

Witness Over Spectacle: A Covenant for Reading in the Centaur Era

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

December 26, 2025

A century of broadcast and a decade of algorithmic feeds have quietly taught a metaphysics: reality is what is seen, meaning is what spikes quickly, and truth is what repeats. In that environment, “free TV” is never free; it is funded by attention capture, emotional escalation, and the conversion of persons and events into consumable episodes. The result is not merely distraction. It is a deformation of the reader’s posture toward reality—an easy slide from witness to voyeur, from understanding to stimulation, from truth-seeking to identity theater. This paper argues for a metaphysical sea change that does not depend on new laws, better platforms, or moral scolding. It proposes a practical, portable alternative: a covenant for reading—short, explicit commitments that re-order attention as a moral act and make auditability feel sacred rather than bureaucratic. The covenant is designed for “centaur” cognition: human–AI collaboration, high-volume text production, and a world where persuasive synthesis can arrive faster than conscience. The goal is simple: restore slow authority, preserve personhood, and make truth-seeking resilient against spectacle, hedonism, and propaganda.

Keywords: witness, voyeurism, spectacle, attention economy, FAST TV, propaganda, hedonism, auditability, epistemic humility, no-oracle rule, centaur cognition, reader covenant, formation, disconfirmation, provenance, uncertainty discipline, moral psychology, media ecology, truth posture, Logostics.

A collaboration with GPT-5.2-Thinking. 59 pages. CC4.0.

Outline

1. The Problem Statement: The Broadcast Century as Metaphysical Training

- 1.1 The normalization of “looking” as a civic habit
- 1.2 From public information to private arousal engineering
- 1.3 Why “free” channels are a new kind of contract
- 1.4 The hidden curriculum: what repeated exposure forms in the soul

2. Definitions That Stop the Slippage

- 2.1 Voyeurism as consumption of proximity
- 2.2 Witness as accountable seeing and speech
- 2.3 Hedonism as governance by immediate appetite
- 2.4 Propaganda as salience control, not only state messaging
- 2.5 Spectacle as an attention technology that substitutes for meaning

3. The Coupled System: How the Four Reinforce Each Other

- 3.1 How “free TV” finances hedonism
- 3.2 How hedonism recruits voyeurism
- 3.3 How voyeurism lowers resistance to propaganda
- 3.4 How propaganda stabilizes the machine by narrating the machine
- 3.5 The escalation ladder: why “ordinary” stops working

4. The Metaphysical Sea Change

- 4.1 Two rival hierarchies of goods: stimulation-first vs truth-first
- 4.2 Attention as moral act rather than neutral resource
- 4.3 Personhood as non-consumable category
- 4.4 Slow authority: why auditability is a spiritual virtue in public life
- 4.5 The cost of speech: reclaiming responsibility for what we repeat

5. From Information to Formation

- 5.1 Why data does not cure deformation
- 5.2 Formation primitives: discipline, test, fruit
- 5.3 The difference between being informed and becoming able
- 5.4 The role of restraint: choosing not to watch as an ethical competence
- 5.5 How formation scales without turning into moralism

- 6. A Covenant for Reading in the Centaur Era
 - 6.1 The covenant's purpose and design constraints
 - 6.2 The witness rule: no spectacle, no humiliation harvest
 - 6.3 The audit rule: provenance, uncertainty, disagreement, disconfirmation
 - 6.4 The no-oracle rule: non-delegation of conscience to models or feeds
 - 6.5 The lag rule: replacing "live" with "later" for tragedy and conflict
 - 6.6 The fruit rule: long-horizon outcomes over virality and heat
 - 6.7 The dignity rule: persons remain persons, even when wrong
- 7. Implementation as a Portable Practice
 - 7.1 One-page front matter for papers, newsletters, and videos
 - 7.2 Checklists that do not infantilize the reader
 - 7.3 Community norms: how to enforce without coercion
 - 7.4 Personal protocols: attention fasting, re-entry rules, stop rules
 - 7.5 Institutional protocols: editors, curators, and audit trails
- 8. Centaur Dynamics: Human-AI Collaboration Under the Covenant
 - 8.1 When synthesis becomes too easy: persuasion without responsibility
 - 8.2 "Model as witness" vs "model as idol"
 - 8.3 Guardrails that preserve agency: prompts, logs, and downgrade conditions
 - 8.4 Disconfirmation at scale: testing claims in high-throughput writing
 - 8.5 Avoiding synthetic authority: citation discipline and humility parameters
- 9. Objections and Failure Modes (Make Them Strong)
 - 9.1 "This is prudish or anti-media"
 - 9.2 "People have always enjoyed spectacle"
 - 9.3 "Auditability is cold and inhuman"
 - 9.4 "Covenants become performative badges"
 - 9.5 "This can't compete with the feed"
 - 9.6 "Covenant language is religious coercion"
 - 9.7 Failure modes: legalism, elitism, paranoia, and purity spirals
- 10. Case Studies and Thought Experiments
 - 10.1 The tragedy clip: what must be shown vs what must not be replayed
 - 10.2 The tourist wreck: history, memory, and the selfie temptation
 - 10.3 The "free channel" interface: how UI scripts metaphysics
 - 10.4 The viral claim: how the covenant handles uncertainty and disagreement
 - 10.5 A good-faith propaganda test: distinguishing persuasion from manipulation

11. The Practical Payoff: What Changes If the Covenant Takes Root

11.1 A quieter public square with stronger claims

11.2 Reduced susceptibility to engineered outrage

11.3 Restored dignity for victims and opponents

11.4 A new norm for AI-assisted discourse: audit first, eloquence second

11.5 Why this is compatible with joy, art, and legitimate entertainment

12. Conclusion: The Charge of Witnesshood

12.1 Restating the thesis in one paragraph

12.2 The minimal ask: one covenant, practiced daily

12.3 The sea change: attention re-ordered toward truth and love

12.4 Closing charge: build witnesses, refuse spectacle, document everything

References

Art

1. The Problem Statement: The Broadcast Century as Metaphysical Training

The claim of this paper is not that radio and television were “bad,” nor that the internet uniquely corrupted an otherwise healthy public square. The claim is more surgical: a century of broadcast, followed by algorithmic distribution, has trained millions of people in a largely unspoken metaphysics. It has taught a posture toward reality—what counts as real, what counts as important, what deserves attention, and what kind of person you become by paying attention.

This training did not arrive as a philosophy course. It arrived as habit. Habit becomes reflex. Reflex becomes character. Character becomes the default “world” within which evidence, argument, compassion, and even worship are interpreted. When people say, “It’s just TV,” they often mean “it has no authority over me.” But what forms us most reliably is not what we consciously submit to; it is what we repeatedly rehearse without noticing. The broadcast century made “looking” effortless, constant, and socially normal. The feed century made “looking” personalized, optimized, and financially compulsory.

In that environment, the attention economy is not merely a market. It is a formation system. It trains a public to become a certain kind of perceiver: fast, reactive, appetite-led, emotionally programmable, and increasingly tolerant of the conversion of persons into content. Calling this “voyeurism” is not moral panic; it is a descriptive claim about what the medium rewards and what the audience becomes.

1.1 The normalization of “looking” as a civic habit

Broadcast changed “the public” into an audience. In older civic models, public life was something you participated in—town meetings, churches, clubs, neighborhood disputes, and local institutions where your speech and reputation were attached to your person. Broadcast introduced a new civic role: the spectator who is present without being present, involved without obligation, informed without accountability.

That role has benefits. It can widen awareness beyond one’s village. It can warn, educate, and inspire. But it also introduces a subtle moral shift: the center of gravity moves from action to observation. And observation, when made effortless and constant, becomes a default substitute for responsibility. The viewer learns to feel “I have done something” simply by watching, reacting, and sharing. The body remains still; the emotions move; the conscience feels engaged. This is not “fake.” It is simply incomplete. It is participation without cost.

Over time, the medium trains the appetite for sights that break ordinary life open—war, disaster, scandal, sexualized imagery, humiliations, confrontations. Not because viewers are uniquely wicked, but because novelty and intensity are the easiest way to keep attention. The camera arrives, the world becomes a stage, and the audience learns a basic rule: what is worth

seeing is what is extreme. The mundane becomes invisible, even though the mundane is where most moral life actually happens.

Once “looking” is normalized as a civic habit, the public square becomes a theater with a constant demand for scenes. The question “What should we do?” is gradually displaced by “What happened?” and then by “Show me.” The viewer becomes a connoisseur of moments rather than a steward of reality. This is the seedbed of voyeurism: not lust for sex alone, but lust for access—access to other people’s worst day, other people’s private collapse, other people’s fear, other people’s shame, and other people’s death.

1.2 From public information to private arousal engineering

Early broadcast largely operated as one-to-many. The same signal went to everyone. That structure carried at least one restraint: if you had to program for a whole population, you could not perfectly tailor content to each person’s triggers. The viewer might be entertained, but the medium could not easily adapt itself to the private nervous system of each individual.

The feed era removed that restraint. Content became selectable, recommendable, and optimizable. The medium stopped being merely a channel and became a mirror that learns. It learns what you pause on, what you replay, what you finish, what you skip, what makes you angry, what makes you afraid, and what makes you stay. “Public information” is no longer the primary product. The primary product is retention—and retention is easiest when content is tuned to arousal: outrage, fear, lust, moral disgust, superiority, and tribal belonging.

This is why “propaganda” becomes too small a word if it is restricted to governments. When any system optimizes attention by steering emotion, the line between persuasion and manipulation becomes structural, not occasional. The system need not “believe” anything. It only needs to keep you watching. But keeping you watching has predictable psychological requirements. The result is a private theater that is curated to your vulnerabilities. It is not merely that you see “more.” It is that you are trained into a patterned emotional life that you did not explicitly choose.

This is where hedonism becomes more than a personal vice. It becomes a governance principle embedded in the architecture of information. The self is treated as an appetite engine. The content stream becomes a sequence of small rewards. The citizen is reframed as a consumer of sensation. Once this shift is established, truth is no longer the organizing value; stimulation is. Truth may appear, but it appears as one ingredient among others—and often as a costume worn by more profitable emotions.

1.3 Why “free” channels are a new kind of contract

The phrase “free channels” sounds like the old promise of broadcast: turn it on and you get programming at no direct cost. But “free” in the feed era is not the old civic bargain; it is a new contract with new terms.

The old broadcast bargain was roughly: advertisers fund programming to reach a public audience. The new bargain is closer to: advertisers fund programming to reach a *measured, profiled, targetable* audience, and the channel lineup exists to keep the measurement flowing. The payment is not money; the payment is attention plus the behavioral exhaust of your attention. Even when no one consciously thinks in those terms, that is the economic logic. And economic logic shapes editorial logic. Editorial logic shapes what is repeatedly shown. Repetition shapes what feels normal. What feels normal becomes “reality.”

This is why “free channels” deserve metaphysical scrutiny. They are not merely content offerings. They are a default environment offered at the moment of device setup, when habits are easiest to form. They arrive as pre-installed “common sense.” The contract is signed not with a pen but with a routine: you turn it on, it starts playing, it suggests the next thing, and soon your evenings are no longer chosen but allocated.

In such a contract, the viewer is not just the audience; the viewer is the commodity being delivered. That does not mean every program is evil, nor that every viewer is deceived. It means the environment has a built-in direction: toward more watching, more arousal, more personalization, and less reflective agency. If you want a metaphysical sea change, you have to name this contract honestly—because what is unnamed cannot be resisted.

1.4 The hidden curriculum: what repeated exposure forms in the soul

The most important effects of the broadcast/feed century are not informational. They are formational. They are the hidden curriculum: the lessons the medium teaches about reality, persons, and speech.

Repeated exposure trains the soul (and here “soul” can be read as conscience, character, or the deep self) in at least five ways.

First, it trains the assumption that seeing is a kind of possession. If I can watch it, I have access to it; if I have access to it, I am entitled to it. This is the root of voyeurism: the sense that other people’s privacy, pain, and catastrophe are “available” because they are visible.

Second, it trains emotional immediacy as a substitute for moral discernment. The system rewards fast feeling. Fast feeling displaces slow judgment. Over time, people come to trust the intensity of their reaction more than the integrity of their reasoning. This is a profound metaphysical inversion: truth becomes what moves me, not what withstands testing.

Third, it trains the reduction of persons into roles: victim, villain, hero, idiot, celebrity, scapegoat. Roles are simpler than persons. Roles are easier to monetize than persons. But when persons are reduced to roles, dignity erodes. And when dignity erodes, cruelty becomes plausible, because the target is no longer “a person like me” but a character in a story that I am permitted to hate.

Fourth, it trains the appetite for escalation. Ordinary life becomes flat. Normal conversations feel boring. Subtlety becomes a tax. One needs hotter takes, sharper edits, more extreme scenes. This is how hedonism becomes the operating system: not constant pleasure, but constant stimulation as the condition for staying present.

Fifth, it trains speech to become performance. In a mediated environment, the reward is not careful truth but social display—wit, dominance, ridicule, certainty, tribal cues. The self learns to speak for applause rather than for accuracy. And this is where propaganda thrives, because propaganda is not just a set of claims. It is a style of speech that rewards loyalty signals over reality contact.

If that is the hidden curriculum, then the “problem statement” becomes clear: we are not merely facing misinformation. We are facing mis-formation. We are facing a civilization trained to look without covenant, to feel without testing, to speak without cost, and to treat attention as morally neutral. A metaphysical sea change must therefore begin where formation begins: with a new discipline of attention, a new dignity for persons, and a renewed sense that witness is accountable seeing—not consumable access.

2. Definitions That Stop the Slippage

The most destructive drift in modern discourse is not disagreement; it is category confusion. We use the same words to name different things, and then we argue past each other while the machine keeps running. “News,” “entertainment,” “documentation,” “awareness,” “exposure,” “truth,” “propaganda,” “free,” and even “witness” collapse into one big undifferentiated stream. When that happens, the attention economy wins automatically, because it can rebrand itself as whatever the moment requires.

This section does not aim to be exhaustive or academic. It aims to be functional. Each definition is shaped as a tool: something you can hold in the hand when a clip goes viral, when a tragedy is packaged as content, when a “free channel” autoplays, when a feed invites you to stare, and when the public square becomes a theater. The goal is to stop slippage: the subtle moral substitution by which we call voyeurism “being informed,” call hedonism “self-care,” call propaganda “just my side,” call spectacle “raising awareness,” and call witness “I saw it with my own eyes.”

These definitions also have a discipline built into them: they separate the act (what is done), the aim (why it is done), and the effect (what it forms in the doer). The same footage can be witness in one setting and voyeurism in another. The same persuasion can be legitimate in one setting and propaganda in another. The difference is not only content; it is posture, contract, and consequence.

2.1 Voyeurism as consumption of proximity

Voyeurism is not limited to sex. Sexual voyeurism is simply the clearest case because it reveals the structure in high contrast. The broader definition is this: voyeurism is the consumption of proximity to another person's intimacy, vulnerability, suffering, humiliation, or catastrophe, for the sake of stimulation, curiosity, or identity reinforcement rather than responsibility.

Three elements matter.

First, voyeurism seeks access without obligation. It wants the closeness of "being there" without the cost of being a neighbor. It wants the emotional charge of another person's worst day without the duties that closeness would normally impose (help, restraint, privacy, compassion, accuracy).

Second, voyeurism turns persons into surfaces. The person becomes a scene, an object of inspection, a delivery system for a feeling. Once that conversion happens, the viewer is no longer oriented toward the person as a moral equal; the viewer is oriented toward the experience the person provides.

Third, voyeurism is repetitive. It is not "I needed to see this once to understand." It is "I rewatch, zoom, freeze-frame, slow-motion, speculate, share, and discuss as entertainment." This is why the phrase "clearest view yet" is often a warning sign. Clarity is morally neutral; the use of clarity is not. The moment of impact, the moment of collapse, the moment of death, the moment of humiliation—these are not "content upgrades." They are thresholds where dignity demands restraint.

A practical test: voyeurism increases consumption while decreasing responsibility. If watching makes you more excitable, more certain, more hungry for more footage, and less careful about the person's dignity and the facts, you are in voyeurism territory even if you tell yourself you are "staying informed."

2.2 Witness as accountable seeing and speech

Witness is not just seeing. Witness is seeing under obligation. It is the posture that treats perception as a moral act and speech as a form of accountability.

Witness has four core features.

First, witness is bounded by necessity. A witness shows what must be shown to understand and to act responsibly. A witness does not maximize intimacy. A witness does not harvest the private. A witness does not replay the wound to extract more emotion. Witness asks: what is required for truth, justice, safety, repair, or remembrance? Everything beyond that is suspect.

Second, witness carries a cost. The witness accepts responsibility for accuracy, for restraint, and for the consequences of repeating a claim. This is why witness is naturally allied to auditability: provenance, uncertainty labeling, willingness to correct, and openness to disconfirmation.

Third, witness preserves personhood. A witness does not reduce the subject to a role. Even when the subject is guilty, foolish, violent, or deceived, witness refuses dehumanization. The witness may condemn an act and still refuse to consume a person as a spectacle.

Fourth, witness aims toward action that is fitting. Witness is ordered toward repair: informing appropriate authorities, supporting victims, preventing recurrence, clarifying truth, or building a more truthful public record. If a piece of “witnessing” does not cash out into anything beyond “look at this,” it is likely slipping toward spectacle.

A practical test: witness decreases consumption while increasing responsibility. After exposure, you become slower, more careful, more humble about what you know, more protective of dignity, and more oriented toward what should be done.

2.3 Hedonism as governance by immediate appetite

Hedonism here is not simply pleasure, and it is not a puritanical rejection of joy. Pleasure is part of creaturely life. Joy can be holy. Rest can be obedience. The definition in this paper is narrower and more diagnostic: hedonism is governance by immediate appetite—when the highest practical authority in the self is “what I want right now.”

This governance can be emotional rather than sensual. Many modern forms of hedonism are hedonisms of affect: the desire for outrage, the desire for righteous superiority, the desire for fear relief, the desire for the warm bath of belonging, the desire to punish an enemy with a share button, the desire to feel informed without the labor of understanding.

Hedonism operates like a moral shortcut. It replaces “what is true?” and “what is right?” with “what feels good (or intense) now?” Over time, it trains the self to treat discomfort as illegitimate, delay as oppression, and restraint as loss. It makes the soul brittle. Anything that cannot be consumed quickly becomes “boring,” and boredom becomes a reason to abandon reality contact.

In a media ecosystem, hedonism is particularly profitable because it is predictable. Immediate appetites can be measured, targeted, and escalated. This is why the system naturally drifts toward arousal engineering. If you govern yourself by immediate appetite, the feed can govern you by the same principle.

A practical test: in hedonism, the self's steering wheel is pulled forward into the present moment. It becomes hard to wait, hard to verify, hard to read long-form, hard to confess uncertainty, hard to sit with ambiguity, hard to do the unglamorous work of truth.

2.4 Propaganda as salience control, not only state messaging

If “propaganda” is defined only as state lies, modern systems will evade the word easily. The better definition is structural: propaganda is the shaping of beliefs and attitudes by controlling salience, repetition, and emotional framing in ways that bypass independent judgment.

This definition includes governments, but it also includes parties, corporations, movements, influencers, and even informal tribes. The key is not who speaks; the key is how the environment is arranged.

Propaganda works less by inventing claims than by arranging attention. It decides what you see, how often you see it, what emotional key it is played in, what alternatives are ridiculed, what questions are taboo, and what kind of person you become if you dissent. It is a management of the plausible. It shrinks the space in which independent thought can breathe.

In the feed era, propaganda can be produced without a central propagandist. An ecosystem can converge on propaganda-like effects because the attention market rewards what mobilizes. Mobilization requires emotional certainty, simplified enemies, and rapid narratives. The platform does not need a political agenda to become a propaganda engine; it only needs a retention agenda. Retention naturally selects for the kind of content that propaganda has always used: fear, disgust, tribal identity, and moral urgency.

A practical test: propaganda reduces the reader's freedom to revise. If consuming a channel, movement, or “free TV” environment makes it socially costly or psychologically difficult to say “I was wrong,” “I don't know,” “I need more evidence,” or “the other side has a point,” you are inside a salience-managed world.

2.5 Spectacle as an attention technology that substitutes for meaning

Spectacle is not merely “something dramatic.” Spectacle is a technology of attention that substitutes for meaning. It creates the feeling of significance without the substance of understanding.

Spectacle has a distinct signature.

First, it is optimized for viewing, not for comprehension. It privileges what can be shown over what must be explained. It prefers images to arguments, conflict to context, immediacy to causality.

Second, it creates pseudo-participation. The viewer feels involved by watching. The system offers moral emotions as a replacement for moral action. Outrage becomes a sacrament. The share button becomes the altar. The viewer leaves with the feeling of having “taken a stand” without having taken a risk.

Third, spectacle eats time while producing little wisdom. It exhausts attention and leaves the mind with fragments: a face, a clip, a slogan, a villain, a vibe. Wisdom requires integration, patience, and proportion. Spectacle delivers intensity and then moves on.

Fourth, spectacle converts suffering into currency. It does not necessarily deny suffering. It often amplifies it. But it amplifies it as a commodity, and commodity logic is indifferent to dignity.

This is why spectacle is the great solvent of witness. A single event can begin as witness and become spectacle through repetition, editing, monetization, and “reaction” culture. The same footage that might be necessary for accountability becomes unnecessary for dignity once the basic facts are established. Yet spectacle insists on replay, commentary, hot takes, and emotional extraction.

A practical test: spectacle makes you feel full while leaving you empty. You have “seen so much,” yet you know little. You have “felt so strongly,” yet you have not become more truthful, more humble, more compassionate, or more capable of repair.

These definitions are meant to be portable. When something goes viral, you can ask: am I consuming proximity, or bearing witness? Is my appetite governing me, or is truth governing me? Is this salience being managed to close my freedom to revise? Is the spectacle giving me significance without meaning? If the definitions can be used in real time, they have done their job: they stop the slippage before it becomes a habit, and before habit becomes a metaphysics.

3. The Coupled System: How the Four Reinforce Each Other

If voyeurism, “free TV,” hedonism, and propaganda were separate vices, the remedy would be straightforward: turn off the channel, cultivate better habits, and move on. The harder truth is that these four often form a coupled system, a stable loop in which each element strengthens the others. That loop does not require a mastermind. It can emerge from incentives,

measurement, and habituation. Once it emerges, it becomes self-protecting: it defends itself by making alternatives feel boring, naïve, or morally suspect.

A coupled system matters because it changes how you fight it. You cannot solve a coupled system by attacking one node while feeding the others. You cannot “be less voyeuristic” while letting hedonism govern attention. You cannot “avoid propaganda” while living inside an environment that continuously tunes arousal. You cannot “enjoy free TV responsibly” while allowing the medium to set the default posture of looking without obligation. The system is recursive: it trains the desire that finances the platform that delivers the content that trains the desire.

This section maps the loop in plain terms. The aim is not cynicism; it is clarity. Once you can see the mechanism, you can design a counter-practice that is strong where the loop is strong: at defaults, at habits, at repetition, and at the management of salience.

3.1 How “free TV” finances hedonism

“Free TV” finances hedonism by making stimulation the economic center of gravity. In a traditional paid model, revenue is at least partly tied to satisfaction, quality, or loyalty. In an ad-supported, measurement-rich model, revenue is tied to attention volume and attention persistence. Attention persistence is most reliably purchased with arousal. This is not a moral judgment; it is a behavioral observation. People stay longer when the content makes them feel something strong and quickly.

Once “free” is funded by ads and measurement, the environment develops four pressures that tilt toward hedonism.

First, the pressure toward immediacy. The system must hook you fast, because attention is scarce and easily interrupted. This favors content that triggers immediate appetite: curiosity, outrage, fear, lust, or tribal affirmation. Slow-building arguments and quiet nuance are penalized.

Second, the pressure toward personalization. The more the system learns your triggers, the more it can feed you content that governs you by immediate appetite. Even without explicit malice, the machine discovers the shortest path to “keep watching,” and “keep watching” is often the same as “keep craving.”

Third, the pressure toward abundance. “Hundreds of free channels” sounds like generosity. But abundance functions like a casino floor: it reduces the felt cost of switching and increases the incentives for the platform to compete by increasing stimulus intensity. Where there is abundance, the easiest differentiator is the emotional spike.

Fourth, the pressure toward continuous play. Autoplay, previews, “up next,” and default launchers are not neutral conveniences. They are an architecture of momentum. Momentum is governance. Momentum hands the steering wheel to the stream. And when the stream governs, immediate appetite becomes law.

The result is hedonism not as a personal moral failing but as a platform-level operating principle: the system runs best when viewers are governable by quick reward.

3.2 How hedonism recruits voyeurism

Hedonism recruits voyeurism because appetite does not remain content with the ordinary. An appetite-governed self wants stronger signals over time. The same content that once delivered a charge begins to feel flat. This is not because the person becomes “worse” in a dramatic sense; it is because the nervous system habituates. Habituation is a law of sensation. It pushes the viewer toward escalation.

Voyeurism is one of the cheapest escalations available. It provides intensity without requiring artistry, deep insight, or long production cycles. It can be captured with a phone. It can be framed as “exclusive.” It can be sold as “real.” It can be delivered as a short clip. It carries the thrill of forbidden access: the sense that you are seeing what you are not supposed to see.

Hedonism also recruits voyeurism through a second mechanism: it turns persons into instruments. If immediate appetite governs you, other people become means to your stimulation. Their suffering becomes “content.” Their humiliation becomes “drama.” Their death becomes “impact footage.” Their private tears become “raw authenticity.” The viewer can still claim compassion, but compassion becomes a feeling harvested from another person’s pain rather than a responsibility undertaken for another person’s good.

A third mechanism is moral camouflage. Voyeurism rarely presents itself as voyeurism. It presents itself as “awareness,” “keeping up,” “bearing witness,” “seeing the truth,” “watching what the mainstream hides,” or “staying vigilant.” These frames allow appetite to dress itself in virtue. The more the self is governed by appetite, the more it needs moral cover, because no one wants to think of themselves as a consumer of harm.

So hedonism recruits voyeurism by making the ordinary boring, by turning persons into means, and by offering virtue-flavored labels that protect the self-image while intensifying consumption.

3.3 How voyeurism lowers resistance to propaganda

Voyeurism lowers resistance to propaganda by training the mind to prefer proximity over verification and intensity over proportion. Propaganda does not need you to be stupid. It needs you to be aroused, hurried, and identity-locked. Voyeurism supplies exactly that posture.

First, voyeurism teaches “seeing is believing.” A clip feels like proof. A vivid scene feels like reality itself. But a clip is always a frame, and a frame is always selective. When viewers are trained to treat proximity as truth, they become easier to steer with curated proximity. Propaganda thrives on curated images that bypass slow reasoning.

Second, voyeurism trains emotional certainty. When you consume high-intensity scenes, your nervous system produces strong conclusions quickly. The mind confuses emotional clarity with factual clarity. Propaganda exploits this: it supplies scenes that deliver a conclusion without argument.

Third, voyeurism trains dehumanization through role reduction. If you repeatedly consume people as “villains,” “idiots,” “threats,” or “monsters,” you become more tolerant of narratives that justify cruelty, censorship, or scapegoating. Propaganda loves simplified enemies. Voyeurism provides the raw emotional footage needed to keep the enemy vivid.

Fourth, voyeurism fragments context. Short clips break causal chains. They isolate a moment from a history. Fragmentation makes it difficult to build a stable model of reality. In that vacuum, propaganda offers ready-made models: simple stories with clear heroes and villains. The more fragmented your perception, the more you crave narrative closure, and the more willing you are to accept closure that feels satisfying.

So voyeurism lowers resistance to propaganda by producing a mind that is vivid, fast, certain, and context-poor.

3.4 How propaganda stabilizes the machine by narrating the machine

A system that feeds on arousal must defend itself against the suspicion that it is feeding on arousal. It must also defend itself against the rising fatigue and shame that eventually accompany compulsive consumption. Propaganda stabilizes the machine by providing narratives that justify the machine.

This is one of the most important couplings: propaganda does not only push a political line; it supplies moral stories that make consumption feel righteous.

It tells the viewer: “You are not watching; you are resisting.”

It tells the viewer: “You are not addicted; you are awake.”

It tells the viewer: “You are not being manipulated; you are the few who see.”

It tells the viewer: “Your anger is proof of your virtue.”

It tells the viewer: “Anyone who questions this stream is part of the lie.”

These narratives create an enclosure. They turn doubt into betrayal. They turn boredom with the feed into weakness. They turn restraint into complicity. They turn slow verification into naivete. They immunize the system against critique by making critique psychologically costly.

Propaganda also stabilizes the machine by redefining its outputs as “truth.” If the feed is framed as the only honest source, then every sensational clip becomes sacred evidence and every correction becomes suppression. In that world, “auditing” looks like collaboration with enemies. “Humility” looks like cowardice. “Waiting for facts” looks like surrender. The machine becomes self-sealing.

This is why propaganda is not merely content; it is an anti-exit mechanism. It keeps the viewer inside the loop by providing an identity that requires the loop.

3.5 The escalation ladder: why “ordinary” stops working

The coupled system has a predictable trajectory: escalation. Ordinary life stops working as a stimulus. The viewer’s baseline shifts. The medium must raise intensity to maintain engagement. This is not a conspiracy; it is the natural consequence of habituation plus competition.

The escalation ladder often moves in stages.

First stage: novelty. The viewer wants new topics, new faces, new formats.

Second stage: intensity. The viewer wants conflict, outrage, fear, scandal, humiliation.

Third stage: intimacy. The viewer wants “raw,” “unfiltered,” private moments, behind-the-scenes access, the collapse caught on camera.

Fourth stage: certainty. The viewer wants narratives that remove ambiguity and deliver moral clarity fast.

Fifth stage: identity lock. The viewer’s consumption becomes part of who they are. The feed becomes a homeland. Leaving feels like exile.

Sixth stage: appetite inversion. The viewer begins to seek not only pleasure but the pleasure of outrage, the satisfaction of hatred, the relief of fear confirmed, the comfort of enemies explained. Negative affect becomes addictive because it provides meaning and belonging.

At this point, ordinary conversation, ordinary reading, ordinary prayer, ordinary beauty, and ordinary duty feel thin. The soul becomes dependent on a high-arousal stream to feel alive. This is the deepest tragedy of the system: it does not merely misinform; it makes reality feel inadequate.

This is why the remedy must be more than “fact-checking.” You cannot correct a formation system with a spreadsheet. You need counter-formation: practices that lower the baseline, restore the capacity for slowness, rebuild the pleasure of the ordinary, and reattach seeing to obligation. Only then can propaganda lose its hold, because propaganda thrives where boredom is intolerable and ambiguity is unbearable.

The coupled system, then, is the enemy of witness. It turns watching into a moral substitute, appetite into a governor, persons into content, and narratives into cages. Seeing this coupling is the first step toward breaking it—not by moral grandstanding, but by changing defaults, restoring covenant, and relearning what it means to look at the world as a steward rather than as a consumer.

4. The Metaphysical Sea Change

A “sea change” is not a new opinion; it is a new baseline. It is the moment when a society’s default assumptions about reality, value, and personhood shift so deeply that even disagreements occur on different ground. The argument here is that the broadcast/feed century did not merely produce new technologies; it produced a new metaphysical posture: stimulation-first. It trained people to treat attention as neutral, persons as consumable, and speech as costless. In that posture, the system described in Section 3 is not an aberration. It is the expected outcome.

A metaphysical sea change, then, cannot be achieved by saying “be better people” while leaving the architecture of attention untouched. It must re-order the hierarchy of goods, redefine attention as a moral act, restore personhood as inviolable, and rebuild a norm of slow authority grounded in auditability. These are not technical reforms alone. They are a moral and spiritual re-foundation of public life, including the public life that occurs through screens.

The sea change proposed here is not anti-technology. It is anti-idolatry: it refuses the substitution of speed for truth, proximity for meaning, and algorithmic convenience for conscience.

4.1 Two rival hierarchies of goods: stimulation-first vs truth-first

Every public culture has a hierarchy of goods, whether it admits it or not. A hierarchy of goods answers the question: what do we treat as most valuable in practice? Not in slogans, but in what our systems reward.

The stimulation-first hierarchy is easy to state:

- The highest good is felt intensity (pleasure, outrage, fear relief, belonging).
- Truth is a tool used when it helps intensity, and discarded when it slows intensity.
- Persons are means of producing affect and identity.
- Speech is performance, and its main aim is social positioning.
- Time is an enemy; speed is an unquestioned virtue.

This hierarchy can wear respectable clothes. It can call itself “engagement,” “awareness,” “authenticity,” “staying informed,” “fighting the lies,” “speaking truth to power.” But the operating principle remains: the next emotional spike rules.

The truth-first hierarchy reverses the order:

- The highest good is reality contact: what is true, even when it is inconvenient.
- Persons are ends, never mere means; dignity is not negotiable.
- Speech is accountable; it must answer to consequences.
- Attention is stewardship; what we look at forms us and harms or helps others.
- Time is not an enemy; slowness is sometimes the cost of honesty.

This is not an argument for joylessness. Truth-first does not forbid pleasure. It demotes pleasure from governor to guest. Pleasure becomes something received within a moral order, not something that defines the moral order.

A society cannot run both hierarchies at once. It can allow pluralism of taste, but it cannot keep stimulation-first as the default and then expect truth-first outcomes. The sea change begins by naming the competition: these are rival religions of attention.

4.2 Attention as moral act rather than neutral resource

Modern systems talk about attention as if it were money: a scarce resource to allocate efficiently. That language is not entirely wrong. But it is incomplete and therefore dangerous. Attention is not only economic; it is formative. What you repeatedly attend to becomes a kind of liturgy. It trains your reflexes of compassion, fear, disgust, curiosity, and trust. It shapes what counts as “real” and what feels “important.” Over time, attention becomes character.

Treating attention as neutral produces two false beliefs.

First, the belief that “watching doesn’t change me.” In reality, repeated exposure changes you even if you disagree with what you see. You become accustomed. You become desensitized or hyper-reactive. You learn new scripts of contempt, panic, lust, or mockery. You learn what to desire.

Second, the belief that “watching harms no one.” In reality, attention is demand. Demand funds supply. Supply shapes what is done to produce it. If humiliation pays, more humiliation will be produced. If “moment of impact” footage pays, more such footage will be sought, replayed, and monetized. Attention is not passive; it is participation in a market of meaning.

Calling attention a moral act is simply an honest description of its consequences. It means that “What do I look at?” is adjacent to “What do I worship?” because both questions ask what you give your inner life to. Even in secular terms, attention is the gateway through which values enter the nervous system. A sea change means restoring the question of attention to the realm of conscience.

This is why the covenant language matters. A covenant is not a mood. It is a binding. It says: I will not treat my attention as morally weightless. I will not let my screen be a loophole for my soul.

4.3 Personhood as non-consumable category

The attention economy’s deepest violation is not “bias” or “misinformation.” It is the conversion of persons into consumables. It trains the viewer to relate to human beings primarily as sources of sensation, moral signaling, or tribal satisfaction.

Personhood as non-consumable means something specific: a person is not “available” simply because they are visible. Their pain is not a product. Their fear is not entertainment. Their humiliation is not a learning experience for strangers. Their death is not a replayable moment. Their private life is not public property. Their repentance is not content. Even their wrongdoing, if it must be exposed, must be exposed with restraint.

This principle collides directly with the logic of voyeurism. Voyeurism says: if it can be shown, it can be consumed. The sea change says: visibility does not confer entitlement.

In practical terms, personhood as non-consumable produces norms like these:

- Do not replay suffering for emotional extraction.
- Do not zoom in on private degradation under the banner of “truth.”
- Do not treat victims as props for your moral identity.
- Do not treat enemies as cartoons to be mocked into non-personhood.
- Do not harvest someone’s worst moment as your entertainment.

This principle is not “nice.” It is stabilizing. Societies that allow people to be consumed eventually permit anyone to be consumed. When personhood becomes consumable, the public square becomes predatory. The sea change insists on a boundary that the market cannot cross.

4.4 Slow authority: why auditability is a spiritual virtue in public life

Broadcast created “fast authority”: the voice on the air, the anchor, the breaking chyron, the live feed. The logic is: it is happening now, therefore it is real, therefore it is authoritative. In the

feed era, fast authority becomes personalized: the influencer, the thread, the “exclusive,” the “clearest view yet,” the “everyone is lying but us” channel.

Fast authority feels like certainty. It feels like being protected from chaos. But it is structurally vulnerable to manipulation because it bypasses the patient disciplines that make knowledge reliable: provenance, cross-checking, context, error correction, and time.

Slow authority is the opposite. It does not ask you to submit to a voice. It asks you to submit to procedures that can be inspected.

Slow authority is built from auditability:

- Provenance: where did this come from?
- Uncertainty: what is known, what is inferred, what is unknown?
- Disagreement: what do serious opponents say?
- Disconfirmation: what would change the conclusion?
- Correction: how does this system handle error?

Calling auditability a spiritual virtue is not rhetorical excess. Auditability trains humility. It says: I am not God; my story about the world must answer to the world. It also trains honesty. It refuses to launder weak claims through confidence or charisma. It trains patience. It refuses to treat urgency as permission to lie.

In a religious key, auditability is the opposite of idolatry. Idolatry demands submission to an image or voice. Auditability demands submission to reality and to truthfulness. In a secular key, auditability is the opposite of propaganda. Propaganda asks you to accept a narrative enclosure. Auditability gives you exit ramps.

The sea change therefore involves a re-sacralization of method: not in the sense that method replaces God, but in the sense that truthfulness is treated as holy work. Public life becomes healthier when citizens treat “Can this be audited?” as more important than “Does this feel right?”

4.5 The cost of speech: reclaiming responsibility for what we repeat

The final element of the sea change is speech. A century of mass media has trained speech to feel costless. You can repeat a claim without owning it. You can share a clip without verifying it. You can retweet a rumor and later say “I was just asking questions.” You can perform outrage without repairing the damage. You can speculate in public and never pay for the consequences.

Costless speech is a paradise for propaganda and a breeding ground for cruelty. When speech is unmoored from responsibility, it becomes a weapon used for thrills, status, and tribal bonding.

Reclaiming the cost of speech does not mean censorship. It means restoring personal accountability to repetition. It means treating “I said it” and “I shared it” as morally adjacent to “I did it,” because in a mediated environment, repeating is a form of doing. Repetition spreads harm, hardens narratives, and trains crowds.

A truth-first posture changes the default:

- Before repeating: classify the claim as observation, inference, or speculation.
- If speculative: label it, and do not launder it as certainty.
- If uncertain: slow down, or do not amplify.
- If about a person’s private suffering: ask whether repeating violates dignity.
- If about enemies: refuse dehumanizing language that makes cruelty easier.
- If wrong: correct publicly with the same energy with which you spread it.

The sea change is not “be nicer.” It is “stop treating speech as consequence-free.” When speech has cost, propaganda loses velocity. When propaganda loses velocity, the coupled system loses fuel. When the coupled system loses fuel, the soul begins to recover the capacity for ordinary life, slow reading, honest doubt, and genuine witness.

This is the heart of the proposal: a re-ordered hierarchy of goods, expressed not as a sermon but as daily practice. Truth-first metaphysics is not proved by rhetoric. It is proved by what it forms: slower certainty, stronger dignity, greater humility, and a public square less addicted to spectacle.

5. From Information to Formation

Modern culture’s instinctive remedy for almost every public problem is informational: more data, more exposure, more transparency, more content, more “awareness.” This instinct is understandable. In many domains, ignorance is harmful, and information can save lives. But when the injury is not merely ignorance but deformation of the perceiver, information alone becomes an ineffective medicine. You can pour true facts into a warped vessel and still get a warped output. The vessel must be reshaped. That reshaping is formation.

Formation is not a vague religious word. It is the way repeated practices train perception, desire, reflex, and conscience. It is how a musician becomes able to hear intonation, how a

scientist becomes able to spot confounds, how a parent becomes able to stay calm, how a pilot becomes able to respond under stress. The claim here is that the attention economy is not primarily producing an uninformed public; it is producing a differently formed public: quicker to react than to verify, more hungry for intensity than for proportion, more loyal to narrative identity than to reality contact. That is why the remedy must be formation: a disciplined posture toward seeing, speech, and the treatment of persons.

This section proposes a small set of formation primitives that can travel: discipline, test, and fruit. Together they shift the reader from “consumer of information” to “practitioner of truth.”

5.1 Why data does not cure deformation

There are at least five reasons why “more data” fails to cure the specific deformation described earlier.

First, data competes in the same arena as spectacle. If the environment rewards arousal, then data is selected and framed to produce arousal. True data can be used as ammunition. Partial data can be used as bait. Even accurate information can be deployed as propaganda if it is arranged to control salience and lock identity.

Second, data does not automatically produce proportion. A deformed perceiver can treat a rare event as ubiquitous because it is vivid and repeatedly shown. The statistical rarity does not matter if the nervous system is trained by repetition. Proportion requires cultivated judgment: the ability to relate the vivid to the typical, the clip to the base rate, the anecdote to the distribution.

Third, data does not automatically produce patience. A person trained by the feed can want “the facts” in the same way they want candy: immediately, in a satisfying shape, with no ambiguity. But serious facts often arrive slowly and with uncertainty. Without a formed capacity for delay and ambiguity, data becomes either frustrating (and therefore ignored) or distorted into premature certainty.

Fourth, data does not automatically produce humility. A person can become more confident and more cruel with more information if they treat information as a weapon for status and victory. In fact, a high-information person inside a propaganda enclosure can be more dangerous than a low-information person, because they can rationalize better.

Fifth, data does not automatically repair the moral injury of voyeurism. Seeing more images of suffering does not necessarily produce compassion; it can produce numbness, cynicism, and appetite for escalation. The heart is not a neutral receptacle. It habituates.

So the point is not “facts don’t matter.” The point is that facts are downstream of posture. If the posture is wrong, facts become fuel for the wrong machine.

5.2 Formation primitives: discipline, test, fruit

To move from information to formation, we need a minimal toolkit—small enough to be adopted by ordinary people and robust enough to resist manipulation. The proposed primitives are discipline, test, and fruit.

Discipline is a repeatable practice that trains the perceiver. It is not an ideology. It is a habit that reshapes reflex. Examples include: labeling claims as observation/inference/speculation; delaying consumption of breaking tragedy footage; reading long-form before reacting; limiting replay; refusing dehumanizing language; practicing “steelman the best objection” before sharing.

Test is a built-in disconfirmation pathway. It is the exit ramp that makes a belief accountable to reality rather than to identity. A test asks: what would change my mind? What evidence would downgrade this claim? What alternative explanation fits the data? What is the strongest counter-evidence? Without tests, beliefs become ornaments of belonging. With tests, beliefs become hypotheses held with humility.

Fruit is the long-horizon evaluation criterion. It asks: what does this way of seeing and speaking produce in a person over time? Does it yield patience, honesty, compassion, courage, and stable sanity? Or does it yield addiction to outrage, contempt, paranoia, and fragile certainty? Fruit is not a feeling in the moment; it is character over time. It is the moral and spiritual “outcome metric” that refuses short-term arousal as a proxy for truth.

These three primitives work together. Discipline trains the posture. Tests keep the posture honest. Fruit keeps the posture oriented toward human good rather than toward merely being right.

5.3 The difference between being informed and becoming able

Being informed is possession of facts. Becoming able is possession of capacities.

A person can be informed about aviation and still be unable to judge a crash video responsibly. A person can be informed about politics and still be unable to resist propaganda. A person can be informed about war and still be unable to speak about enemies without dehumanizing them. The question is not only “Do you know?” but “Can you do?” Can you slow down? Can you verify? Can you hold uncertainty? Can you preserve dignity? Can you revise publicly? Can you refuse to consume suffering as entertainment? These are abilities, not facts.

Formation shifts the goal from accumulating information to acquiring competence in truthfulness.

Competence in truthfulness includes:

- Temporal competence: the ability to wait without panic.
- Epistemic competence: the ability to distinguish observation from inference.
- Emotional competence: the ability to feel without being governed by feeling.
- Moral competence: the ability to preserve personhood under anger or fear.
- Social competence: the ability to disagree without turning disagreement into war.

The feed trains the opposite: immediate reaction, emotional certainty, role-based dehumanization, and identity-locked narratives. That is why a corpus of papers, no matter how true, will not “win” merely by existing. It must become a training ground—an apprenticeship in able truth.

5.4 The role of restraint: choosing not to watch as an ethical competence

Modern culture often treats restraint as either weakness or prudishness. But restraint is a positive skill: the ability to refuse stimuli that would deform the self or harm others, even when that refusal costs curiosity, social inclusion, or the pleasure of intensity.

Choosing not to watch is not always the right choice. Sometimes witness requires looking. The competence is knowing the difference.

Restraint has several dimensions.

First, restraint protects dignity. There are moments where looking is already a violation: the private grief of a victim’s family, the humiliating collapse of a stranger, the replay of death, the zoomed-in wound. Even if such footage exists, repeating it can be an act of consumption rather than witness.

Second, restraint protects the perceiver. There are images and loops that etch themselves into the mind and train it in fear, contempt, or lust. The question is not “Can I handle it?” but “What will this practice make me?” Mature people treat their attention like a garden: they do not plant weeds and then boast that they can tolerate weeds.

Third, restraint protects truth. The rush to see and share often outpaces verification. In a high-arousal environment, people become confident before they become accurate. Restraint creates time for the basic disciplines of truth to operate.

Fourth, restraint protects community. Many propaganda engines thrive because people repeatedly amplify the same emotionally charged fragments. Refusing to watch, refusing to

replay, and refusing to share are forms of civic hygiene. They slow the spread of manipulative content.

So restraint is not abstinence for its own sake. It is stewardship. It is the cultivated ability to keep appetite from becoming governor.

5.5 How formation scales without turning into moralism

A central risk in any formation project is moralism: turning the program into a badge system, a purity contest, or a new class hierarchy of “the enlightened” versus “the sheeple.” That would simply reproduce propaganda dynamics with a different aesthetic. The sea change must therefore scale through humility, not through superiority.

Formation scales in at least five non-moralistic ways.

First, through defaults rather than demands. Make the covenant a front matter, a shared checklist, a standard practice. Do not make it a club. People adopt defaults more readily than they obey lectures.

Second, through procedures rather than virtue signaling. Teach practices like labeling uncertainty, offering disconfirmation conditions, and correcting publicly. These are observable behaviors that do not require reading hearts.

Third, through invitation rather than accusation. The covenant is offered as a tool for sanity and dignity, not as a weapon to shame those who fail. Shame rarely produces stable formation; it produces hiding, performance, and resentment.

Fourth, through fruit rather than instant judgment. Formation is measured over time: are people calmer, truer, kinder, more capable of revision? This removes the temptation to police every slip and instead focuses on trajectory.

Fifth, through symmetry: apply the rules to “your side” first. Propaganda thrives when standards are asymmetric. Formation survives when standards are applied to the self before being imposed on others.

The aim is a public square in which restraint is honored, auditability is normal, and speech carries cost—not because everyone is suddenly virtuous, but because the practices are designed to make virtue easier and manipulation harder.

In sum, moving from information to formation is not a retreat from truth. It is a recommitment to truth at the level where truth is either sustained or sabotaged: the level of habits. Facts matter. But the ability to live with facts—without becoming a spectator, without becoming an addict, without becoming a propagandist—matters more. Formation is what makes a truth-first metaphysics operational in a world that profits from the opposite.

6. A Covenant for Reading in the Centaur Era

A covenant is a voluntary binding: a set of promises that shape practice when appetite, fear, or tribal heat would otherwise seize control. In older worlds, covenants were explicit because communities knew that humans drift. In the modern media world, drift is not only human; it is engineered. The environment is designed to keep you watching, to keep you reactive, to keep you certain, to keep you sharing. That is why the remedy must be more than “try harder.” It must be a binding that can survive pressure.

This covenant is not a religious test and not a purity ritual. It is a minimal constitution for attention. It is designed for the centaur era: a time when AI systems can summarize, persuade, and generate high-volume text faster than any human; when “evidence” can be clipped, enhanced, and circulated at industrial scale; when the line between witness and spectacle is constantly exploited; and when the temptation to treat the model or the feed as an oracle is pervasive.

The covenant’s aim is not to stop you from seeing the world. Its aim is to restore the moral shape of seeing and to protect persons from being consumed as content. It makes auditability normal, restraint honorable, and revision possible. It gives you exit ramps when a system tries to lock your identity to a stream.

6.1 The covenant’s purpose and design constraints

A covenant that is too long becomes a manifesto. A covenant that is too abstract becomes a mood. A covenant that is too strict becomes legalism. The design constraints here are practical:

- It must be short enough to remember.
- It must be actionable at the moment of temptation.
- It must be symmetric: it applies to “my side” first.
- It must be compatible with pluralism: believers and non-believers can adopt it without adopting each other’s metaphysics.
- It must preserve joy: it is not an anti-art or anti-humor rulebook.
- It must handle AI explicitly: because in the centaur era, the most persuasive voice in the room may not be human.

The purpose is likewise practical:

- To prevent the slide from witness to voyeurism.
- To prevent appetite from governing interpretation.

- To prevent propaganda from enclosing the mind.
- To protect dignity in a world that monetizes humiliation.
- To rebuild slow authority by normalizing auditability.
- To keep conscience in human hands, not delegated to systems.

The covenant is not a replacement for wisdom. It is a scaffolding for wisdom—especially when wisdom is expensive.

6.2 The witness rule: no spectacle, no humiliation harvest

The witness rule is the first boundary because it addresses the core moral danger of mediated life: treating the suffering or intimacy of others as consumable.

The witness rule states:

- I will look only as far as is necessary for understanding and responsibility.
- I will not maximize intimacy, replay wounds, or zoom into humiliation.
- I will not consume a person’s worst moment as entertainment.
- I will not treat “clearest view yet” as a moral permission slip.

This rule does not forbid documentation. It forbids extraction.

In practice, the witness rule changes behavior in two ways.

First, it limits repetition. A single viewing may be necessary to understand what happened. Rewatching for thrill, speculation, or “micro-analysis entertainment” is almost always voyeuristic, especially when the subject is harmed.

Second, it changes what you share. You do not share raw suffering to provoke outrage or to gather attention. If a public record is necessary, you share the least intimate form that still supports accountability: a summary, a transcript, a still frame that avoids gore, an official statement, or a carefully described account that preserves dignity.

The witness rule restores a lost distinction: seeing can be required, but consuming is not required. The covenant insists on that distinction.

6.3 The audit rule: provenance, uncertainty, disagreement, disconfirmation

The audit rule is the heart of slow authority. It treats knowledge as accountable rather than charismatic.

The audit rule states:

- I will not treat vividness as proof.
- I will ask where a claim came from and how it was produced.
- I will label what I know, what I infer, and what I do not know.
- I will actively seek serious disagreement.
- I will state what would change my mind.

This rule has four parts that travel well.

Provenance means origin and chain of custody. Who recorded this? Where is the original file? Was it edited? What is the time and place? What is the source's incentive? In the centaur era, provenance also includes AI: was this text generated, summarized, or translated by a model? If so, what materials were used? What was omitted? Was anything synthesized?

Uncertainty means refusing to launder inference as observation. It is the discipline of saying: "I saw X," "I think X implies Y," "I don't know whether Z." Uncertainty is not weakness. It is honesty under incomplete access.

Disagreement means treating the best opposing view as a necessary instrument. Propaganda collapses disagreement into contempt. The covenant restores disagreement as a tool for truth.

Disconfirmation means exit ramps. What evidence would downgrade this story? What observation would falsify the interpretation? What alternative explanation could fit? Without disconfirmation, beliefs become identity jewelry. With disconfirmation, beliefs remain accountable to reality.

The audit rule is the practical antidote to both propaganda and "AI oracle" temptation: it refuses to let persuasion substitute for method.

6.4 The no-oracle rule: non-delegation of conscience to models or feeds

In the centaur era, the strongest temptation is not to be fooled by a lie. It is to be persuaded by a voice that sounds wiser than you. Models can generate coherent narratives at speed. Feeds can provide constant confirmation. Together they can create a synthetic authority that feels like inevitability.

The no-oracle rule states:

- I will not delegate final belief, moral judgment, or conscience to any model, platform, or crowd.

- I will treat outputs as claims to be audited, not commands to be obeyed.
- I will not confuse prediction accuracy with moral authority.
- I will preserve the right and duty to say “no,” “wait,” and “I don’t know.”

This rule protects a clean hierarchy: tools assist; conscience rules. Even when a model is correct, it should not become a crown.

Practically, the no-oracle rule changes how you use AI:

- You ask for alternatives and objections, not only for confirmation.
- You request uncertainty labeling and disconfirmation conditions.
- You keep a personal stop rule: if the system is increasing arousal, urgency, or certainty without corresponding evidence, you pause.
- You refuse “voice-of-God” summaries that compress away the very ambiguities that require humility.

The no-oracle rule is the anti-idolatry rule in epistemic form: no synthetic voice receives submission.

6.5 The lag rule: replacing “live” with “later” for tragedy and conflict

The lag rule is the covenant’s most countercultural move because it attacks the market’s strongest weapon: immediacy. “Live” creates urgency. Urgency creates compliance. Compliance creates manipulation opportunities.

The lag rule states:

- For tragedy and conflict, I will prefer later summaries over live feeds.
- I will not watch breaking footage when facts are unstable and dignity is most at risk.
- I will allow time for verification and for the human cost to be honored.
- I will treat “breaking” as a warning label, not a call to binge.

This is not denial. It is discipline. It recognizes a simple truth: in the first hours of a tragedy, the highest probability is confusion, misattribution, rumor, and emotional exploitation. Watching live rarely increases understanding. It often increases arousal and error.

The lag rule also protects victims. The most humiliating, raw, and private footage often appears early, before norms of restraint can reassert themselves. Refusing the live window is a way of refusing to be part of the initial harvest.

In practice, the lag rule can be as concrete as “24 hours” for major tragedies, or “until official statements and multiple independent accounts exist.” The exact number is less important than the posture: time is your ally, not your enemy.

6.6 The fruit rule: long-horizon outcomes over virality and heat

The fruit rule is the covenant’s evaluation metric. It is the question that breaks the spell of short-term arousal: what does this produce in me and in my community over time?

The fruit rule states:

- I will evaluate narratives and habits by their long-horizon outcomes.
- I will not treat virality, intensity, or group affirmation as evidence of truth.
- I will prefer practices that produce humility, patience, courage, and love.
- I will downgrade practices that produce addiction to outrage, contempt, paranoia, and dehumanization.

Fruit is not merely personal mood. It is character and community health. A channel can be “right” about some facts and still produce bad fruit: constant rage, suspicion of everyone, inability to revise, joylessness, fixation on enemies, and a steady erosion of mercy. The covenant insists that such fruit is diagnostic: it signals that the person is being governed by the coupled system, even if some claims are true.

In the centaur context, the fruit rule also guards against “synthetic holiness,” where beautifully written moral language is used to justify manipulation. The test is not aesthetic. The test is outcome.

6.7 The dignity rule: persons remain persons, even when wrong

The final rule is the foundation beneath all the others. Without it, the covenant becomes a clever technique rather than a moral reorientation.

The dignity rule states:

- I will treat every person as a person, not as content.
- I will not dehumanize opponents to gain rhetorical speed.
- I will not use humiliation as entertainment or as proof.
- I will separate critique of actions from consumption of persons.
- I will preserve the possibility of repentance, correction, and repair.

This rule is essential because propaganda always requires enemies, and enemies are easiest to mobilize when they are less than human. The dignity rule refuses that shortcut. It makes cruelty harder, which is precisely why it is resisted by the attention economy.

Practically, the dignity rule disciplines language: no contempt nicknames, no “subhuman” framing, no eroticized hatred, no delight in downfall, no pleasure in humiliation. It disciplines sharing: you do not spread a person’s disgrace simply because they “deserve it.” It disciplines interpretation: you do not assume the worst motive when a more ordinary motive fits.

The dignity rule does not forbid accountability. It forbids consumption. It is the difference between justice and entertainment.

Taken together, these rules are not “nice ideas.” They are a counter-architecture for attention. They are a portable constitution that can be attached to papers, newsletters, classrooms, churches, and AI-assisted writing workflows. They are meant to be practiced, not admired. And if practiced, they produce the sea change this paper calls for: a shift from stimulation-first metaphysics to truth-first metaphysics, where witness replaces voyeurism, auditability replaces charisma, and persons remain non-consumable even in a world that profits from treating them otherwise.

7. Implementation as a Portable Practice

A covenant that remains a beautiful statement is not a covenant; it is a poster. The centaur era does not need more posters. It needs portable practices: small forms that can be carried across platforms and repeated under pressure. Implementation must therefore be designed around three realities.

First, most attention decisions are made at the speed of habit. When you are tired, angry, lonely, afraid, or bored, you do not “reason your way” into good attention; you default into it. Implementation must reshape defaults.

Second, the coupled system resists reform by using virtue language. It will label restraint as cowardice, auditability as complicity, and lag as “letting them get away with it.” Implementation must therefore be socially durable, not merely intellectually correct.

Third, the centaur environment amplifies both good and bad. A single person can now publish at industrial scale; a single model can generate persuasion at industrial scale. That means small errors can propagate rapidly, and small disciplines can also scale rapidly. Implementation should assume high throughput and build in checks that are lightweight but real.

This section offers practical forms: front matter, checklists, community norms, personal protocols, and institutional protocols. The aim is to make truth-first practice easier than drift.

7.1 One-page front matter for papers, newsletters, and videos

The simplest way to make a covenant portable is to put it where attention begins: at the entrance. Front matter is a threshold object. It sets posture before content. In the centaur era, posture is half the battle.

A one-page front matter should do five things quickly.

First, declare the purpose. Not “to be right,” but to be a witness: to preserve dignity, truth, and auditability.

Second, declare the rule-set in a compact form. Ideally one line per rule: witness, audit, no-oracle, lag, fruit, dignity.

Third, declare the claim grammar. A small legend that distinguishes observation, inference, and speculation. This becomes a shared language that readers can use to critique without contempt.

Fourth, declare the reader’s agency. The front matter should explicitly bless the reader’s right to pause, to doubt, to request provenance, and to refuse manipulation. This is a subtle but powerful anti-propaganda move: it makes revision honorable.

Fifth, declare the correction pathway. Where do corrections live? How are updates handled? What is the expected behavior if an error is found? The existence of a correction pathway changes the moral atmosphere of a project from “performance” to “accountability.”

Front matter is not merely moral decoration. It is a design element that trains the reader to approach the material as a practice. Over time, readers begin to expect such thresholds. That expectation is itself formation.

A useful implementation strategy is consistency: the same front matter appears in every paper, every newsletter issue, and every video description or pinned comment. Repetition is not redundancy; it is training.

7.2 Checklists that do not infantilize the reader

Checklists fail when they treat readers as children and authors as police. The point is not compliance; the point is competence. A mature checklist feels like a pilot’s preflight list: a professional tool that protects judgment under stress.

A non-infantilizing checklist has three features.

First, it is short enough to use. If it takes longer than a minute, it will be skipped in the very moments it is needed most.

Second, it is phrased as questions rather than commands. Questions preserve agency and invite honest self-assessment.

Third, it distinguishes levels of seriousness. Not every claim needs the same scrutiny. A meme is not a medical instruction. A rumor about a private person is more sensitive than a critique of a public policy. The checklist should scale effort to stakes.

A practical way to structure checklists is by moment.

A “Before you watch” checklist can ask: is this tragedy or conflict? Is this likely to contain humiliation harvest? What is my motive for clicking? Do I need this now?

A “Before you share” checklist can ask: what is the provenance? Am I sharing observation or inference? Have I seen serious disagreement? Would I be willing to correct publicly? Would this violate dignity?

A “After you consume” checklist can ask: what fruit did this produce in me? Did it make me calmer and truer, or hotter and more certain? Do I need a stop rule right now?

Checklists should also include one explicit permission: it is allowed to exit. A mature checklist assumes that the highest skill is not “watching wisely” but sometimes “not watching at all.”

7.3 Community norms: how to enforce without coercion

Communities either enforce norms or they are enforced by the loudest manipulators. The question is not whether norms will exist; the question is whether they will be truth-first or stimulation-first. Yet enforcement can easily slide into coercion, shame rituals, or ideological policing. Implementation must therefore aim for enforcement that is procedural, gentle, and reversible.

A community can enforce a covenant without coercion by using five practices.

First, make norms explicit and opt-in. The covenant is a house rule for a space, not a moral verdict on outsiders. People can leave without being labeled enemies.

Second, enforce behavior rather than motives. Motives are inaccessible and easy to weaponize. Behaviors are observable. A community can say, “We require provenance for claims,” without saying, “You are a liar.” It can say, “We do not repost humiliation footage,” without saying, “You are wicked.”

Third, create correction as a honor practice. In a truth-first space, correction is not humiliation. It is stewardship. Public corrections should be treated as strength, not defeat. This single norm, if real, destroys much propaganda energy.

Fourth, use graduated responses. Most violations are drift, not malice. A gentle nudge, a request for provenance, or a label correction often suffices. Only repeat offenders who refuse correction require stronger boundaries.

Fifth, prohibit dehumanizing speech as a non-negotiable. This is where the dignity rule becomes operational. Contempt language is not merely rude; it is the seedbed of cruelty and propaganda. A community that tolerates dehumanization will eventually normalize cruelty.

Enforcement without coercion is possible when the community sees itself as practicing a craft together, not as maintaining a purity wall.

7.4 Personal protocols: attention fasting, re-entry rules, stop rules

The covenant becomes durable when it is translated into personal protocols that function under stress. Protocols are pre-decisions. They remove moral deliberation from moments when you are least capable of deliberation.

Attention fasting is one such protocol. It is not a ban on media. It is a periodic reset of the nervous system's baseline. A fast can be time-based (a day a week), content-based (no breaking news, no tragedy footage), or interface-based (no autoplay environments). The point is to restore the capacity to enjoy ordinary life and to read long-form without craving spikes. Attention fasting is also diagnostic: it reveals what you are dependent on.

Re-entry rules are equally important. People often leave a feed and then re-enter with the same vulnerabilities and immediately relapse into the coupled system. A re-entry rule might be: when I return, I begin with long-form, not clips; I begin with primary sources, not reactions; I do not watch live tragedy; I use lag by default. Re-entry rules prevent a fast from becoming a temporary pause that changes nothing.

Stop rules are the sharpest tool. A stop rule is a trigger that forces you to pause when the system is most likely to hijack you. Examples include: if I am replaying a harm clip, I stop; if I feel righteous rage rising, I stop; if I want to share before verifying, I stop; if I am reading only headlines, I stop; if a model output makes me feel "inevitable certainty," I stop.

Stop rules are not moral shame. They are emergency brakes. They treat the human being as a system that can be overloaded. They protect dignity and sanity.

The most important personal protocol in the centaur era is a model-use protocol. It should include: ask for counterarguments; ask for uncertainty; ask for sources later in the Reference

section; demand a disconfirmation condition; never treat the model's coherence as authority; log what you asked and what you accepted. This keeps the no-oracle rule from being a slogan.

7.5 Institutional protocols: editors, curators, and audit trails

Individual discipline is necessary, but the coupled system is industrial. If institutions do not adopt truth-first protocols, individual virtue will be overwhelmed by the scale of manipulation. The sea change therefore needs institutional forms that normalize auditability and protect dignity.

Editors are the first line. In the centaur era, editing is not only grammar; it is epistemic hygiene. Editors can require claim labeling (observation/inference/speculation), enforce dignity rules around depiction of victims, and insist on correction pathways. Editors can also define thresholds: when a claim is sensitive, it requires provenance before publication.

Curators are equally important. The modern public square is not short of content; it is short of trustworthy selection. Curators who operate under the covenant can provide an alternative to algorithmic curation. They can bundle material into coherent paths, highlight uncertainty, and prevent sensational fragments from becoming the public's primary diet.

Audit trails are the backbone. An audit trail is not a bureaucratic burden; it is a public record of honesty. It can include: what sources were used; what was inferred; what was speculative; what changed; what corrections were made; and what disagreements exist. In the centaur context, an audit trail can also include the role of AI: what was generated, what was rewritten, what was checked, and where human judgment was applied.

Institutions should also define "dignity thresholds." Some material should not be published even if it would draw attention. A truth-first institution will sometimes choose lower engagement to preserve personhood. That choice is a moral signal to the culture that the market does not own everything.

Finally, institutions need a revision culture. A revision culture is the opposite of propaganda culture. It rewards corrections, publishes updates, and treats "we were wrong" as integrity rather than as defeat. The easiest way to kill propaganda is to build spaces where revision is honored.

In sum, implementation must be portable, repeatable, and durable under pressure. One-page front matter sets posture. Checklists protect judgment without insult. Community norms enforce dignity without coercion. Personal protocols build emergency brakes into daily life. Institutional protocols make auditability and revision normal at scale. This is how a covenant becomes practice, and practice becomes formation, and formation becomes the sea change: a culture that no longer treats attention as neutral, persons as consumable, or speed as authority.

8. Centaur Dynamics: Human–AI Collaboration Under the Covenant

The centaur era is not simply “humans using tools.” It is a new cognitive environment in which persuasive synthesis can be produced faster than conscience can metabolize it. This is the crucial novelty: not that AI can generate text, but that it can generate *coherence at speed*—and coherence is one of the strongest triggers of human trust. A coherent narrative feels like understanding. It feels like control. It feels like being “on the right side of reality.” That felt control is precisely what propaganda has always sold. The difference now is scale and personalization. The model can deliver the feeling of understanding on demand, tailored to the user’s preferred metaphysics, vocabulary, and moral tone.

The covenant is therefore not merely a media discipline. It is an AI-use constitution. It must govern how humans relate to synthetic synthesis. Otherwise, the covenant will be applied to “TV and the internet” while the greatest new oracle temptation grows unchecked inside the very instrument that produces our writing.

This section names the risks and then offers covenant-shaped practices that preserve agency, humility, and auditability in human–AI collaboration.

8.1 When synthesis becomes too easy: persuasion without responsibility

Synthesis is the act of integrating scattered information into a coherent account. In science, synthesis is valuable because it creates models that can be tested. In journalism, synthesis can provide context that a single clip cannot. In philosophy and theology, synthesis can clarify commitments. The danger in the centaur era is that synthesis becomes effortless in a way that decouples persuasion from responsibility.

There are at least four ways this decoupling occurs.

First, speed outruns verification. When the model produces a clean story quickly, the human is tempted to accept the story as “probably right” and move on. Verification feels like friction. Friction feels like weakness. The result is a new kind of epistemic negligence: not deliberate lying, but habitual acceptance of coherence as a substitute for checking.

Second, eloquence launders uncertainty. A model can phrase an uncertain inference in a manner that feels settled. The human reader experiences “closure” and forgets that the closure was stylistic. This is especially dangerous in moral or political domains, where certainty carries social rewards.

Third, responsibility is diffused. In a centaur workflow, it becomes psychologically easy to say, “The model wrote it,” or “I was just summarizing,” or “It’s only a draft.” But published speech has effects regardless of its authorship. The covenant insists that responsibility cannot be outsourced. If you used the instrument, you own the outcome.

Fourth, the model can adapt to the user's appetite. If the user wants a spicy take, the model can deliver. If the user wants a holy tone, the model can deliver. If the user wants a conspiratorial vibe, the model can deliver. This is not "evil AI." It is the system's flexibility meeting the human's governable appetite. The result can be persuasion that feels like revelation. The covenant's no-oracle rule exists precisely to interrupt that.

So the risk is not that models will always be wrong. The risk is that they will often be *convincing enough* to bypass the practices that make truth trustworthy.

8.2 "Model as witness" vs "model as idol"

The covenant's central distinction—witness versus spectacle—has an AI analog: model as witness versus model as idol.

A model as witness is treated as a fallible describer that assists human agency. It is a tool that helps gather, organize, restate, and explore. Its outputs are claims to be audited, not verdicts to be obeyed. It helps a human become more honest, more careful, and more able to revise.

A model as idol is treated as a synthetic authority. Its coherence becomes a crown. Its confidence becomes a substitute for evidence. Its speed becomes a substitute for wisdom. The user begins to treat the model's outputs as what reality "must be," and dissent becomes psychologically costly because it threatens the comfort of closure.

Idolatry in this context does not require religious language. It is functional. If you begin to surrender the right to doubt, to pause, to verify, or to disagree—because the output feels too good, too elegant, too "obviously true"—the instrument has become an idol.

The covenant pushes collaboration toward witness by enforcing three habits:

- Always preserve uncertainty boundaries.
- Always preserve disconfirmation pathways.
- Always preserve the primacy of conscience.

That is what makes the model a witness: it increases human responsibility rather than replacing it.

8.3 Guardrails that preserve agency: prompts, logs, and downgrade conditions

A covenant is operational only when it becomes concrete procedures. In centaur work, the most effective guardrails are not abstract warnings; they are design patterns you build into your prompting and documentation.

Prompts that preserve agency have several features.

They explicitly request claim typing. Ask the model to label statements as observation, inference, speculation, and to keep those labels visible in the draft. This prevents coherence from laundering uncertainty.

They explicitly request counterarguments and alternatives. If a model can generate a strong case for one view, it can usually generate a strong case for the opposite. Making this routine prevents identity lock.

They explicitly request disconfirmation conditions. A good prompt asks: what evidence would weaken this claim? What would falsify it? What would change the conclusion? This is the anti-propaganda move inside the writing process.

They explicitly request restraint rules. When dealing with tragedies or personal harm, prompts can instruct: do not dwell on graphic detail; do not replay the moment; preserve dignity; focus on necessary facts and responsible framing. This builds the witness rule into generation.

Logs are the second guardrail. A log is not mere bureaucracy. It is memory and accountability. In high-throughput writing, you need a minimal record of: what question was asked, what assumptions were made, what sources were considered, and what uncertainties remained. Without logs, the work becomes a stream of persuasive outputs with no provenance. With logs, the work becomes auditable.

Downgrade conditions are the emergency brakes. A downgrade condition is a pre-committed rule that triggers when outputs show certain failure signatures. Examples include:

- If the output becomes more certain than the evidence supports, downgrade confidence and slow publication.
- If the output becomes emotionally escalatory, pause and apply lag.
- If the output relies on unnamed sources or vague authority, require provenance or withhold.
- If the output increases contempt toward a group or person, apply the dignity rule and revise.
- If the output makes you feel “inevitable clarity” on a controversial claim, force a disagreement pass.

These guardrails preserve human agency by preventing the instrument from becoming a momentum engine that carries the author away.

8.4 Disconfirmation at scale: testing claims in high-throughput writing

High-throughput writing is the new normal. The problem is not that it exists; the problem is that it can produce high volumes of plausible text with insufficient disconfirmation. Disconfirmation is the discipline that makes a belief answerable to reality. Without it, high throughput becomes high-velocity error.

Disconfirmation at scale requires procedures that are lightweight and repeatable.

One method is the “two-pass contradiction check.” After drafting, require a second pass that attempts to falsify the key claims. The model is asked to generate the strongest objections, alternative interpretations, and counter-evidence categories. The author then either revises or explicitly marks the claim as speculative.

Another method is the “source placeholder discipline.” Even if references are reserved for the end, the draft can still mark where citations will be required: claims about dates, numbers, attributions, and contested facts are flagged for later sourcing. This keeps the author from forgetting which statements need grounding.

A third method is the “triangulation rule.” For a high-stakes claim, require independent corroboration categories rather than just multiple repetitions of the same source. Triangulation can be conceptual (different kinds of evidence) as well as documentary (different outlets). The point is to prevent the illusion of consensus created by copy-paste journalism or by social media echo.

A fourth method is the “uncertainty budget.” Treat certainty as a scarce resource. The draft is allowed a limited number of high-certainty statements on contested topics unless they are strongly sourced. This forces the author to reserve certainty for what is truly supportable.

Disconfirmation at scale is not about perfection. It is about creating systematic exit ramps so the work remains corrigible rather than self-sealing.

8.5 Avoiding synthetic authority: citation discipline and humility parameters

Synthetic authority is the phenomenon where AI-generated coherence is mistaken for grounded truth. Avoiding it requires two complementary disciplines: citation discipline and humility parameters.

Citation discipline does not mean in-line references in the expansion sections. It means disciplined behavior about what will later be referenced and how. It includes:

- Distinguishing “general background” from “specific factual claims.”

- Flagging where a source will be required: dates, statistics, quotations, official actions, and contested claims.
- Refusing “vibe citations,” where the text implies authority without specifying it.
- Refusing “consensus claims” without evidence that a true consensus exists.

Humility parameters are the internal constraints you impose on the system and on yourself.

One humility parameter is “no false precision.” Models often produce exact numbers, exact dates, and exact attributions because precision sounds confident. The covenant requires that precision be earned. If you do not know, say you do not know.

Another humility parameter is “no moral laundering.” Models can produce righteous language that makes a conclusion feel morally mandatory. The covenant requires separating moral aspiration from factual certainty. You may condemn evil without pretending you have perfect information about every actor and motive.

A third humility parameter is “no single-stream dependence.” If your conclusions repeatedly depend on one channel, one narrative community, or one model output, you are drifting toward an oracle. The covenant requires diversification of inputs and openness to correction.

A fourth humility parameter is “revise in public.” Synthetic authority thrives where revision is shameful. When correction is honored, synthetic authority loses its aura. The author’s credibility becomes grounded not in always being right but in being reliably corrigible.

Taken together, citation discipline and humility parameters keep the centaur work from becoming a factory of persuasive hallucination or a vehicle for propaganda. They preserve the model as witness rather than idol.

In sum, the covenant’s function in the centaur era is to preserve agency in an environment where coherence is cheap and persuasion is abundant. It insists that speed does not confer authority, that eloquence does not confer truth, and that conscience cannot be delegated. It converts human–AI collaboration from a persuasion engine into a witnessing practice: auditable, humble, corrigible, and protective of personhood.

9. Objections and Failure Modes (Make Them Strong)

A covenant that cannot survive its best objections is not a covenant; it is a mood written in confident prose. The attention economy is not defeated by sincerity. It is defeated by practices that remain durable when mocked, misunderstood, or exploited. This section therefore treats the strongest objections as instruments. If the covenant can absorb these objections—without

self-deception and without becoming brittle—then it has a chance of being portable in the real world.

Just as important, every reform carries its own failure modes. A truth-first project can accidentally recreate the pathologies it opposes: it can become a badge system, a purity contest, or a paranoid enclave. Those are not side issues; they are the easiest ways for the coupled system to “re-enter” through the back door. The covenant must therefore include self-critique and self-limits, or it will become another form of propaganda.

9.1 “This is prudish or anti-media”

The objection: This covenant sounds like an aesthetic of withdrawal. It pathologizes watching. It treats entertainment as morally suspect. It implies that mature people should avoid vivid footage, conflict-driven stories, and cultural media. It feels like old-school prudery dressed in new language. It risks making ordinary human curiosity into a sin. It may also become politically selective: people will call “voyeurism” whatever footage embarrasses their side and call “witness” whatever helps their side.

The steelman: Much of human learning is mediated. Stories teach empathy. Images can document injustice. Entertainment can restore joy. Even tragedy coverage can galvanize aid, expose corruption, and protect the vulnerable. A covenant that discourages people from looking could protect bad actors and reduce public accountability. Worse, the covenant could become a veneer for comfort: a way for privileged people to avoid witnessing suffering they would rather not face.

The answer the covenant must give: The covenant is not anti-media. It is anti-consumption of persons. It does not forbid looking; it binds looking to obligation, restraint, and auditability. It specifically targets replay, humiliation harvest, and arousal-driven browsing. It also requires symmetry: the covenant applies to “my side” first, including the temptation to call only the other side’s media “spectacle.” Finally, it must explicitly affirm legitimate joy and art, and it must include a clause that witness sometimes requires looking at hard realities—but looking in a way that preserves dignity and increases responsibility rather than merely increasing stimulation.

9.2 “People have always enjoyed spectacle”

The objection: There is nothing new here. Humans have always gathered around the gallows, the arena, the scandal, and the rumor mill. The covenant treats a permanent feature of human nature as if it were a novel crisis. Even if we implement these practices, appetite for spectacle will remain. Why pretend we can change what humans are?

The steelman: This objection is historically strong. Societies have long used spectacle to entertain, to enforce norms, and to build social cohesion. The urge to stare is ancient. Reform

movements that ignore this often collapse into hypocrisy. Moreover, the covenant risks over-reading the moral significance of “watching,” when watching is sometimes just leisure in a hard life.

The answer the covenant must give: Yes, the appetite is old. But the scale, ubiquity, and personalization are not. A stable vice becomes a crisis when it is industrialized, frictionless, and financially optimized. The covenant does not claim to eradicate the appetite; it claims to reintroduce friction, restraint, and dignity where industrialization removed them. It aims to change defaults and train competence, not to create a utopia. It treats formation as realistic precisely because it assumes human drift. A covenant is not an ideology of perfectibility; it is a discipline for imperfect creatures living inside a powerful machine.

9.3 “Auditability is cold and inhuman”

The objection: The audit rule sounds like bureaucratic rationalism. It risks turning human experience into checklists and procedures. In many moral situations, what matters is compassion, courage, and presence, not “provenance.” People can be audited and still be cruel. Excessive auditing can also be weaponized: endless demands for proof can become a way to dismiss testimony, especially from vulnerable people. The audit aesthetic can drain life of empathy and reduce truth to paperwork.

The steelman: This is a serious critique. A culture that worships method can become heartless. “Just show me the evidence” can become a way to avoid responsibility or to ignore what cannot be easily quantified. Additionally, propaganda can also be “audited” in the narrow sense—propagandists can cite facts selectively, provide plausible provenance, and still manipulate through framing and omission.

The answer the covenant must give: Auditability is necessary but not sufficient. It must be paired with the dignity rule and fruit rule. The covenant does not replace compassion with method; it protects compassion from manipulation. It also requires proportional auditing: the level of scrutiny should match the stakes and the claim type. Finally, the covenant must include a guardrail against “audit as dismissal”: testimony can be honored as testimony, while still distinguishing it from broad generalizations and still refusing dehumanizing conclusions. Auditability is framed here as a spiritual virtue because it trains humility and corrigibility, not because it is the whole of the good.

9.4 “Covenants become performative badges”

The objection: The covenant will become a brand. People will paste it on their content as a signal: “I’m one of the good ones.” They will use covenant language to shame outsiders and to elevate themselves. The covenant will be gamed as virtue theater. Once gamed, it becomes

another propaganda style, another identity marker, another way to claim moral superiority while doing the same manipulations with nicer words.

The steelman: This is not hypothetical; it is the default fate of many moral frameworks online. Anything that can be used as a badge will be used as a badge. The covenant may also attract people who enjoy rules because rules provide control, status, and a target for contempt.

The answer the covenant must give: The covenant must be anti-badge by design. That means: emphasize practices over slogans; emphasize correction over self-congratulation; emphasize symmetry and self-application; emphasize fruit over performance. A community can operationalize this by treating “covenant compliance claims” as irrelevant and treating actual behaviors as the only evidence: do you label uncertainty, correct publicly, refuse humiliation clips, and preserve dignity for enemies? Most importantly, the covenant must include an explicit “no virtue signaling clause”: the covenant is not to be used to claim purity, only to accept accountability.

9.5 “This can’t compete with the feed”

The objection: Even if the covenant is true, it will lose. The feed is engineered by billion-dollar incentives. It will outcompete slow authority, lag rules, and careful auditing. People are tired, stressed, and lonely. They will choose easy stimulation. The covenant asks for self-control at precisely the moment when self-control is most depleted. This is like proposing meditation to stop a hurricane.

The steelman: This is probably the strongest practical objection. It names the asymmetry: one side is an industrial system tuned for addiction; the other side is voluntary discipline. Voluntary discipline often loses at scale. Even if a minority adopts the covenant, it may not meaningfully change the public square.

The answer the covenant must give: The covenant is not a plan to “win the feed.” It is a plan to preserve witnesshood and sanity within it, and to build alternative micro-environments where truth-first practices are normal. Societal change often begins as subculture, then becomes institution, then becomes default. The covenant’s strategy is therefore not mass conversion by argument; it is durable practice by communities, editors, and institutions that can create islands of slow authority. Additionally, the covenant leverages one advantage: it produces better fruit. Arousal systems burn people out. The covenant aims to become attractive through stability: less anxiety, less paranoia, less contempt, more clarity. The “competition” is not only for clicks; it is for the kind of life people can bear.

9.6 “Covenant language is religious coercion”

The objection: “Covenant” is religious language. It implies binding, moral authority, and perhaps divine sanction. In pluralistic public life, covenant talk can feel like coercion or like smuggling theology into media critique. People may reject the framework not because they love voyeurism but because they dislike religious framing. Worse, some may weaponize covenant language to impose their metaphysics on others.

The steelman: This is fair. Public discourse needs shared terms that do not require conversion. Even for believers, covenant language can be abused: it can be used to claim divine endorsement for one’s media choices and to shame outsiders.

The answer the covenant must give: The covenant is explicitly voluntary and procedural. It does not demand religious assent. It names universal goods: dignity, truthfulness, corrigibility, restraint under tragedy, and non-delegation of conscience. If the word “covenant” is a barrier, communities can translate it into secular equivalents: “constitution,” “reader’s charter,” “attention code,” “practice pledge.” The substance is what matters. The covenant also forbids coercion by design: it is a rule for the self and for opt-in communities, not a mandate for outsiders. The dignity rule includes ideological dignity: disagreement about metaphysics is not treated as moral inferiority.

9.7 Failure modes: legalism, elitism, paranoia, and purity spirals

The covenant’s greatest danger is that it will recreate the pathologies of the coupled system in a “truth-first” costume. The following failure modes must be anticipated.

Legalism: People treat the covenant as a scorecard. They focus on rule compliance rather than on truth, dignity, and fruit. Legalism produces brittle communities that punish minor deviations and miss major harms. It also produces hypocrisy: people learn to appear compliant while harboring contempt.

Antidote: keep the covenant short, emphasize fruit, require mercy, and treat corrections as normal. The covenant is a guardrail, not a ladder to moral superiority.

Elitism: People treat covenant practice as proof of intelligence or moral rank. They begin to despise those “still trapped in the feed.” Elitism reproduces dehumanization, which is the core enemy.

Antidote: symmetry and humility. Apply the covenant to your own side first. Speak as a fellow struggler, not as a superior. Remember that attention capture exploits human vulnerability, not human stupidity.

Paranoia: Because the covenant is alert to propaganda, it can drift into seeing propaganda everywhere. The practitioner can become suspicious of all institutions, all narratives, all disagreement. This creates a self-sealing worldview, which is propaganda's structure.

Antidote: disconfirmation discipline and proportionality. Not every error is manipulation. Not every disagreement is malice. Keep the exit ramps open. Maintain the possibility that the other side is partly right.

Purity spirals: Communities can escalate standards to prove seriousness, then punish those who cannot keep up, then splinter. Purity spirals are a known dynamics in moral communities, and they are especially dangerous online.

Antidote: prioritize essential rules (dignity, no-oracle, auditability) and allow flexibility in non-essentials. Use graduated responses. Treat restraint as honorable without turning it into a badge competition.

A final meta-failure is substitution: using covenant talk as a replacement for action. A person may become excellent at auditing media and poor at loving neighbors. That would be a tragic irony: the covenant would become another form of spectator life.

Antidote: the fruit rule must include outward-facing goods: greater patience, greater mercy, greater courage, and more fitting action in the real world.

If these objections and failure modes are taken seriously, the covenant becomes stronger. It becomes less a moral performance and more a practical constitution for attention. It acknowledges human weakness, resists manipulative systems without becoming a manipulative system, and aims for a modest but real achievement: preserving witnesshood under conditions designed to dissolve it.

10. Case Studies and Thought Experiments

A covenant proves itself not in principle but in friction: the moments when a person is tempted to click, share, rage, mock, or surrender their judgment to a stream. Case studies function as training simulations. They help the reader practice the covenant on concrete scenarios where the line between witness and spectacle is blurred by emotion and by incentives.

The five cases below are structured as thought experiments with decision points. The aim is not to dictate a single correct choice in every instance, but to show how the covenant reshapes posture: it introduces restraint, auditability, and dignity into moments that the attention economy wants to turn into consumption.

10.1 The tragedy clip: what must be shown vs what must not be replayed

Scenario: A public tragedy occurs. Within minutes, clips circulate. Some show the moment of harm. Others show the aftermath. Others are shaky, partial, or edited. Commentary arrives instantly. Viewers are urged to “watch this” to understand what really happened.

The covenant decision points:

A. Do I watch now or apply lag?

The lag rule says: default to later when facts are unstable and dignity risk is highest. Early footage is often misattributed, selectively framed, or monetized precisely because it is raw. Watching live rarely improves understanding more than a later, verified account. So the covenant’s default is: do not watch immediately. Wait for stabilization.

Exceptions exist. If you are directly responsible for safety (local official, investigator, onsite responder), or if your viewing is needed for immediate protective action, witness may require looking. But for most viewers, the honest truth is: it is not needed. Curiosity is not necessity.

B. What must be shown?

Sometimes an event must be documented to establish facts: location, sequence, responsible parties, failures of security, evidence of wrongdoing. In such cases, “must be shown” means the least intimate, least humiliating representation that still supports accountability.

That often means:

- A still image rather than replaying a moment of death.
- A redacted clip with the most graphic or humiliating segments removed.
- A timeline and description rather than raw footage.
- An official report or multiple corroborated witness summaries rather than a single viral clip.

The witness rule reframes the goal: not “show everything,” but “show enough for truth and repair.”

C. What must not be replayed?

The line is crossed when repetition turns suffering into currency. Replay is rarely necessary after basic facts are established. Rewatching for “analysis,” “micro-timing,” or emotional arousal is usually voyeurism.

A simple covenant heuristic:

- One view to understand, perhaps.

- Replays as entertainment, no.
- Zooming into harm or humiliation, no.
- Sharing the raw clip to rally outrage, no.

D. What do I say after I see it?

The audit rule forces claim typing: what did I observe, what do I infer, what do I not know?

Many viral reactions are built on the conversion of inference into certainty. The covenant also forces the dignity rule: do not turn victims into props or enemies into subhuman caricatures.

Outcome: The covenant does not “keep you uninformed.” It keeps you uninformed by voyeurism. It preserves witness and slows the conversion of tragedy into spectacle.

10.2 The tourist wreck: history, memory, and the selfie temptation

Scenario: A decades-old wreck becomes an “unlikely tourist attraction.” People visit, take photos, post dramatic captions, and treat the site as a scenic backdrop. The wreck is real history—sometimes tied to survival, death, or displacement.

This case tests something subtle: not propaganda, but the normalization of consuming damage as aesthetic experience.

Covenant decision points:

A. What is the site: memorial, artifact, or theme park?

The covenant begins with classification. Some sites function as memorials: they demand reverence. Others are artifacts: they can be visited with educational intent. The same physical object can be treated in different ways depending on posture.

The key covenant question: does my behavior increase understanding and respect, or does it use the object as a prop for my personal brand?

B. The selfie temptation

The selfie is not inherently immoral. But it is often a tell. It can turn “I learned something” into “I displayed myself.” A wreck—especially if linked to suffering—creates a strong aesthetic backdrop that can seduce the visitor into spectacle posture.

The dignity rule applies even when no living victim is present. History has dignity. Human loss has dignity. If the site is linked to death or harm, the covenant asks for restraint: fewer performative poses, fewer jokey captions, more context.

C. What is owed to memory?

A covenant-shaped visit would include:

- Basic historical context (what happened, why it matters).
- A tone of reverence if loss is involved.
- Avoidance of “thrill” framing that treats harm as entertainment.
- A refusal to monetize the site as pure aesthetic content.

Thought experiment: Imagine the same behavior at a gravesite. If it would feel grotesque there, the posture may be wrong here. This is not to collapse all wrecks into graveyards, but to test the moral reflex: do you treat damage as a playground?

Outcome: The covenant does not forbid tourism. It forbids conversion of history and harm into a hedonistic backdrop. It trains a visitor to be a steward of memory rather than a consumer of ruins.

10.3 The “free channel” interface: how UI scripts metaphysics

Scenario: A new television arrives with “free channels.” The interface launches into a grid of options, previews autoplay, “breaking” banners pop, recommended channels appear, and the system nudges continuous watching.

This case reveals an often-ignored truth: UI is moral pedagogy. The interface does not merely offer content; it scripts the posture of consumption.

Covenant decision points:

A. Defaults are theology

“Autoplay,” “preview,” “up next,” and “recommended” carry a metaphysical claim: the stream should govern your time. The covenant treats this as a contract offer. Accepting defaults is accepting governance by appetite.

A covenant-shaped response begins with reclaiming agency: disable autoplay, disable previews, remove “free channel” apps from the home screen if possible, and force intentional entry rather than passive drift.

B. Salience management

The interface highlights “trending,” “breaking,” and “most watched.” That is propaganda-adjacent salience control even when no ideology is involved. It teaches: what matters is what others watch. The covenant teaches: what matters is what is true and dignified.

A practical covenant move: reorganize the interface around intentional choices—education, art, local information, long-form—rather than “what’s hot.”

C. Momentum as governance

The simplest implementation is friction: make it slightly harder to drift. Put the remote away. Require a deliberate search. Set a time boundary. Use “watch lists” that you curate rather than the feed curating your attention.

Outcome: The covenant exposes interface as formation. It turns “free channels” from a neutral gift into a negotiable contract that can be declined or rewritten.

10.4 The viral claim: how the covenant handles uncertainty and disagreement

Scenario: A claim goes viral. It may involve a political figure, a crime, a scandal, a scientific finding, or a shocking statistic. It comes with confident framing and social pressure to react.

This case tests whether the covenant can survive urgency and tribe.

Covenant decision points:

A. Apply claim grammar immediately

Before arguing, label: observation, inference, speculation. Most viral claims are inference masquerading as observation. The covenant insists on saying: “I have not verified this,” or “This is a claim circulating,” rather than “This happened.”

B. Provenance before amplification

The audit rule says: do not amplify without provenance. If you cannot trace the claim to a reliable origin, you withhold. Withholding is not cowardice; it is stewardship.

C. Disagreement as instrument

The covenant requires seeking serious disagreement. Not mockable dissent, but the best opposing explanation. If you cannot state the strongest counterargument, you are not ready to share.

D. Disconfirmation condition

State what would change your mind. This prevents identity lock. It also makes revision possible without humiliation.

E. Dignity under uncertainty

Even if the claim is true, early discussion often turns into dehumanization. The covenant forbids contempt as a substitute for evidence. It also forbids treating victims or accused persons as entertainment.

Outcome: The covenant produces slower speech, fewer false positives, and a social norm that “waiting” is honorable. It directly degrades the velocity of propaganda and rumor.

10.5 A good-faith propaganda test: distinguishing persuasion from manipulation

Scenario: You encounter a channel, a documentary, a thread, or a model-generated essay. It feels persuasive. It might even contain real facts. How do you tell whether it is legitimate persuasion or manipulation?

A good-faith test must avoid paranoia. The goal is not to call everything propaganda. The goal is to preserve agency.

The covenant offers a multi-part test:

A. Does it preserve your freedom to revise?

Manipulation makes revision socially or psychologically costly. It mocks doubt, punishes questions, and frames dissent as betrayal. Legitimate persuasion invites scrutiny and allows revision without exile.

B. Does it disclose uncertainty honestly?

Manipulation launders inference as certainty, and uses emotional tone to cover gaps. Legitimate persuasion distinguishes what is known from what is proposed.

C. Does it treat opponents as persons?

Manipulation simplifies enemies into cartoons, monsters, or fools. Legitimate persuasion can critique sharply while preserving personhood.

D. Does it rely on salience tricks?

Manipulation uses repetition, selective footage, and constant urgency to keep you aroused. Legitimate persuasion can be compelling without constant heat.

E. Does it provide disconfirmation pathways?

Manipulation is self-sealing: it explains away counterevidence as part of the conspiracy. Legitimate persuasion names what would falsify it and shows how it would change under new evidence.

F. What fruit does it produce?

Manipulation often produces addiction to outrage, contempt, paranoia, and identity lock. Legitimate persuasion tends to produce clearer thinking, humility, and increased ability to act fittingly.

This test is “good-faith” because it does not assume malice. A creator can be sincere and still be propagandistic in effect. The covenant is less concerned with motive than with the structural outcome: does this environment keep you free and corrigible, or does it enclose you?

These case studies show the covenant at work as a practical instrument. It does not promise to eliminate spectacle, propaganda, or hedonism. It promises something more realistic and more necessary: it preserves witnesshood, protects dignity, and keeps conscience from being outsourced—exactly where modern media and modern AI try hardest to dissolve those goods.

11. The Practical Payoff: What Changes If the Covenant Takes Root

If the covenant is treated as a private virtue, its benefits remain private. But the covenant's ambition is public: it proposes a shift in default posture that changes what is rewarded, what is tolerated, and what is possible in shared life. The question is not whether every person will adopt it. The question is whether enough people—writers, readers, editors, curators, teachers, pastors, journalists, researchers, and ordinary citizens—adopt it strongly enough that it becomes a recognizable norm. Norms are powerful because they alter incentives. Once a norm exists, manipulative systems must work harder to bypass it. Once a norm exists, truth can travel farther without being turned into spectacle.

The covenant's payoff is not utopia. It is a set of realistic improvements: less velocity for propaganda, less humiliation harvest, more corrigibility, more patience, and a public square in which “I don't know yet” regains legitimacy. That would be a sea change precisely because the current environment punishes those stances.

11.1 A quieter public square with stronger claims

The first visible change would be quieter speech—not in the sense of silence, but in the sense of reduced heat. Heat is not the same as seriousness. A public square can be loud and still be shallow; it can be quieter and still be braver.

If the covenant takes root, three dynamics shift.

First, fewer claims are amplified at peak uncertainty. The lag rule and audit rule reduce the volume of instant reactions. That alone lowers noise.

Second, claims that do circulate carry more structure: provenance markers, uncertainty labeling, and disconfirmation conditions. This makes claims stronger. Not more dramatic—stronger. A strong claim is one that can survive scrutiny and be revised without collapse.

Third, correction becomes ordinary. In a quieter public square, people can correct without humiliation because the environment is not constantly demanding instant certainty. The norm becomes: publish, update, correct, and move forward. This reduces the need for rhetorical overkill and the fear-driven refusal to admit error.

The net result is paradoxical: less content, more meaning. Less reaction, more responsibility. Less virality, more durability.

11.2 Reduced susceptibility to engineered outrage

Engineered outrage works because it hijacks predictable human reflexes: the desire for moral clarity, the satisfaction of belonging, the rush of righteous anger, and the relief of having an enemy. The covenant reduces susceptibility not by demanding emotional numbness but by installing brakes and reframing triggers.

Several covenant rules attack outrage engineering directly.

The lag rule breaks the timing advantage of manipulators. Outrage campaigns often depend on the first wave: the moment before facts stabilize, when the narrative can be set by vivid fragments. If viewers wait, the campaign loses its highest-leverage window.

The audit rule interrupts “clip-to-conclusion” reasoning. Outrage engineering thrives on incomplete frames. Auditability forces context and provenance.

The no-oracle rule breaks identity lock to a stream. Outrage becomes hardest to resist when your belonging depends on staying hot. The covenant protects the freedom to revise and to pause.

The fruit rule adds a long-horizon test that outrage cannot pass. Outrage feels good short-term, but it often produces bitterness, paranoia, contempt, and social fracture. When communities learn to judge by fruit, outrage loses its glamour.

This does not eliminate anger. Some anger is appropriate. But it changes anger’s shape. Anger becomes a signal for scrutiny and action rather than a cue to binge and share. Outrage becomes less governable.

11.3 Restored dignity for victims and opponents

The most humane payoff is also the most urgent: restoration of dignity.

For victims, the covenant reduces humiliation harvest. It discourages replaying the worst moment, discourages sharing raw footage for emotional extraction, and discourages treating a person’s suffering as a public spectacle. In a covenant-shaped culture, a tragedy is not a content opportunity. It is a wound that demands restraint.

For opponents, the covenant reduces dehumanization. The dignity rule makes contempt language socially costly. The audit rule makes “enemy stories” harder to launder without evidence. The fruit rule exposes communities that feed on hatred as producing bad fruit.

This restoration matters because dignity is not a sentimental extra; it is a civilizational stabilizer. Societies that normalize humiliation become cruel. Societies that normalize dehumanization become violent. The covenant is a nonviolent technology: it makes cruelty harder by changing what is normal.

A subtle but crucial change would be that “the other side” becomes speakable again. Not lovable in an instant, not trusted automatically, but speakable without immediate contempt. That alone would reduce the memetic arms race that currently drives polarization.

11.4 A new norm for AI-assisted discourse: audit first, eloquence second

The centaur era’s greatest epistemic risk is synthetic authority: coherence mistaken for truth. If the covenant takes root, the norm changes from “Is it compelling?” to “Is it auditable?”

That shift would have several practical consequences.

First, AI outputs would be treated as drafts with explicit status labels. A reader would expect claim typing, uncertainty marking, and disconfirmation conditions—not as optional humility theater, but as standard form.

Second, models would be used more often for adversarial testing: generating counterarguments, listing what would falsify a claim, identifying missing provenance, and surfacing disagreement. This makes AI a tool for humility rather than for confirmation.

Third, publishing workflows would include logs and audit trails. Not endless paperwork, but minimal provenance and revision records. This would reduce the ease with which synthetic text can become synthetic authority.

Fourth, correction would become a first-class feature of AI-assisted writing. Because models can generate updates quickly, the danger is “fast error.” The covenant turns that into “fast correction” by making revision honorable and expected.

Over time, the public would learn a new literacy: they would become suspicious of rhetorical perfection without audit structure. Eloquence would no longer be enough to win. That is an enormous change.

11.5 Why this is compatible with joy, art, and legitimate entertainment

A predictable fear is that a truth-first covenant will make life gray: fewer jokes, fewer stories, fewer dramatic films, fewer guilty pleasures, fewer human moments. That fear misunderstands what the covenant is against.

The covenant is not against pleasure. It is against appetite as governor. It is not against art. It is against humiliation as entertainment and dehumanization as sport. It is not against drama. It is

against turning real suffering into a replayable thrill. It is not against persuasion. It is against manipulation that encloses revision and exploits salience.

Joy and art are not enemies of truth-first living. In many lives, they are the primary restorers of sanity. Beauty can re-train attention away from the feed's franticness. Humor can deflate propaganda's gravity and restore proportion. Stories can cultivate empathy and moral imagination. Legitimate entertainment can rest a weary mind and renew capacity for duty.

The covenant simply asks that entertainment remain within a moral frame:

- Fiction is not used as an excuse to consume real humiliation.
- Humor is not used as a license for cruelty.
- Art is not used as a laundering machine for dehumanization.
- Leisure is not used as a narcotic that destroys attention capacity.

In fact, the covenant likely increases joy over time, because it lowers the baseline of arousal and restores the pleasure of ordinary life: reading, conversation, music, prayer, craft, nature, friendship, and work done with care. The feed makes ordinary life feel insufficient. The covenant makes ordinary life possible again.

So the payoff is not a scolding culture. It is a steadier culture: less addicted to heat, more capable of truth, more protective of persons, and more able to use AI without bowing to it. If that takes root, the public square becomes quieter—but also stronger, kinder, and harder to hijack.

12. Conclusion: The Charge of Witnesshood

A century of broadcast and a decade of feed have not merely changed what we know; they have changed what we *are like while knowing*. They have trained reflexes of looking without obligation, speaking without cost, and consuming persons as content. In that training, voyeurism, “free TV,” hedonism, and propaganda become a coupled machine—self-financing, self-narrating, and escalating. The paper's claim has been simple but severe: the crisis is not only informational; it is formational. Therefore the remedy cannot be more content alone. It must be a metaphysical sea change: a reordered hierarchy of goods in which truth, dignity, and auditability outrank stimulation, speed, and tribal heat. The instrument for that change is a covenant for reading and writing in the centaur era—portable, voluntary, and practiced under pressure.

12.1 Restating the thesis in one paragraph

The broadcast century and the feed era function as metaphysical training: they normalize a posture of “looking” that drifts into voyeurism, make hedonistic appetite an implicit governor of attention, and enable propaganda through salience control and identity lock. “Free channels” and algorithmic interfaces are not neutral gifts; they are contracts that monetize arousal and reward spectacle. In the centaur era, AI accelerates the same dynamic by making persuasive synthesis cheap and abundant, raising the temptation to treat coherence as authority and models as oracles. The remedy is a truth-first reordering of attention—treating attention as a moral act, personhood as non-consumable, auditability as a public virtue, and speech as responsible repetition. A practical covenant—witness, audit, no-oracle, lag, fruit, dignity—translates these commitments into daily habits that reduce humiliation harvest, slow propaganda velocity, and preserve human agency under high-throughput persuasion.

12.2 The minimal ask: one covenant, practiced daily

The minimal ask is not that everyone become a philosopher, an editor, or an ascetic. It is that ordinary people adopt one small constitution and practice it daily—especially at the moments when the machine tries hardest to take custody of their attention.

The covenant is not a badge and not a weapon. It is a pre-commitment that can fit in one page and live at the front of your writing, your newsletters, your videos, your reading practice, and your AI workflows.

Practiced daily, it looks like this:

- Resist spectacle: do not replay humiliation; do not convert tragedy into entertainment.
- Demand auditability: provenance, uncertainty, disagreement, disconfirmation.
- Refuse oracles: keep conscience in human hands, models as assistants.
- Prefer lag: replace “live” with “later” for tragedy and conflict.
- Judge by fruit: long-horizon outcomes over heat and virality.
- Preserve dignity: persons remain persons, even when wrong.

That is it. Not a thousand rules. One covenant, repeated until it becomes reflex.

12.3 The sea change: attention re-ordered toward truth and love

If the covenant takes root, the sea change is not primarily political. It is spiritual and civic at the level where politics is downstream: attention.

Attention re-ordered toward truth means that coherence no longer confers authority, vividness no longer substitutes for evidence, and speed no longer bullies the conscience. It means that “I don’t know yet” becomes honorable again, because the public square recognizes that fast certainty is often manufactured.

Attention re-ordered toward love means that the human being on the screen is no longer raw material for entertainment or tribal bonding. Love, here, is not sentiment. It is the refusal to consume. It is restraint where the market demands extraction. It is criticism that preserves personhood. It is the discipline to treat enemies as morally accountable humans rather than as cartoon objects.

This reordering changes everything without claiming to fix everything. It does not eliminate conflict, evil, or deception. It restores the capacity to respond to them without becoming them.

12.4 Closing charge: build witnesses, refuse spectacle, document everything

The closing charge is practical.

Build witnesses: in yourself, in your writing, in your institutions. Train the ability to look responsibly, speak carefully, and revise honestly. Use AI as a witness-tool—summarize, compare, generate objections, surface uncertainty—without enthroning it as an oracle.

Refuse spectacle: stop feeding the humiliation harvest. Apply lag to tragedy. Do not give your attention to what deforms your heart and trains your contempt. Let the market learn that it cannot buy your soul for free.

Document everything: because the centaur era will be saturated with synthetic authority and high-velocity persuasion. Your defense is auditability—logs, provenance, correction pathways, and disconfirmation conditions. Not because paperwork saves, but because honesty is operational. A truth-first public life must be corrigible or it will become another enclosure.

The machine will keep offering stimulation-first metaphysics as the default. The only way a different world emerges is if enough people bind themselves to a different default: truth-first, dignity-first, witness-first. The work is not glamorous. It is daily. But that is how sea changes happen: one practiced covenant at a time, until “looking without obligation” is no longer normal, and witnesshood becomes the public art again.

References

Media ecology, spectacle, and the broadcast century

- Adorno, T. W., and Horkheimer, M. "The Culture Industry: Enlightenment as Mass Deception." In *Dialectic of Enlightenment*. 1944/1947.
- Boorstin, D. J. *The Image: A Guide to Pseudo-Events in America*. 1961.
- Debord, G. *The Society of the Spectacle* (La société du spectacle). 1967. [Wikipedia](#)
- Ellul, J. *The Technological Society* (La technique). 1954 (French); 1964 (English).
- Ellul, J. *Propaganda: The Formation of Men's Attitudes* (Propagandes). 1962 (French); 1965 (1st American ed.). [Wikipedia+1](#)
- McLuhan, M. *Understanding Media: The Extensions of Man*. 1964.
- Postman, N. *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. 1985. [Wikipedia](#)
- Sontag, S. *Regarding the Pain of Others*. 2003.

Attention economy, interface design, and behavioral capture

- Carr, N. *The Shallows: What the Internet Is Doing to Our Brains*. 2010.
- Eyal, N. *Hooked: How to Build Habit-Forming Products*. 2014.
- Fogg, B. J. *Persuasive Technology: Using Computers to Change What We Think and Do*. 2003.
- Hari, J. *Stolen Focus: Why You Can't Pay Attention—and How to Think Deeply Again*. 2022.
- Norman, D. A. *The Design of Everyday Things*. 1988 (rev. ed. 2013).
- Wu, T. *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. 2016. [Columbia Law School Scholarship Archive](#)
- Zuboff, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. 2019. [Harvard Business School](#)

Propaganda, disinformation, and salience control

- Allcott, H., and Gentzkow, M. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives*. 2017.
- Allport, G. W., and Postman, L. *The Psychology of Rumor*. 1947.
- Arendt, H. "Truth and Politics." In *Between Past and Future*. 1967.
- Bernays, E. L. *Propaganda*. 1928.
- Benkler, Y., Faris, R., and Roberts, H. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. 2018.
- Herman, E. S., and Chomsky, N. *Manufacturing Consent: The Political Economy of the Mass Media*. 1988.
- Lippmann, W. *Public Opinion*. 1922.
- Oreskes, N., and Conway, E. M. *Merchants of Doubt*. 2010.
- Tufekci, Z. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. 2017.

Vosoughi, S., Roy, D., and Aral, S. "The spread of true and false news online." *Science*. 2018. DOI: 10.1126/science.aap9559. [Science](#)

Wardle, C., and Derakhshan, H. *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. Council of Europe report. 2017. [Council of Europe+1](#)

Epistemic responsibility, virtue, and correction culture

Frankfurt, H. G. *On Bullshit*. 2005.

Haidt, J. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. 2012.

Kahneman, D. *Thinking, Fast and Slow*. 2011.

Milgram, S. *Obedience to Authority: An Experimental View*. 1974.

Polanyi, M. *Personal Knowledge: Towards a Post-Critical Philosophy*. 1958.

Popper, K. *Conjectures and Refutations: The Growth of Scientific Knowledge*. 1963.

Roberts, R. C., and Wood, W. J. *Intellectual Virtues: An Essay in Regulative Epistemology*. 2007.

Weil, S. "Reflections on the Right Use of School Studies with a View to the Love of God." In *Waiting for God*. 1951.

Lewis, C. S. *The Abolition of Man*. 1943.

AI, auditability, and governance in the centaur era

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., and Mané, D. "Concrete Problems in AI Safety." arXiv:1606.06565. 2016.

Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *FAccT*. 2021. DOI: 10.1145/3442188.3445922. [ACM Digital Library](#)

Christian, B. *The Alignment Problem: Machine Learning and Human Values*. 2020.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., and Crawford, K. "Datasheets for Datasets." arXiv:1803.09010. 2018. [arXiv](#)

Gebru, T., et al. "Datasheets for Datasets." *Communications of the ACM*. 2021. DOI: 10.1145/3458723. [ACM Digital Library](#)

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., and Gebru, T. "Model Cards for Model Reporting." *FAT* / FAccT*. 2019. DOI: 10.1145/3287560.3287596. [ACM Digital Library](#)

Noble, S. U. *Algorithms of Oppression: How Search Engines Reinforce Racism*. 2018.

O'Neil, C. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. 2016.

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., and Barnes, P. "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing." *FAccT*. 2020. [arXiv+1](#)

Russell, S. *Human Compatible: Artificial Intelligence and the Problem of Control*. 2019.

W3C. *PROV-DM: The PROV Data Model* (W3C Recommendation). 2013. [W3C+1](#)

Journalism practice, victims, and public responsibility

Kovach, B., and Rosenstiel, T. *The Elements of Journalism*. 2001 (rev. eds. later).

Plaisance, P. L. *Media Ethics: Key Principles for Responsible Practice*. (Any standard edition suitable for your style; cite edition used.)

