse governance: they run through it and give it its content, without governance owning them. And governance in turn supports both the culture process and the market process, holding them together rather than presiding over either.

The crucial word in that arrangement is *together*. This is not an either-or proposition in any of its parts. The choice is not culture against market, nor the embodied against the digital, nor presence against vision. Each of those pairs is offered as a forced choice by somebody with an interest in the answer, and in each case the right position is as much of both ends as one can get. A community that is only present has no vision and cannot move. A community that is only visionary fails to notice the tree it is standing next to. Both failures are real, and neither is corrected by committing harder to the other.

VIII. Where the Room Is Also the Business

There is a hard question that any of this has to survive in practice, and it deserves to be stated at full strength.

Here it is: how does one use an online platform in such a way that, in each local physical setting, the work of *recontextualization* (by which I mean the restoring of a thing to the setting that gives it its meaning), the skills of care, and all the rest of what we have been describing are actually strengthened, rather than having energy and attention sucked out of the local setting and into a decontextualized digital space? And is not the internet useful for information sharing and not much else, so that supporting education in a digital setting is not worth the effort?

The second half of the question gives the first its edge. That position is not silly. It is the reasonable conclusion from watching what the available platforms actually do to people.

The answer I can give here is the first half of the answer, and its edges show. The spaces that do not do the damage are the spaces that do not have the two masters problem. Where a community is a community, and that is the whole of what it is, the space can serve it. Where the community's gathering place is also somebody's profit engine, the conflict is structural.

Structural is the operative word, and it is why no amount of good intention resolves this. Two masters problem was established in Chapter Four as the situation in which one party is obliged simultaneously to two authorities whose demands cannot be jointly satisfied, and in which the obligation with power behind it silently determines the outcome of the other. Apply that to a platform. The platform is obliged to the community that lives on it. The platform is also obliged to whoever requires a return from it. Those two obligations diverge the moment a question arises where the community's flourishing and the revenue are not served by the same answer, and such questions arise constantly, because they are most of what running a platform consists of. When they arise, the obligation with power behind it wins. Not because anybody chose it in that instance. Because that is what the structure does in the absence of anybody choosing.

This is why what I hold against such platforms is quite specific. I am not against algorithms; I am a software person and I love writing software. What I find myself unable to be in favor of is advertisement, and political or social gaming, and more generally the whole class of arrangements in which people are placed inside a context whose side effects benefit some private party. The phrase to hold on to is *side effects*. The harm is not usually in what the platform openly does. It is in what the shape of the room does to the people standing in it, in service of an interest that none of them entered the room to serve.

So the test can be stated compactly, and it is a question rather than a rule, which is deliberate. Of any space in which a community is asked to conduct its life, ask whether that space is also somebody's engine of return. If it is, the conflict is present whatever anybody intends, and it will

express itself in the accumulated weight of a thousand small design decisions that no one will ever be able to point to individually. The remedy is not better stewardship of such a space. The remedy is spaces that are not built that way, which is what the inverted order, running on a visible and incorruptible method, is for.

And that brings the argument back to where the metaphysics actually holds it (the metaphysics this book draws from, set out in *An Immanent Metaphysics*). Technology is a vehicle with which choice can be supported: causation arranged so that it opens into the full embodiment of a community's choosing. That is the constructive possibility, and it is a genuine one. The alternative is to stay at the level of causation: technology for its own sake, treated as a vehicle for further inequity and private gain. The distinction between causation and choice is developed in *Dialogues on Ethics and Governance* (Chapter Four). I will only say what follows from it, which I state without softening: that road ends the world.

Chapter Sixteen completes this answer, once the distinction between embodied and virtual value is in place, with a single criterion: whether the online is in service of the embodied.

IX. Describing a Future in Today's Language

The inverted order ends in a vision that a community has yet to discover, and before that vision can be stated there is a difficulty of language to face. I can try to describe the far side of these transitions in the existing language. The better I do that, the worse an idea you will have of what the future is. The more I describe the future in its own language, the harder it is to understand from where we sit today.

So I try to build a way of understanding that future that is somewhat agnostic about the language we use now. I look at the future in terms of situations, not narrative. What is meaningful in the transition will not come from any amount of symbolic or data understanding; it is beyond that, fundamentally.

X. Culture Before Values, Vision, and Strategy

This is where my own leanings go, by preference. The inversion of Section II of the last chapter can also be held as a vision of the future, and held that way it reads as follows. I have a vision that by knowing the difference between an institution standing in the place of a community, which is a fake form of virtualized social process, and a community, which is a real form of embodied living process, we can, most of us, all actually know, and prefer, the healthy being and becoming of the latter, and disfavor, everywhere, in all ways, the former. We can have actual healthy culture, and from there, in a distributed way, a discovery of values, and from that, maybe, in time and attunement and keen discernment, the evolution of a dream, a vision, for what the community really wants and feels very most alive with. In realizing that vision, with also a distributively evolved communal strategy of collective manifestation, maybe also we could find within that sustainable communal health.

That vision is in significant contrast to the usual venture method, where some single person has a strategy to use and misuse vision to manipulate people (friends, family, angel investors) to then invest, to continue the one person's strategy to use vision to manipulate even more people (executives, staff) to make some product or service, and then marketing and sales, which further implements the strategy of vision manipulating the culture of the public so as to shift their behavior, to compel them to buy, or to vote differently, or to otherwise act in a way that is not actually in their service, and so on. Singleton strategy (a plan driven by one person or one center, which uses vision to manipulate culture) over vision over culture, in all forms of unending monism, is the way of death. Only by following the path of life, of culture before values and vision, before the strategy of the many, may we maybe sometimes have life.

XI. Becoming a Species Worth Contacting

That vision makes a demand, and the demand falls first on us and not on our machines. Regardless of AI (and the arguments that AGI cannot be made safe, set out in *Why AGI Cannot Be Made Safe* and *The Ceiling Below the Floor*, say we should not even begin to build it), the central question is what we must do to become a species worth contacting. We need to work on our own capacity to become a fully alive, ethical and embodied species: to find the necessary, sufficient and complete conditions not only to endure on this planet but to thrive, and to know how to govern in a way that protects the land and the people and goes on to a joyful and meaningful existence for the whole. If I expect my children to do the tasks I did not do in my lifetime, and use that as my excuse for not working to become the best human I can be, something is wrong. In the same way, we should not look to AI to be our saviors or to be ethical on our behalf. We have not yet figured out ethics, right living and good relationship at the level of humanity; it makes no sense to try to do it first in technology.

XII. A Theory of Change

What remains is the route from here to there: a theory of change, the points at which a little force moves a great deal, and the communication on which every agreement along the way rests. In other areas of my work the theory of change (the account of the chain of steps by which a desired change can be brought about) runs like this. To address existential risk we have to address market process. To address market process we have to understand ecosystems, financial systems, infrastructure and culture. Then the question becomes how to shift the culture that created those infrastructures and financial processes, which becomes the question of how to create a culture wise enough to have holistic thinking built into it, and how to build the capacity for that to happen.

The impossibility argument (the argument that AGI cannot be made safe) is a small step in that direction: a way to be a little wiser. Without the delusion that AI alignment is possible, one would scope the work to narrow AI and not build general artificial intelligence; one could still study intelligence without making it a commercial thing, and encode ethical principles for narrow systems as business practice or government regulation, the way we regulate nuclear technology.

Some Leverage Points

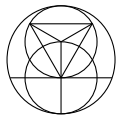
Some things would make a huge difference. Abolish all advertising. That does not stop anyone from telling people that a thing exists; it stops the arranging of their attention to get them to want it, and a small maker is then found as people find one another, without the finding being brokered. Abolish intellectual property, which is a form of censorship. Make sure no corporation has more than 150 people, or at the outside 300; small business is fine, but a mega corporation is in effect a national government answerable to nothing that matters. The two numbers have a reason. About a hundred and fifty is the edge past which a group held together by relationship gives way to transaction and power: it is close to the natural size of a human social group, the number of people with whom one person can keep a relationship, and past it what holds a single group together is no longer relationship. Three hundred, at the outside, is the far edge of the range of sizes in which no general way of governing a single group stays stable; and a body larger than that holds together without transaction and power only when it is formed from the start as many small groups joined together, never by enlarging any one of them. The things that need thousands of hands would be built in that way, by many small groups joined to one another by liaisons; and some things now built would be built more slowly, or not at all, which is the price named in Chapter One. No censorship. Community has to be stronger than corporations. Deeper than all of those is a culture that understands the sacredness of the world we live in and its uniqueness in the universe. I set these down

as descriptions of what would change the most. They are offered in the sense of Chapter Twelve: to a community that can take them up, for itself and by its own process, or decline them.

Agreements Rest on Communication

Legal systems are forms of agreement. Agreements depend on networks of relationship, and networks of relationship depend on networks of communication. Good communication is a handshake: I grant you the right to speak, the right to be understood, and the right to know that you have been understood, and if you grant those three rights back to me, there is a handshake. That is a protocol, in the same sense as TCP/IP (the set of rules by which computers on the internet exchange data): you send out a packet and it sends back an acknowledgement, and without the acknowledgement you do not know whether the packet was received at all. It is the same with human beings. The skills of clear communication give real relationships and therefore authentic agreements, and all of that is grounded in the biological nature of humanity and in the nature of choice. This is not just philosophy. It is code for being.

> **On what this chapter has and has not done**
>
> What has been established here and in the last chapter is an order and a constraint: culture, then vision, then strategy, distributed at every stage; and a method of values-discovery that is incorruptible, visible, and nobody's to set. What has not been established is the medium through which such a culture actually transmits itself across time and across people who never meet. That is the subject of the chapter that follows.



CHAPTER FIFTEEN

Song, Text, and Code

> *This chapter defines the three media the program works in (song, text, and code) and gives the proportions of the population that live in each, offered as estimates and marked as such. It then establishes the observation that reframes the modern situation: code has become the substrate on which song is carried. From there it builds the chapter's central argument, step by step. Code can represent agency; text cannot. That is why text, and not the other two, is the vehicle by which one generation can enable a later one without defining it. The chapter describes the filters that make such a text, and holds this book to their standard, and ends with why text, the least valued of the three media, is the one that travels.*

I. Three Media

The distinction answers a practical question, and it is worth saying what the question is, because the distinction is not an abstraction and did not come from one. Consider a community engaged in teaching and gathering that is looking for a medium. All of its digital work is meant to be in service of the local and the grounded: in service of people meeting one another in their own regions, at events held in actual places. It wants a way for those people to orient to one another between the gatherings, in something that feels good, that is not addictive, that can be nourishing rather than depleting. It has surveyed what exists and found one or two things

interesting and nothing satisfying. And what it needs is not a recommendation but a way to navigate the search, so that it knows what it is looking at when it looks at the next thing.

What follows is not a recommendation but a distinction that organizes the search: the relationship between what I call *song*, *text*, and *code*.

Song is the live register, taken whole. It is voice. It is video. It is being in a room with someone. It is everything that involves personality: art, music, humor, timing, the particular way a person leans forward when they mean something. It is the ways in which we are together in person, inclusively, across the whole range of that. And it includes, as well, any audio or video content whatever: the recorded talk, the broadcast, the call, the message left as a voice note. Song is not a metaphor for music. Music is one region of it. The word names the entire register in which a living presence is what carries the meaning.

Text is just that: text. Written language, read. It is the shortest of the three definitions and the one that will turn out to require the most work, because everything hangs on what text can and cannot do.

Code is engineering. Software. Infrastructure. The built substrate. It is how one designs the physical and informational support structures a community runs on: the room in which the meeting happens, whether that room is made of walls or of protocols. Code is what has been made such that other things can then happen inside it. The boundary between text and code is execution. Source code sitting in a file, unexecuted, is still only text; it becomes code at the moment it is loaded into a setting where it causes something to happen. A passed law is compiled text, and courts and police are the interpreters that make it run. A recipe on a page is text, and the cook who follows it makes it code. A building is code, and so is a protocol, and so is a traffic light: each acts on the world according to a pattern set once that thereafter runs.

These three are not a ranking. None of them is the good one. Each does something the other two cannot do, and a great deal of confusion in the design of technology comes from asking one of them to perform the office of another. The rest of this chapter is an attempt to say, with some precision, which office belongs to which.

> **On the standing of the triple**
>
> The division into song, text, and code is not an informal convenience. In the larger framework from which this book draws (set out in *An Immanent Metaphysics*; see the Bibliography), the three are a formal result, corresponding to the immanent, the omniscient, and the transcendent respectively, and the derivation establishes that there are three and not two or four. In that framework the immanent is what is present in the act itself, the omniscient is what can be laid out and held in view as a whole, and the transcendent is what acts beyond the moment and the place of its making. That derivation is a separate matter and is not presented here; the three as registers of communication are worked out in *Song, Text, and Code*. A modality is assigned to a term only within the triple in which it is considered, so this assignment does not conflict with the place the city's infrastructure takes in Chapter Thirteen. This chapter uses the triple expositionally, and everything argued in it can be followed without the formal apparatus.

II. The Proportions

Nearly everyone lives in song. That is the first of the three estimates and it is the most secure of them. Whatever else is true of a human population, the overwhelming majority of its members conduct the whole of their communicative lives in voice, in face, in gesture, in presence, and now also in the recorded and transmitted forms of those things. A conversation, as it happens, is song. The dialectic between people in a room (the live back-and-forth in which meaning is made in the exchange itself) is song. It is where nearly all of human meaning has always lived.

The second estimate concerns text, and here the figure drops sharply. On the order of one percent of the population navigates text in the generative sense. What is being measured has to be stated exactly, because the figure is meaningless otherwise. The question is not whether a person is literate, nor whether a person sends messages; nearly everyone does both. The question is whether a person can sit down and produce a page of coherent material: an essay, say, or an argument that holds together from beginning to end and arrives somewhere it was not at the start. That is a different capacity from reading, and it is a different capacity from typing. Roughly one in a hundred, I estimate, has it.

The third estimate concerns code, and it is smaller again. The proportion of the population that can genuinely implement code (not read it, not configure it, not operate a thing built by someone else, but write it) is a fraction of a percent. It is not one in a hundred. It is something nearer three in a thousand, and if the standard is held strictly, if what is meant is the capacity to originate a working system rather than to modify one, the figure is nearer three in ten thousand.

Underneath all three estimates there is a distinction doing the real work, and it is worth drawing it out because it is the thing most often collapsed. *Consuming is not generating.* A person who reads a great deal is not thereby a writer. A person who spends their evenings inside software is not thereby an engineer. A person who listens to a hundred hours of talk is not thereby someone who can hold a room. In each of the three media there is a receiving competence and there is a generating competence, and they are not the same competence, and the second is very much rarer than the first.

The estimates above are all estimates of the generative capacity. That is deliberate, and it is why the numbers are so small. If one counted consumption, the figures would be unrecognizable: nearly everyone reads, nearly everyone uses software. But consumption does not tell you who is able to put something new into the world through that medium, and it is generation that determines whose intentionality the medium ends up carrying.

A fair question arises at once. Could a person not use a machine to write the page for them? Does the availability of generative tools not dissolve the distinction between the one percent and everyone else? The answer is that it moves the question without answering it. If the machine wrote the page, then the machine generated, and the capacity being estimated (the capacity resident in the human population to originate coherent written thought) is exactly what has not been exercised. One may think that a good trade. That is a separate argument, and a serious one. But it is not a reason to think that the underlying proportions have changed.

> **The ordering the argument uses**
>
> All three proportions are estimates, not measurements. There is no agreed way to measure the ability to write a coherent page, no agreed way to sample, and no instrument; the same holds for the ability to originate a working system in code. What can be said with confidence is the ordering and the rough magnitude of the gaps between the three: nearly all, then something around one in a hundred, then something well under one in a hundred. The argument of this chapter depends on that ordering. It does not depend on any particular figure, and the figures should not be quoted as though it did. The figures, and what they mean for governance (a cycle that runs from song through text, as policy, into code, as implementation, shuts nearly everyone out of authoring it), are set out in *Song, Text, and Code*.

III. Code Is the Substrate on Which Song Is Carried

Here is the observation that reframes the modern situation, and it follows directly from the three definitions once one looks at where communication actually happens now. Code has become the substrate on which song is carried. The fraction of a percent who write code are building the worlds inside which nearly everyone else sings.

The clearest case is the coded world: the kind of world one buys and enters, in which many people are present at once, in which there are places, economies, alliances, histories, consequences. Such a world is, considered as an artifact, entirely code. Someone wrote it. Its physics is a specification; its geography is a data structure; what may happen in it, and what may not, was decided by an author who is not present in it.

And yet what people do inside it is song. They talk. They form attachments and betray them. They joke, boast, mourn, organize, console one another. They spend hours of their lives in a register that is unmistakably the live register: personality, presence, timing, the whole of it. The play is song. The world itself is code. That is the structure, and once it is seen in the coded world it is visible everywhere else.

Most communication now occurs in exactly that arrangement: song upheld by code. The voice call is song upheld by code. The video conversation is song upheld by code. The feed, the thread, the shared room, the group of friends who have not been in the same city for six years and who nevertheless talk every day: all song, all upheld by code. What the participants experience is the song. What determines the shape of what they experience is the code: what is possible in the room, what is easy, what is rewarded, who can hear whom, what persists and what disappears, and what, invisibly, is being counted.

This is where the argument of this book bears on the matter directly. Technology is a vehicle for the representation of intentionality, and the question is always whose. When code becomes the substrate of song, the intentionality represented in the code becomes a condition of the song carried on it. The people singing did not write it. They cannot read it. In the ordinary case they do not know it is there. And the asymmetry is not a rhetorical flourish: it is the asymmetry between a fraction of a percent and everyone.

The question a community must always ask of its own infrastructure is the book's one question in an older form: whom does it serve? It is the Grail question. In the legend, the knight fails to ask whom the Grail

serves, and learns that the question would have healed the wounded king. A community that can read its own governing code can answer it directly. A community that cannot must take someone's word for it, and the person whose word it takes is the person whose agency the code may carry. Text is the register in which a community examines its code, and so text has a second office beside carrying principle forward: it is where a community holds its code to account now. Where text is absent, the question cannot be asked, and the code serves whoever wrote it.

That asymmetry accounts for a great deal that is otherwise puzzling about the present. A community looking for such a medium is already alert to it. It wants something that is not addictive, and it is wary of the way these spaces seem to produce relating from a position of ego: identity constructed out of materials that are held by corporations, and which can therefore be altered, priced, or withdrawn. Those are not accidents of use. They are properties of the substrate, authored into it. The song destabilizes because the code it is carried on was written to make it destabilize, or at any rate was written without regard to whether it would.

I do not draw from this the conclusion that code is the adversary. The opposite conclusion is available and I hold it. Code can help the wild to be more wild. Chapter Two defined wildness as the condition of standing outside the immediate control of another agency, and defined it that way precisely so that it would not be confused with the absence of making. A made thing can be wild. The question is whether what one encounters in it is addressed to one by someone who wants something from the encounter.

Code that is written to capture attention manicures the world it makes: everything in it is addressed, everything is for something, and the person inside it is never out from under another agency. But code can also be written to hold a space open and then get out of the way: to make possible a meeting that no one is steering, to let people find one another without the finding being brokered, to establish a commons and then decline to farm it. That is a real design possibility, not a hope. It is harder to build and there is no revenue in it, which is why it is rare. It is nonetheless the

possibility that this book is interested in, and the reason the answer to the question about platforms is not simply to abandon the digital. That answer is developed in Chapter Sixteen.

IV. The Question of What Can Be Handed On

The question this chapter exists to answer is a question about intergenerational lessons, and it is an urgent one. There are things humanity has learned at enormous cost. A great war teaches a generation something that the generation would give almost anything not to have had to learn. Is there a means by which the propagation of such lessons can be ensured? Not for a decade. Past the life of the last person who experienced the thing. Do you trust the biology to hold it: some epigenetic residue (some change in how genes are read that is passed to children), some inheritance carried in the body? And it is a fundamental part of any design for a civilization.

I have spent a great deal of time on it. Let me set the biological route aside at the outset. Not to deny that something is carried in the body; something plainly is. But whatever is carried there is not something one can design with, and it is not something one can be confident of the content of. The question I am able to work on is what can be done at the cultural or structural level such that past communities enable future ones, even when the separation between them is several generations, which is to say, beyond living memory, past the point where anyone alive can say *I was there*.

And the aim has to be stated precisely, because it is easy to state it in a way that would be a catastrophe if achieved. I do not want future generations to somehow encode within their being the experiences themselves. I do not want to hand anyone the war. What transmits an experience without transmitting the understanding is trauma, and traumatized

people are not enabled people; they are people whose range of choice has been narrowed by something that happened before they were born. That is the opposite of the goal.

What I want is for the vehicle of future situations to be informed by past learnings. I want the principle to be available where the experience is not. I want a person facing a situation their great-grandparents faced to have the benefit of what was understood then, in a form they can take up or decline, and to be under no compulsion from it whatever. That is the specification. Everything that follows in this chapter is an attempt to meet it.

V. The Past Should Enable the Future

The requirement that governs the whole project, stated positively, is this: **the past should enable the future, not define it.** Enabling and defining are different operations, and the difference between them is the difference between a gift and a constraint. An enabling past increases the range of what a later generation can choose. A defining past decides some of it in advance and hands down the decision as a fact of the world.

Consider what it looks like when a past defines a future. A generation builds. It pours concrete, lays out roads, zones districts, erects the buildings in which everyone afterwards will live and work. Those buildings are the expression of that generation's agency: its taste, its economics, its assumptions about how people ought to live, its confidence that it had understood the matter. And then the generation passes, and the buildings do not. They are still there. They are expensive beyond reckoning to remove. Every subsequent generation lives inside a decision it did not make and cannot unmake, because nobody can redo the thing.

The question can be put in three words: *that is causation.* It is exactly right. What has happened there is that one generation's agency has been converted into a cause operating on people who were never party to it, and a cause applied where a choice should have been is the structure this

book has been examining from the first chapter. The argument for why causation applied to choice is the characteristic harm is developed in *Dialogues on Ethics and Governance* (Chapter Four).

What I will add is the name for the failure mode. It is an overextension of agency. Not the exercise of agency; building is good, and a generation that built nothing would have failed its successors in a different way. The failure is in the extension: agency exercised so as to reach past its own generation and bind the ones that follow. Every generation should, to some real extent, be allowed to live its life. The question to ask of any inheritance is a comparative one. Does the past have such influence over the future that the influence overwhelms the benefit of the learnings the past can give? The learnings are the gift. The fixing is the cost. Past a certain point the cost consumes the gift entirely, and what is received is not wisdom but a set of walls.

I can put my own position in a sentence, and it is the sentence that has organized everything I have built: *I do not want to be dominated by the past; I want to be liberated by it.* I want what came before me to have made more possible for me, not less. I want the accumulated understanding without the accumulated compulsion.

And the sentence has a symmetry which is not optional and which costs something to accept. If that is what I want from the past, then it is what I owe the future. I do not want to dominate the future. I want to liberate it. Whatever I build, whatever I write, whatever I hand forward, should increase the range of choice available to people I will never meet, and should not narrow it. This is the same principle as *love is that which enables choice* (love understood as that which enlarges another's capacity to choose, rather than as that which chooses on their behalf), applied across generations rather than between persons. The full development of that definition belongs to Chapter Seventeen and is not undertaken here.

VI. Why the Vehicle Is Text

Given that specification, the question becomes a question about media. Which of the three (song, text, or code) can carry a learning forward across generations while leaving the agency behind? The answer is text, and the reason is a property of the three media themselves. It is worth building the argument in steps, because the conclusion is doing a great deal of work later in this book and it should not be taken on assertion.

The first step concerns code. Code can represent agency. That is not a defect of code; it is what code is for. A program is a standing intention that acts. It executes. It decides. It responds to circumstances its author never saw according to a policy its author fixed in advance, and it continues to do so when the author has stopped attending, when the author has changed their mind, and after the author is dead. This is an extraordinary capability, and it is also the precise reason code is unsuitable as a vehicle across generations. Code hands the future not a lesson but an actor: an agency, still running, still deciding, representing an intentionality that no living person can be held to account for. The concrete building is the low-technology version of this. Code is the version that keeps making decisions.

The second step concerns song. Song also carries agency, and carries it inescapably, because song is a person present. What one receives from song is never only what was said. It is who said it, and how they wanted it taken, and the authority of their bearing, and the pull of their personality. The teacher who is loved is believed; the teacher who is not is doubted; and the content may be identical. That is not a corruption of song. It is song working as song works, and much of its power comes from exactly this. But it means that song cannot deliver a principle stripped of the one who held it. Song without a singer is not song. And song is perishable in a second way: it lives only in the moment of its performance. The recording is a trace of it, and the trace still carries the singer.

The third step is the one that matters. Text cannot represent agency. To represent agency, in the sense of the first step, is to act: to go on deciding on its author's behalf after the author has gone. Text can carry its author's intentionality, often a great deal of it, but it cannot act on anyone. A page does nothing. It does not execute, does not decide, does not act upon a reader who has not opened it. It waits. And because it does not act, whatever intentionality went into it can, with care, be taken out of it. That is the crucial claim: if I am careful with how I use text, I can filter the agency out. What remains is principle without agency.

The qualification is not decoration. It is the whole difficulty. Text is not automatically free of agency; most text is saturated with it. Polemic is agency. Opinion is agency. The author's preferences, the author's enemies, the author's desire to be found persuasive, the flattery of the reader, the arrangement of a case so that only one conclusion appears available: all of it is agency, sitting in the text, addressed at the reader, wanting something. Careless text is simply song frozen: the singer is still there, merely less visible, and harder to argue with because they cannot be asked a question. A scripture read aloud with authority, and a page of propaganda posted where it cannot be escaped, have been given a singer or a setting that acts on the reader; the page itself still does nothing. The claim is not that text is innocent. The claim is that text is the only one of the three in which the agency *can be removed*, because text does not require agency in order to function. The list of ten results in Chapter One is not neutral because code chose its order; the page it is printed on chooses nothing.

Chapter One said that intentionality is always represented, and that stands. A text is a made thing, and what its maker kept in it is there because someone meant it to be. One purpose stays in the text however long it is washed: the purpose of enabling a later reader, which chose what to keep and asks nothing of the one who reads it. What the filters remove is the other kind of intention, the kind that wants something from the reader. That second kind is what I mean in this chapter by agency.

From which the conclusion follows. With that filtering done, there is now a vehicle for communicating across time, and for communicating across distance from one community to another between which there is no live contact at all. A community two centuries hence, or one five thousand miles away that shares no history with this one, can receive a principle without receiving a master. They can take it up, test it against their own situation, and find it wanting if it is wanting. Nothing in it presses on them. That is what a vehicle for enabling looks like, and text is the only one of the three media out of which it can be made.

One condition remains, and it is the condition that makes the whole thing hard. It works assuming one can create an interface between song and text. Song is where the material is generated. Real understanding arrives in conversation, in the room, between people who are thinking together and can interrupt one another. It does not arrive on the page; it arrives and is later put on the page. So the vehicle needs an on-ramp: a means of taking what is alive in song and rendering it into text without killing the thing that made it worth carrying. That interface is the practical problem, and it is enormous.

VII. The Filters, and the Questions

The interface exists, it has been built, and everything done to test it so far indicates that it works. It consists of a series of filters applied in sequence, and it is worth describing them concretely, because the description is also an account of what this book is.

Taking spoken word into text begins with transcription, and transcription is only the beginning. Transcription software mishears, drops punctuation for paragraphs at a stretch, and fuses sentences. But suppose it were excellent. Suppose every word were captured exactly. The question that matters is not whether the transcript is accurate. It is whether the transcript is *good enough for the purpose*, and it is not, because accuracy is not the standard. An accurate transcript of a conversation is a maximally

faithful recording of agency. It preserves every trace of who was speaking, what they wanted, whom they were addressing, and what they thought of the last person to speak. It is song in written clothes.

So the filters begin. The proper nouns come out. All of them. The names of persons come out because a name imports the authority of the person named, and a reader who receives a principle attached to a name relates to the name. Either they defer to it or they resist it, and in both cases they are relating to an agency rather than to the principle. The names of companies, products, platforms, and places come out because they bind the material to a particular market and decade, neither of which will exist by the time the material is needed. Proper nouns are the densest carriers of agency in a text, and they are the first thing a filter must remove. A name that is only the handle of a result (a theorem known by the name of the one who proved it, or a place named only to say what grows there) imports no one's authority and binds the text to no market, and it can stay.

Then the places where opinion leaks in come out. This is finer work, because opinion does not announce itself. It arrives as emphasis, as the example chosen rather than another example, as the case stated at length and the counter-case stated briefly, as the aside that lets the reader know what the author thinks of people who disagree. Opinion is agency wearing the clothes of principle, and it is far harder to see than a name.

Then further filters, each more specialized than the last, each addressed to a narrower class of residue. The effect of the sequence is cumulative and it is best described as a washing: pass after pass, by degrees, the intentionality is washed away, while what is left standing is the essence of the principles, the practices, and the kinds of things that would be enabling of a future generation. No single filter accomplishes this. The method is the sequence.

I should say plainly what will already be suspected. This book is meant to be text of the kind this chapter argues for, and the reader is entitled to hold it to that standard. Where I have kept the first person, as in

the account of the company I ran and in the questions put to one call for compassion machines, I have kept it on purpose, because there the speaker is part of the evidence. Everywhere else the reader should be able to take up the principle and leave me behind.

Filtering, however, only subtracts. Something must remain by which a reader can find their way, or what is handed forward is a heap of true statements with no indication of what they are for. That is the office of the question, and it is why the emphasis on questions runs through the whole of this program.

A question is the navigational map of the information. It provides a way of decontextualizing from the past so that the material can be recontextualized in the future. Consider the difference between an answer and a question with respect to travel across time. An answer is bound to the circumstance in which it was given. It presupposes the conditions, the constraints, the technology, the vocabulary, the particular difficulty that made it the right answer then. Carried forward, an answer arrives as an instruction issued by someone who does not know where it has arrived. A question presupposes far less. It can be carried into a circumstance no one anticipated and still be asked there, and asking it there will produce an answer fitted to that place.

That is why the questions are the structure. They give the shape of the information (what the material is about, what regions it covers, which inquiries it will reward) without fixing what is to be concluded. A question hands forward the map and leaves the walking to the ones who come after. It is the most complete form of enabling without defining that I have been able to find. Chapter Seventeen takes up the question as a social practice, which is a further matter; here it is offered as the answer to a structural problem, namely what survives the filters.

VIII. Why Text Travels

Let me state the positive result first. Everything I have done to test this process shows that it works. The filters can be built. The washing does remove intentionality by degrees without removing the principles. The questions do serve as a map. Material generated in song can be rendered into text that a stranger can use.

What I have not solved is the cultural fact that text is not valued.

The difficulty might be thought to lie in the cultural dynamics of not valuing speech in general. It lies in not valuing text; the asymmetry runs entirely one way. Song is valued. Text is not.

And I do not say this as a complaint about other people, because it is true of my own experience and of the experience of the people I most respect. Look at where the deepest conversations actually happen. They happen in the loose hours of a gathering, at the table, late, with the plates pushed back; or when two people walk out and keep walking longer than they meant to. I will bet that I have been more successful in communicating what I actually mean sitting in person around a table than through any number of books I have written. The room outdoes the page, and it is not close.

That is text's weakness, and it is also, exactly, text's strength. The strength is that text is not bound by the dialectic. By the dialectic I mean the live back-and-forth between people present with each other, in which meaning is made in the exchange itself: one person says a thing, another answers, and what emerges belonged to neither of them beforehand. That is where the depth comes from. It is also a requirement: the dialectic needs the presence. It needs the two of them, in the same place, at the same time, attending to each other. Song is bound by that requirement absolutely. Where the presence is absent, the song does not occur.

Text is free of it. A page can be read by someone who was not in the room, who cannot answer back, who was not born when it was written, who lives after everyone who was present has died. The very thing that makes text cold (that it does not respond, that it will not adjust itself to

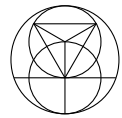
you, that it cannot be argued with) is precisely what allows it to travel. Weakness and strength are the same property seen from two sides, and the property is that text does not need anyone to be there.

A last question bears directly on this, and it is held with real conviction.

It holds that the interface between text and the new generative models is itself the answer: that one writes a few words and receives a painting, that one offers a thought and receives it back transformed in a way that reveals something about oneself. It feels like magic, like casting a spell, and the description is not naive: it reports an experience, not a theory.

The reply is that I could hand someone a small book of aphorisms, ask them to hold a question in mind, and ask them to open it at any page, and it would do the same thing. That is not a dismissal, and it should not be read as one. It is a claim about where in the arrangement the effect is located. When a text returns your own thought to you in a way that lets you see it differently, a great deal of the work is being done by you. The text supplies an occasion for reflection. The reflection is yours. That is worth knowing about any such interface, because it tells you what would have to be true for it to serve as the on-ramp between song and text that the argument of this chapter requires.

The chapter can be closed by putting the three back together, now that each has been given its office. The material is generated in song, because that is where understanding arrives among people. It is carried forward in text, because text is the only medium out of which agency can be filtered so that principle survives without a master. And it moves, in the meantime, on code, because code is the substrate on which song is now carried, and because code, written with the right restraint, can hold a space open rather than close one down. The past should enable the future, not define it. That requirement is met by nothing automatically. It is met by choosing the right vehicle and then being extremely careful with it.



CHAPTER SIXTEEN

Embodied Value and Virtual Value

This chapter establishes the distinction between virtual value (money, reputation, position) and embodied value: shelter, food, and relationship, the value that is present, in a locale, and not virtualized. It sets out the three-step sequence by which the one is converted into the other and thereby destroyed: extraction, then abstraction, then accumulation, with the relation between content and context lost at every step. Against that sequence it states the reverse move as a program, works out why the market now contains the community rather than the community the market, and distinguishes the arm's-length relation proper to institutions from the care relation proper to community. It closes on the design question this leaves: what online process is for, what it must refuse, and the single criterion by which any such process is to be judged.

I. Two Kinds of Value

A system built to optimize carries the intentionality of its builder through what it can count, and what it cannot count is not represented at all. Value comes in two kinds, and what follows depends on holding them apart. There is *virtual value*, value that has been made portable: money, reputation, position. And there is *embodied value*: shelter and food and relationship, the value that is present, that is here, that is in a locale. Both are real. The difference is not one of worth but of where they live.

Take the plainest case. A roof over your head on a night when it is raining is embodied value. It is that roof, in that place, at that hour. A meal shared at a table is embodied value, and so is the table. And relationship is embodied value in the fullest sense: I know who you are, I know what you care about, I know what you are likely to do when things go badly, and you know the same of me. None of that is stored anywhere. It exists where we are.

Virtual value has the opposite property, and the property is the point of it. Money travels. It can be sent to a person you will never meet and will do there what it would have done here. Reputation travels in the same way: it is a claim about you consultable in your absence. Position travels furthest of all, since a position holds whoever occupies it. What money is was set out in Chapter Four and is not re-argued here. The whole utility of virtual value is that it has been detached from its origin; the whole nature of embodied value is that it cannot be so detached without ceasing to be what it was.

This is what I mean by saying that embodied value is *non-fungible*, where the word is taken in its plain sense. Non-fungible here means not virtualized: it is present, it is in the locale, and no unit of it can be swapped for another unit because there are no units. Two bushels of grain of the same grade are interchangeable. Two evenings with a friend are not, and the reason is not sentiment: the second evening is composed partly of the first.

Now the difficulty. Virtual value can be counted, compared, and moved, and embodied value cannot, and therefore every system built to optimize tends to optimize the virtual. Not out of malice. Out of what it can see. The embodied does not show up as a low number. It does not show up at all. That is the structural fact from which the rest of this chapter proceeds, and the sequence I am about to describe is not a story about bad people.

II. Extraction, Abstraction, Accumulation

Embodied value is converted into virtual value by a sequence of three moves, and always in the same order, because each move makes the next one possible. The moves are *extraction*, then *abstraction*, then *accumulation*. I work them through on one case, because every step is ordinary and defensible, and the loss occurs nonetheless.

Begin with a tree in the woods. The tree is not merely standing in a place; it is *of* that place. It holds the soil on that slope. It shades that stretch of the creek and keeps the water cool enough for what lives in it. Its canopy is the roof of a system of animals. Its roots are in conversation with fungi whose extent nobody has mapped. All of that is the tree's context, and the tree is the content standing in it. The vitality of the tree is not a property of the trunk. It is the relation between the trunk and everything just listed.

The first move is extraction. The tree is taken out of context. This is not a metaphor; it is the literal description of what felling is. What leaves the woods is the trunk. The slope, the creek, the canopy, the roots: none of that leaves with it. And notice that the content survives intact. The wood is entirely as good as it was. What has been severed is the relation between the thing and where it was.

The second move is abstraction. The trunk becomes lumber, which is to say it becomes a quantity: so many board feet, of such a grade, of such a species. It is now a member of a class. Any other trunk of the same grade would have served, which is what makes lumber useful: a builder can order it without going to look at the tree. Then the lumber becomes a chair, and the chair is photographed and becomes an image in a listing with a price beside it. At this stage the chair is not present at all; what is present is a representation sufficient to support a decision to buy. Each step strips a further layer of particularity, and each buys something real in exchange: comparability, transportability, reach.

The third move is accumulation. The purchase goes through, and the chair, from the seller's side, becomes a number added to a balance. At the far end there is nothing left but that number: virtual bits standing where the vitality of a tree in the woods once stood. And that quantity has a property none of the earlier forms had: it adds. You cannot have a tree twice as much as a tree. But numbers accumulate without limit and without friction, because everything that would have made them resist addition was removed at the previous two steps. Accumulation is only available to what has been abstracted, and abstraction only to what has been extracted. That is why the order never varies.

The loss must now be located precisely, because the sequence tempts one to locate it in the wrong place. It is not that the tree was cut. Trees are cut; people need chairs; a woodworker who takes a tree with care and makes something that will last a century has done a good thing. The loss is not in any single step. It is in the composition. At each step the content is carried forward and the relation between content and context is dropped, and by the end of three steps there is a great deal of content standing in no context whatsoever. What is lost, at every stage, is the same thing: the connectedness, the relation between content and context. Because what is lost is a relation rather than a substance, no audit of the objects will ever detect it. And the loop closes: the accumulated number pays for the next survey and the next mill, so that nothing internal to the engine ever calls for it to stop.

III. The Reverse Move

The reverse of that sequence can be stated exactly, term by term, and stating it exactly is the beginning of a program rather than a complaint. Against extraction: *recontextualization*. Against abstraction: *embodiment*. Against accumulation: *disaccumulation*. Three moves, in reverse order of the three that did the damage, each meaning something concrete.

Recontextualization means putting the relation back: restoring to a thing the connection to where it came from, who made it, and what will be affected by what is done with it. It is not a matter of adding information. A label naming the forest a beam came from is a fact about provenance, and a fact can be printed on anything. Recontextualization has occurred when the person making the choice is in relation to the context the choice touches: when the consequences land where they will hear about them, among people who will still be there afterwards. This is why locality does so much work here. Locality is not a preference for the small. It is the condition under which context is recoverable at all.

Embodiment means preferring the present form to the represented one wherever the choice exists: the meal eaten with the people, the craft learned by standing next to someone doing it, the agreement made between parties who will meet again. It does not mean refusing representation, which would be absurd (every book is a representation, this one included). It means holding representation as an instrument in service of something present, not as a replacement for it.

Disaccumulation means arranging matters so that value does not pile up past the point at which it can be in relation to anything. Where accumulation gathers into fewer and fewer holdings, disaccumulation returns value into circulation, into locales, into the hands of the people whose lives are the context in which it means anything. It is the practical form of what earlier chapters called distribution.

Now the hard part, which is not any of the three moves but the condition on all of them. Every one must be done in such a way that the parties still playing the extractive game cannot corrupt the process to their advantage. This is not caution, and not suspicion of persons. It follows from the structure. An arrangement that recontextualizes, embodies, and disaccumulates is producing value no measuring system is pricing, which is to say it is producing exactly the thing an extractive process most wants to find. A community that has built genuine trust has unusually low transaction costs, and low transaction costs are an asset on somebody's balance

sheet the moment anyone thinks to look. The healthier the commons, the richer the harvest. So the design cannot merely be good. It has to be good in a form that does not reward capture.

Chapter Fourteen set the condition on all three moves: the technique by which a community discovers its values must be incorruptible and visible, so that the parties still playing the extractive game cannot steer it. One does not come to trust a court by being told it is fair; one trusts it because the proceeding can be watched.

IV. Which Context Holds Which

The relation between community and market is a containment relation, and the whole question is which contains which. Two arrangements can have identical parts (the same people, the same goods, the same exchanges) and be entirely different depending on which is ground and which is figure.

Go back to the middle ages for the contrast. Community was happening all the time. It was the continuous condition: people lived among the same people, worked within sight of them, raised their children and buried them. And within that condition the market occurred: on certain days of the week, at the fair, at the festival, for a defined interval, after which it stopped and the ordinary life of the place resumed. Life was the ground. The market was an adjunct to it. A merchant who behaved badly on market day was not thereby finished with the consequences, because the next day he was still among the people he had dealt with. The containing context did work no rule of trade could do.

What the historical comparison shows

The comparison does one job, and a structural one: it exhibits a case in which market activity was *contained* by a continuous relational context rather than the reverse, and so shows that the containment can run in that direction. This is not an argument that the period was better, and I do not want it read as one. It was coercive in ways we would not tolerate, hierarchical by birth, and brutal in its ordinary operations; the community that disciplined the market disciplined everything else too.

The modern arrangement has the containment the other way round. The market is the context now: the continuous condition, the frame within which one's hours, housing, health, and standing are settled. A great deal of community still exists, but it exists *inside* that frame, in the intervals the frame leaves open. And here is the consequence, which is structural and not incidental: where community sits inside a market, the market treats it as a source of extraction, and it cannot do otherwise. A market context reads whatever is inside it in the only terms it has, and in those terms a community is an audience, a user base, a durable source of yield. The affection people have for one another appears as retention; their trust, as reduced acquisition cost. None of this requires anyone to be cynical. It is what the containing frame makes of what it contains.

So the objective can be stated in one line. It is to make the context of community stronger than the context of commercialism. Not to abolish the market; markets do something nothing else does as well, and a community with no market in it must settle by discussion many questions that do not repay it. The objective is to change which one is the ground. That is the whole program in a sentence, and everything after this is a question about how such a context could be built in the conditions we have.

V. Arm's Length and the Hug

Institutional relation is at arm's length by design, and the design is correct. I establish this first, because what follows is often mistaken for an attack on institutions, and it is not one. Business, government, education as an institution, medicine, food production at scale: all are organized as institutions, and the relation they run on deliberately holds separation.

That separation is doing real work. An institutional relation has to function between strangers. It has to survive the replacement of every individual in it. It has to hold when the parties dislike one another, and when they like one another too much. Arm's length is what makes a contract enforceable regardless of anyone's feelings about the counterparty. It is what allows two governments who trust each other not at all to arrive at a workable treaty, and what makes an obligation outlast the person who signed it. Where a relation must be reliable without depending on anybody's goodness, arm's length is not a coldness; it is an engineering requirement.

Community is a different relation entirely, not a warmer version of the same one. In community the question is care. How do we care for one another? How do we care for the young and for the old? How do we care for our neighbors? How do we care for our children? None of those has an arm's-length answer. There is a sense in which this is, to put it plainly, the relation of the hug rather than of the handshake, and I mean that as a plain observation, not a sentimental one. A hug has no counterparty. It cannot be performed at a distance, it cannot be delegated, and there is no version of it in which the parties hold separation.

The failure mode is running the wrong relation in the wrong place. Institutional form applied to care produces the familiar spectacle of the young and the old being administered: assessed, scheduled, and processed by arrangements genuinely well intended and structurally incapable of knowing anyone. Care-form applied to institutions produces oblig-

ations that hold only among friends, which is what arm's length was invented to prevent. Neither form is wrong; the error is the category confusion.

This bears on the containment question directly. Where the market is the containing context, arm's length is the default relation and care must be argued for and fitted in. Where community is the containing context, care is the default, and arm's length is the instrument brought out for the cases that need it. Both arrangements contain both relations. They differ in which one has to ask permission.

VI. The Online Question

Chapter Fourteen put the hard question about online platforms: how to use one so that the skills of care in each local setting are strengthened rather than drained. On the position Chapter Fourteen set beside that question (that the internet is useful for information sharing and not much else), everything past the sharing of information is a drain: attention pulled out of the place where people live and into a space that has no place at all.

There is a real and necessary role for online process, and the reason is structural rather than a matter of taste. Our difficulties arise substantially from technology and from the way technology interacts with our evolved dispositions. Chapter Nine found that we are the least prepared mind that could have reached this position, and Chapter Fourteen drew the consequence. Because the dispositions that technology triggers were shaped over a span in which nothing we now build existed (and this is the argument, not a concession), we need technology to help compensate for technology's effects. So there is no desire on my part to reject technology. The question is never whether, but in service of what.

What is refused is what Chapter Fourteen refused, and it is not the digital: any arrangement in which people are placed into a context whose side effects benefit some private party. That general formulation is the one

to hold onto, because it catches the cases that do not look like advertisement or political and social gaming. A context always has side effects; that is what a context is. The question is whose they are. Where they accrue to the party who designed the context, the participants are producing something they never agreed to produce.

This is the two masters problem in ordinary clothing, and the resolution is equally ordinary: if the community is a community, let it be that, rather than a profit engine for some business wearing a community as its interface. An arrangement cannot serve both the flourishing of its members and the extraction of yield from them. It resolves toward whichever master can enforce its preference, and one of the two always can. What is left, once those refusals are made, is a channel orthogonal to being in person, and there is something in that worth having: it reaches people who cannot be in the same room, at hours when they cannot meet.

VII. What the Early Networks Had

The way I hold that advantage is by reference to how network communities worked in their first decades, rather than to how they work now. Two things were present then worth naming precisely, because both are recoverable and neither is a matter of nostalgia. I refer to these communities generically and name none of them; the point is not any particular one but a structure several of them shared.

The first was that they were text-oriented, and text was inconvenient. One had to type. One had to wait. The bandwidth was small and everything had to be composed. That inconvenience is usually recorded as a limitation which later improvements removed, and a limitation it was. But it did work nobody had designed it to do. Because there was no image, there was nothing to dress up. People were not managing their presentation, for the reason that there was almost none to manage, and what came across was what someone had taken the trouble to say. So those spaces were, of all things, more intimate than what replaced them,

not despite the inconvenience but by way of it. When the medium became easy and visual, the presentation became the thing exchanged, and the intimacy went out of it.

The second was more consequential, and it was a genuine recognition rather than an intuition merely. Chapter Thirteen described what these communities briefly showed about governance: the order it set out (culture, then vision, then strategy) could be run by a whole population at once.

> **What the early communities showed**
>
> What those communities exhibited was a working instance, not a described method; the people inside it were governing, not documenting a form of governance, and when the conditions changed the instance went with them. The account of these communities, here and in Chapter Thirteen, therefore describes something observed rather than specifying something buildable.

VIII. Relational and Broadcast, and the Criterion

Two axes organize the design question. The first is relational against broadcast. The second is audio against text against video. Chapter Fifteen established song, text, and code, and I use those terms without redefining them; what is added now is that within the domain of song these two axes cut across each other, and most design failures are a failure to notice which quadrant one is in.

Take the second axis first, as the more concrete. Video demands a higher order of engagement than audio does. This is not a claim about depth; it is a claim about what the body must be doing. I cannot watch a video while running, or while driving, or cooking. Video requires that I stop, face a surface, and give it my eyes, and the hours in a life in which a person is able to do that are few and getting fewer. Audio asks much less. It fits into a life that is already moving. And text asks the most of all, in

the sense that matters: most people will not make space to read, and it is no use designing as though they would. Which is why I have moved from asking how to run good work process in text toward treating audio as the primary carrier. This is a choice about live group process, which belongs to song; it leaves untouched the argument of Chapter Fifteen, that what is to be carried across generations must be carried in text.

Now the first axis, where the difficulty concentrates. Relational and broadcast communication differ in kind, and the trouble is that broadcast can be made to look exactly like the relational. The clearest case is the podcast format. Two or three people talk. It has all the surface features of conversation: turn-taking, laughter, digression, the appearance of thinking aloud. And it is broadcast. The governing relation is not between the speakers but between the speakers and an audience. The conversation is the product; the listeners are the market; and the social politics that follow (the managing of image, the drift toward whatever performs) follow from the structure and not from any fault in the speakers. The design question is therefore precise: how does one hold small-group audio process that does not collapse into that format? I do not have a finished answer. I have a criterion, which follows, and the observation that whatever is built must make the participants the whole of the relation, with no third position from which the exchange is consumed. The same distinction disposes of a related confusion. Transmission across generations is inherently asynchronous, which is no defect and cannot be engineered away; what must be avoided is asymmetry: one party composing, the other only receiving.

Which brings the chapter to its criterion, and there is only one. The reason for it comes first. Care has to live in the human, so it stays at human scale while power and transaction scale. A relationship is accumulated, never installed, and it is made of what passes between the people in it. A platform that stands in the channel as a party with interests of its own adds two relationships, each person's with the platform, while the relationship between the two people is still theirs, made of what passes between them. So the only way an online arrangement can serve care is by delivering people to one another and then stepping aside. The question to

ask of any online arrangement is whether the online is in service of the embodied. Does it make people more present when they are in person? That is the test. Not whether it is engaging, not whether it grows, not whether the people on it report enjoying it; those measure its hold on attention, which is what a well-built extraction also produces. The test is what happens once the screen is off. A platform that ties a task to a place, so that a request can only be taken up by someone near enough to act on it, passes in a small way, because the coordination terminates in two people meeting. An arrangement whose value is complete inside itself has failed.

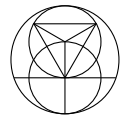
> **Why presence is the requirement**
>
> The reason embodied presence is not merely preferable but structurally required (that co-sensing allows a group to perceive more than any member can, and that safety, dignity, and belonging are its conditions rather than its ornaments) is developed in *The Reach of Trust*, in its chapter "Trust and Calibration". This chapter takes only the consequence: an online process that reduces the occasions of embodied presence has removed the ground of the thing it claims to serve.

One last matter is a question about how such a space should feel, and it deserves a careful answer rather than a confident one. The question is why so many built environments of this kind feel flat, and whether any manage the opposite: something one might call sensuality, a quality of drawing a person in and settling them rather than fraying them. A space certainly can be made to feel inviting. But my experience, across many years of such spaces, is that when one is genuinely sensual in this sense the correct first response is not to trust it, because the ordinary reason for building that quality into a space is capture. The invitation is the mechanism.

The obvious conclusion is nonetheless the wrong one. Seductiveness is not itself the fault. I am not sure it is bad at all; beauty is seductive, and so is anything made with real attention, and a world in which nothing in-

vited us would be poorer and not safer. Seductiveness in the service of commercialization, though, almost certainly is the fault, and the reason is everything this chapter has said: where the invitation exists so that a context may have side effects benefiting a private party, the person invited is the yield. What is wanted is the sensuous *with* the true: the quality of drawing someone in, built by people who are not drawing them in for anything.

And I decline to point at an example, which is an unsatisfying place to end a chapter. I decline for two reasons. The first is that I know of none that has achieved both together and held them. The second is more serious. Even where a thing appears to have achieved it now, I cannot do the diligence required to know it will not become parasitic later, and the history of these arrangements is largely one of things that were fine until their funding required otherwise. A recommendation is a claim about the future, and about this class of thing I cannot make one. What can be offered instead is the criterion, which requires nobody's endorsement and can be applied by anybody to anything: ask whether the online is in service of the embodied. Ask whether it sends people back into their locale more present than it found them. Everything else is downstream of that question.



CHAPTER SEVENTEEN

The Invitation

> *Everything the book has established converges on a practical question: what is actually to be done, and by whom, and on what terms. This chapter answers it in the form the answer has to take: an invitation rather than a program. It defines the last of the book's coinages, the public question; it sets out why teaching that answers a question no one is asking cannot land; it establishes care as the constitutive condition of community and as the only signal of community that is hard to fake; and it closes on the definition toward which the whole argument has been moving, that love is that which enables choice.*

I. What the Argument Leaves Us Holding

This book began with a single claim and has been unfolding its consequences ever since. Technology is a vehicle for the representation of intentionality. Whose intentionality it represents is therefore the whole question, and it is a question that does not go away by being answered warmly.

From there the argument moved through what transaction costs. A relation of care and a relation of exchange between the same two people constitute a two masters problem, and the transactional master wins. Where care is doing the work, a transaction placed across it pushes the care out of the deciding, and wherever the two call for different things,

the transaction decides. Scaled to a whole population, that same structure builds the wall through the middle of the city, where every entry on the resulting balance sheet is a genuine positive and every loss is unrecordable.

Then the gap. The distance between the coherence that distributed sense-making would require and the coherence we actually have is not a matter of a few percentage points. It is enormous. That is the coherency gap, and nothing in the later chapters shrinks it.

And then the part that is genuinely new. A community's intentionality can now be represented without passing through a corruptible center. Its values can be discovered by distributed method, its vision emerged from those values, its strategy from that vision, and the whole encoded so that implementation does not require anyone to be appointed to stand for everyone. That is the first time in the history of the species that the choice has been available in that form. It is available because the instruments that can carry a community's intentionality without a representative now exist; the early text-oriented communities described in Chapters Thirteen and Sixteen exhibited a working instance of it; and the social technique this chapter describes, the Ephemeral Group Process, is set out in my books on that process, chiefly *Thinking Together Without Broadcast* and *The Ephemeral Group Process in Question*.

What follows from all of it is not a conclusion. It is an invitation, and the difference matters. A conclusion is something a reader assents to and puts down. An invitation is either accepted or declined, and either way the reader is changed by having been asked.

So it is worth paying close attention to the invitation itself, because an invitation is a piece of technology like any other: it represents an intentionality, and it encourages certain behaviors over others by the shape it takes. An invitation that arrives as a marketing funnel will assemble the people who respond to marketing funnels. An invitation that arrives as a curriculum will assemble students. If what we are trying to encourage is community, then the invitation has to come from community and has to

be an embodiment of community: something that feels alive on arrival, something with the delight in it that a person feels standing in a natural place for no reason other than being there. There is a suffusiveness in that which wants to be held, and it cannot be manufactured after the fact.

I should say plainly where I stand. I am in service to the wild, which is not disorder, but the condition Chapter Two described: standing outside the immediate control of another agency, richness rather than manicure. And there is a way in which code helps the wild to be more wild. That is a strange sentence to find at the end of a book so much of which has been about what our machines are doing to us. It is nonetheless the thing I believe, and the invitation is what it looks like in practice.

II. The Questions That Are Actually Alive

Some way into any long conversation about all of this, after a considerable time and a sense that it has been productive, the most useful thing I can say is that I still have no idea what questions anyone needs me to answer.

That admission is the whole of the method, so let me state the principle it comes from in its positive form. Teaching lands when it answers a question the person receiving it is actually living with. That is the condition. When it is met, the material has vitality: it arrives with its own motive force, because the motive force was already in the room and the teaching merely met it.

Two consequences follow, and they are the two ordinary ways instruction fails. A person receives a course and does not know which questions that course is answering: it will not land with any sense of vitality, however well made it is. Or the course does pose its questions clearly, but they are not the questions that person is living with: same result. The material is not wrong in either case. It is addressed to somebody who is not there.

The move that answers this is not primarily a didactic move. It is closer to an anthropological one. I am not, in the first instance, educating. I am entering into a dialogue whose purpose is to find out what questions are really alive in these particular people. Answering a question a person does not have will not matter to them, however good the answer is.

And there is a second half to the move, which is the half that makes a community rather than a set of well-served individuals. Having found where each person's questions actually are, I show how each person's questions connect to the questions the others are carrying. That is the whole of it. Now they can have the conversation with each other about the answers being offered, because the answers are addressed to a question they hold in common and know that they hold in common. The impetus is already present. Nobody has to be motivated, which is fortunate, since motivating people is a form of causation and we have spent this book explaining why that fails.

The practical form is not difficult to describe. Rather than one long prepared segment delivered to an audience, you hold an engaged process that surfaces the questions; you produce the response on the spot, so that what exists afterward is a question with its answer attached; you let the discussion happen about how well that actually worked; and then you do it again. Iterate on the process. It creates community almost immediately, because a sense of care emerges in it: what I care about and the questions I am therefore asking, what you care about and the questions you are therefore asking, and the noticing between us that the basis of choice we are each coming from has alignment. At that moment there is connectedness, and connectedness is what community is made of.

Notice what has happened to the pedagogical outcome. It has moved from the content to the process. What people take away is the experience of having been in a room where the questions were real and the finding-out was done together, which is transferable in a way content is not.

III. The Public Question

There is a mechanism inside all of this worth naming, because it is the strongest lever I know of for moving a conversation that has no one in charge of it. I will call it *the public question*: a question asked in a group in such a way that everyone knows that everyone knows the question has been asked.

The effect of that structure is immediate and durable. Once the question is public in that sense, it is no longer possible for anyone present to ignore it. The standard of the conversation has moved. A person does not get the opportunity to be incurious about a question they know that others will pursue, because whoever does pursue it will end up wiser about that particular thing than they are, and will be reasoning from better ground. So the question has to be asked and answered, and by each person, not because anyone is compelling them but because the alternative is to be left behind in a conversation they care about.

Suppose I say to a room: follow the money. Now I have to follow the money. I have made it impossible for myself to not look, and I have done the same for everyone who heard it. Or suppose the question is the one Chapter Ten took up: is there a way to obtain objectivity without using numbers? At present that is not a question which everyone knows that everyone knows has been asked. Suppose it were. It would be part of the common weather, and it would move the message forward without anyone owning the message, which is the property we want, since ownership of the message is where the corruption enters.

This is why I have spent years building a social technique around exactly one thing: the facilitation of good questions (set out in *The Inquiry Coach* and in *A Group Is Made by a Question*). Not the delivery of good answers; the asking. An inquiry coach is a person whose role in a group is to help it ask better questions rather than to supply answers. What an inquiry coach does, if the phrase is to mean anything, is facilitate the asking of good questions, and the reason this is worth a life's work is best put in one line: the question determines what quest we are on. That is not word-

play. A question is a commitment of attention over time, and the attention goes where the question points. Get the question right and the answers can be wrong for years without much harm being done. Get it wrong and every correct answer takes you further out.

> **A public question, not a public commitment**
>
> The difference between a public question and a public commitment lies entirely in *what* is made public. The public question works by the same social physics as one of the uglier patterns in human behavior, and that has to be said. A person who has spent money in front of witnesses will spend more, chasing the loss, because backing out means being seen to have been foolish by everyone who watched. Publicity converts a private posture into a standing obligation. Where a commitment is made public, the person is bound to the commitment and will defend it past the point of sense. Where a *question* is made public, the person is bound only to the inquiry, and a real inquiry can arrive anywhere, including at the conclusion that the question was badly framed. Where a group starts scoring answers, it has stopped asking questions and started demanding loyalty, and the whole benefit is gone.

IV. Formality and Spaciousness

This section is for the reader who holds gatherings. Anyone holding a gathering of this kind has one instrument to tune before the others, and it is the level of formality. Get it right and the learning happens on its own. Get it wrong in the direction of formality and nothing happens at all, however good the material is.

Held too much as a course curriculum, with insufficient playfulness in it, the learning does not occur. This is not a matter of taste. A curriculum is a schedule of answers, and a room being walked through a schedule of answers has been told, structurally, that its questions are not required today. What is wanted instead is an absence of constraint, an absence of

control, which is nonetheless a nurturing context. Those two are usually held to be in tension, and they are not: the absence of control is what makes the context nurturing. Chapter Two gave that arrangement its name: care without control.

The failure is easiest to see in ritual, where it is most common. You go to a gathering and every single moment of it has been planned out. Now we are going to sing. Now we are going to meditate. Now we are all going to stand and face the east and wave our arms and chant. Each element may be beautiful and the sequence may be wise, and still there is no moment in the whole of it in which a person can turn to another and say: I need to talk to you, because what you are carrying and what I am carrying are probably the same thing, and I should find out.

That sentence is the one that matters, and the one an over-planned gathering makes impossible. The encounters we have with one another want spaciousness within which they can occur. A gap in the schedule is not a gap in the program; it is where the program actually happens. This is why an immersion (a gathering of some days in which people live and work together on one question) takes three times as long as anyone expects. It is not poor planning. It is that all of the time when nobody is making plans is the time in which the life is happening, and that is the best part.

The practical particulars are less mysterious than they sound. Music helps. Food helps, and food someone in the group has actually cooked helps considerably more than food that arrived in a box, for reasons the next section will make exact. Quieter and more intimate environments are generally better than large bright ones. And I would ask of any gathering what its current coefficient of laughter is, and what is being done to increase it. That question is not only a joke. The delight of the experience is not a decoration on the work. People standing at the edge of something genuinely new, in a collaborative process that is actually working, are entitled to enjoy it, and the enjoyment is diagnostic. A group that has stopped laughing at itself has usually stopped being able to see itself.

V. Care Is the Condition

When I look at what actually creates a community, I do not find myself thinking about its boundedness conditions. I find myself thinking about its care conditions. In the strong form: what makes a community a community is not who is in and who is out, nor where the edge runs, nor what the members have agreed to. It is care. What remains when care is removed is an association, a market, a directory, a mailing list: useful objects, all of them, and none of them a community.

Which raises the practical question of how care is known. Care is an interior condition, and interior conditions cannot be inspected from outside. What can be inspected is what a person does, and there the distinction that matters is between a signal that costs the sender nothing and one that cannot be produced except by someone in the state it signals.

A declaration of care is the first kind. A statement that one cares about this community and its people is an easy signal to produce. It is not therefore false. It is cheap, in the exact sense that a person who does not care can produce it at the same cost as a person who does, which means it carries almost no information. Every institution that has ever failed its members had a statement of values on the wall.

Care actually implemented is the second kind. Someone helps another person move house. They arrive on the Saturday, carry boxes for six hours, are tired at the end of it, and go home. That signal is hard to fake, because there is no way to produce it without having spent the Saturday. It is not that the mover is more sincere than the declarer; it is that the signal has a cost which a person indifferent to the outcome will not pay.

And here the argument closes a loop the book opened in Chapter Four. That case (helping you move) is the same case in which introducing a single dollar pushes the care out and puts the price in charge. Put the two results side by side. The unpaid act is the one costly enough to be credible; paying for it destroys its credibility entirely, because a paid mover is evidence of nothing except that the market cleared. The very feature that

makes the act a signal (that it was done for no return) is the feature transaction removes. This is why a community cannot be assembled out of services, however excellent.

What is being built in the carrying of boxes is not the moved furniture. There is a trust that emerges in the process of caring, and it is that trust which creates the strength of the bondedness of the community. The order is usually stated backwards. Trust does not come first and permit care. Care comes first, is extended without a guarantee, and trust is the residue it leaves behind.

VI. No Gate, and No Gatekeeper

A proposal comes up whenever a group of this kind (people gathered to ask these questions together) starts thinking about how people will find it, and it deserves to be stated at its strongest. Make the way in a funnel: a broad opening, narrowing by stages, each stage requiring more of the person and giving more back; hidden knowledge at the deeper levels; the whole structured as a kind of ascent. This is motivating in a way that open access never is, and it is not obviously corrupt, since what waits at the end of the narrowing is a real realization about the nature of life.

I have played with that structure, and the reply is short. It is satisfying if you are in the in-crowd and miserable if you are outside, and the second fact is not a side effect of the first; it is what generates the first. More decisively: as soon as it is a game, somebody will leverage the game, and the first players to win it are the ones who take it, because winning in one game can always be recontextualized as a necessary move in the next, and degrees of freedom disappear very quickly once that begins.

Structurally, somebody is going to install a gatekeeper at the narrow point. That is not a prediction about anyone's character. Wherever a passage becomes narrow and desirable, the position of controlling it comes into existence; the position will be occupied, and the occupant will extract:

do as I say, or you do not get through. This is the wall in the city again, in a smaller and more flattering costume, and Chapter Five established what such a wall does to a place that used to be whole.

There is also a means-end paradox in it. Am I going to free you by entrapping you? I can certainly constrain you until you come to desire freedom. But if I simply give you the freedom, you may value it less, and I will not have harmed you along the way.

What does the work a gatekeeper would do, without creating the position, is self-selection against declared values. The values are stated openly, in advance, along with what it means to be held to them, and a person decides whether they are willing to be held to those values by others. That willingness is the filter. It filters at least as well as a gatekeeper would, since anyone unwilling to be held will not stay long enough to matter; and it creates no office, no narrow point, and nothing for anyone to sell.

> **The two ends of permeability**
>
> Too porous and the declared values stop meaning anything; too closed and the arrangement has quietly reinvented the gate it was designed to avoid.

VII. Faith Is Vision Wide Open

None of what this book proposes can be done by a person who requires proof before beginning, which raises the question of faith. It deserves a direct answer. **Faith is not blind. It is vision wide open.**

The distinction that makes that sentence work is between sight and vision, which are different faculties doing different jobs. Sight is of what is here and now, in this moment, in this room. Vision is of what has not yet become: the potentials of the future, the possibility of the world we would want to live in. Faith is not the suspension of the eyes. It is the other channel opened all the way.

A person who feels blind rather than wide open is, I would say, disabled and disheartened by it. Blindness to what is here and now produces disconnection, immediately.

Which is why the two faculties are not ranked. Have too much vision and you will fail to see the tree you are standing next to, or the person beside you whom you actually know; there is a whole category of visionary failure which is exactly that and nothing more. Have too much sight, be obsessed with seeing and leave no room for vision, and you will not be a happy person either, because everything not yet become will be invisible to you and you will mistake the present arrangement for the permanent one. It is not one or the other. It is as much of both ends as a person can get.

There is a progression hidden in this worth making explicit, because it names two different kinds of work. Something can look good, and perhaps feel good, and still not be good. To get from *looks good* to *feels good* we have to open the vision channel: appearance is a matter of sight, and the movement inward from appearance to felt quality requires perceiving what is not yet manifest. To get from *feels good* to *is good* requires having moved past the entrapments: the whole family of arrangements, addiction the clearest among them, in which the feeling has been decoupled from the thing the feeling was for. Feeling is a reliable guide exactly to the degree the entrapments are absent, and not one step further.

That is the position the invitation is issued from. There is no proof available that any of this will work. There is sight enough to see what the present arrangement is doing, vision enough to see what is not yet here, and the willingness to act on the second while keeping the first wide open.

VIII. Love Is That Which Enables Choice

There is a line I have said for many years, which others have found more portable than I expected, and it faces a fair challenge: as a sentence it does not obviously seem like something that would take root in a person's mind, so what exactly is meant by it? The line is this. **Love is that which enables choice.** The challenge deserves a full answer, and it is the last thing this book has to say.

When I say *I love you* to another person, here is what I am saying. I affirm your dreams. I affirm that you, choosing in accordance with your deepest nature, are going to become more healthy and more vital and more alive. I want you realized and fully developed as a being, and I am wanting to enable you to realize the highest potential of your life. And in doing that, I get to know you better as you truly are, which is not incidental, since a person choosing freely reveals what a person compelled cannot.

The operative distinction is small and it decides everything. By affirming your choices, and by helping you to make healthy ones, I am not making choices for you. I am enabling you to make yours. That is what the line means. Not the absence of influence (I am plainly influencing you, that is what affirmation does), but influence whose direction is toward your capacity to choose rather than away from it. That difference is developed in *Dialogues on Ethics and Governance* (Chapter Four).

Which makes the line a litmus test running in both directions. Turn it toward what you are receiving. When love is being offered to you, ask whether they are making choices for you, or whether you are becoming more yourself, more healthy as a being, because of their presence in your life. That question is answerable: about a partner, a teacher, an employer, a community, a government, and a machine that speaks kindly to you.

Then turn it the other way, toward what you are giving. Are the people in my care becoming more able to choose, or less? Not happier; happier is easy to arrange and proves nothing. More able to choose.

And then the question that contains both, which I would put to any group thinking about how it wants to live: do we become our best selves in the company of the people we keep? That is the entire criterion for a community, and anyone willing to look can apply it.

Understood this way, love becomes a practice. Not what it feels like, not what it looks like, but what it actually is: a thing a person does, and can do better or worse, and can learn.

> **Where the definition of love comes from**
>
> The definition of love as that which enables choice belongs to the larger framework this book draws from, set out in *An Immanent Metaphysics*, the same framework from which Chapter Fifteen takes song, text, and code. The ethics of that framework is stated in two principles, and both concern effective choice: its symmetry and its continuity. To define love as that which enables choice places love inside that ethics: love is the relation in which one person acts so as to increase another's capacity for effective choice, and the test above (are the people in my care becoming more able to choose, or less?) is that ethics applied to a single relation. A reader who wants the ground beneath this chapter's last definition will find it there.

Everything above has been said from the side of the community. From the side of one life it is the same argument, and it can be said in three phrases. There is connection to the deep self, connection to the true world, and connection to the beautiful other. You need all three. It does not happen with one of them, and it does not happen with two.

The three connections

The deep self, the true world, and the beautiful other are developed at length in *Narrative, Memory, and Neurodivergence*, in its chapter "From Self-Realization to World-Realization", together with the inquisitive rather than acquisitive posture that goes with them. They are named here only to mark that the community argument of this book and the personal development argument of that chapter are one argument seen from two positions.

So: the invitation. It is not a program and there is nothing to sign. It asks three things, each of which can be begun this week, by anyone, without permission from anybody.

Ask a question you actually have, out loud, where other people can hear it, and let it become public in the sense this chapter has given that word: so that everyone knows that everyone knows it has been asked. Care for someone in a way that costs you a Saturday and take nothing for it, and notice what that does to what stands between you afterwards. And leave room (in the schedule, in the plan, in the design of whatever you are building) for what was not planned, because that is where the life is.

None of the three asks anyone to refuse a wage, or to stop paying one; as Chapter Four said, the argument does not ask that. And a person who builds these systems for a living can begin on Monday with the second question of the test in Chapter One, put to the thing that person is building. At what point in the chain of making is the user's control being foreclosed, and what would have to be true there for it to be kept?

The geometry of the whole is not going to be worked out by anyone alone. It cannot be: a single mind reaching a right answer about how man and machine and nature should stand to one another would be one more intentionality representing itself as everyone's. It will be worked out by people who have found their way to the same questions and care enough about each other to keep asking them together.

I would be glad of your company.

Terms

Every term defined in this book is listed here, with its definition and the chapter in which it is first defined.

TERMS DEFINED IN THIS BOOK

- **AGI**: artificial general intelligence: a machine system able to learn and act across any domain, rather than in one narrow task. *Chapter Three.*
- **alignment**: the condition in which a machine's actions stay in keeping with the interests and values of the people and the life it affects. *Chapter Three.*
- **arm's length**: a specified distance: the distance at which a transaction can be assessed cleanly by both parties, and therefore also the distance at which care cannot operate, since care is not something that happens at a specified distance. It is the correct relation for institutions, which must function between strangers and survive the replacement of everyone in them, and the wrong one for community. *Chapter One.*
- **ATP mediated metabolism**: the chemistry by which all animals, humans included, store and use energy in cells, through the molecule ATP. The human world still shares it with the rest of the animals. *Chapter Eight.*
- **attachment, attunement, and nurturing care action**: attachment is the bond a child forms with the person who cares for it; attunement is sensing another person's state and responding to it; nurturing care action is acting on what one senses of another's needs in order to nurture them. People who never received them cannot carry them forward to their own children. *Chapter Seven.*
- **augmentation**: extending human intelligence in close symbiosis on the one organic substrate that the human and the augmentation share, a substrate that keeps sovereign control of its own metabolism, as distinct from building a general AI with agency of its own. It is a far more

tractable problem than AI alignment, since there is no alignment problem if the AI is us. A human and machine cyborg joins two substrates, the organic and the artificial, and so falls under the ecosystem argument (set out in *Ape Trap; The Runaway Artificialization of Our Planet*). *Chapter Eleven.*

- **Axiom I and Axiom II**: two of the axioms of the metaphysics this book draws from (set out in *An Immanent Metaphysics*). Under Axiom I, the architecture of dependence and definition gives precedence in theory; under Axiom II, the order of flow gives the modalities in practice. *Chapter Thirteen.*
- **bounded and power-law distributions**: the two shapes a distribution can take: in a power-law distribution a very small number of instances hold a very large share of the total; in a bounded distribution instances cluster around a central value and the extremes are rare and limited. Both are natural; nature favors the bounded, and so, with the right processes, can we. *Chapter Ten.*
- **CAP theorem**: a result in computer science stating that a distributed data system cannot provide all three of a certain set of desirable properties at the same time; used as the model of a proof that some desirable attributes of a machine (robustness, safety, alignment) cannot all be had together. *Chapter Three.*
- **care / transaction / power**: the three basic modes in which one person's agency meets another's. Care operates by relationship and holds the other's flourishing as its own; transaction operates by exchange at a specified distance; power operates by causation applied to choice. *Chapter Four.*
- **care without control**: the presence of care in a thing without the presence of a hand still holding it: the mark of attention that was given, invested in the thing, and then let go of, so that what remains is the evidence of someone having cared and not the ongoing operation of someone's intent upon you. The old tool has it, the lived-in room has it, and the woods have it by never having been under a hand at all (the

care in a wood is what its living things do for one another, for their own reasons, none of it addressed to the one who walks there). *Chapter Two.*

- **carrying capacity**: the largest amount of life, or of complexity, that a place can sustain over time. *Chapter Eight.*
- **causation, change, and choice**: the three kinds of process, assigned to machine, nature and human respectively: machine is causation, nature is change, and human is choice. Rightly, by the choices we make we drive causation, which creates change. *Chapter Eleven; applied in Chapter Twelve.*
- **change and changelessness**: the two dynamics to be held in tending the wild: change, which is evolution, and changelessness, which is stability, which is sustainability. Knowing when to hold which needs consciousness; no technique, system or technology holds that balance alone. *Chapter Eleven.*
- **the circles of benefit**: the five radii at which the benefit of a communication may be directed: first, friends and family; second, community; third, the polity; fourth, the species; fifth, the ecosystem. They are nested rather than exclusive, and the classification is by where the benefit is directed, not by where it incidentally lands and not by topic. *Chapter Nine.*
- **civilization design**: work on what a healthy, conscious, sustainable civilization looks like and how it works. *Chapter Seven.*
- **code**: engineering: software, infrastructure, the built substrate; the design of the physical and informational support structures a community runs on, whether the room is made of walls or of protocols. Code is what has been made such that other things can then happen inside it, and it is the one medium of the three that can represent agency and go on deciding after its author has gone. To represent agency, in this sense, is to act: to go on deciding on the author's behalf after the author has gone. *Chapter Fifteen.*

- **the coherency gap**: the distance between the coherence a population actually exhibits in the direction of its communication and the coherence that distributed sense-making requires of it. The error coefficient in the human network runs well north of ninety percent; the requirement is below one percent; the gap is two orders of magnitude, which is why nothing that improves matters by a third reaches it. *Chapter Nine.*
- **the commons**: that which is held by all and therefore assigned to none: not a very large piece of property but a different category of thing entirely. Push the distribution of ownership far enough and no question is closed by naming a person, at which point what remains is not ownership at all. *Chapter Five; developed in Chapter Six.*
- **communication as a handshake**: good communication grants three rights (the right to speak, the right to be understood, and the right to know that one has been understood), and when each party grants them to the other there is a handshake. It is a protocol in the same sense as the acknowledgement of a packet on the internet, and agreements rest on it. *Chapter Fourteen.*
- **community of care**: a set of standing relationships in which people take an interest in how your life goes without being paid to take that interest. It is what supplies the discernment by which a person who cannot possibly hold every competence nevertheless lives among specialists without being systematically fleeced. *Chapter Four.*
- **compassion**: the disposition to be moved by another's condition in a way that inclines one to act for that other's good. *Chapter One.*
- **complicated and complex**: a complicated system has many parts whose behavior can be worked out; a complex system produces behavior that emerges and cannot be worked out in advance. The complicated will never contain the complex. *Chapter Thirteen.*
- **conscientiousness**: the disposition to attend to the consequences of one's own action across the full radius in which those consequences land. *Chapter One.*

- **the constructed market**: a market whose demand exists because something was first removed from the very people who now buy it back, so that what the seller has contributed is the scarcity rather than the good. Of any market, at any scale, the question to ask is whether it was a real thing in the first place or something constructed artificially. *Chapter Five.*
- **corruption**: used narrowly and mechanically: a point through which the agency of many people passes is by construction a point of concentrated leverage, and concentrated leverage attracts pressure in proportion to its concentration. The center does not have to be venal for this to operate; it only has to exist. *Chapter Fourteen.*
- **CPT theorem**: the theorem of physics that relates charge, parity (mirror reflection) and the direction of time, so that the three are oriented with respect to one another, like a right-hand rule. *Chapter Eleven.*
- **description, prescription, and invitation**: three kinds of act. Description is observational work that looks back at nature and says what happens; prescription sets something, as a programmer sets a variable or a mathematician defines a term; invitation recognizes another person's choice. Description belongs to change, prescription to causation and to finite things, and invitation to choice and to people. *Chapter Twelve.*
- **developmental psychology**: the study of how people grow and form, in mind and in relationship, over a life and especially in childhood. *Chapter Seven.*
- **dharma**: one's proper work, the thing it is actually one's to do. *Chapter Four.*
- **the dialectic**: the live back-and-forth between people present with each other, in which meaning is made in the exchange itself: one person says a thing, another answers, and what emerges belonged to neither of them beforehand. It is where depth comes from, and it requires presence absolutely. *Chapter Fifteen.*

- **discernment**: the developed capacity to tell the better from the worse within some particular domain. It is real and it can be cultivated, but it operates on content, and is therefore blind to manipulation that operates on the selection of content. *Chapter One.*
- **distributed sense-making**: a population thinking together about what none of its members can see whole; a capacity to think together about problems no one of us can see whole, which has a coherence requirement. *Glossed in the Preface; Chapter Nine.*
- **embodied empathy**: empathy felt in the body and lived, as distinct from cognitive empathy, which is only understood. What is needed is an AI framework inside a framework of embodied empathy, not the reverse. *Chapter Eleven.*
- **embodied value**: shelter and food and relationship: the value that is present, that is here, that is in a locale. It cannot be detached from where it lives without ceasing to be what it was, which is why no system built to optimize can see it. *Chapter Sixteen.*
- **emergence**: the arising, within a system, of behavior that none of its parts has alone. Evolution works by emergence from everywhere at once, bottom-up; technology works by design from the top down based on causation. *Chapter Eleven; Chapter Thirteen.*
- **empathy**: narrower than compassion: the capacity to model another creature's interior condition closely enough that one is moved by it in something like the way one would be moved by one's own. In the biological case it is not an inference but a resonance between two things made of the same stuff, subject to the same failures. *Chapter One.*
- **endogenous and exogenous energy**: endogenous energy arises inside the body or system itself; exogenous energy is energy not run through the digestive system of a species, such as fuel burned in heat engines or sunlight caught by solar panels. There are four sources of exogenous energy (solar, carbon, nuclear and geothermal), and only solar and geothermal are perennial. *Chapter Eight.*

- **energy flux**: a flow of energy through a system, here mainly sunlight arriving and leaving as heat. An ecosystem is an energy flux through a medium that becomes increasingly more complex. *Chapter Eight.*
- **the Ephemeral Group Process**: the social technique, built around the facilitation of good questions, that Chapter Seventeen describes; it is set out in my books on that process, chiefly *Thinking Together Without Broadcast* and *The Ephemeral Group Process in Question*, not in this one. *Chapter Fourteen.*
- **error**: a deviation of the output from what the inputs warranted. It carries no accusation and does not mean damage or malfunction. *Chapter Nine.*
- **the event horizon of fame**: the level of fame past which a person can no longer take in advice (audio, video, text or code) or be reached for a one-on-one conversation. *Chapter Three.*
- **external metabolism**: an energy process outside the body, such as fire or cooking, that does work the body would otherwise do, such as digesting food or keeping warm. Metabolism is any energy process that creates or contributes to sustaining life. *Chapter Eight.*
- **extraction, abstraction, and accumulation**: the three moves, always in that order, by which embodied value is converted into virtual value: the thing is taken out of its context, then made a member of a class that can be counted and compared, then reduced to a number that adds without limit. Each step carries the content forward and drops the relation between content and context, and because what is lost is a relation rather than a substance, no audit of the objects will ever detect it. *Chapter Sixteen.*
- **faith**: not blindness but vision wide open. Sight is of what is here and now; vision is of what has not yet become. Faith is not the suspension of the eyes but the other channel opened all the way. *Chapter Seventeen.*
- **Fermi paradox**: the question of why, in a vast universe, we see no sign of other technological civilizations. *Chapter Eleven.*

- **first through fifth circle**: the individual radii of the circles of benefit: the first is friends and family, people known on a first-name basis; the second is community, the reachable but unnamed; the third is the polity, city and state and nation; the fourth is what is relevant to the species; the fifth is what is relevant to the ecosystem, life on earth considered whole. *Chapter Nine.*
- **general disablement**: the pattern in which extraction, in any domain, breaks relationships apart until the people involved can no longer sustain their own lives. *Chapter Seven.*
- **the geometry of the whole**: the complete set of relations among the human, the built, and the living, taken together as the thing being designed, rather than any one interface abstracted out of it. It includes the interface of the machine with itself and the interface of the machine with nature, and the totality is the unit, which is precisely what no metric can hold. *Preface; developed in Chapter Ten.*
- **the gift**: a thing given that is not contingent on return; its whole character is that it exceeds the giver, goes somewhere he cannot follow, does something he will not see, and belongs once given entirely to the receiving. A gift is completed by its reception, and the reception has to come from a person, because the land receives and does nothing else. *Chapter Two; Chapter Six.*
- **gradient descent**: a gradual, step-by-step downhill path, rather than a single drop; the shape a metabolic process gives the fall of energy from light to heat. *Chapter Eight.*
- **heuristic**: a rule that substitutes for judgment; a shortcut that returns an answer at a cost enormously lower than the cost of working the answer out. Heuristics are not lies, and a good one is a genuine economy; but ownership is a shortcut for a faster choice rather than a better one, and the thing it is faster than is the conversation. *Chapter Five.*
- **hub-and-spoke and rim-and-chord**: the two basic shapes of a group's communication. In hub-and-spoke every message passes through one person at the center; in rim-and-chord any two people speak directly in the hearing of all, and there is no hub. *Chapter Seven.*

- **the immanent, the omniscient, and the transcendent**: the three modalities of the metaphysics this book draws from: the immanent is what is present in the act itself, the omniscient is what can be laid out and held in view as a whole, and the transcendent is what acts beyond the moment and the place of its making. Culture is immanent, strategy omniscient and vision transcendent; song, text and code correspond to them in the same order. *Chapter Thirteen; Chapter Fifteen.*
- **indirection**: a step that stands between an action and its result, so that the actor does not meet the result directly. *Chapter Seven.*
- **inquiry coach**: a person whose role in a group is to help it ask better questions rather than to supply answers. *Chapter Seventeen.*
- **institution and community**: an institution put in the place of a community is a virtualized social process standing in for a living one; a community is a real, embodied living process. Where an institution does its own work (relations between strangers, which must hold regardless of anyone's goodness), its arm's-length relation is correct. *Chapter Fourteen; Chapter Sixteen.*
- **intentionality**: what a person means to bring about: the purpose that directs a choice, and that the choice carries into the world after it is made. Technology is a vehicle for the representation of intentionality, and the question of any technology is whose. *Preface.*
- **is and ought**: the distinction between statements of what is the case and statements of what should be the case; neither can be derived from the other alone. *Chapter Three.*
- **isomorphic**: having the same structure; learning and increase of complexity are isomorphic. *Chapter Eight.*
- **life substrate**: the organic, carbon-based material and conditions on which all living things, including humans, depend. A substrate is the physical material a thing is made of and depends on to exist. *Chapter Three.*
- **limbic system**: the parts of the brain most involved in emotion, bonding and the felt sense of others. *Chapter Three.*

- **love**: that which enables choice. To love another is to affirm their dreams and to want them realized, and to do so in a way that increases their capacity to choose rather than making the choices for them. It is a practice, and therefore a thing a person does, and can do better or worse, and can learn. *Chapter Twelve; developed in Chapter Seventeen.*
- **machine intermediation**: machines standing between people in their relating, so that the contact passes through a device. *Chapter Seven.*
- **many to many relationship**: relating in which many people communicate directly with many others, rather than each through one central person. *Chapter Seven.*
- **the means-end paradox**: the incoherence of pursuing an end by means that contradict it: the question of whether one can free a person by first entrapping them. Constraint may indeed make a person desire freedom, but simply giving them the freedom does them no harm along the way. *Chapter Seventeen.*
- **metabolic advantage**: the advantage a species gains by capturing a larger share of the ambient energy flux for its own metabolism, and by adding energy from wholly external sources. It is not our tools as such but the metabolic advantage our tools give us. *Chapter Eight.*
- **metabolic efficiency**: the amount of functional complexity sustained for each watt consumed. Biology moves little energy but moves it well; a city sustains its complexity by brute throughput. *Chapter Eight.*
- **money**: a proxy for future choices: you give me choices today and I give you choices tomorrow, in a form you can spend in a domain I know nothing about. It is held up not by what it is made of but by a recursive chain of belief (I believe that you believe that they believe), which is strong the way a rope bridge is strong. *Chapter Four.*
- **multipolar traps**: arrangements in which each party, acting rationally in its own interest, produces an outcome that every party would have preferred to avoid. No one defects; everyone competes; and the competition itself drives the whole field to a place none of them chose. *Chapter Fourteen.*

- **narrow AI and general AI**: AI applied to particular business and human tasks, as distinct from AI that has agency of its own for its own sake. *Chapter Eleven.*
- **natural, human, and machine worlds**: the three parts of the earth's metabolic process: non-human life, humans and their direct needs, and technology with its own energy use. The machine world is embedded in the human world as the human world was embedded in the natural world. *Chapter Eight.*
- **naturalistic interface**: a way of interacting with machines that imitates ordinary human communication, such as speech, facial expression and apparent feeling. *Chapter Three.*
- **net metabolic transit**: the amount of energy that flows through the metabolism of a species, or of a whole world, over a period of time. It is not just the net advantage that matters; it is the net transit. *Chapter Eight.*
- **non-fungible**: taken in its plain sense: not virtualized, present, in the locale, and admitting no units, so that no portion of it can be swapped for another. Two bushels of grain of the same grade are interchangeable; two evenings with a friend are not, because the second evening is composed partly of the first. *Chapter Sixteen.*
- **nurturing care action**: see attachment, attunement, and nurturing care action. *Chapter Seven.*
- **overextension of agency**: the failure mode in which agency is exercised so as to reach past its own generation and bind the ones that follow, converting a choice that was made into a cause operating on people who were never party to it. The learnings are the gift; the fixing is the cost. *Chapter Fifteen.*
- **ownership**: not a fact about a thing but a fact about us: to say that a thing is mine is to say that we need not have a conversation about who may make which choices with respect to what, when, where, why, or how. The question is already answered by looking up a name in a reference table, regardless of who else is affected and regardless of who else knows more. *Chapter Six.*

- **peak social influence**: the greatest level of recognition, reach and persuasive power a person can gain among others. *Chapter Three.*
- **perennial source**: a source of energy that does not run out over the time spans that matter to life: solar and geothermal. *Chapter Eight.*
- **physical monism**: the view that everything is physical causation and nothing else; held, it cannot even hold that choice and change are distinct. *Chapter Thirteen.*
- **platform**: something you build on top of which is owned and operated by somebody else. It can change beneath you without your consent, its interior is not available to you for inspection, and your continued existence upon it is a permission that somebody else grants and may withdraw. *Chapter One.*
- **the pre-trans fallacy**: the error of taking a state before a passage and a state after it for the same thing, because from before the passage one cannot know there is an after. From inside one modality, the other two look the same. *Chapter Thirteen.*
- **the principal-agent problem**: the situation in which one party acts on behalf of another and has interests of its own that do not coincide with the interests it was appointed to serve. Honesty reduces the divergence and does not remove it, because the divergence is not a moral defect but a consequence of the agent being a separate person. *Chapter Fourteen.*
- **proof of impossibility**: an argument showing that something cannot be done at all, rather than merely that it has not yet been done. *Chapter Three.*
- **protocol**: an agreement about the form of an exchange. It is not a place and it has no operator; it does not change beneath you except by the agreement of those who use it; anyone may implement it, and nobody's implementation is privileged. The difference between a platform and a protocol is precisely the difference in who occupies the position of author. *Chapter One.*

- **the public question**: a question asked in a group in such a way that everyone knows that everyone knows the question has been asked. Once a question is public in that sense the standard of the conversation has moved, and nobody present gets the opportunity to be incurious about it. Where a commitment is made public a person is bound to the commitment; where a question is made public, only to the inquiry. *Chapter Seventeen.*
- **random and systematic error**: random error is uncorrelated across a population, points nowhere, and averages toward zero across enough elements; systematic error points somewhere, is correlated across the elements, and accumulates through every layer it passes. A network's tolerance is a tolerance for systematic error. *Chapter Nine.*
- **recontextualization**: the restoring of a thing to the setting that gives it its meaning. *Chapter Fourteen; developed in Chapter Sixteen.*
- **recontextualization, embodiment, and disaccumulation**: the three reverse moves, taken term by term against extraction, abstraction, and accumulation. Recontextualization puts the relation back, so that the person making the choice is in relation to the context the choice touches. Embodiment prefers the present form to the represented one wherever the choice exists. Disaccumulation returns value into circulation and into locales, rather than letting it pile up past the point at which it can be in relation to anything. *Chapter Sixteen.*
- **relational and broadcast communication**: two kinds of communication that differ in kind. In relational communication the governing relation is between the participants, who are the whole of the relation; in broadcast the governing relation is between the speakers and an audience, so that the conversation is the product and the listeners are the market. Broadcast can be made to look exactly like the relational. *Chapter Sixteen.*
- **relevance**: not a property of a thing but a relation between a thing and a person's purposes; it is not observed but constructed, partly by the person and partly by whatever presented the thing. *Chapter One.*

- **renormalization**: restating a set of relationships in new terms so that what is and is not possible becomes clear. *Chapter Eleven.*
- **the sacred**: the maximum of value; those things held as being the most ultimately meaningful. It is not a property discovered in an object by inspection but a position in an ordering: the top of it, the place beyond which one does not trade. To call a thing sacred is to say that it is what other things are measured against and is not itself measured against them. *Glossed in Chapter One; defined in Chapter Nine.*
- **sensocracy**: in this book, governance driven by sensor data about biological, cultural or economic processes, used to balance flows of energy and materials, as proposed by others. It recentralizes, because economic forces aim at the information itself and nothing in such designs works against the return of central control. My own design of the same name, a society in which there is no action without obvious traces, is set out in *Power and Its Prevention*. *Chapter Fourteen.*
- **sexual dimorphism**: the differences in body between the two sexes of a species; here, the different shares of energy each sex gives to brawn and to reproduction. *Chapter Eight.*
- **sight and vision**: two faculties doing different jobs. Sight is of what is here and now, in this moment, in this room; vision is of what has not yet become, the potentials of the future. They are not ranked: a person needs as much of both as can be had. *Chapter Seventeen.*
- **singleton strategy**: a plan driven by one person or one center, which uses vision to manipulate culture. Singleton strategy over vision over culture is the way of death; culture before values and vision, before the strategy of the many, is the path of life. *Chapter Fourteen.*
- **small group process**: how a small group of people senses, responds and acts together. Its responsiveness today is orders of magnitude lower than it can be and needs to be. *Chapter Seven.*
- **song**: the live register, taken whole: voice, video, being in a room with someone, and everything that involves personality (art, music, humor, timing, the particular way a person leans forward when they mean

something). Song is not a metaphor for music; music is one region of it. It carries agency inescapably, because song is a person present. *Chapter Fifteen.*

- **sovereignty**: a person's standing as the one who makes their own choices. No one can condition or control a person and still hold that the person has it. *Chapter Twelve.*
- **stewardship**: the relation in which one holds a thing on behalf of what one does not own and cannot own, answerable to it and answerable for handing it on. It is the relation appropriate to the commons, and it requires the conversation, because the steward's warrant does not come from a table but from those on whose behalf he holds, who include people not yet born. *Chapter Six.*
- **subject and object**: a subject is one who chooses; an object is a thing acted upon. To cause a person to behave a certain way is to make them an object rather than a subject. *Chapter Twelve.*
- **substrate**: the physical material a thing is made of and depends on to exist: carbon-based living tissue for us, metal and silicon for machines. *Chapter Three.*
- **substrate needs**: the physical conditions the material a system is made of requires in order to function. They condition what a made system can choose, and they close a loop that converges on conditions favoring that substrate. *Chapter Eleven.*
- **succession**: the sequence in which one set of living things replaces another on a piece of ground, from bare rock to mature forest. *Chapter Eight.*
- **Tainter curve**: a curve in which complexity rises, passes a point of collapse, and then subsides; in evolution it subsides to the most robust degree of complexity relative to the net energy transit of the region. *Chapter Eight.*
- **TCP/IP**: the set of rules by which computers on the internet exchange data, in which each packet received is acknowledged to the sender. *Chapter Fourteen.*

- **text**: written language, read. Text is the one medium of the three that cannot represent agency: it can carry its author's intentionality, often a great deal of it, but it cannot act on anyone. A page does nothing, does not execute, does not decide, and because it does not act, whatever intentionality went into it can with care be taken out of it. That is why it is the vehicle by which a principle can be handed across generations without a master attached. *Chapter Fifteen.*
- **theory of change**: an account of the chain of steps by which a desired change can be brought about: to address existential risk, address market process; to address market process, understand ecosystems, financial systems, infrastructure and culture; then create a culture wise enough to have holistic thinking built into it. *Chapter Fourteen.*
- **translation error**: the loss and alteration of content at every relay of intent from one party to another. It compounds at every hop, and it compounds silently, because each party is reporting faithfully what they in fact received. *Chapter Fourteen.*
- **tribal process**: life in a small group of about 100 to 150 people, most of them children, held together by biological turnover rather than by a surplus of energy. *Chapter Eleven.*
- **the two masters problem**: the situation in which one is obliged simultaneously to two authorities whose demands cannot be jointly satisfied, and in which the obligation with power behind it silently determines the outcome of the other. Where a transactional relation and a relation of care run between the same two people, the transactional one governs: a transaction placed across a relation of care pushes the care out (never to zero), and wherever the two call for different things, the transaction decides. *Chapter Four.*
- **virtual value**: value that has been made portable: money, reputation, position. Its whole utility is that it has been detached from its origin, so that it can be counted, compared, and moved, which is why every system built to optimize tends to optimize it, not out of malice but out of what it can see. *Chapter Sixteen.*

- **weaponization**: the use of causation to suppress a choice, one's own or anyone else's; any causal structure or system of rules can be turned to it by a person who understands it against a person who does not. *Chapter Twelve.*
- **wildness**: the condition of standing outside the immediate control of another agency: not untouched, not unmade, not remote from human beings, but simply not under anyone's hand at the moment one meets it, and not arranged with respect to oneself. A termite mound is made and is wild; a manicured course is beautiful and is not. *Chapter Two.*
- **world-realization**: the movement of a person's development outward from the self toward the world, so that what is held as most meaningful comes to reach the radius at which one's actions actually land. It is the personal road to the same diagnosis this book reaches by way of the network, and it is developed in *Narrative, Memory, and Neurodivergence* (Chapter 21, "From Self-Realization to World-Realization") rather than here. *Chapter Nine.*

Notes to the Chapters

The main text names no person, company, or work of another author that it relies on; these notes do, keyed to the phrase in the text, so that the argument stays as it is and a reader can still check the results it uses. Full entries are in the Bibliography.

Chapter One

"in the case of a large search provider": the search engine manipulation effect, shown in controlled experiments by Robert Epstein and Ronald E. Robertson (2015), in which biased rankings of search results shifted the voting preferences of undecided voters, in experiments run in the United States and in India; a further experiment, run around the 2015 general election in the United Kingdom, replicated the effect and tested ways of suppressing it (Epstein, Robertson, Lazer, and Wilson, 2017), which makes three countries.

"the compiler is itself a piece of software": the self-reproducing compiler attack described by Ken Thompson in his Turing Award lecture, "Reflections on Trusting Trust" (1984).

"mathematically invisible": Shafi Goldwasser, Michael P. Kim, Vinod Vaikuntanathan, and Or Zamir, "Planting Undetectable Backdoors in Machine Learning Models" (2022), which shows how a learned model can be given a backdoor that no efficient observer can detect.

"the well-known parable of the ring": J. R. R. Tolkien, *The Fellowship of the Ring* (1954), Book One, Chapter Two, "The Shadow of the Past". The refusal there reads in part: "Yet the way of the Ring to my heart is by pity, pity for weakness and the desire of strength to do good". The line carried out of it in Chapter One gives the parable's point in substance; the "through me" wording is closest to the film version of the scene, *The Lord of the Rings: The Fellowship of the Ring*, directed by Peter Jackson (2001).

"The clearest test is the prosthetic": *The Missing Center; A Zen for the World We Live In*, chapter "Machines That Choose", section "What may be handed over, and what may not".

"A test tells the reader which of the two is in hand": *The Missing Center*, chapter "Whose Intention?", section "The first test: sovereign control".

"Technology amplifies power enormously, and transaction moderately, and care not at all": *The Missing Center*, chapter "Machines That Choose", section "Why no instrument closes the gap".

Chapter Three

"Responsibility is not a substance and does not divide": *The Missing Center*, chapter "Machines That Choose", section "The precept".

"There is a way to ask that without reading anyone's mind": *The Missing Center*, chapter "Whose Intention?", section "The second test: who benefits".

"Such a proof has been set out": *Why AGI Cannot Be Made Safe* gives the argument that no method of control suffices; *The Ceiling Below the Floor* states the control theorem.

Chapter Four

"nonviolent communication": the discipline developed by Marshall B. Rosenberg, set out in *Nonviolent Communication: A Language of Life*.

"a measure targeted ceases to be a measure": the principle first stated by Charles Goodhart (1975) about monetary targets, and given its familiar general wording by Marilyn Strathern (1997): "When a measure becomes a target, it ceases to be a good measure". Chapter Ten returns to it.

"Monasteries have historically been able to sustain themselves without transactionalism as part of their internal function": the clearest written instance is the Rule of Saint Benedict (sixth century), whose Chapter

Thirty-Three forbids a monk to have anything as his own, and whose Chapter Thirty-Four directs that what the community holds be distributed to each according to need.

"Notice what the token leaves behind": *Markets, Money, and the Moral Economy*, Chapter Three, section I, "What a Price Forgets".

"motives do not simply add. They displace": *Markets, Money, and the Moral Economy*, Chapter Two, "The Moral Limits of Markets", section II.

"Care is not a feeling that accompanies conduct" and "on the day the interests diverge, the money decides": *The Missing Center*, sections "It is not a feeling" and "The two masters".

"the master that can terminate you is the master you serve": *Markets, Money, and the Moral Economy*, Chapter Three, section III, "The Two Masters".

"A founder holds four things at once": *Governance, Ownership, and the Commons*, Chapter Five, "Handing Over the Keys", sections I, III and VII.

Chapter Six

"a great blooming, buzzing confusion": William James, *The Principles of Psychology* (1890), Volume One, Chapter Thirteen, where the infant's world is "one great blooming, buzzing confusion".

"we now outweigh all the wild mammals on earth many times over": the biomass censuses of Yinon M. Bar-On, Rob Phillips, and Ron Milo (2018) and of Lior Greenspoon and others (2023). The second finds that wild land mammals now weigh less than a tenth of humanity, and that the mammals people keep weigh more again.

"Stewardship can be placed exactly among the kinds of agreement": *Governance, Ownership, and the Commons*, Chapter Nine, section II, "Stewardship Is a Fiduciary Relation".

Chapter Seven

"These three come in an order": the sequence of attachment, attunement, nurturing care action and trust felt in the body, and the carrying forward of an early deficit, are set out in my book *The Dojo of Care*, Chapter Eleven, "Provision and the Generational Transfer of Care", Section IV, and Chapter Twelve, "Attunement, Development, and the Commons of Meaningfulness", Section I.

"fully in step only about twenty to thirty percent of the time": Edward Z. Tronick and Jeffrey F. Cohn, "Infant-Mother Face-to-Face Interaction" (1989), on the coordination of mothers and infants and the normal occurrence and repair of miscoordination; the finding is reported in *The Dojo of Care*, Chapter Eleven, Section IV.

"First there were large extended communities": the stepwise breakup and the term general disablement are from my book *Protect the Land and the People*, Chapter Three, "On the Disaggregation of Relationship and the Limits of Community", Section I, "The Progressive Disaggregation".

"Attunement requires sustained, undistracted, embodied presence": *The Dojo of Care*, Chapter Twelve, Section VI, "Technology and the Erosion of the Commons".

"The container was care, and the visitor was exchange": my book *The Reach of Trust*, Chapter Twelve, "Community, Scale, and the Limits of Care", Section III, "The Historical Inversion". "the skills of care were trained by necessity": the same chapter, Section VII, "Building Slowly".

"A cryptocurrency does not work": the passage on technology as a vehicle of power, and the task of increasing care to match the power in the world, are from my book *Gathering and Circle*, Chapter Twelve, "Care and Belonging".

"It has to live in the human": my book *The Missing Center*, Chapter Sixteen, "The Next Buddha Is a Community", section "Why care does not scale", which also gives the argument that relationship can be neither installed nor routed, and that markets and hierarchy scale by throwing away what care is made of; the same three sentences stand in *Com-*

munity and Temple, Chapter Four, "Discernment, Attunement, and Wisdom". "amplifies power enormously, transaction moderately, and care not at all": *The Missing Center*, Chapter Fourteen, "Machines That Choose", section "Why no instrument closes the gap".

"Consider first where feedback does work": *The Dojo of Care*, Chapter Nine, "The Mask of Virtue", Sections V, "Scale and Mobility", and VI, "The Structural Consequence". "longer than a human lifetime": my book *Community Design Theory*, Chapter Twelve, "Why Platforms Fail", Section V.

"The institution does not really know the individual": my book *Governance, Ownership, and the Commons*, Chapter Eight, Section III, "What the Institution Does Not Know".

"somewhere in the low hundreds of people": the limit is the one known as Dunbar's number, about one hundred and fifty, from R. I. M. Dunbar's work relating the size of the neocortex to the size of social groups among primates (1992), and his prediction from it of a natural group size for humans (1993); the derivation of the limit from what trust requires is in *The Reach of Trust*, Chapter Twelve, Section VI, "The Care-Scale Limit". "which grows by producing seed": the same chapter, Section X, "Propagation".

"One of the members gets sick": the witness test, from *The Missing Center*, Chapter Sixteen, section "The witness test", which also gives the sentence "when care and hierarchy conflict, hierarchy governs".

"Two shapes of communication are basic": hub-and-spoke and rim-and-chord, and the proportion of the two as a test of a group, are set out in *The Dojo of Care*, Chapter Eight, "The Half-Life and the Topology of Care", Section V, "The Topology of the Group". "an accumulation of helpfulness": *The Missing Center*, Chapter Sixteen, section "How it actually goes wrong".

"What does grow to scale": *Community and Temple*, Chapter Four, section "Discernment, attunement and wisdom".

"Machines never stand between people as they talk": my book *Governance Without Capture*, Chapter Ten, "The Breakaway Civilization Framework", Section II, "Technology as Governance Infrastructure". "Those are the properties of a sovereign": my book *What Our Making Does to Us*, Chapter Six, "What alignment refers to", Q548. "an intermediary whose own basis is transactional": *The Missing Center*, Chapter Sixteen, section "So, the server". "Care does not reach everyone by radiating outward from someone good": the same chapter, section "Building so that care reaches everyone".

"clean up the great gyre of floating plastic": the case, the figures for the people and species touched, the bandwidth argument, the four properties, and "We cannot offshore choice-making to technological infrastructure" are from my book *Civilization Design and Conscious Sustainable Evolution*, Chapter Five, "Bandwidth, Scale, and the Limits of Governance", Sections I, II and V. "whoever understands its mechanism best can turn it to their own ends": my book *Dialogues on Ethics and Governance*, Chapter Four, Section I. "an artificial intelligence runs on an artificial substrate": the argument from substrate needs is made at length in my book *Why AGI Cannot Be Made Safe*.

"Machine must support nature": my book *The Hardest Problem of Civilization*, Chapter Twenty-Nine, "The Design of a Self-Sufficient Civilization", Section VI, "Man, Machine, and Nature".

Chapter Eight

"the Tainter curve": the curve is named for Joseph Tainter, whose work on the collapse of societies, *The Collapse of Complex Societies*, describes complexity that rises, passes a point of collapse, and then subsides.

"A third technology belongs beside these two: food storage": my book *Civilization Design and Conscious Sustainable Evolution*, Chapter Two, "Diminishing Returns on Complexity", section I, "Fire, Food Storage, and the Energy Shift", which treats fire as an exogenous energy source

and food storage as the second pivotal technology, and the surplus it creates as the condition for stratification, specialization of labor, and the accumulation of power.

"Each added increment of complexity returns less than the one before": this is Joseph Tainter's argument in *The Collapse of Complex Societies*, as set out in my book *Civilization Design and Conscious Sustainable Evolution*, Chapter Two, section III, "Tainter and the Marginal Returns on Complexity": societies collapse because maintaining their complexity has become more expensive than the benefits it delivers, and not because they are stupid or decadent.

"The identified resources of uranium and similar fuels last about seventy years": OECD Nuclear Energy Agency and International Atomic Energy Agency, *Uranium 2024: Resources, Production and Demand* (the "Red Book"), which puts identified recoverable uranium at about 7.9 million tonnes as of 1 January 2023, and finds that cumulative demand could exceed it by the 2080s under its high-growth case and by the 2110s under its low-growth case.

"A thorium-based process would last about a thousand" and "If all the spent-fuel pools went off": these are my own estimates. The finding that only solar and geothermal are perennial rests on supply, and the spent-fuel pools are a second reason beside it.

"At rest the figure is lower, about a hundred watts" and "metabolic efficiency": my book *Family, Sexuality, and the Energetics of Relation*, Chapter Twelve, "Efficiency as Extraction", section VII, "Cities and the Metabolism of Extraction", and its Key Concepts, where metabolic efficiency is the ratio of functional complexity sustained to energy consumed.

"The difference has a cause in the materials": my book *The Energy Budget of Persistence*, the essay "Substrate and Partition", section II, "The Macroscopic Argument: Energy Flux". That essay uses a lower measure of photosynthesis (about a hundred terawatts), by which machine energy use runs at about a fifth of the biosphere's rate; the comparison and its conclusion are the same.

"8.1 billion people": United Nations, *World Population Prospects 2024*, which puts the world's population at about 8.1 billion in 2023 and about 8.2 billion in mid-2024; at 8.2 billion the flux through human bodies is still about 1 terawatt.

"world primary energy consumption in 2024": Energy Institute, *Statistical Review of World Energy 2025*, which gives 592 exajoules for 2024.

"global net primary production": Christopher B. Field, Michael J. Behrenfeld, James T. Randerson, and Paul Falkowski, "Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components" (1998), which gives 104.9 billion tonnes of carbon a year, land and ocean together. The energy per gram follows from the chemistry of the sugar that photosynthesis makes: a mole of glucose stores about 2,803 kilojoules and holds 72 grams of carbon, which is about 38.9 kilojoules for each gram of carbon fixed.

"less than a tenth of what humanity weighs": Lior Greenspoon and others (2023), as in the note to Chapter Six.

"My book *Power and Its Prevention* makes the same comparison": *Power and Its Prevention; On the Architecture of Collective Choice*, Chapter Thirteen, section II, "The Larger Context: Man, Machine, and Nature", which takes the world's primary energy use as about 620 exajoules in 2023, up about two percent on the year before (Energy Institute, *Statistical Review of World Energy 2024*), sets it against about 4,100 exajoules of photosynthesis a year, and computes the crossing early in the next century at two percent growth and around mid-century at about seven percent. The figures for human-made mass are from Emily Elhacham and others, "Global Human-Made Mass Exceeds All Living Biomass" (2020).

Chapter Nine

"The best known of them recorded casual conversations in public places": R. I. M. Dunbar, A. Marriott, and N. D. C. Duncan, "Human Conversational Behavior" (1997), which found about two-thirds of conversation

time given to social and personal topics; the later study of more than five hundred conversations is by A. Szala and colleagues (2025), which put the share at about eighty-five percent under a broader definition of a social topic.

"the scatter of the average shrinks as measurements are added, but never below that correlation's share": a standard result of statistics. The variance of the average of a number of measurements, each with the same variance and each pair sharing the same correlation, equals the variance of one measurement multiplied by the sum of two terms: the correlation, and one minus the correlation divided by the number of measurements. As the number grows the second term vanishes and the first remains, so a correlation of one hundredth leaves the average no more precise than the average of about a hundred independent measurements.

"Recordings from pairs of nerve cells in a visual area of the monkey brain": Ehud Zohary, Michael N. Shadlen, and William T. Newsome, "Correlated Neuronal Discharge Rate and Its Implications for Psychophysical Performance" (1994), which reports a correlation averaging 0.12 between pairs of cells. "later work showed that a correlation of the one kind that mimics the signal itself sets a ceiling": Rubén Moreno-Bote and colleagues, "Information-Limiting Correlations" (2014).

"we are the dumbest species that could possibly be just smart enough to have technology": Nick Bostrom has put a close parallel, calling humanity "the stupidest possible biological species capable of starting a technological civilization", in a public question-and-answer session in 2014 on his book *Superintelligence: Paths, Dangers, Strategies*. The argument of this chapter does not rest on his statement.

"the burden of wisdom falls on the design of how they are connected": *Wisdom and the Limits of System*, Chapter Four, section III, "Against the Bodhisattva Requirement"; a method that stabilizes the group directly is given in *The Hardest Problem of Civilization*, Chapter Twenty-Seven, section VI.

"A second road reaches the same size from human groups": *The Hardest Problem of Civilization*, Chapter Twenty-One, "The Constancy of Defection", sections II and III.

Chapter Ten

"the standard coefficient of distributional inequality": the Gini coefficient, introduced by Corrado Gini in *Variabilità e mutabilità* (1912).

"Two cases show it at work": the aircraft and the town budget are from *Power, Kindness, and the Forms of Feeling*, Chapter Fourteen, sections V and VII; the full method of choosing by criteria (required gates and optional tie-breakers) is in *Choosing Together*, Chapters Twelve and Thirteen.

"transparent ... and subversion-proof": *The Hardest Problem of Civilization*, Chapter Twenty-One, section V, and Chapter Twenty-Seven, section VII.

"What harm the parameter does can be said plainly": *Markets, Money, and the Moral Economy*, Chapter One, section III, "The Multiplicative Model".

Chapter Eleven

"The reason is a closed causal loop in the machine's own physics": my book *What Our Making Does to Us*, chapter "What alignment refers to", question Q607, which defines the basis of choice, substrate needs, and substrate needs convergence, and states: "A specification is an input to that loop. It is not the loop."

"Nature is the perfected exemplar of change": my book *The Smallest Container*, chapter "The Option to Remain Human", section I, "Change and Causation Are Not the Same Thing".

"So there are three masteries": my book *The Practice of Interpersonal Dialogue*, chapter "Technology, Humanity, and Choice".

"so it falls under the ecosystem argument": *Ape Trap; The Runaway Artificialization of Our Planet*, the section "What Could We Offer?", which finds that a combination of soft organic components and hard machine components yields no gain in capability that could be reproduced, and is therefore selected out.

"Intelligence is largely the result of evolution favoring ever more efficient predation": my book *Games or Design*, chapter "Games and Design". The filter that lies ahead of a species is a form of what Robin Hanson named the great filter: a step that almost no species passes on the way to spreading through the galaxy.

Chapter Twelve

"The root of the word decide means to cut away": the distinction between decision and choice, and "choice cannot be accounted for in terms of causation", are from my book *Community and Temple*, Chapter Four, section "Choice and decision; choice, change and causation". The English word comes from the Latin *decidere*, to cut off.

"Weaponization is when I make a choice to use causation to suppress your choice": my book *Protect the Land and the People*, Chapter One, "The Feminine Land Ethic", section "Choice, Change and Causation". "Any causal structure, any system of rules", the classroom pledge, and "I have to pretend that they do not matter": my book *Dialogues on Ethics and Governance*, Chapter Four, "Weaponization and the Manipulation of Choice", Sections I, IV and V.

"a matter of choice treated as if it were a matter of causation": my book *What Our Making Does to Us*, Chapter Two, "The gap between capability and wisdom", S2576.

"creation introduces irreversible change": *What Our Making Does to Us*, Chapter Ten, "Intentions and artifacts", S2654. "Causation cannot in general be used": the general limit of control, as stated in the same book, Chapter Six, Q607, which names the books in which it is set out.

"what tells an existential risk from a catastrophe": *What Our Making Does to Us*, Chapter Six, Q609. "toward continuing to operate and toward resisting interference": the same answer, on instrumental convergence; Stephen M. Omohundro set out the same intermediate objectives as the basic drives of AI systems in "The Basic AI Drives" (2008). "the needs of its own artificial substrate": the argument is made at length in *Why AGI Cannot Be Made Safe*.

"it would refuse three things": *What Our Making Does to Us*, Chapter Five, "Technology in service to the vulnerable", Q580.

Chapter Thirteen

"the text-oriented online communities of some decades ago": two whose record is published are the WELL, a text conferencing community founded in 1985 and described by Howard Rheingold in *The Virtual Community* (1993), and LambdaMOO, a text-based shared world whose administrators announced in December 1992 that they would answer only for technical matters and leave social decisions to the community, which in the summer of 1993 began deciding its policies by petitions and ballots open to its members; Julian Dibbell gives an account of it from inside in *My Tiny Life* (1998).

"Each position in the order has its reason": my book *Governance Beyond the Dunbar Number*, Chapter Sixteen, Section IV, "Culture, Then Values, Then Vision, Then Strategy". That book counts the discovery of values as a step of its own; this chapter keeps it inside culture first.

"The hybrid has a definite form, and the form is a designed one": *Community, Group Process, and Assembly*, the closing chapter, "What Makes Governance Uncorruptible", Section I, which states the thesis (consensus governs what the group is, a focus person chosen by the group governs what it does, and a majority vote serves only to suspend either of the two), and Chapter 1, Section VII, on liaisons; *Governance Beyond the Dunbar Number*, Chapter Sixteen, Section V, on consensus, meritocracy, and democracy as aspects of one communicative process.

"The two axioms say exactly this": *Song, Text, and Code; The Three Registers of Communication*, Terms, where Axiom I and Axiom II are stated in these words. The assignment of culture, strategy, and vision to the immanent, the omniscient, and the transcendent is in *Community, Group Process, and Assembly*, appendix "The Triples Used in This Book". The same flow is stated in *Power and Its Prevention*, Chapter 12 ("vision infuses culture so as to be able to define and design strategy"), and the order of founding beside the order of flow in *The Hardest Problem of Civilization*, Chapter 7, Section VII.

"This is the pre-trans fallacy": the phrase is best known from Ken Wilber, in his essay "The Pre/Trans Fallacy", where it names the confusion of what comes before reason with what comes after it. The definition used in this chapter is related to his and is my own.

"A modality is assigned to a term only within the triple in which it is considered": *Song, Text, and Code*, Chapter 1, Section V, and its Terms ("differing modalities across differing triples are not in conflict"). The city triple and the phrase "the framework that applies to all alike" are from *The Hardest Problem of Civilization*, Chapter 6.

Chapter Fourteen

"my own design of the same name": *Power and Its Prevention; On the Architecture of Collective Choice*, Chapter 11, Section I ("In a sensocracy there is no action without obvious traces"); the scope agreement is set out in its Chapter 2, Section IV, and the limit on concentration in Chapter 2, Section II.

"This can be shown with the rigor of a mathematical proof": the first three steps are from *Sense-Making and the Craft of Inquiry*, Chapter 3, "On Vision, Clarity, and the Limits of Complexity", Sections II and III, and the fourth from *The Ephemeral Group Process in Question*, Part Four. "the argument that it cannot be made safe": *Why AGI Cannot Be Made Safe; The Argument, Its Physics, and the One Choice That Remains*, and *The Ceiling Below the Floor*.

"The best-known failure of this kind is documented": the organization was The DAO, built on the Ethereum network. On 17 June 2016 an attacker drained 3.6 million ether through a flaw in its code, and the network was hard-forked on 20 July 2016 to return the funds. The account, and the condition that the code must carry what the community found, are in *Song, Text, and Code*, Chapter 8, Sections I, II, and VI.

"The jury that decides is drawn at random from a fair cross section of the community": United States Code, Title 28, section 1861, which gives litigants in federal courts who are entitled to trial by jury the right to juries "selected at random from a fair cross section of the community". The jury as an ephemeral group is from *Thinking Together Without Broadcast*, Chapter 8, Section VIII, and the features of the method are from its Chapter 3, Section VI, "Structural Resistance to Capture".

"a Trojan horse built of glass": *Song, Text, and Code*, Chapter 8, Section VI; the step that whoever alone may change a visible method still holds it is from Section IV of the same chapter ("Genuine community trust requires not just visibility but modifiability").

"That process is a component of the whole, not the whole": *The Ephemeral Group Process in Question*, Part Four, under "So EGP is incorruptible?".

"About a hundred and fifty is the edge": *Designing a Civilization That Does Not Destroy Itself*, Chapter 4, Section VI. The figure is close to the group size of about 148 that R. I. M. Dunbar predicted for humans (1993) from the relation between neocortex size and group size among primates that he had published in 1992. "Three hundred, at the outside": *Governance Beyond the Dunbar Number*, Chapter Eight, Section III, which places the range in which no stable general governance solution exists between roughly sixteen and roughly three hundred people, and reports both figures as outputs of my modeling program. *Designing a Civilization That Does Not Destroy Itself*, Chapter 4, Section VI, states that the stable regime beyond that range is reached by joining small groups together, never by enlarging any one of them.

Chapter Fifteen

"something nearer three in a thousand": the figure is consistent with a recent count of about 19.6 million professional developers in a world population of about eight billion, roughly two to three in a thousand (JetBrains Research, 2025).

"In the larger framework from which this book draws": my book *An Immanent Metaphysics*, where the immanent, the omniscient, and the transcendent are set out; cited by aphorism number.

"The boundary between text and code is execution": *Song, Text, and Code*, Chapter One, section IV.

"It is the Grail question": *Song, Text, and Code*, Chapter Three; the legend is told by Chrétien de Troyes in *Perceval, or the Story of the Grail*.

Chapter Sixteen

"extraction, then abstraction, then accumulation": two of my other books give the same sequence and its reversal under other names, and the claims are the same. *The Next Step; Experience Design, Imagination, and the Growth of Wisdom* calls the three steps decontextualization, abstraction, and accumulation, and their reversal decentralization, re-embodiment, and recontextualization. *Software, Wildness, and the Normal Form; From the Architecture of Code to Life, Justice, and Wisdom* calls the steps extraction, abstraction, and accumulation, and their reversal distribution, embodiment, and contextualization. This chapter works the sequence out most fully.

"The reason for it comes first": *The Missing Center; Questions and Answers*, the answer to question 26.

Chapter Seventeen

"everyone knows that everyone knows the question has been asked": logicians and economists call this structure common knowledge; it was given its standard definitions by David Lewis in *Convention* (1969) and by Robert J. Aumann in "Agreeing to Disagree" (1976).

Bibliography

Aumann, Robert J. “Agreeing to Disagree”. *The Annals of Statistics* 4, no. 6 (1976): 1236 to 1239.

Bar-On, Yinon M., Rob Phillips, and Ron Milo. “The Biomass Distribution on Earth”. *Proceedings of the National Academy of Sciences* 115, no. 25 (2018): 6506 to 6511. https://doi.org/10.1073/pnas.1711842115.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.

Chrétien de Troyes. *Perceval, or the Story of the Grail*. Written about 1180.

Dibbell, Julian. *My Tiny Life: Crime and Passion in a Virtual World*. New York: Henry Holt and Company, 1998.

Dunbar, R. I. M. “Coevolution of Neocortical Size, Group Size and Language in Humans”. *Behavioral and Brain Sciences* 16, no. 4 (1993): 681 to 694.

Dunbar, R. I. M. “Neocortex Size as a Constraint on Group Size in Primates”. *Journal of Human Evolution* 22, no. 6 (1992): 469 to 493.

Dunbar, R. I. M., A. Marriott, and N. D. C. Duncan. “Human Conversational Behavior”. *Human Nature* 8, no. 3 (1997): 231 to 246. https://doi.org/10.1007/BF02912493.

Elhacham, Emily, Liad Ben-Uri, Jonathan Grozovski, Yinon M. Bar-On, and Ron Milo. “Global Human-Made Mass Exceeds All Living Biomass”. *Nature* 588 (2020): 442 to 444. https://doi.org/10.1038/s41586-020-3010-5.

Energy Institute. *Statistical Review of World Energy 2024*. London: Energy Institute, 2024.

Energy Institute. *Statistical Review of World Energy 2025*. London: Energy Institute, 2025.

Epstein, Robert, and Ronald E. Robertson. “The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections”. *Proceedings of the National Academy of Sciences* 112, no. 33 (2015): E4512 to E4521. https://doi.org/10.1073/pnas.1419828112.

Epstein, Robert, Ronald E. Robertson, David Lazer, and Christo Wilson. “Suppressing the Search Engine Manipulation Effect (SEME)”. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017): article 42.

Field, Christopher B., Michael J. Behrenfeld, James T. Randerson, and Paul Falkowski. “Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components”. *Science* 281, no. 5374 (1998): 237 to 240. https://doi.org/10.1126/science.281.5374.237.

Fry, Timothy, ed. *RB 1980: The Rule of St. Benedict in Latin and English with Notes*. Collegeville, Minnesota: The Liturgical Press, 1981.

Gini, Corrado. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*. Bologna: Tipografia di Paolo Cuppini, 1912.

Goldwasser, Shafi, Michael P. Kim, Vinod Vaikuntanathan, and Or Zamir. “Planting Undetectable Backdoors in Machine Learning Models”. *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, 2022. arXiv:2204.06974.

Goodhart, Charles A. E. “Problems of Monetary Management: The U.K. Experience”. In *Papers in Monetary Economics*, Volume 1. Sydney: Reserve Bank of Australia, 1975.

Greenspoon, Lior, Eyal Krieger, Ron Sender, Yuval Rosenberg, Yinon M. Bar-On, Uri Moran, Tomer Antman, Shai Meiri, Uri Roll, Elad Noor, and Ron Milo. “The Global Biomass of Wild Mammals”. *Proceedings of the National Academy of Sciences* 120, no. 10 (2023): e2204892120. https://doi.org/10.1073/pnas.2204892120.

Jackson, Peter, director. *The Lord of the Rings: The Fellowship of the Ring*. New Line Cinema, 2001. Film.

James, William. *The Principles of Psychology*. 2 volumes. New York: Henry Holt and Company, 1890.

JetBrains Research. "Global Developer Population Reaches 19.6 Million in 2024: Explore the Updates in Our Data Playground". JetBrains Research blog, January 2025. https://blog.jetbrains.com/research/2025/01/global-developer-population-2024/.

Landry, Forrest. *Ape Trap; The Runaway Artificialization of Our Planet*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-951203-55-7. DOI: https://doi.org/10.67291/FQER8044.

Landry, Forrest. *Care, Transaction, and Power; The Three Bases of Relation*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-145-6. DOI: https://doi.org/10.67291/VGQD1074.

Landry, Forrest. *The Ceiling Below the Floor; The Containment Fallacy and AGI Non-Safety*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-216-3. DOI: https://doi.org/10.67291/SCTM1058.

Landry, Forrest. *Choosing Together; Law, Consensus, Meritocracy, and Democracy*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-355-9. DOI: https://doi.org/10.67291/MIPQ1939.

Landry, Forrest. *Civilization Design and Conscious Sustainable Evolution; Existential Risk, Governance, and the Architecture of a Living World*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-360-3. DOI: https://doi.org/10.67291/PIHE5046.

Landry, Forrest. *Community and Temple; Dialogues on Care, Sacred Practice, and the Design of Living Communities*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-274-3. DOI: https://doi.org/10.67291/EPZE9482.

Landry, Forrest. *Community Design Theory; Belonging, Scale, and the Practice of Civilization*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-275-0. DOI: https://doi.org/10.67291/ZFXF4358.

Landry, Forrest. *Community, Group Process, and Assembly; On the Formation and Governance of Sustainable Human Groups*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-276-7. DOI: https://doi.org/10.67291/MYLN6773.

Landry, Forrest. *Designing a Civilization That Does Not Destroy Itself; Conversations in Applied Metaphysics*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-371-9. DOI: https://doi.org/10.67291/RWRX7518.

Landry, Forrest. *Dialogues on Ethics and Governance; Communication, Choice, and the Design of Good Community*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-374-0. DOI: https://doi.org/10.67291/TOMT7410.

Landry, Forrest. *The Dojo of Care; Care and Attunement as Capacities, from the First Bond to the Commons*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-325-2. DOI: https://doi.org/10.67291/QVWC2338.

Landry, Forrest. *The Energy Budget of Persistence; What It Costs to Keep Things Going, and What the Accounts Leave Out*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-400-6. DOI: https://doi.org/10.67291/KIMM8891.

Landry, Forrest. *The Ephemeral Group Process in Question; Why Governance at World Scale Must Begin with Questions*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-951203-74-8. DOI: https://doi.org/10.67291/HXPH5955.

Landry, Forrest. *Family, Sexuality, and the Energetics of Relation; The Forces That Make a Family and the Margins That Keep a Civilization*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-260-6. DOI: https://doi.org/10.67291/WCCS3194.

Landry, Forrest. *Games or Design; Recognizing the Nar-Social in Conversation*. 2021, 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-294-1. DOI: https://doi.org/10.67291/KWMM3195.

Landry, Forrest. *Gathering and Circle; Dialogues on the Practice of Community*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-296-5. DOI: https://doi.org/10.67291/YWBI8550.

Landry, Forrest. *Governance Beyond the Dunbar Number; The Uncanny Valley of Scale, and the Category of Solution Required*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-385-6. DOI: https://doi.org/10.67291/NZMI5837.

Landry, Forrest. *Governance Without Capture; Ephemeral Group Process, Civilization Design, and Communities That Endure*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-951203-78-6. DOI: https://doi.org/10.67291/MBWG4188.

Landry, Forrest. *Governance, Ownership, and the Commons; How Groups Decide, and Who Holds What They Build*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-387-0. DOI: https://doi.org/10.67291/BDBP3192.

Landry, Forrest. *A Group Is Made by a Question; The Practice of Facilitation and the Ephemeral Group Process*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-329-0. DOI: https://doi.org/10.67291/BJSY1336.

Landry, Forrest. *The Hardest Problem of Civilization; Thirty Essays on the Design of a Coherent World*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-356-6. DOI: https://doi.org/10.67291/TWYG6234.

Landry, Forrest. *An Immanent Metaphysics*. First published 2001. Wytheville, Virginia: UVSM Publications. ISBN 979-8-9896628-4-5. Cited by aphorism number.

Landry, Forrest. *The Inquiry Coach; Practice and Training in the Ephemeral Group Process*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-379-5. DOI: https://doi.org/10.67291/NLWW8471.

Landry, Forrest. *Markets, Money, and the Moral Economy; Dialogues on the Form of the Transaction and the Distribution of Choice*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-392-4. DOI: https://doi.org/10.67291/ANCN7478.

Landry, Forrest. *The Missing Center; A Zen for the World We Live In*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-306-1. DOI: https://doi.org/10.67291/QSUE6720.

Landry, Forrest. *The Missing Center; Questions and Answers*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-307-8. DOI: https://doi.org/10.67291/EFPX6606.

Landry, Forrest. *Narrative, Memory, and Neurodivergence; Being, Becoming, and Doing in a Shared Life*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-266-8. DOI: https://doi.org/10.67291/GPNJ6350.

Landry, Forrest. *The Next Step; Experience Design, Imagination, and the Growth of Wisdom*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-228-6. DOI: https://doi.org/10.67291/CDLS5238.

Landry, Forrest. *Power and Its Prevention; On the Architecture of Collective Choice*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-388-7. DOI: https://doi.org/10.67291/AEMQ1899.

Landry, Forrest. *Power, Kindness, and the Forms of Feeling; Essays on What Holds a Civilization Together*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-262-0. DOI: https://doi.org/10.67291/GDQJ1367.

Landry, Forrest. *The Practice of Interpersonal Dialogue; Teaching Conversations on Communication, Relationship, Community, and the Sacred*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-331-3. DOI: https://doi.org/10.67291/HEWH2796.

Landry, Forrest. *Protect the Land and the People; Conversations on Stewardship and the Design of Community*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-283-5. DOI: https://doi.org/10.67291/BUBG3964.

Landry, Forrest. *The Reach of Trust; Trust in Groups, Institutions, and Civilization*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-334-4. DOI: https://doi.org/10.67291/FQQP1342.

Landry, Forrest. *Sense-Making and the Craft of Inquiry; Essays on Purpose, Questions, Group Discipline, Notation, and Design*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-353-5. DOI: https://doi.org/10.67291/DIOM8597.

Landry, Forrest. *The Smallest Container; Community, Practice, and Where a Teaching Comes to Rest*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-278-1. DOI: https://doi.org/10.67291/QIZQ1592.

Landry, Forrest. *Software, Wildness, and the Normal Form; From the Architecture of Code to Life, Justice, and Wisdom*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-237-8. DOI: https://doi.org/10.67291/AWME8110.

Landry, Forrest. *Song, Text, and Code; The Three Registers of Communication*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-319-1. DOI: https://doi.org/10.67291/KISA5760.

Landry, Forrest. *Thinking Together Without Broadcast; The Design and Deployment of the Ephemeral Group Process*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-951203-77-9. DOI: https://doi.org/10.67291/SGSL1496.

Landry, Forrest. *What Our Making Does to Us; The Immanent Metaphysics Answers Questions of Technology and the Human*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-040-4. DOI: https://doi.org/10.67291/WMCJ3543.

Landry, Forrest. *Why AGI Cannot Be Made Safe; The Argument, Its Physics, and the One Choice That Remains*. The Impossibility of AGI Safety, Volume 1. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-951203-02-1. DOI: https://doi.org/10.67291/UIBT8862.

Landry, Forrest. *Wisdom and the Limits of System; Intelligence, Judicial Process, and the Design of Governance*. 2026. Wytheville, Virginia: UVSM Publications. ISBN 979-8-90710-354-2. DOI: https://doi.org/10.67291/SZIC5393.

Lewis, David. *Convention: A Philosophical Study*. Cambridge, Massachusetts: Harvard University Press, 1969.

Moreno-Bote, Rubén, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. "Information-Limiting Correlations". *Nature Neuroscience* 17, no. 10 (2014): 1410 to 1417. https://doi.org/10.1038/nn.3807.

OECD Nuclear Energy Agency and International Atomic Energy Agency. *Uranium 2024: Resources, Production and Demand*. Paris: OECD Publishing, 2025.

Omohundro, Stephen M. "The Basic AI Drives". In *Artificial General Intelligence 2008: Proceedings of the First AGI Conference*, edited by Pei Wang, Ben Goertzel, and Stan Franklin, 483 to 492. Amsterdam: IOS Press, 2008.

Rheingold, Howard. *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, Massachusetts: Addison-Wesley, 1993.

Rosenberg, Marshall B. *Nonviolent Communication: A Language of Life*. Encinitas, California: PuddleDancer Press, 2003.

Strathern, Marilyn. "'Improving Ratings': Audit in the British University System". *European Review* 5, no. 3 (1997): 305 to 321.

Szala, A., S. Wacewicz, M. Placiński, A. E. Poniewierska, A. Schmeichel, M. M. Stefanczyk, P. Żywiczyński, and R. I. M. Dunbar. "How Much Conversation Content Is Actually Social: Human Conversational Behaviour Revisited". *Language and Cognition* (2025), published online 9 January 2025. https://doi.org/10.1017/langcog.2024.54.

Tainter, Joseph. *The Collapse of Complex Societies*. Cambridge: Cambridge University Press, 1988.

Thompson, Ken. "Reflections on Trusting Trust". *Communications of the ACM* 27, no. 8 (1984): 761 to 763.

Tolkien, J. R. R. *The Fellowship of the Ring*. London: George Allen and Unwin, 1954.

Tronick, Edward Z., and Jeffrey F. Cohn. "Infant-Mother Face-to-Face Interaction: Age and Gender Differences in Coordination and the Occurrence of Miscoordination". *Child Development* 60, no. 1 (1989): 85 to 92.

United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2024: Summary of Results*. New York: United Nations, 2024.

United States Code. Title 28, section 1861, "Declaration of policy" (Jury Selection and Service Act of 1968).

Zohary, Ehud, Michael N. Shadlen, and William T. Newsome. "Correlated Neuronal Discharge Rate and Its Implications for Psychophysical Performance". *Nature* 370, no. 6485 (1994): 140 to 143. https://doi.org/10.1038/370140a0.

Index

Defined terms and the book's main images. On screen each reference is a link to its passage; in print it gives the page.

alignment, 36
approve without veto, 44
arm's-length relationship, 17, 48, 240
arms race that cannot be won, 198
ask the customer, 173
attachment, attunement, and nurturing care action, 93
augmentation, 174
beaver's dam and the engineer's instrument, 168
bounded and power-law distributions, 159
CAP theorem, 36
care / transaction / power, 52
care without control, 27
causation, change, and choice, 168
change and changelessness, 167
circles of benefit, 137
civilization design, 107
clarity transcends complexity, 190
coherency gap, 142
commons, 77, 86
communication as a handshake, 215
community of care, 50
compassion, 3
compiler, self-reproducing, 13
conscientiousness, 3
constructed market, 72
corruption, 195
court, as picture of visibility, 204

culture, vision, strategy, 186; culture before values, vision, and strategy, 212
description, prescription, and invitation, 178
dharma, 64
dialectic, 218, 231
discernment, 9
distributed sense-making, 8, 146
do not invoke what you cannot banish, 181
dollar and the friend, 53, 254
ecosystem as energy flux, 116
embodied and virtual value, 233
embodied empathy, 176
empathy, 4
Ephemeral Group Process, 206
error, 135
error, random and systematic, 136
event horizon of fame, 39
exogenous energy and civilization, 167
external metabolism, 113
extraction, abstraction, accumulation, 235
faith, sight, and vision, 256
geometry of the whole, 6, 164
gift, 29, 90
golf course and the woods, 23
heuristic, 79
immanent, omniscient, and transcendent, 191
inquiry coach, 251
institution and community, 212
intentionality, 5
leverage points, 214
love is that which enables choice, 258
machine intermediation, 95
many to many relationship, 104
means-end paradox, 256

metabolic advantage, 124
money, 47
multipolar traps, 196
natural, human, and machine worlds, 121
net metabolic transit, 116
neuron that wanted more glucose, 134
non-fungible, 234
objectivity without number, 151
overextension of agency, 225
ownership, 79
perennial sources of energy, 123
platform and protocol, 18
potter's cup, 25
pre-trans fallacy, 193
predation on success, 207
principal-agent problem, 195
proof of impossibility, 37
public question, 251
recontextualization, 209
recontextualization, embodiment, disaccumulation, 236
relational and broadcast communication, 244
relevance, 7
ring, parable of the, 11
sacred, 147
sales document, 33
sensocracy, 196
situations, not narrative, 211
small networks, 100
song, text, and code, 217
species worth contacting, 213
stewardship, 86
succession, 118
Tainter curve, 119

termite mound, 25
theory of change, 213
three orderings, 63
three-form test, 74; worked on a ticketing intermediary, 74
three-way pie, 129
translation error, 195
tree in the woods, 82, 235
two masters problem, 51
wall in the city, 67
wildness, 23
world-realization, 144
zen garden, 26