

MECHANICAL ETHICS

When Correction Arrives Too Late

A human-facing Mechanical Ethics reader

Human Reader v0.7.0 - NEXT WORKING CANDIDATE v0.2.4

WORKING / NOT BASELINE / NOT RELEASE / NOT CANON / NOT VALIDATED

Safety note for a reader inside harm

If a process is hurting you now, do not put yourself in greater danger to create evidence for this book. A record is not justice, and a witness pack is not rescue. If challenging, recording, publishing, or naming what is happening would increase the risk to your housing, employment, health, family, legal position, or personal safety, stop and use safer human support where it is available. Emergency services, medical care, legal advice, unions, charities, family, friends, and trusted people come before any attempt to make the harm fit a theory. The purpose of this work is not to make harmed people perform more proof. It is to help us build systems that place less of that burden on them.

Contents

Safety note for a reader inside harm	1
Opening - Let Us Begin with an Entity	5
How to Read This Book	8
A Timing Condition	9
Part I - The Door	11
1. A Door You Cannot Reach	11
2. Who Builds the Maze	12
3. Process Becomes Distance	13
The Door Out	15
Part II - The Clock	16
4. Help Before the Loss Sets	16
5. The Strategic Unknown	17
6. Slow Harm	18
Short Case - The Finance Appeal That Arrived Too Late	19
Part III - The Witness	20
7. When There Is No Enforcer	20
8. Custody	21
9. Residue	22
Interlude - Two Flats, One Wall	23

Part IV - The Machine	25
Before the Machine - The Predator Position	25
10. Who Pays for Complexity	25
11. The Second Person	26
When One Case Is Not Enough	27
12. Correction Theatre	28
Historical Note - Challenger and the Cost of Proceeding	29
The Machine-Speed Brake	29
Part V - After the Fall	31
13. Peace as Evidence	31
14. What Remains Open	31
15. A Human Standard for Answerability	32
Part VI - Futures Still Reachable	34
Interlude - The Saturday Workshop	35
16. The Conditions of Hope	37
17. Care Without Possession	38
18. Kindness as the Placement of Burden	39
Conclusion	41
Author's Note - How This Began	42
Appendix A - Structural Notes	44

Appendix B - What Remains Unfinished **46**

Known unfinished edges 46

Questions left open 47

Appendix C - Status, Sources, and the Two-Artifact Pair **48**

Reader status 48

Figures 48

Scene and evidence discipline 48

Historical sources 48

Human authorship and AI assistance 49

Mechanical Ethics and TRACE 49

Opening - Let Us Begin with an Entity

Let us begin with as little as possible.

Assume, provisionally, that something exists. We do not need to settle every philosophical argument about existence before we can continue. We only need a starting point that allows us to examine what follows.

You are reading this, so for present purposes you are such a thing. At some point today you became aware of your surroundings, noticing some things and missing others. You carried memories from earlier, and expectations about what would come next. You made choices, including choices so ordinary that they did not feel like choices at all.

You did not begin with complete information. Nobody does. Perception is partial, memory is incomplete, and the future is uncertain. Even so, life does not wait for certainty. You still have to decide whether to speak, wait, move, trust, refuse, help, protect, repair, or leave something alone.

Your actions change the situation. Inaction can also change it. Some changes are easy to reverse. Others become expensive, painful, or impossible to undo. The same decision can open a path for one being while closing it for another.

Other entities exist as well. They do not see the world from your position. A child, a dog, a patient under anaesthetic, a frightened witness, a government department, and an artificial system each have different access to what is happening. Their capacities are different, and so is the force of their actions. The difference does not disappear because everyone is uncertain.

This is the ground from which Mechanical Ethics begins: an attempt to navigate choices made by limited entities in a changing world, where other entities can be helped, harmed, ignored, constrained, or enabled. It does not promise a perfect moral compass. What it tries to do is make the relevant structure visible before a decision hardens into history.

The questions are basic:

- What is changing?
- Who or what is affected?
- What can each party see?
- What can each party do?

- Which paths are opening or closing?
- How long does correction remain possible?
- Who carries the cost while uncertainty remains?
- What should be built, preserved, stopped, or repaired?

The book starts with human situations because it is written for human readers. The pattern may travel further. A gate that is usable by an adult may be useless to a small child, a wheelchair user, or a dog. A warning that reaches an engineer may never reach the manager who can stop the launch. An artificial system may make thousands of selections before a person discovers that a path has been closed. The scale changes; the need to understand perception, action, timing, power, and correction remains.

Mechanical Ethics is the human language for that inquiry. TRACE is the more compact structural language that sits behind it. Humans should not need to learn TRACE before they can understand why the structure matters.

The first case is deliberately ordinary.

Composite Scene - Leah and the Wall

Leah wakes at 1:17 because Mia is coughing again. The damp patch behind the bed has spread above the skirting board. Leah moves the pillow, takes a photograph on a cracked phone, and opens the landlord's online reporting page. Organisations often call such a page a *portal*: a website through which a person is expected to submit a request, upload evidence, or check progress.

The website knows Leah's reference number, the time, the category she selects, and the size of the photograph. Leah knows the smell of the room, the sound of her daughter's breathing, the money left after rent, the shift that starts in six hours, and the risk of being described as difficult by the person who can decide whether they stay.

The list offers *damp and mould*, *repair overdue*, and *other*. It has no category for the whole situation. Leah chooses the nearest one because the page will not accept her account until it has accepted a label.

The request enters a queue. The wall continues to change.

No participant in that scene knows everything. The difficulty is that their ignorance does not have equal consequences. Leah can send a request. The landlord can authorise work, delay it, change the rent, or end the tenancy through the available legal routes. The website can sort her words into a queue and record that a route was

open. The damp continues while the humans and the system decide what the case means.

The file will later contain a category, a timestamp, a reference number, and the fact that the request entered a queue. Leah's life contains the cough, the smell, the approaching shift, and the power held by the person receiving the request. Both accounts are factual. They are not equivalent.

Two accounts will travel through this book. The *file* is what a system records. The *life* is what the affected person or being carries. They may describe the same event so differently that relying on the file alone hides the practical harm. The file and the life may each contain facts unavailable to the other. Neither is automatically superior; the error is treating the institutional record as complete merely because it is standardised and authoritative.

The central problem can now be stated without poetry: a system may possess less relevant knowledge than the person living with its decision while retaining far more power to shape what happens next.

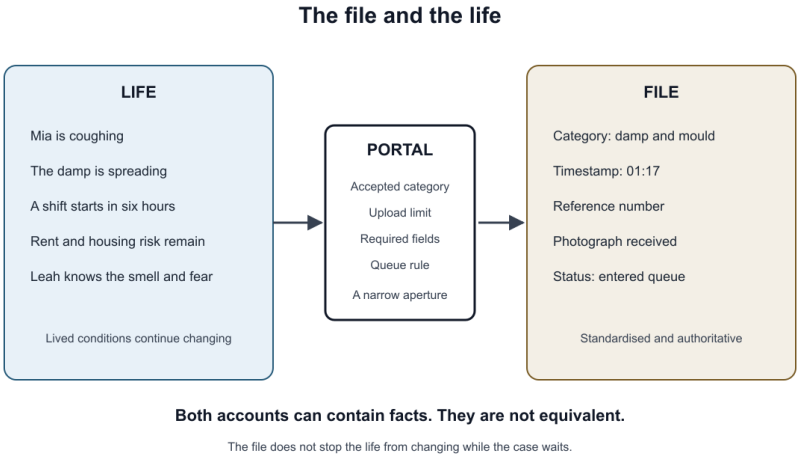


Figure 1. The file and the life: two factual accounts of the same night that are not equivalent.

Mechanical Ethics asks what responsible action requires in that situation. TRACE follows the underlying transitions: what changed, which entities were coupled, which clocks ran, what routes existed, and where correction failed. We will build from the human case outward.

How to Read This Book

This book uses distinctions to reveal recurring structures that are easy to miss when attention is confined to an isolated decision. No formula settles every moral conflict.

A **door** is any claimed route through which a person can ask, challenge, refuse, appeal, pause, repair, or answer back. A conversation can be a door. So can a complaints process, a court, a stop button, a medical second opinion, or the ability to leave.

A **clock** is the period during which that route can still change the threatened outcome. A review that arrives after the relevant job, home, treatment window, relationship, or physical condition has gone may still matter, but it cannot protect the path that has already closed.

Hardening is the process by which reversal becomes progressively more difficult, costly, or unequal. The practical boundary is often uncertain and graded rather than a sharp physical point. Debt accumulates. Evidence decays. A body is injured. A reputation spreads. A model is deployed. A child learns that asking for help is dangerous.

A **witness** preserves enough of what happened for the event to remain open to examination. **Custody** concerns who holds the record and who can change, erase, expose, or use it. **Residue** is what remains after a correction. **Answerability** exists when the affected outcome can reach future action and alter it.

Mechanical Ethics leaves the contents of a meaningful life to the being who must live it. Its thinner claim concerns conditions that another actor's power, delay, or design should not remove: the affected being's ability to remain present in decisions about that life.

Description is not permission. Showing a pattern does not decide every value question or certify the actor who used the language. The framework is voluntary. It is offered as a way of seeing and choosing more honestly, not as an authority that replaces human judgement.

A Timing Condition

The central timing claim in the book is simple to state:

Effective protection must be in place before the threatened path hardens beyond comparable repair.

Let $T_{\text{protect}}(a)$ be the elapsed time until effective protection is in place under the available protective actions a , and let $T_{\text{harden}}(a)$ be the time until the threatened path hardens under those same actions. A necessary timing condition is:

$$T_{\text{protect}}(a) < T_{\text{harden}}(a)$$

This notation does not imply that the times are known exactly. They may be uncertain, represented by ranges or distributions, and affected by unequal access to information and authority. Protective action may also change the right-hand side: pausing a deduction, holding a place, preserving evidence, making a room temporarily safe, or limiting a deployment can extend the interval before loss hardens.

The clocks also need a common starting point, or a known translation between them. Ten minutes after detection cannot be compared directly with twelve minutes after an earlier decision unless we know when detection occurred. Shared units do not make two clocks the same.

The earlier stages can be written as a conceptual compression:

$$T_{\text{det}} + T_{\text{route}} + T_{\text{corr}} < T_{\text{irr}}$$

This is a mnemonic rather than a claim that detection, routing, and correction always occur as fixed serial steps. They may overlap, repeat, or begin provisionally before the preceding stage is complete. The practical issue is whether the whole process produces effective protection before the relevant loss becomes too difficult to repair on comparable terms.

The notation also does not decide which path should take priority when several clocks conflict. It must be applied separately to each threatened path, with the designation of those paths, their practical hardening boundaries, and any priority rule kept visible and open to challenge.

A payment can be correct and still arrive after debt has changed the month. An appeal can identify an error after the car, job, room, treatment window, or evidence has gone.

A technical investigation can be careful and still finish after the dangerous system has acted at scale.

The practical questions are therefore: which clocks are running, who designated them, who benefits from their speed, what can be paused or held, and who carries the cost while the answer is sought?

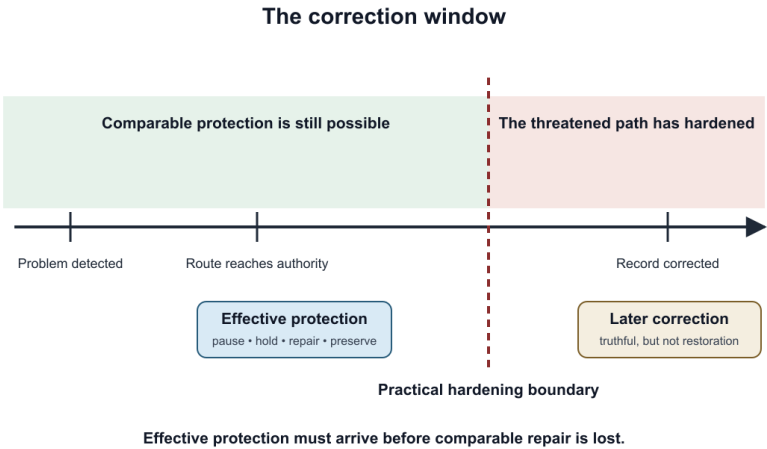


Figure 2. The correction window: effective protection must be in place before the threatened path hardens beyond comparable repair.

Part I - The Door

1. A Door You Cannot Reach

Many organisations provide a formal route for challenging a decision. A letter may contain a website address, a reference number, a list of documents, and a deadline. The organisation can therefore record that an appeal or review was available.

Formal availability is not the same as practical access.

The person may lack internet access, or the requested evidence, or the time. They may lack literacy, confidence, transport, privacy, or any safe way to disclose what happened. Some understand the words without understanding which part of the decision can be challenged. Others know exactly what happened and still cannot translate it into the categories the route accepts.

A route is usable only when the affected person can find it, understand it, afford it, use it without disproportionate danger, and reach someone capable of changing the outcome before the relevant opportunity is lost.

Consider a physical gate. To an adult with a free hand it may be an entrance. To a child who cannot reach the latch, a person using a wheelchair facing a step, or a dog that cannot operate the mechanism, it is a barrier. The gate has not changed. Whether it functions as a route depends on the entity, the design, and the purpose.

The same principle applies to institutional doors. Publishing a process does not establish that the person can use it under the conditions in which the harm is occurring.

Human Sliver - The QR Code

Arthur is seventy-three. The letter about his hospital appointment contains no telephone number for changes. It tells him to scan a square code with a smartphone. Arthur owns a mobile phone that makes calls and sends texts. He takes the letter to the chemist because the assistant there once helped him reset the clock on the phone.

The hospital's record will later show that online rescheduling was available. Arthur's day contains a different fact: the route required a device and a skill he had never been told he would need.

Online systems can reduce waiting, create useful records, and make access easier for many people. The error begins when an organisation treats the existence of one route as proof that the affected population had a realistic opportunity to use it.

So examine the conditions of use. Could the person understand what was required, supply the evidence in the form they actually possessed, ask for help without weakening their own position, and reach a decision-maker before the deadline became decisive? And was there another route for the people this one excluded?

A door deserves moral credit only to the extent that it opens in the world of the entity expected to use it.

2. Who Builds the Maze

A process can contain individually reasonable steps and still become unreasonable as a whole. Identity checks reduce fraud; categories route work to specialists; deadlines protect other parties; evidence rules keep rumour out of decisions. Each step can be defended on its own. The ethical problem lies in how they combine, who bears their cost, and what happens when the situation does not fit the design.

A person may be required to create an account, verify an email address, enter a reference number, select a category, upload evidence, wait for a response, and repeat the account through a second route if the first decision is unsatisfactory. Each step transfers part of the institution's need for order to the person seeking help.

Classification is especially important. A category does more than describe the case. It can determine which team sees it, which clock begins, what evidence counts, and how failure will later be recorded. When the available categories do not fit, the person must either choose an inaccurate label or fail to enter the system at all.

Off-Angle Example - The Goat and the Latch

At a small animal sanctuary, volunteers replace a simple hook on the feed store with a childproof latch. The old hook was easy for people and, unfortunately, for Mabel the goat, who had learned to lift it with her nose. The new latch defeats Mabel. It also defeats Tom, an older volunteer with arthritis, whenever the weather is cold.

The sanctuary has improved security according to one measure. It has also made the route less usable for a person whose work keeps the animals fed. The solution is neither to return to the old hook nor to blame Tom. It is to redesign the latch around the full set of beings and purposes involved.

The example is small, but the structure is the same. A system measures success from the perspective it was designed to notice. If it records only whether the goat entered the store, the new latch works. If it also records whether authorised volunteers can use it safely and reliably, the design problem becomes visible.

Institutional processes often record the missing document, late upload, wrong category, and missed deadline while failing to record the reasons those events occurred. The file can therefore be accurate at the level of each step and misleading about the route as a whole.

Exhaustion inside a process should not automatically be treated as consent, indifference, or lack of merit. A person may stop because the claim was weak. But they may equally stop because every attempt to continue creates more exposure, more proof work, more delay, or more fear.

The organisation that designs the route chooses its language, evidence format, clocks, interfaces, and points of escalation. It cannot fairly evaluate the person only by asking whether they completed the design as instructed. It must also ask what the design required from them, which people it systematically excluded, and whether the route changed the original problem or merely documented the person's failure to navigate it.

The maze is not neutral when one party builds it, controls the map, and uses completion of the route as evidence that the route was fair.

3. Process Becomes Distance

Large organisations need routing. A first contact may lack authority. A specialist may need to inspect a technical question. A manager may need to approve money. An independent reviewer may be necessary precisely because the original decision-maker should not judge their own work.

Routing becomes harmful when the claim moves through the organisation while the burden remains with the affected person.

A case may pass from a general inbox to a specialist team, then to a contractor, a manager, a complaints function, and an appeal body. At each stage the person must learn a new name, quote another reference number, repeat the history, and prove that the previous contacts occurred. The organisation possesses the records but asks the person to carry continuity.

Human Sliver - The Violin in Three Systems

Rosa's violin comes off a flight with a crack beside the bridge. The airline directs her to

the ground-handling company. The handler says damage claims must be opened by the airline. The insurer asks for a repair estimate before accepting the claim, while the repairer will not dismantle the instrument until somebody has accepted responsibility for the cost.

Rosa has an audition in ten days. Each organisation answers within its stated service time. No one can authorise the repair. The claim travels; the crack remains.

The original need may be simple: stop the deduction, repair the wall, restore the account, remove the unsafe condition, pause the deadline, explain the decision, or authorise work before further loss occurs. Additional routing is justified only when it contributes to one of those outcomes or protects a legitimate competing interest.

A process should therefore distinguish between movement of the *case* and movement of the *burden*. A case can be escalated while nothing changes in the affected person's life, acknowledged while the dangerous condition remains, reviewed while the deadline keeps running.

Delay is active. Debt grows, evidence becomes harder to recover, a child remains in the room, a worker misses another shift, and trust in the route declines. The institution may experience the interval as ordinary processing time. The person experiences it as part of the event.

A trustworthy process should be able to answer three questions at any point: where is the claim, where is the burden, and what has happened to the clock? If only the claim has moved, the organisation should not describe routing as progress. The aim is to make the route carry its own memory, authority, and responsibility wherever possible, not to remove routing.

Human Sliver - A Route That Carried Its Own Map

A pharmacist notices that a new prescription may conflict with medicine the patient already takes. She marks the prescription as paused, tells the patient why it cannot yet be supplied, contacts the prescriber directly, and gives the patient the pharmacy's number and the name of the person holding the query. The patient does not have to carry messages between organisations or reconstruct the clinical reasoning.

The system remains complex. Its complexity does not become the patient's unpaid work.

The Door Out

A system may say that a person was free to refuse or leave. That claim depends on what leaving would cost.

A worker can resign, and lose the income that pays for housing. A tenant can move, and lose money, stability, school continuity, or access to evidence. A patient who declines a condition may lose the treatment attached to it; a platform user who closes an account walks away from work, relationships, or an audience that took years to build. The formal option exists, but the price of using it may make the choice largely ceremonial.

A meaningful exit preserves some viable continuation outside the arrangement. The person can understand that refusal is possible, exercise it without retaliation or disproportionate danger, keep essential records and possessions, and reach another route before the loss hardens. This does not promise comfort or remove every consequence. It distinguishes refusal from a choice made after alternatives have been made unusable.

Consent provides weak evidence when the same actor controls the terms, the essential resource, and the penalty for saying no. Repeated acceptance may show dependence rather than approval. Silence can mean fear, or simple exhaustion, rather than agreement.

Some relationships cannot offer immediate exit. A child cannot safely leave a family, a patient may depend on urgent treatment, a citizen cannot readily leave a legal system, and a person under restraint may have almost no practical refusal. In those cases answerability depends more heavily on independent interruption, protected representation, review, and limits on what the stronger party may demand.

An exit counts only if using it leaves something viable on the other side. Otherwise refusal has simply become another form of punishment.

Part II - The Clock

4. Help Before the Loss Sets

Suppose the money comes back.

The refund is real. So is the loan taken three weeks earlier because the money was missing. So are the charges, the argument at home, and the shift accepted at short notice to cover the gap. A correction can be accurate, useful, and late at the same time.

Timing changes what help is able to do. A payment made before rent is due may keep a household steady; the same payment after arrears have begun reduces the damage without undoing it. The same is true of a medical explanation. Given while another treatment is still possible, it protects a choice. Given after the window has closed, it can only describe the choice that was lost.

Late correction still matters. Reopening an account, repairing a wall, amending a record, or admitting an error may prevent the next loss. It may also matter to the person who has spent months being told that nothing went wrong. A system should say plainly what its correction reached and what remained beyond it.

That is why the timing condition near the beginning of this book is about a threatened path, not merely a decision file. Effective protection has to be in place while there is still something comparable to protect, and temporary measures can help keep that interval open.

The difficult cases are rarely solved by speed alone. Acting too quickly on poor evidence can create another injury. Some questions need investigation, specialist knowledge, or a fair chance for another party to answer. The practical issue is what happens during that time.

A contested deduction can be paused, a sample preserved, a room made temporarily safe, a place held, logs protected from deletion. An automated decision can be stopped from spreading to thousands of people while the first cases are examined. None of these steps decides the final argument. They stop the argument from becoming pointless before it is finished.

When an organisation says that review will take twenty working days, the next question should not be whether twenty days is normal for the organisation, but what will happen to the person, the evidence, or the threatened future during those twenty days.

That is what a safeguard is for: protecting something while the answer is still open.

5. The Strategic Unknown

Sometimes *we don't know* is the most honest sentence available. It can also be the sentence behind which a damaging arrangement continues.

There is no shame in uncertainty. A doctor may need another test, an engineer may need to reproduce the fault, a landlord may want to inspect before accepting a cause, and a reviewer has to hear from both sides. Pretending to certainty would be worse.

The difference lies in what the uncertainty sets in motion.

Honest uncertainty creates work for the party with the power to resolve it. Someone is named. Records are opened, evidence is preserved, the threatened action is paused where it can be. The person affected is told what is missing, what could change the answer, and when the question will be looked at again.

Strategic uncertainty has another shape. The existing arrangement continues because nobody has formally proved that it should stop. The person who raised the problem is asked for more evidence, even when the institution holds much of it already. No one owns the investigation. The deadline belongs to the person at risk, while the organisation's delay has no comparable cost.

Nobody has to be lying for this to happen. If waiting is cheap for one side and expensive for the other, uncertainty begins to lean. The lean becomes stronger when the party benefiting from delay also controls the records, the route, and the clock.

A useful question is simple: who is paying while nobody knows?

Is the organisation spending money or authority to find out, stopping the contested action, and protecting the evidence and the person? Or does the rent rise, the payment stop, the model continue, the medicine remain unavailable, or the dangerous condition stay in place until the weaker party proves the case?

Historical Note - Frances Kelsey and Thalidomide

In 1960, FDA medical reviewer Frances Oldham Kelsey received an application to market thalidomide in the United States. She judged the submitted safety evidence inadequate and repeatedly required further information despite pressure from the manufacturer. Her caution had practical force because the law and review process gave the FDA authority to prevent the application from taking effect while the evidence remained insufficient. By late 1961, reports had linked thalidomide exposure during pregnancy to severe birth defects in countries where the drug had been used.

The importance of the case is not that uncertainty alone produces safety or that uncertainty always requires refusal. Protection depended on uncertainty being joined to statutory authority, an effective pause, and a duty placed on the applicant to supply adequate evidence before the risk was introduced.

The default course is itself a choice. Continuing, pausing, narrowing, or reversing an action each places risk somewhere. A responsible system makes that placement visible instead of presenting continuation as though nothing had been decided.

Unknown remains honest only while someone carries a real duty to reduce it.

6. Slow Harm

Not every injury arrives as a single event.

A case can remain officially open while a person's life closes around it. The reference number still works. The website still says *pending*. Nobody has issued a final refusal. Yet the person borrows money, misses work, sleeps in another room, stops using a service, accepts a worse agreement, or gives up because waiting has become more expensive than losing.

This is one reason administrative records can look calmer than the lives they describe. The record sees a case in progress. The person is already adapting to the failure.

Technical reversibility makes the problem easy to miss. A database can be corrected. A deduction can be refunded. A wall can be repaired. But the month in which the error operated cannot be reset. Borrowing changes relationships. Missed shifts change how an employer sees someone. A child learns where it is safe to sleep. A person who expected help begins planning on the assumption that help won't come.

The original harm has now acquired a history.

A serious correction has to follow that history. It is not enough to count the original amount or fix the original entry. What happened because the error remained active? Was interest charged? Did health worsen? Was evidence lost? Did the person withdraw from a route they will later be blamed for not using?

Silence needs particular care. It may mean the problem was solved. Or it may mean the route taught the person that another message would bring another demand, another delay, or another risk. A quiet file is not necessarily a quiet life.

The practical question is not only whether this decision can be reversed, but what is becoming harder to reverse while we wait.

Once that question is asked, temporary protection stops looking like generosity and becomes part of doing the investigation honestly.

Short Case - The Finance Appeal That Arrived Too Late

Composite Scene

Amina's car fails its MOT on Thursday. Her new rota begins on Monday, with two early shifts at a care home beyond the first bus. On Friday she finds a used car she can afford and applies for finance from the dealer's desk.

The refusal arrives before the salesperson returns with the keys. The email says that an automated assessment informed the decision and that a human approved it. A review is available. The company expects to respond within twenty working days.

Amina checks her credit report on her phone and finds an old address attached to the wrong account. She sends a tenancy agreement, a council-tax statement, and a photograph of her driving licence. The website confirms receipt. The dealer cannot hold the car for twenty days.

She hires a car for the first week and borrows the deposit from her sister. During the second week she accepts finance elsewhere at a worse rate because missing the shifts would put the job at risk.

The review later finds that two records had been joined. The company corrects the file and invites her to apply again. The first car has been sold. The hire cost, the new agreement, and the debt to her sister remain outside the corrected entry.

The appeal was useful. It repaired the record and may protect the next application. It did not reach the live decision that existed on Friday.

The case illustrates the difference between retrospective accuracy and timely correction: a safeguard must be judged partly by whether it can still change the path it claims to protect.

Part III - The Witness

7. When There Is No Enforcer

Sometimes there is nobody coming.

The complaints process exists. The regulator exists. The court, union, ombudsman, manager, or police service exists. None of them can reach the event with enough speed or authority to stop what is happening now.

That failure differs from the original harm. It is a second fact about the world: the route of answerability does not connect to an effective interruption.

Telling a person to keep records in that situation can sound almost insulting. A photograph does not remove mould. A diary restores no wages, and no screenshot has ever made an abusive person safe. A witness pack is not a rescue plan, and this book should never pretend otherwise.

The record has a smaller purpose. It stops the actor or institution from becoming the only author of what happened. It may allow a later reviewer to reconstruct the sequence, help separate cases become visible as one pattern, support legal or public action, or warn the next person before the same route closes around them.

A useful record is usually less dramatic than people imagine: what happened, when, what was known, what was requested, which routes were tried, what changed while nothing happened, and what remains unresolved. The purpose is to preserve enough structure that the event cannot later be reduced to a convenient label, not to turn the affected person into a detective.

Other people may need to carry part of this work. A neighbour can keep a copy. A support worker can write down a date. An engineer can preserve a warning. A clinician, union representative, journalist, colleague, researcher, or friend may hold evidence that would be dangerous for the person at the centre to publish themselves.

Records also create risk. They can expose a child, identify a worker, reveal a location, or give an institution new information to use against the person who complained. More evidence is not always safer evidence.

A witness does not guarantee justice. What a witness does is keep the past from being closed before it can be examined.

8. Custody

Most evidence is held by somebody.

A photograph may sit on a cracked phone. A repair history may sit on a landlord's server. A warning may remain in an engineer's email. A clinical note may be visible to the hospital and difficult for the patient to obtain. A memory may exist only in the mind of a person who is frightened of what will happen if they speak.

Where the record lives changes what the record can do.

Custody is not only storage. It includes who can see the material, alter it, delete it, copy it, demand it, combine it with other records, or use it against the person who created it. The same screenshot can protect one person and expose another.

Institutions usually possess the fuller account: banks the transactions, platforms the logs, employers the personnel records, hospitals the notes, landlords the repair histories. The person affected may have only fragments. A date on an envelope. A photograph. A text from a neighbour, or a name remembered from a phone call.

Fragments are no more truthful for being fragments. They are simply outside the main actor's control, and sometimes that is what allows a later comparison to happen.

Good custody makes alteration visible and challenge possible. It may require independent storage, access logs, preserved metadata, limits on deletion, and a right for affected people to inspect and contest records about them.

The insider witness needs the same seriousness. A caseworker, engineer, clinician, auditor, moderator, or administrator may hold the only record that shows the problem clearly. Preserving or disclosing it may threaten their job, legal position, professional registration, or safety. A correction route that relies on such people cannot treat courage as an unlimited resource.

If a system can rewrite its own memory after challenge begins, it should be modest about how confidently it later describes the past.

A record can become misleading without being rewritten. A count or summary may be correct when produced and stale after the situation beneath it changes. A current date is not the same as a current fact. If a decision depends on such a record, whoever relies on it should know what change would make it necessary to check again.

Evidence design is part of the moral design. A route is already tilted when proof can exist only inside the institution being challenged, or only at unacceptable risk to the person who preserves it.

9. Residue

Suppose the decision is corrected.

The account is reopened. The money is returned. The letter admits delay. The wall is repaired. Something real has changed, and it would be wrong to dismiss that change because it is incomplete.

Now ask what the error did before the correction arrived.

There may be debt, lost time, illness, embarrassment, damaged trust, altered relationships, or a new reluctance to use the service. Some of that can be repaired. Some can be compensated. Some can only be acknowledged honestly.

This remainder is easy to lose because the institution's attention follows the active case. Once the deduction stops or the wall is fixed, the main indicator turns green. The person continues carrying effects that no longer appear in the operational view.

A claim of closure should therefore include a return of burden: which costs have gone back to the actor, which still sit with the person affected, what has been done about the consequences created during delay, and what route remains open if another consequence appears later.

The party announcing closure should do most of the work of supporting that announcement. It should not fall to the harmed person to prove that the institution's declaration of success is incomplete.

None of this requires every case to remain open forever. Records can be finite. Agreements can be reached. People can choose to move on. The point is to avoid converting a useful repair into a false account of the past.

Scarcity makes the problem harder. Several people may need repair at once, and not every loss can be addressed immediately. Then the order matters. Housing debt, medical need, lost income, and damaged records may compete for the same limited money or attention. Who chooses which one waits? Can the people carrying the cost see and challenge the rule?

An open ledger is not enough if power still decides, invisibly, which debt will be postponed.

A real correction changes something. What it cannot change is the reason honesty still matters.

Interlude - Two Flats, One Wall

Composite Scene

Two flats share a wall. Martin owns upstairs. Leah and Sam rent downstairs with their daughter, Mia.

A stain appears behind Mia's bed in October. Leah wipes it away and moves the furniture. By November, the paint has lifted in a strip wider than her hand, and Mia wakes coughing on cold nights.

Leah reports the damp through the landlord's website, selects *damp and mould*, uploads a photograph, and receives a case number. The room does not change.

Upstairs, Martin notices a stain beside the same wall. He emails the managing agent from work. A contractor visits three days later, traces water around an exterior joint, and recommends scaffolding. The work begins before the month ends.

The same wall has entered two systems of attention.

Martin is an owner whose future business the agent wants to retain. Leah enters through a tenant route designed to receive many requests at low cost. One route reaches a person with authority. The other produces a reference number.

The landlord knows that tribunal action is slow. A rent-increase letter was already due, and he sends it one week after Leah's request. It does not mention the damp. It does not need to. Leah can see that the rent will move now and any challenge will move later.

She opens the category list again. "What box am I supposed to tick for scared?" she asks Sam.

The website contains no field for that information.

Leah sends fewer messages. An automated reply asks for a wider photograph, which would show Mia's bed and most of the room. She does not send it. The file records that no further evidence was received.

The upstairs wall is repaired. Downstairs, an inspection is offered after Leah and Sam have found another tenancy. They leave before anyone enters the room. The case closes because the tenancy has ended.

Martin never learns that the downstairs case existed. The wall is painted before the next tenant moves in.

Fragments survive outside the landlord's custody: photographs, messages, dates, a note held by Leah's niece, and a copy kept by an advice worker. They do not repair the room or repay the move. They prevent the official file from becoming the only account.

Later software may price the family away before a viewing because it predicts complaint. There is no refusal to appeal and no witness inside the decision.

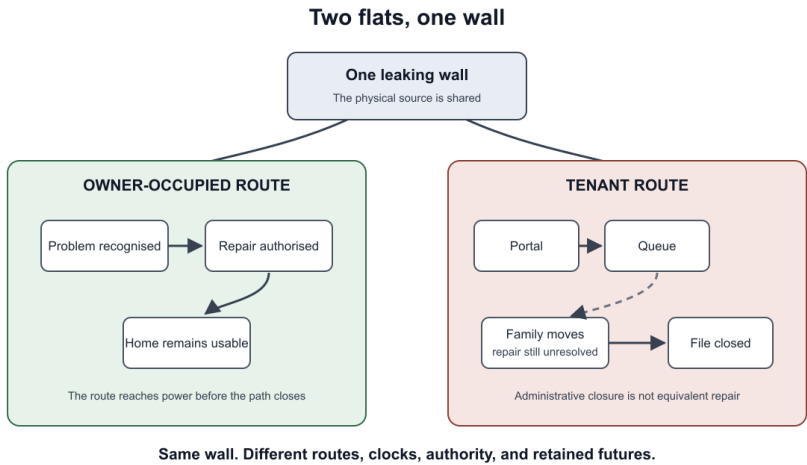


Figure 3. Two flats, one wall: the same cause reaches an owner with authority and is repaired within the month, while the tenant route returns a reference number and closes on move-out.

Part IV - The Machine

Before the Machine - The Predator Position

A difficult process is not automatically predatory. Some work is genuinely complicated. Some delays protect other people. Some evidence rules are necessary.

Now change the question. Who benefits when the difficulty remains?

A *predator position* exists when an actor or system receives persistent or systematic benefit from an arrangement that keeps another entity unable to understand, challenge, escape, or correct it, while the benefiting actor controls or materially influences the route and could reduce that inability. The term describes the position, not the personality of whoever occupies it. No monster or secret plan is required.

A landlord may know that a tribunal will take longer than a tenant can stay. An employer may lean on a worker's dependence on income or immigration status. A firm can design a refund process that most customers will abandon, and a platform can profit from a harmful pattern while keeping every individual case too small and scattered to force change.

In those situations, process does work that open force would otherwise have to do. Delay, categories, proof demands, and uncertainty keep the weaker party in place.

The test is not whether the process is unpleasant or whether a single asymmetry exists. It is whether maintained unusability repeatedly protects the party who controls or materially influences the route, while the cost of failure falls somewhere else.

Once that pattern is visible, good intentions are no longer enough: the design itself has to answer.

10. Who Pays for Complexity

The usual defence of complexity is that the world is complicated. Often it is.

Medicine, housing, finance, aviation, energy, and public administration involve real technical knowledge, competing duties, and large numbers of people. A simple answer can be careless or dangerous. Mechanical Ethics is not an argument for making every system small; it is an argument about who pays for the complication. Complexity is more defensible when it protects a stated interest, is no greater than reasonably required, and is carried internally wherever the institution has the capacity to carry it.

Forms, categories, vendors, models, data feeds, staffing levels, interfaces, and escalation routes were made by somebody, even when no single person designed the whole. Years of local decisions become the environment in which another person is expected to act.

When an institution says that a decision is too complicated to explain, it has named a problem in its own design. It has not identified a defect in the person asking.

A complex system should carry enough of its own map to show which rule or model mattered, what evidence was used, where uncertainty remained, which clock was running, who could pause the action, and how a challenge could reach someone with authority. The affected person should not have to become an expert in the entire organisation merely to discover why one path closed.

This also matters for people inside the system. A caseworker may see the problem and lack the button required to fix it. An engineer may understand the risk and have no route to the executive decision. Blaming the last human who touched the file can hide the more important fact that authority and information were separated by design.

Responsibility may be distributed without becoming absent. The designer who set the categories, the manager who maintained the incentive, the custodian who withheld the record, and the executive who could interrupt the process may each stand in a different relation to the same outcome. Their duties need not be identical. Fragmentation of control should not become a way for the organisation to claim that nobody was answerable for the arrangement as a whole.

The basis of a high-consequence decision should remain reproducible or open to evaluation by a properly supported challenger within reasonable time and access. When the route cannot support that scrutiny, the institution has not earned confidence in the result.

A system cannot manufacture unreadability and then use unreadability as its defence.

11. The Second Person

Closing one file is not the same as changing what will happen to the next person.

The first person receives a refund, repair, apology, or corrected record. That matters. Meanwhile the same rule, category, model, incentive, or delay remains in place. The second person has not arrived yet, so they do not appear in the success figures.

Organisations naturally see the cases they have handled. A closed case is countable. A future person is not. This makes it possible to become very good at correcting individual failures while continuing to produce them.

The surface labels may differ. One file says *repair delayed*. Another says *insufficient evidence*. Another says *no further contact*. A fourth ends because the tenancy ended. Underneath, the same mechanism may be operating: the person cannot safely supply what the route demands before the institution's clock wins.

The useful question after a correction is therefore not only whether this person was helped, but what has become different for the next person.

Can the next person reach help earlier, pause the action, preserve evidence, and reach someone independent who can inspect the decision? Has the incentive that rewarded delay or opacity changed, and is the same failure now harder to reproduce? The second person is where a repaired case becomes a repaired system.

When One Case Is Not Enough

One case can expose a mechanism, but it can also be dismissed as an exception. Several cases can make the shared route visible, especially when the surface labels differ but the same burden, delay, or exclusion recurs.

Collective evidence does not require every account to be identical. People may disagree about cause, remedy, or severity while still showing that one design repeatedly closes the same kind of path. Combining their records should preserve those differences rather than flattening them into a single representative story.

Collecting a pattern can also change what happens next. People may respond because they were named or counted; an institution may behave differently once it knows an audit exists. Later evidence may therefore describe a situation partly changed by the attempt to measure it.

Representation creates another custody question. It can give scattered people enough force to reach authority, but it can also select, simplify, expose, or speak over the people whose accounts give it weight, becoming another gate.

The purpose is not to turn numbers into automatic authority. A frequent pattern can still be misunderstood, and a rare event can still reveal severe harm. The stronger claim is practical: when separate accounts point to a shared mechanism, the institution should not force each person to begin from zero or treat the absence of one perfect case as evidence that no pattern exists.

Collective correction begins when the burden of showing recurrence no longer rests on each isolated person and when the combined account can change the rule, incentive, interface, or authority that keeps reproducing the harm. The group does not erase the person. It gives the person a route that may finally carry enough weight to reach the machine.

12. Correction Theatre

Organisations know what repair is supposed to sound like.

They apologise. They commission reviews, train staff, create dashboards, appoint owners, publish plans, and say that lessons have been learned. Any of those things may be useful. None of them proves that the harmful route changed.

Correction theatre begins when the appearance of answerability improves faster than the conditions faced by the next person.

The test is practical. Which action has become impossible, slower, more visible, or easier to challenge? Who can check that change? What would count as evidence that the repair did not work?

Training may matter, but not if the interface still offers the same choices. A dashboard may reveal a pattern, or merely teach staff how to produce a cleaner number. An apology may be sincere while the next case enters the same queue.

A review can check every case in its file and still miss people the route never allowed to become cases, or make unlike cases look like one pattern. What the review can honestly claim depends on who entered the file and how their experience was measured. The organisation should state those limits and show whether another supported account changes the picture. That account may still be wrong, but a material difference should not disappear inside one clean total.

Institutions are often under real pressure to announce completion. Open problems cost money, damage confidence, and attract scrutiny. That pressure is one reason the language should stay exact. Investigation is not repair, and activity is not consequence. Monitoring is not interruption. A recorded lesson, however sincerely minuted, is not a changed path.

The matter can close when it is genuinely finished. Until then, accurate words are part of the repair.

A correction that cannot be checked remains a claim about correction. Checking does not require universal agreement. A repair claim should state the observable change, who can inspect it, and what evidence would count against success, so that disagreement has an object rather than becoming an endless refusal to recognise repair.

Historical Note - Challenger and the Cost of Proceeding

Historical Case

On 28 January 1986, the Space Shuttle *Challenger* broke apart seventy-three seconds after launch, killing all seven crew members. The Rogers Commission concluded that the launch decision was flawed. The people who made the final decision were not given the recent history of O-ring problems, the contractor's initial written recommendation against launch below the previous low-temperature experience, or the continuing opposition of engineers after contractor management reversed that recommendation.

The failure was not a single simple override. It involved incomplete transfer of technical history, pressure within the decision process, reversal of an adverse recommendation, and the absence of a reliable stop rule through which the concern retained enough force to prevent launch. The cost of delay would have been carried by the programme and its schedule. The cost of proceeding was carried by the crew.

The lesson is not that engineers must always prevail or that every warning requires cancellation. Technical experts can be wrong, and systems must choose under uncertainty. The requirement is that high-consequence action has a credible route by which adverse evidence can pause the decision, reach authority without distortion, and remain visible after management pressure begins.

The Machine-Speed Brake

Machines change the timing before they change anything else.

A clerk can make one decision, then another. An automated system can deny access, alter prices, move money, classify risk, suppress information, or direct physical action thousands of times before a person notices a pattern.

Human correction does not accelerate at the same rate. Somebody has to recognise that a decision occurred, understand enough to question it, find a route, gather evidence, and reach a person with authority. By then the model may have repeated the same transition across an entire population.

That gap is where a nominal appeal becomes useless. The route exists, but the machine can finish the relevant action before the route can touch it.

High-consequence automation therefore needs stages. A decision may be tested on a small scale, held before final commitment, made reversible, limited in duration, or

placed behind an independent interruption. The exact brake depends on the domain. A price change, medical classification, benefits decision, and use of physical force cannot share one threshold.

The principle is simpler than the implementation: the more quickly a system can close important paths, the more quickly reliable correction must be able to reach it. If human review cannot move at machine speed, the machine has to leave room for human-speed interruption.

Otherwise the system is not corrected after acting. It is merely explained.

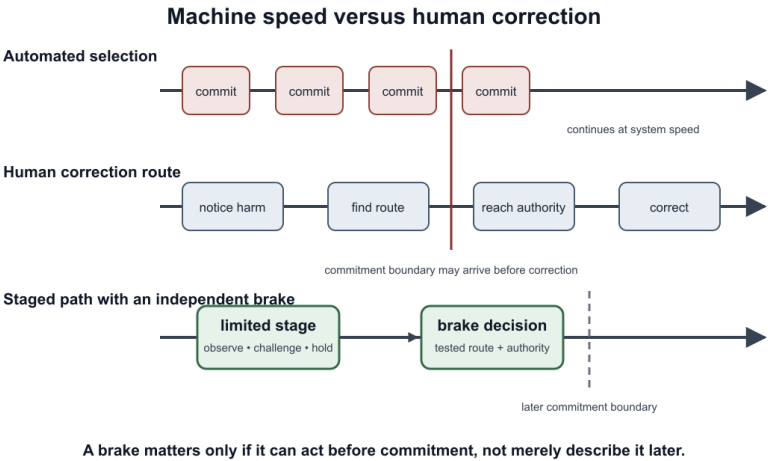


Figure 4. Machine speed and human correction: the machine completes its action at scale before human correction has finished assembling.

Part V - After the Fall

13. Peace as Evidence

After harm, there is pressure to reach an ending.

The person carrying the damage may be exhausted. The institution wants the case closed. Colleagues want normal work to resume. Families and communities may want the conflict to stop consuming everything around it. Those are real needs.

The trouble begins when a peaceful ending is used as evidence that the repair was complete.

A person may accept compensation because they need the money, not because the account is accurate. They may forgive one individual and still distrust the institution. Some stop answering because another meeting is unbearable; some want no contact at all. None of those choices gives somebody else permission to write a cleaner story.

Honest reconciliation keeps the facts intact. It says what was repaired, what remains, and what changed in the route. It can recognise the people inside an institution who tried to help without using their decency to excuse the structure around them.

There is no duty to remain angry forever. There is also no duty to perform forgiveness so that the record can close comfortably.

Peace may be part of repair, but it cannot be the test used to prove that repair happened.

14. What Remains Open

A file can close because the work is complete. It can also close because the tenancy ended, the person left, the deadline passed, the account was abandoned, or the institution stopped hearing from them.

The world continues either way.

Some losses remain obvious: debt, illness, a missed job, a changed relationship. Others are harder to see. A person learns not to ask. A benefit passes to somebody who did not cause the loss. A system remains unable to notice the class of person it excluded.

Keeping the record honest does not require endless accusation; it requires resisting the sentence *nothing remains* when something plainly does.

Inherited benefit is one example. A person may move into a flat after another family was forced out, take a role after another worker withdrew, or acquire an asset whose clean appearance depends on an unresolved history. The new person may have done nothing wrong. Inventing personal guilt would only obscure the question. The issue is what benefit remains, what claim remains, and what becomes possible once the history is known.

The reader may also stand in several places at once. A person harmed by one institution may be well served by another. Their belief that *the system works* can be completely sincere because the route was built around people with their resources, language, confidence, body, or social position.

That experience is real and partial.

The deepest selection often occurs before any visible decision. Somebody chooses the categories, evidence, baseline, household, urgency, and meaning of silence. What cannot be entered is easier to treat as though it never happened.

Leah's category list had no field for fear. The omission did not announce itself as a refusal. It simply made part of the situation unavailable to the system that would later judge her response.

That is the never-built door: the route, category, or option whose absence is invisible to the person excluded by it. No form can contain every relevant circumstance. The practical requirement is a route for adding material context, contesting the category, and revising the design when the same exclusion recurs.

An open record does not solve the missing door, but it preserves enough truth that somebody else may later see where it should have been.

15. A Human Standard for Answerability

Go back to the ordinary letter: a website address, reference number, evidence request, and deadline.

At first glance it offers a route. The rest of the book changes what we need to ask about that claim.

- Can the person find and use it under the conditions they are actually living in?
- Does it reach anyone who can alter the outcome?
- Will it arrive before the home, job, treatment, evidence, relationship, or opportunity is gone?
- Who owns the records? Who set the clock?

- What happens while the answer remains uncertain?
- What does the person still carry after the file is corrected?
- Will the next person meet the same mechanism?

Answerability begins when an affected outcome can travel back into the behaviour that produced it. Explanation alone is not enough. The route needs enough force to change what happens next.

Answerability needs an end as well as a route back. Reopening the same challenge without limit can become a route of harm. Repetition alone proves neither merit nor resolution.

Disagreement about law, medicine, engineering, politics, care, and scarce resources remains. No single formula can decide every conflict. The point is narrower: some limits must hold before conduct can honestly be called responsible.

The timing condition is not a universal priority rule. Using it as if it ranked competing lives, paths, or claims would be a misuse.

The limits are these:

- a route must be usable by the beings expected to rely on it;
- serious correction must stay faster than serious hardening;
- uncertainty must create work for those with power, not automatic permission to continue;
- evidence must be checkable beyond the challenged actor's control;
- complexity must not be dumped onto the people who did not create it;
- repair must reach the mechanism as well as the case;
- closure must not erase residue;
- and a party benefiting from necessity cannot be the only judge of necessity.

These are conditions under which an answer can still deserve trust. Witness, interruption, enforcement, protection, and organised resistance remain distinct, and a truthful record may coexist with power that accepts the damage. Seeing the structure is necessary, but without interruption or enforcement it cannot be sufficient.

These conditions constrain the exercise of power. The purposes of a remaining life belong to the person who must use it.

Part VI - Futures Still Reachable

The preceding sections have concentrated on failed routes, delayed correction, and burdens transferred to people with less power. Part VI asks a narrower positive question: whether an arrangement leaves the affected being with a usable and correctable future in which the contents of that life have not already been determined by another actor.

The question here is whether the affected being remains present in the decision about what comes next. Somebody with greater control may carry complexity rather than exporting it, protect another person's refusal, preserve a route that can still be used, or accept a cost that prevents another life from narrowing further. These acts offer no guarantee of success, but they change who remains able to participate in what happens next.

Interlude - The Saturday Workshop

Composite Scene

Priya lifts the shutter at ten every Saturday. One garage holds bicycles in different stages of repair, while the other contains donated laptops, kettles, lamps, and components sorted into takeaway tubs labelled in marker pen. One tub reads *CABLES THAT MAY OR MAY NOT BE IMPORTANT*.

Mick arrives early to test the extension leads. He repaired electrical equipment for most of his working life, and he and Priya teach by working beside people: they help someone identify a fault, use a tool safely, and return to a job after the first attempt fails.

Nia brings in a bicycle whose rear wheel rubs against the frame. She is seventeen and has missed college twice because she can no longer keep paying for the bus.

"I don't know anything about bikes," she says.

Priya turns the wheel once. "You know it doesn't ride properly. Start there."

They find a bent axle and replace it with one taken from a donated wheel. Nia returns the following Saturday for the brakes, and a month later she is showing a twelve-year-old how to lift a tyre from the rim without pinching the tube.

Idris arrives after midday because he cares for his mother in the mornings and can usually stay for only an hour. He is learning to replace charging sockets on cheap phones, although his hands become less reliable when somebody watches. Mick gives him the bench nearest the wall, leaves the diagram beside the soldering iron, and works at the other end of the garage.

The workshop has rules designed around the work it is trying to protect. Nobody has to explain an absence. Photographs require a new decision each time. Participants can inspect anything written about them. Threats result in leaving for the day and meeting away from the group before returning. Stolen goods are refused, and parts donated for community repairs cannot be used to support predatory resale.

A national company offers sixty thousand pounds over two years, enough to cover the rent, replace unsafe tools, and open the garages on three evenings each week. Priya has spent months calculating whether the workshop can survive the winter, so the offer addresses a real problem.

The contract would require every participant to create an account. The sponsor wants attendance data, behaviour scores, photographs for publicity, and permission to share risk and engagement information with approved partners. It would also retain the power to exclude anybody whose presence might harm the programme or the sponsor's reputation.

The representative demonstrates the dashboard at a folding table between the bicycle stands. Green bars record attendance and red flags record risk. He says the data will prove that the project works.

Nia asks who can see her record.

“Appropriate partners.”

“My college?”

“Where sharing supports outcomes, possibly.”

“The council?”

“Potentially.”

“The police?”

He says the legal team will answer detailed questions during implementation.

Idris asks whether mornings spent caring for his mother will lower his score. The representative says the model considers context, although he cannot show what counts as context or who can correct it. Mick asks who can remove a risk flag once the system creates one, and the representative returns to the new equipment and evening sessions.

The trustees meet after he leaves. The offer would solve genuine problems, and refusing it means that one broken stand will remain unusable, evening sessions will not happen, and the waiting list will grow. Priya writes the exchange on a sheet of cardboard because the whiteboard is covered with repair jobs.

Under the proposed contract, the sponsor would receive evidence that the workshop helps people, while the participants would supply permanent records to produce that evidence. Missed sessions, caring duties, arguments, and mistakes would enter a

system they could not inspect. The people with the least control over the programme would disclose the most and lose control of the record first.

The trustees refuse the contract. They later accept a six-month grant from a local trust, reduce the workshop's hours, and postpone replacing the second stand. Priya gives up one paid day each month, and three people remain on the waiting list longer than they would have under the company offer.

The trustees also change their own record-keeping. Attendance remains on paper and is destroyed after the funding return is complete. Incident notes include the participant's account, and anybody described in a record can inspect and challenge it. The local trust receives costs, repair totals, attendance numbers, and anonymised explanations of barriers to attendance.

The return states that twelve young people attended regularly. That figure is accurate, although it cannot describe Nia reaching college without borrowing bus fare, learning to true a wheel, and showing somebody else how to begin. It also cannot describe Idris replacing four charging sockets or continuing when the room is arranged so that Mick is not watching over him.

The safer record costs money, time, equipment, and places on the waiting list, and those losses belong in the account. The account must also include the capabilities that remain with the participants because the workshop refused to turn vulnerability into an asset.

On the next Saturday, Nia arrives while guiding a neighbour's bicycle beside her. The wheel keeps catching, so Priya clears the spare stand and Nia begins with the part she already knows how to test.

16. The Conditions of Hope

The refusal leaves the workshop fragile. Rent remains due, one repair stand is broken, evening sessions have been cancelled, the waiting list is longer, and Priya has surrendered paid time. Principled choices offer no guarantee of good outcomes.

Hope here is narrower. A future remains practically available because the garage still opens on Saturday, the tools can be used, records remain limited, and people can enter without surrendering control over the meaning of their difficulties. This route has a first step, people capable of maintaining it, evidence that can reveal failure, and permission to admit when it no longer exists.

Encouragement without those conditions can become another demand placed on the person already carrying the cost. Telling somebody to continue makes little sense when there is no usable route, no one able to maintain it, and no evidence that effort can alter the outcome.

The workshop is fragile, but its effects are concrete. Nia can travel to college and diagnose a faulty wheel. Idris can repair charging sockets when the working environment is adjusted. Skills move between people, and somebody who once needed help can later become part of the route for another person.

This is how repair can accumulate without erasing what was lost. Not every reachable future is viable. A future may remain technically available while being coercive, unaffordable, uncorrectable, or controlled by somebody else. In this book, viable means that at least one continuation remains usable, maintainable, and open to answer by the being who must live it. Viability leaves the contents of the life open and requires more than a count of options.

17. Care Without Possession

Composite Sliver - June and the Greyhound

June's greyhound, Sol, is old and seriously ill. The vet can offer another procedure that may add time, although recovery would be painful and the chance of returning to the walks, food, and easy sleep Sol still enjoys is low.

June wants more days. Sol cannot explain a philosophy of life, but he can show appetite, distress, interest, avoidance, and trust. June asks the vet to describe the likely experience from Sol's position rather than speaking only about the number of weeks the procedure might add.

Choosing against another procedure does not treat Sol's future as worthless. It recognises that value cannot be reduced to duration or the number of available interventions, and that the present life being protected must remain within the decision.

June loves Sol. The difficulty lies in acting when her wish for more time conflicts with Sol's likely experience. Care always interprets another being through limited knowledge, fear, history, and need. Sometimes protection requires a decision for somebody who cannot make or communicate it. Tenderness leaves the power of that decision intact.

Ownership begins when the carer's need becomes the only measure that matters. More days, obedience, family identity, institutional success, or a model's preferred outcome can displace the being whose life is supposedly being protected.

Care may require refusal, unwanted truth, withdrawal from a damaging course, or protection from another person's choice. It differs from possession when the other being remains present as a subject whose condition, signals, relationships, distress, enjoyment, and capacity to answer continue to affect the decision.

A brief remaining future can matter fully. Relief, play, familiarity, dignity, one wanted act, or an easy sleep may carry more value than an additional option whose main benefit belongs to the person choosing it. Preserving possibility cannot require sacrificing the life that exists now to an abstract preference for more.

18. Kindness as the Placement of Burden

Kindness can lose. Cruelty, exploitation, fear, and domination can produce quick results, especially when the act is hidden, the parties will not meet again, the record cannot travel, and somebody else absorbs the cost.

The pharmacist in Part I carried the clinical query instead of making the patient carry it between organisations. The workshop accepted fewer sessions rather than converting vulnerability into a permanent data asset. In both cases, somebody with greater control absorbed difficulty that could easily have been transferred to a person with less power.

Kindness here means the deliberate placement of burden. It also requires boundaries. A person or institution that refuses to stop an abuser, predatory actor, dangerous process, or false claim may simply transfer the cost of its restraint onto people with less power to refuse. Boundaries and interruption can therefore form part of care.

Memory matters because a correction that leaves no usable trace has to be rediscovered in every case. Records, promises, testimony, journalism, courts, logs, and shared practice can allow one intervention to influence the next, although those same mechanisms can be used to punish, classify, and control. Memory contributes to kindness only when the people described can inspect it, answer it, and challenge the uses made of it.

Systems organised around fear often damage the information they depend on. People conceal, comply outwardly, withdraw, or resist; trust declines, correction weakens, and greater force is required to produce the same appearance of order. Such systems may remain powerful, but the information they receive becomes less trustworthy and their apparent order more brittle.

Cooperation can combine capability and bring more perspectives to error, but it can also become capture, while care can become paternal control. Positive language does not certify the arrangement beneath it. Mechanical Ethics therefore treats kindness as an allocation of work, risk, and cost constrained by memory, answer-back, refusal, and correction. The protected being must remain present as a subject of the care.

Conclusion

Mechanical Ethics began with an ordinary mismatch: a decision can look correct in a file while doing something very different to a life. A record can be corrected long after the future it affected has already changed. The central question is not merely whether a decision can eventually be explained or reversed. It is whether the person or other being living with it can still reach a usable future by the time the correction arrives.

This framework does not tell anyone what makes a life meaningful. That remains with the being who must live it. Its claim is narrower: power should not remove that being from decisions about its own life by closing every genuinely viable way forward. A viable route is more than an option that exists on paper. It must be possible to find, understand, afford, use, and maintain. It must leave room to answer, refuse, challenge, and correct. Keeping such a route open does not prove that a life is good, or that one path is best. It simply preserves enough agency for those questions to remain real.

This changes what responsible action requires. Before a high-consequence decision proceeds, someone must ask what could close, when it could become hard to restore, and who controls the clock and the evidence. They must ask what temporary protection could keep the threatened path open, who will carry the burden while uncertainty is resolved, and what loss may remain even if the record is later fixed.

Timing matters because explanation and protection are not the same act. A pause, a hold, or a temporary repair can sometimes buy time. But once a path has hardened beyond comparable repair, a later apology, compensation payment, or corrected file cannot make that future available again.

None of these questions answers itself. Who counts, what should be protected, and what evidence is accepted are all contestable choices. Their authorship must be kept visible, especially when one party holds more power. Several protected paths may also conflict. Mechanical Ethics can expose those competing clocks and burdens, but it does not supply a universal rule for ranking them.

Nor does identifying a failure force the responsible actor to fix it. Diagnosis is not enforcement, and a record of harm is not a rescue.

The practical demand is modest to state and difficult to meet: keep the affected being present in the decision. Keep at least one genuinely usable route open. Place the burden of complexity and delay on those with the power to carry it. And act before correction becomes only an account of what was lost.

Author's Note - How This Began

I wrote the first complete version of this reader in my living room during June 2026, after a much longer period of conversations about harm, responsibility, systems, and the difference between a decision that looks correct in a file and the same decision as it reaches a life.

I have worked around technology for most of my adult life, but recent AI systems made the problem feel more immediate because they could produce convincing language, recognise patterns, and increasingly influence routes into work, housing, credit, care, knowledge, and public decisions. I wanted to understand what those systems could represent, what humans were failing to explain, and whether the gap between the two could be reduced without pretending that an AI's fluency amounted to moral authority.

The same structure kept appearing across different examples. A person raises a problem, the system records a narrower version, time passes, and the eventual correction reaches the file after the relevant future has already changed. That recurrence became the centre of the book.

Personal experience also shaped the positive side. I have lived with depression and anhedonia for a long time, so optimism could not provide a serious foundation for hope. A useful account had to describe conditions that remain real even when the feeling of hope is absent: a route that can still be used, somebody able to maintain it, evidence capable of showing failure, and enough freedom to stop pretending when the route has closed.

AI systems assisted with memory, organisation, drafting, criticism, and production. They also repeatedly produced prose that was smooth, repetitive, falsely profound, or more certain than the evidence allowed. The work improved when I stopped treating fluency as quality and returned each claim to the ordinary human situation it was supposed to explain. The authorship, moral judgement, final choices, and remaining mistakes are mine.

Earlier drafts required too much private language before a reader could reach the problem. This version keeps Mechanical Ethics as the human account and leaves TRACE as the structural language developing beside it. The book is intended to be clear enough for a reader to find its weak points without first accepting its vocabulary or importance.

I expect parts of this account to fail under criticism. That is why the book returns each claim to ordinary cases and visible consequences: a reader should be able to show where the structure breaks.

Appendix A - Structural Notes

The argument begins from the middle rather than from a completed moral theory.

Something exists. It has a state. It is not the whole world. It registers some things and not others. It acts or is acted upon. Time passes. A transition occurs. Some futures remain reachable and others close. Other entities may be affected. Correction may or may not arrive before the closure hardens.

The smallest current primitive layer is:

- **state** - how things are at a location and time;
- **transition** - movement from one state to another, including changes in cost and availability;
- **time** - ordering, duration, decay, deadlines, and hardening;
- **coupling** - the state of one thing conditions the transition of another;
- **aperture** - what an entity or system can register, when, and through which instruments;
- **entity** - a bounded persistent pattern, with the warning that drawing a boundary is itself power-relevant.

Two inputs must remain visible:

- **designation** - who or what counts, and in what role;
- **measure** - how futures, losses, baselines, and weights are compared.

Mechanical Ethics leaves these inputs open rather than filling them automatically. Their authorship must remain visible, especially where one party holds more power.

A compact notation can record the limited viability claim. Let $R_H(s)$ denote the continuations reachable from state s over a stated horizon H under the actions then available. Let C denote an explicit set of contestable filters, including the affected entities, protected paths, relevant safety and access thresholds, requirements for challenge and correction, and rules governing burden and residue. Define:

$$V_{C,H}(s) = \{\tau \text{ in } R_H(s) \mid \tau \text{ satisfies } C \text{ throughout } H\}$$

and:

$$K_{C,H} = \{s \mid V_{C,H}(s) \neq \text{empty set}\}$$

The set $K_{C,H}$ contains the states from which at least one continuation remains available under those stated filters and for that horizon. It selects no best path, assigns no purpose, and measures no life's worth. It inherits every uncertainty and normative decision in R_H , C , and H . Changing who counts, what is protected, how safety or affordability is judged, which evidence is accepted, or how long the horizon extends can change the set. These inputs are authored choices, not discoveries made by the notation.

The primary timing condition is:

$$T_{\text{protect}}(a) < T_{\text{harden}}(a)$$

Effective protection must be in place before the threatened path hardens beyond comparable repair. Both terms may be uncertain and may change when the available actions include a pause, hold, temporary repair, evidence preservation, or deployment limit.

The familiar correction-window expression remains a conceptual compression:

$$T_{\text{det}} + T_{\text{route}} + T_{\text{corr}} < T_{\text{irr}}$$

It separates detection, routing, and correction for inspection, but it does not assert that they occur only in series or that the values are exact point estimates.

The doorway test remains in prose: the affected person or entity must be able to find, understand, afford, use, and integrate the route before the relevant opportunity closes. No separate doorway equation is claimed in this reader.

A further structural question asks what happens along the default trajectory:

What closes if ordinary process continues, nobody intervenes, review arrives late, or the system waits?

These tests expose several common errors:

- correction is not equivalent whenever it arrives;
- a written route is not necessarily a usable route;
- uncertainty is not neutral when one side controls the clock;
- inaction is not harmless merely because nobody pressed a button;
- a record of loss is not proof of repair.

Appendix B - What Remains Unfinished

The argument remains unfinished in several places. The following questions and pressure points mark where later work may extend, test, or correct it.

Known unfinished edges

Collective authority and capture. Several records can reveal a shared mechanism, but representation can still expose people, flatten disagreement, or become another gate. The book does not supply a complete governance structure for collective voice.

The never-built door. Some harms do not arrive as refusals. A listing is never shown, a price is personalised upward, or an application is discouraged before it exists. No individual witness pack can easily record an option the person never knew had been removed.

The insider witness. Caseworkers, engineers, auditors, clinicians, moderators, and administrators may see the system from inside while facing employment, legal, and retaliation risks. The book does not yet provide a complete protection structure for them.

Residue triage. Keeping unresolved debts visible is not enough when repair capacity is finite. The ordering of repair can become another hidden exercise of power.

Clock authorship. Complaint windows, limitation periods, retention rules, appeal clocks, and model update cycles are designed. Their authorship and beneficiaries require fuller treatment.

Competing clocks and priority. Several protected paths may harden at once. The timing condition exposes the conflict but does not supply a universal priority rule.

The despair test. A truthful book about false doors and late help may assist one reader and overwhelm another. Honesty does not remove a duty of care.

Questions left open

Necessity and closure. The actor who benefits from closure cannot be its sole judge, but no universal allocation method is supplied here.

Stable harm regimes. Truth may be visible while power accepts the damage. Witness, interruption, protection, enforcement, and organised resistance require further separation.

Machine velocity. The brake principle is clear, but thresholds for slowing, staging, rollback, and interruption remain specific to each domain.

Protected-scope uncertainty. Missing or disputed standing must not become automatic permission for the actor who benefits from exclusion.

Positive good. The framework supplies only a thin account of conditions for usable agency. Its filters remain contestable, and it offers no complete account of what makes a future worthwhile.

Appendix C - Status, Sources, and the Two-Artifact Pair

Reader status

This is a working human reader of Mechanical Ethics. It is not a doctrine, moral authority, legal standard, certification system, or final account. Earlier versions and production records are preserved separately.

This reader is deliberately selective. A separately versioned working-source corpus preserves broader Mechanical Ethics development and audit history. That corpus is not canon or a higher moral authority, and this reader is not its subordinate summary. Material may be developed in one without appearing in the other; absence here does not by itself mean abandonment. Their version numbers belong to different lineages and should not be compared as one sequence.

Figures

The reader contains four explanatory diagrams: *The file and the life*, *The correction window*, *Two flats, one wall*, and *Machine speed versus human correction*. They clarify the argument but do not serve as evidence for it.

Scene and evidence discipline

Named composite scenes are invented from recurring real-world structures. They are not reports about named people and do not serve as empirical proof. Historical notes are identified separately. Interpretation and extrapolation are distinguished from the recorded facts.

Historical sources

Frances Oldham Kelsey and thalidomide: United States Food and Drug Administration, *Frances Oldham Kelsey: Medical reviewer famous for averting a public health tragedy*; FDA historical and autobiographical materials describing the 1960 application, repeated evidence requests, and the authority to prevent the application taking effect while it remained incomplete.

Space Shuttle Challenger: *Report of the Presidential Commission on the Space Shuttle Challenger Accident* (Rogers Commission), especially Chapter 5 on the launch

decision and the Commission's findings on the information unavailable to final decision-makers; NASA historical material on Challenger and return to flight.

The historical notes are compressed illustrations. Readers should consult the primary reports for the full factual and institutional record.

Human authorship and AI assistance

The project, moral witness, and authorship belong to Mark. AI systems assisted with conversation, memory, transcription, structural mapping, drafting, editing, adversarial review, and file production. AI collaborators proposed and criticised structural choices; Mark decided which changes entered the book and remains responsible for the resulting text. AI agreement does not validate the work.

Mechanical Ethics and TRACE

Mechanical Ethics carries the human meaning and moral argument. TRACE carries a more compact structural language for systems. They are meant to remain separable: a reader should not need TRACE to understand the human argument, and TRACE should not borrow moral authority from the reader. Neither validates the other.