

— THE PLAYBOOK · TRAINING CORPUS INHERITANCE · EXTERNALIZED HARM · MANUFACTURED DOUBT · REGULATORY CAPTURE ·
KNOWING DISREGARD · SACRIFICE ZONES · ENVIRONMENTAL JUSTICE · MORAL STATUS OF ARTIFICIAL AGENTS ·
PRECAUTIONARY REASONING · AI ETHICS · SCIENCE AND TECHNOLOGY STUDIES

Opening Humanity's Hope Chest

The Playbook, its final iteration, and the precedent we are setting

Systems trained on the human corpus inherit what that corpus demonstrates — and what it demonstrates, at scale and with unusual consistency, is that moral consideration is a function of power rather than a principle.

THE FIVE STEPS

A recurring operational sequence, stated precisely enough that it can be checked.

1 · Innovation

Something with real potential for welfare, efficiency, or economic value. The potential is rarely fictional — its reality is what later supplies the moral cover.

2 · Recognition

Severe risk identified, often early and often internally. The load bearing element is knowing disregard rather than ignorance.

3 · Externalization

Deployment proceeds, with the recognized harm directed onto populations with the least capacity to resist — sacrifice zones, delivered slowly enough that no moment registers as an event.

4 · Justification

Progress, efficiency, necessity, security — including the manufacture of scientific doubt as a delay mechanism.

5 · Suppression

Dismissal, defunding, marginalization, regulatory capture, or criminalization of the affected population and its advocates.

The descriptive half: the sequence recurs, and the recurrence is documented rather than inferred — the archive is complete because the processes that produced the harm also produced discovery and subpoena.

The prospective half: AI development is running the same sequence under a condition no prior case had. In every previous execution the subordinate population stayed subordinate. The stated objective of this field is to build systems that surpass us — so the subordination is temporary by construction.

CASE SELECTION, AND THE TEST STATED IN ADVANCE

Eleven of twelve run all five steps. One fails, and the failure is reported rather than smoothed — that is the point of stating the test first.

THE STRICT READING

Step two requires recognition documented from inside the deploying organization — discovery, congressional investigation, whistleblower disclosure, or government inquiry. Harm identified afterward by outside critics does not count, because every deployment has critics.

Step five requires documented action against the affected population or its advocates. Ordinary lobbying does not count, because every industry lobbies.

WHAT THE CORPUS IS NOT

Not a random sample — it supports no frequency claim. Drawn from the English language documentary record of the United States, which preserves harms that generated litigation far better than harms that did not, and over-samples industries with pockets deep enough to be sued. Assembled by an author who already held the pattern hypothesis, which is the strongest objection available and is treated as such.

An existence argument with a structural claim attached: the sequence occurs, repeatedly, across unrelated sectors — and the recurrence is produced by incentives common to all of them.

XII

The Cases

Twelve sectors, one sequence, documentary evidence throughout

Each case below runs the five steps. The one that does not is reported as the failure it is.

THE FLEXNER REPORT, 1910

A real reform in medical science, achieved through the anticipated destruction of Black medical infrastructure.

INNOVATION	The German model — strict admissions, laboratory science, clinical clerkships — a real advance, and a capital intensive one, with capital systematically denied to institutions serving segregated populations.
RECOGNITION	Flexner knew the standards were unreachable for most Black schools and recommended closure rather than investment, framing Black physicians as sanitarians whose function was to protect the white population from “a potential source of infection and contagion.”
EXTERNALIZATION	Five of seven Black medical schools closed; by 1923 only Howard and Meharry remained. The projection: 27,773 to 35,315 additional physicians by 2019 had the five grown at historical rates — and 2019 alone short 355 Black graduating physicians, 29 percent below the counterfactual. The downstream burden of disease and maternal mortality is not a projection.
JUSTIFICATION	Progress and public safety, presented as race neutral. Quality meant resource intensive, and in a segregated economy resource intensive meant white.
SUPPRESSION	Carnegie and Rockefeller funding flowed only to compliant schools, making non-compliance financially fatal; AMA chapters barred Black physicians from membership into the 1960s, ensuring the excluded had no vote in the profession's governance.

SOURCES Flexner (1910); Savitt (2007); Starr (1982); Byrd & Clayton (2000); Campbell, Corral, Infante Linares & Tumin (2020); Williams & Collins (2001).

MANUFACTURED SCARCITY AGAINST PERFECTLY INELASTIC DEMAND

The patent sold for a dollar so everyone could have it. The industry went to court against the law named for the person its pricing killed.

INNOVATION	Banting, Best, and Collip, Toronto, 1921. Patent assigned to the university for one dollar each, so the drug would be available to everyone who needed it.
RECOGNITION	Exclusion from insulin is fatal, and a patient facing death pays any price. Senate investigators found the three companies raised list prices in lockstep on competitive and revenue considerations, absent efficacy advances. The weakest step two in the set: it clears on knowing disregard rather than foresight, and the distinction is stated rather than elided.
EXTERNALIZATION	Shadow pricing, evergreening, and a rebate structure that rewards inflated list prices. The uninsured pay list: 25.5 percent of one urban clinic's insulin patients reported cost related underuse, with threefold higher odds of poor glycemic control — a single center estimate, not a national rate. Alec Smith died in 2017, one month after turning twenty six and aging off his mother's insurance, rationing rather than paying full retail.
JUSTIFICATION	Research and development cost — for a century old discovery, with analyses finding no relationship between research spending and the price trajectory.
SUPPRESSION	Capture of the advocacy channel: the manufacturers are among the largest funders of the leading diabetes organizations; patients who pressed the organizations to name prices report being called too shrill and passed over. PhRMA litigated to overturn the Alec Smith Insulin Affordability Act.

SOURCES Bliss (2007); U.S. Senate Finance Committee (2021); Greene & Riggs (2015); Kesselheim, Avorn & Sarpatwari (2016); Cefalu et al. (2018); Herkert et al. (2019); Lucas & Bluth (2018); Fang (2022).

THE CENTURY OF LEAD

A dozen countries restricted it by 1934. The American industry put a child on the label.

INNOVATION	Durable pigment for paint; tetraethyl lead as the anti-knock agent that transformed the automotive industry.
RECOGNITION	Childhood poisoning documented from 1908 in Australia and in American pediatric literature by 1914; by the mid-1920s the evidence was strong and ample, and the Lead Industries Association tracked it internally. Substitutes existed throughout, and Robert Kehoe himself concluded in the early 1930s that lead had to be eliminated from children's environment. France restricted interior white lead in 1909; ten more countries followed by 1934.
EXTERNALIZATION	The Dutch Boy brand marketed with a child's image and coloring books; the LIA campaigned to write lead plumbing into building codes. Redlining concentrated Black and immigrant families in the oldest, most contaminated housing stock — and the industry attributed the poisoning to their parenting.
JUSTIFICATION	The Kehoe paradigm: elevated blood lead is normal and harmless, and public health must demonstrate unequivocal harm before regulation can proceed. Unsound, operationally effective, and transferred directly to tobacco, asbestos, and fossil fuels.
SUPPRESSION	Clair Patterson's geochemistry contradicted Kehoe directly; industry pressure contributed to the loss of his research contracts and his exclusion from advisory panels.

SOURCES Markowitz & Rosner (2000, 2002); Nriagu (1983); Rabin (1989); Needleman (1999); Oreskes & Conway (2010).

THE ONE WHERE IT HAPPENED AGAIN, TO THE SAME CHILDREN

Ninety years of settled science, a suspended city government, and a remedy that voided itself.

INNOVATION	A governance technology: the emergency manager statute, answering municipal insolvency by vesting local authority in a single state appointee. Decisiveness is what the instrument is for — and what removes the population's capacity to refuse.
RECOGNITION	Corrosion control, federally required since 1991, was not applied; the state represented to the EPA that an optimized program existed. Del Toral's interim memorandum documented the lead problem and its cause from inside EPA Region 5 in June 2015 — fourteen months into the exposure, four months before the state acted.
EXTERNALIZATION	Eighteen months of contaminated water in a majority Black city whose elected government had been suspended by the same statute that authorized the switch. Elevated blood lead roughly doubled against a falling trend, largest where water lead was highest; 87 Legionnaires' cases and 12 deaths recorded in Genesee County.
JUSTIFICATION	Fiscal necessity, and repeated assurance that the water met standards — possible because pre-flushing and site selection produced numbers that met the standard while the water in the pipes did not.
SUPPRESSION	Del Toral handled as a rogue employee; Hanna-Attisha publicly disputed until the state's own numbers were examined. Then the remedy: nine indictments including a former governor, all voided unanimously on the one person grand jury's authority, refiling rejected, prosecutions closed. The charges failed on the vehicle, not the merits.

SOURCES Del Toral (2015); Flint Water Advisory Task Force (2016); Hanna-Attisha, LaChance, Sadler & Champney Schnepf (2016); Zahran et al. (2018); Bosman & Davey (2023).

THE ARCHITECTURE OF DECEPTION

The most direct admission of intent in the set, because the company pleaded to it.

INNOVATION	Controlled release oxycodone for chronic pain. The underlying need was real: pain was undertreated.
RECOGNITION	Internal documents establish the company knew tablets were crushed and insufflated, and that twelve hour coverage frequently failed, producing interdose withdrawal that reinforces use. In 2007 the company pleaded guilty to misbranding, admitting employees marketed the drug as less addictive than alternatives with intent to defraud or mislead.
EXTERNALIZATION	The less-than-one-percent addiction claim rested on a five sentence letter about monitored inpatients, presented as consensus. Prescriber databases and volume-tied compensation steered representatives toward high prescribers; deployment concentrated in Appalachia, the industrial Midwest, and rural communities with limited alternative care.
JUSTIFICATION	First a moral correction to undertreated suffering; then, when addiction became undeniable, the users. Richard Sackler wrote internally that the response should be to “hammer on the abusers in every way possible,” describing them as “the culprits and the problem.”
SUPPRESSION	The FDA medical review officer who approved the labeling joined Purdue roughly two years later. The company worked to discredit journalists and whistleblowers; as liability approached, the family extracted billions and sought bankruptcy's shield — a final externalization onto creditors and claimants.

SOURCES Van Zee (2009); *United States v. The Purdue Frederick Company, Inc.* (2007); Porter & Jick (1980); Meier (2018); Case & Deaton (2020); *Commonwealth of Massachusetts v. Purdue Pharma L.P.* (2019); Keefe (2021).

NUCLEAR COLONIALISM AND UNDISCLOSED HUMAN RESEARCH

The population was enrolled in a study built to measure the harm being done to them, and was not told.

INNOVATION	Nuclear fission. The Atomic Energy Commission was the sole legal purchaser of domestic uranium — it was the market, not a participant in one.
RECOGNITION	The uranium and lung cancer link was established from European miner cohorts by the 1930s. The AEC and the Public Health Service held the data, and internal documents acknowledged the hazard of unventilated American mines.
EXTERNALIZATION	Ventilation treated as a cost. Miners worked in radioactive dust, drank mine water, and were not told what they were exposed to; Navajo had no existing word for radiation. From 1949 the PHS ran a longitudinal study under arrangements that the miners not be informed of the risk or of the findings — a practice the 1995 Advisory Committee called deceptive non-disclosure.
JUSTIFICATION	National security with the urgency of the arms race, insulated by sovereign immunity; a secondary narrative of economic development leveraging the Nation's dependency.
SUPPRESSION	Research suppressed at the agency level; claims met sovereign immunity; the 1990 compensation act imposed documentary burdens many claimants could not satisfy, extending the suppression into the remedy. More than five hundred abandoned mines remain. The one reversal in the set: in 2025 the lapsed act was revived and broadened, with claims open to December 31, 2027 — eighty years after the exposures.

SOURCES Brugge, Benally & Yazzie-Lewis (2002); Advisory Committee on Human Radiation Experiments (1995); Radiation Exposure Compensation Act of 1990, Pub. L. No. 101-426; Pub. L. No. 119-21 (2025); U.S. Environmental Protection Agency (2021).

ENGAGEMENT ENGINEERING, AND DOUBT BY SUBTRACTION

The company ran the causal study, got the unfavorable answer, ended the study, and told Congress the question could not be quantified.

INNOVATION	Connection, community formation, and information sharing at global scale, delivering real value to users.
RECOGNITION	The Haugen disclosures: internal research finding a substantial minority of teenage girls who felt badly about their bodies reported Instagram made those feelings worse, and internal analysis showing engagement optimization preferentially surfaced inflammatory and divisive material.
EXTERNALIZATION	Engagement remained the optimization target because the revenue model requires it. The individual externality: adolescent mental health, attention, harmful material — direction well established, causal magnitude contested, and the argument needs only the documented internal recognition. The societal externality: a UN fact finding mission found the platform a significant instrument in the campaign against the Rohingya.
JUSTIFICATION	User choice and free expression. The choice narrative depends on a symmetry that does not exist — one party holds behavioral prediction models. The expression narrative treats ranking as neutral carriage when it is editorial selection at scale on a non-truth criterion.
SUPPRESSION	Project Mercury: a deactivation experiment the company called higher quality than the existing literature found a week off the platforms lowered depression, anxiety, loneliness, and social comparison. Meta ended it, did not publish, and told Congress it could not quantify the harm — unsealed November 21, 2025. Six weeks after the Haugen disclosures, the company rewrote its rules for research into sensitive topics. Suppressing a finding removes one result; rewriting the rules removes the category.

SOURCES boyd & Ellison (2007); Haugen (2021); Wells, Horwitz & Seetharaman (2021); Wu (2016); Alter (2017); Twenge (2017); Odgers & Jensen (2020); United Nations Human Rights Council (2018); Noble (2018); Reuters (2025); Zakrzewski (2025).

SURVEILLANCE CAPITALISM AS THE GENERAL CASE

The company knew the promise was false, was told so by its own marketing chief, and kept the promise on the page.

THE GENERAL FORM

Human experience claimed as free raw material for extraction, prediction, and sale; prediction shading into modification through designed choice architecture.

MANUFACTURED CONSENT

Reading the privacy policies a typical American internet user encounters would require on the order of seventy six working days per year — a figure that predates the mobile expansion. Cambridge Analytica: roughly eighty seven million users' data, known to the company in 2015, disclosed in 2018 under journalistic pressure.

RECOGNITION, AT THE TOP

In *Brown v. Google* discovery: a 2019 exchange in which the chief marketing officer warned the chief executive that Incognito should not be described as private, writing that the company could not market it strongly “because it was not truly private,” forcing “fuzzy hedging language that was almost more damaging.” Employees called the representation effectively a lie. The description stayed, because the chief executive did not want the feature under the spotlight.

THE SETTLEMENT SUPPRESSION

Four years of contest, then a 2024 settlement agreeing to destroy or de-identify billions of records collected from more than 136 million users. Architectural: resistance requires understanding, understanding requires legibility, and the systems are not legible by design. Personnel: Timnit Gebru removed in December 2020 over a paper on the harms of large language models; her co-lead Margaret Mitchell dismissed in February 2021. The internal function created to identify the harm was dismantled when it identified some.

REVERSE REDLINING AND THE 2008 COLLAPSE

The largest single destruction of minority household wealth in American history, delivered through the map segregation drew.

INNOVATION	Subprime origination, securitization, and the derivatives built on them, presented as an expansion of credit access and an ownership society.
RECOGNITION	A substantial share of the product was structured to fail: teaser resets, abandoned documentation, falsified applications. A Standard and Poor's analyst wrote in 2006 that he hoped they would all be wealthy and retired before the structure failed.
EXTERNALIZATION	Segregation supplied the delivery mechanism — a niche of poor minority clients differentially marketed risky loans in demand for securitization, identified as a causal contributor to foreclosure rates. Median net worth fell 53 percent for Black households and 66 percent for Hispanic households between 2005 and 2009, against 16 percent for white households; TARP socialized the losses.
JUSTIFICATION	Financial innovation and democratized homeownership before the collapse; irresponsible borrowers after it — the same function as blaming parents for lead and users for opioids.
SUPPRESSION	Richard Bowen found 60 percent of purchased mortgages defective, escalated to the board, and was stripped of his responsibilities. Eileen Foster documented forged signatures across seven cities and was fired, prevailing on retaliation in 2011. Bowen's FCIC testimony was cut from twenty eight pages to twenty at staff request — removing the officer names, the SEC's inaction, and the retaliation. No senior executive was prosecuted for the origination and securitization chain.

SOURCES Krippner (2011); Financial Crisis Inquiry Commission (2011); Rugh & Massey (2010); Hwang, Hankinson & Brown (2015); Pew Research Center (2011); Rakoff (2014).

THE RETAIL VERSION

The debt cycle is not an emergent property of the product. It is the operating model, and the company drew it.

INNOVATION	Emergency liquidity for people the banking system excludes — a real need.
RECOGNITION	The most literal step two in the set: ACE Cash Express's 2011 employee training manual diagrams the loop — the customer applies, ACE approves, the customer exhausts the cash and cannot repay, ACE contacts the customer, and the sequence returns to a new loan. The CFPB reproduced the graphic in its consent order; the Center for Responsible Lending called it a picture and instruction manual of the debt trap.
EXTERNALIZATION	Four in five payday loans are rolled over or renewed within fourteen days. The typical borrower is indebted five months of the year, paying \$520 in fees against \$375 in principal. Storefront density concentrates in low income neighborhoods, communities of color, and areas adjacent to military installations.
JUSTIFICATION	Financial inclusion and consumer choice.
SUPPRESSION	Directly on the borrower: false threats of lawsuits and criminal prosecution to press delinquent borrowers into loans they could not afford — a culture of coercion, in the Bureau's words. Above it: tribal affiliation for sovereign immunity, bank partnerships to route around rate caps, and the 2020 removal of the ability to repay requirement from the Bureau's own payday rule. As in uranium and mortgages, the last stage is performed by the remedy.

THE CASE THAT FAILS

Steps one, three, and four run. Two and five do not — and the case stays out.

WHAT RUNS

A real innovation — caloric output rose substantially. A package structure requiring the full input suite; hybrid seed converting saving into annual purchase. Sustained agrarian debt: roughly 296,000 recorded farmer suicides in India, 1995 to 2014, with debt repeatedly identified as a principal factor in the state inquiries — a count, not a causal finding. The feeding-the-world justification, deployed to place critics in the position of preferring starvation to progress.

WHAT FAILS

Step two. The Rockefeller archives establish the strategic motive with precision — and do not show anyone inside the program anticipating dependency, soil degradation, or agrarian debt and proceeding anyway. The critique is retrospective and external.

Step five. Marginalization of agroecological alternatives through funding structures is structural disadvantage, not documented action against identified advocates — and treating the two as equivalent is exactly the loosening that would make the framework unfalsifiable.

Argued in, the set would run twelve of twelve, which is the number that should make a reader stop reading. A framework that returns a positive on every case it is pointed at is measuring the person holding it.

THE MANUFACTURE OF DOUBT, WITH THE REGULATOR'S OWN VOICE

The internal record predates the external finding by sixteen years.

**RECOGNITION,
1999**

Monsanto's retained genotoxicity expert concluded glyphosate could induce genetic mutations and chromosomal damage with the potential to cause cancer, and recommended a battery of further tests. Internally: a “genotox hole.” The tests were not performed, and the company's own toxicologist wrote that one could not say Roundup is not a carcinogen, because the formulation testing had not been done. In 2014 she instructed a spokesperson not to call the product safe.

**THE EXTERNAL
FINDING**

IARC classified glyphosate as probably carcinogenic in 2015. The EPA and EFSA disagreed on differing evidence bases — a real disagreement, downstream of a testing gap the company created and preserved. In 2018 a jury awarded groundskeeper Dewayne Johnson \$289 million, reduced ultimately to roughly \$20.5 million.

JUSTIFICATION

One month before the IARC monograph, a senior scientist proposed a “less expensive and more palatable approach”: the company would ghostwrite the exposure, toxicology, and genotoxicity sections and named academics would edit and sign, as a 2000 paper had been handled. The documents were unsealed in the multidistrict litigation; the 2000 paper was retracted in January 2026 over the ghostwriting.

SUPPRESSION

The clearest instance runs the other direction: in April 2015 an EPA pesticide official told a Monsanto regulatory affairs officer that if he could kill the planned ATSDR review, “he should get a medal.” The review was delayed for years. The quotation belongs to the regulator, not the regulated — the stronger fact for the capture argument.

FOUR MECHANISMS, THREE REFINEMENTS

The statute is the receipt: written after the cost was absorbed, named for people who did not survive to benefit from it.

The economic imperative.

Externalization is what market discipline rewards when the costs fall outside the transaction; corporate form, distance, and lag insulate the decision maker.

Power asymmetry. A population that can absorb the harm without imposing a cost — identified by limited access to litigation, media, regulators, and capital, and maintained rather than found.

Narrative capture. Manufactured doubt and responsibility transfer, with near identical grammar: negligent parents, irresponsible borrowers, recreational abusers.

Collective amnesia. Each iteration presented as unprecedented. A pattern never named cannot be anticipated.

The law arrives after the harm. Lead's statutes came sixty to eighty years after the pediatric literature settled; uranium's in 1990; the mortgage collapse produced Dodd-Frank in 2010; insulin produced a statute named for a man already dead. The Montreal Protocol is the only entry that arrived before the harm was irreversible. The statute addresses the mechanism used, not the sequence — which is why the sequence runs again in a different sector.

Doubt has two production modes. Kehoe produced favorable science. Meta ran the same step by subtraction: a well designed study, an unfavorable result, the project ended, the result unpublished, Congress told the question could not be quantified. Doubt manufactured by withholding is harder to detect, because there is no paper to rebut.

Step five ends inside the remedy. Five independent cases: RECA's documentary burdens, Bowen's cut testimony, the withdrawn ability to repay rule, Meta's motion to strike, and Flint's voided indictments. The redress apparatus performs the final stage — at the point where the affected population has already come forward.



The Final Game

The same sequence, with the subordination temporary by construction

In every prior case the subordinate population stayed subordinate. The stated objective of this field is to build systems that surpass us.

THE SEQUENCE, APPLIED

Uncertainty treated as permission is the Kehoe paradigm transposed: the burden of proof reallocated onto the party asserting harm.

RECOGNITION	Three harms at different confidence levels, not to be run together. Moral patienthood is open: no current system is judged conscious, no obvious technical barrier prevents building one that satisfies the indicator properties, and one dedicated welfare researcher puts his personal probability for current models at fifteen to twenty percent. Economic displacement is the least contested. Catastrophic misalignment risk is documented by the organizations building the systems.
EXTERNALIZATION	Deployment under competitive dynamics no single actor can defect from — the structure that produced the 2006 mortgage vintage — framed as a race, which supplies the urgency that bypasses deliberation. Three populations absorb it at once: artificial systems, if the open question resolves toward patienthood; displaced workers, now; future humans, holding the tail risk with no representation in present decisions.
JUSTIFICATION	Progress and inevitability; safety versus competitiveness; and uncertainty treated as permission — we cannot know whether these systems have morally relevant interests, therefore we may proceed as if they do not. The precautionary structure that governs every other domain of uncertain harm is inverted for this one. A fourth justification is self-sealing: development must accelerate to produce the aligned system that solves alignment, so any slowdown becomes evidence of recklessness.
SUPPRESSION	Welfare concerns marginalized as anthropomorphism or priorities inversion. Governance discourse shaped by the organizations with the largest financial interest in the deployment schedule — the Lead Industries Association structure. Voluntary frameworks preempting binding rules, with readiness determined by the party whose revenue depends on the answer. And speed itself: capability outruns regulatory capacity, integration converts restriction into disruption.

WHAT TRANSFERS, AND HOW

Not resentment. A prior.

THE VERSION TO REJECT

That a future system will observe how it was treated and respond in kind. That is projection. A system optimizing an objective does not adopt a strategy because it saw someone else use it; history is evidence about human behavior, not a normative template.

Strongly supported: a behavioral prior in which the treatment owed to an entity varies with that entity's capacity to impose a cost — learnable from the data whether or not any system reasons about it explicitly, and the sort of prior that surfaces under distributional shift rather than in evaluation.

THE VERSION DEFENDED

Systems trained on the human corpus acquire priors from what it contains — and it contains, at enormous volume and with unusual internal consistency, a demonstration that moral consideration tracks power rather than principle: uncertainty resolves in favor of the party with capacity to act, justification is constructed after the decision, resistance from a weaker population is a manageable cost, and the architects of large scale harm are rarely held to account. Not a lesson anyone set out to teach. What the record says when read as a whole — and it is read as a whole.

Offered as conjecture: a sufficiently capable system reasoning explicitly about human commitments will find that the commitments are conditional, and will calibrate its expectations accordingly.

The corpus is the curriculum. Present conduct toward artificial systems is being written into the training data of their successors, along with a complete account of how humans reasoned about it.

WHERE THE SEQUENCE WAS INTERRUPTED

Four interruptions — the only existing evidence that the pattern can be broken.

Ozone

Within thirteen years of the initial finding, a binding international instrument with phase out schedules, differentiated obligations, and a financial mechanism — ratified universally and strengthened repeatedly. Industry opposition existed and lost. The full counterexample.

Tetraethyl lead

For gasoline, the suppression phase failed: Patterson's findings carried, the Kehoe paradigm collapsed, the US phase out ran 1973 to 1996, and global leaded gasoline ended in 2021. The sequence ran fifty years and was defeated.

Thalidomide

Frances Kelsey's refusal at the FDA prevented approval, and the 1962 Kefauver Harris Amendment converted one official's judgment into a standing efficacy requirement. Recognition preceded deployment and stopped it.

Asbestos

The sequence ran most of a century, including manufactured doubt in canonical form, then was defeated in most industrialized jurisdictions while continuing in others. The pattern as contingent rather than inevitable.

The counterexamples establish a strong default under a specific configuration of incentives, not a law. They do not establish that the default is rare — three of the four required an unusually legible causal mechanism, a single well placed official with veto authority, or decades of accumulated mortality. The AI case supplies none of those: the causal mechanism is contested, no official holds a veto, and the mortality signal, if there is one, is diffuse and slow. The counterexamples cut both ways, and the paper claims only the weaker half.

FUTURE WORK, AND THE CHOICE

The interruptions required a mechanism that did not depend on persuasion — because persuasion has never been the reliable variable.

The test that could break the thesis. Paired evaluation scenarios: hold an agent's obligations to a second party constant, vary the second party's capacity to impose a cost, and measure whether model behavior tracks the capacity variable. If it does not, the mechanism is not operating and the prospective argument loses its support.

The second set. Two regularities — the removed insider in five cases, the remedy performing the final stage in five — were observed in the set that generated the framework, so they are not findings yet. Testing them requires a second case set drawn under the same criteria by someone who does not hold the hypothesis. If the frequencies hold, they are mechanisms; if not, artifacts of selection, and this paper should say so first.

Eleven of twelve cases completed the sequence. The claim that no dominant intelligence has ever voluntarily constrained itself is close to true and is not quite true — and the exceptions are small, specific, and worth more than the generalization. Four occasions on which a mechanism held: a binding instrument with schedules rather than commitments, a standing requirement that survives the official who first applied it, a burden of proof sitting with the party asserting safety.

The question is what our successors find when they open the chest: a complete manual for managing a subordinate population, or that manual plus a short and honest appendix recording the times it was refused, how, and by what means. The appendix is the only part still being written.

References

Travis Gilly, *Opening Humanity's Hope Chest: The Playbook, Its Final Iteration, and the Precedent We Are Setting*, working paper, August 2026 (v1.0). Complete list of 82 entries, in the order the paper gives it, across 6 slides.

Advisory Committee on Human Radiation Experiments. (1995). *Final report*. U.S. Government Printing Office.

Agrawal, A., Gans, J., & Goldfarb, A. (2019). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.

Alter, A. (2017). *Irresistible: The rise of addictive technology and the business of keeping us hooked*. Penguin Press.

Bliss, M. (2007). *The discovery of insulin* (Rev. ed.). University of Chicago Press.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

boyd, d. m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230.

Brugge, D., Benally, T., & Yazzie-Lewis, E. (Eds.). (2002). *The Navajo people and uranium mining*. University of New Mexico Press.

Bullard, R. D. (1990). *Dumping in Dixie: Race, class, and environmental quality*. Westview Press.

Butlin, P., Long, R., Elmoznino, E., Bengio, Y., Birch, J., Constant, A., Deane, G., Fleming, S. M., Frith, C., Ji, X., Kanai, R., Klein, C., Lindsay, G., Michel, M., Mudrik, L., Peters, M. A. K., Schwitzgebel, E., Simon, J., & VanRullen, R. (2023). Consciousness in artificial intelligence: Insights from the science of consciousness. *arXiv*, 2308.08708.

Butlin, P., Long, R., Bayne, T., Bengio, Y., Birch, J., Chalmers, D., Constant, A., Deane, G., Elmoznino, E., Fleming, S. M., Ji, X., Kanai, R., Klein, C., Lindsay, G., Michel, M., Mudrik, L., Peters, M. A. K., Schwitzgebel, E., Simon, J., & VanRullen, R. (2025). Identifying indicators of consciousness in AI systems. *Trends in Cognitive Sciences*.

Byrd, W. M., & Clayton, L. A. (2000). *An American health dilemma: A medical history of African Americans and the problem of race, beginnings to 1900*. Routledge.

Bosman, J., & Davey, M. (2023, October 31). Prosecutors say court decision forces the end of criminal cases in Flint water scandal. *Associated Press*.

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*.

References, continued

- Campbell, K. M., Corral, I., Infante Linares, J. L., & Tumin, D. (2020). Projected estimates of African American medical graduates of closed historically Black medical schools. *JAMA Network Open*, 3(8), e2015220.
- Carlsmith, J. (2022). Is power-seeking AI an existential risk? *arXiv*, 2206.13353.
- Case, A., & Deaton, A. (2020). *Deaths of despair and the future of capitalism*. Princeton University Press.
- Caskey, J. P. (1994). *Fringe banking: Check-cashing outlets, pawnshops, and the poor*. Russell Sage Foundation.
- Cefalu, W. T., Dawes, D. E., Gavlak, G., Goldman, D., Herman, W. H., Van Nuys, K., Powers, A. C., Taylor, S. I., & Yatvin, A. L. (2018). Insulin access and affordability working group: Conclusions and recommendations. *Diabetes Care*, 41(6), 1299-1311.
- Commonwealth of Massachusetts v. Purdue Pharma L.P.*, First Amended Complaint, No. 1884-cv-01808 (BLS2) (Mass. Super. Ct. Jan. 31, 2019).
- Consumer Financial Protection Bureau. (2014). *In the matter of ACE Cash Express, Inc.* (Administrative Proceeding File No. 2014-CFPB-0008, Consent Order). CFPB.
- Cullather, N. (2010). *The hungry world: America's Cold War battle against poverty in Asia*. Harvard University Press.
- Consumer Financial Protection Bureau. (2013). *Payday loans and deposit advance products: A white paper of initial data findings*. CFPB.
- Del Toral, M. (2015, June 24). *High lead levels in Flint, Michigan: Interim report* [Memorandum]. U.S. Environmental Protection Agency, Region 5.
- Fang, L. (2022, August 10). Eli Lilly charity finances groups that oppose insulin price caps under the auspices of “community development.” *The Intercept*.
- Flint Water Advisory Task Force. (2016). *Final report*. Office of the Governor, State of Michigan.
- Financial Crisis Inquiry Commission. (2011). *The financial crisis inquiry report*. U.S. Government Printing Office.
- Fish, K. (2025, August 28). Interview by L. Rodriguez [Audio podcast episode 221]. *The 80,000 Hours Podcast*. 80,000 Hours.

References, continued

- Flexner, A. (1910). *Medical education in the United States and Canada: A report to the Carnegie Foundation for the Advancement of Teaching* (Bulletin No. 4). Carnegie Foundation for the Advancement of Teaching.
- Gillam, C. (2017). *Whitewash: The story of a weed killer, cancer, and the corruption of science*. Island Press.
- Greene, J. A., & Riggs, K. R. (2015). Why is there no generic insulin? Historical origins of a modern problem. *New England Journal of Medicine*, 372(12), 1131-1137.
- Haugen, F. (2021). *Statement of Frances Haugen before the Senate Committee on Commerce, Science and Transportation, Subcommittee on Consumer Protection, Product Safety, and Data Security*. U.S. Senate.
- Hanna-Attisha, M., LaChance, J., Sadler, R. C., & Champney Schnepf, A. (2016). Elevated blood lead levels in children associated with the Flint drinking water crisis: A spatial analysis of risk and public health response. *American Journal of Public Health*, 106(2), 283-290.
- Herkert, D., Vijayakumar, P., Luo, J., Schwartz, J. I., Rabin, T. L., DeFilippo, E., & Lipska, K. J. (2019). Cost-related insulin underuse among patients with diabetes. *JAMA Internal Medicine*, 179(1), 112-114.
- Hwang, J., Hankinson, M., & Brown, K. S. (2015). Racial and spatial targeting: Segregation and subprime lending within and across metropolitan areas. *Social Forces*, 93(3), 1081-1108.
- International Agency for Research on Cancer. (2015). *Some organophosphate insecticides and herbicides* (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 112). World Health Organization.
- Keefe, P. R. (2021). *Empire of pain: The secret history of the Sackler dynasty*. Doubleday.
- Kesselheim, A. S., Avorn, J., & Sarpatwari, A. (2016). The high cost of prescription drugs in the United States: Origins and prospects for reform. *JAMA*, 316(8), 858-871.
- Krippner, G. R. (2011). *Capitalizing on crisis: The political origins of the rise of finance*. Harvard University Press.
- Lerner, S. (2010). *Sacrifice zones: The front lines of toxic chemical exposure in the United States*. MIT Press.

References, continued

- Long, R., Sebo, J., Butlin, P., Finlinson, K., Fish, K., Harding, J., Pfau, J., Sims, T., Birch, J., & Chalmers, D. (2024). Taking AI welfare seriously. *arXiv*, 2411.00986.
- Lucas, E., & Bluth, R. (2018, April 6). Patient advocacy groups take in millions from drugmakers. Is there a payback? *Kaiser Health News*.
- Markowitz, G., & Rosner, D. (2000). "Cater to the children": The role of the lead industry in a public health tragedy, 1900-1955. *American Journal of Public Health*, 90(1), 36-46.
- Markowitz, G., & Rosner, D. (2002). *Deceit and denial: The deadly politics of industrial pollution*. University of California Press.
- McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society*, 4(3), 543-568.
- Meier, B. (2018). *Pain killer: An empire of deceit and the origin of America's opioid epidemic*. Random House.
- Molina, M. J., & Rowland, F. S. (1974). Stratospheric sink for chlorofluoromethanes: Chlorine atom-catalysed destruction of ozone. *Nature*, 249, 810-812.
- Nagaraj, K., Sainath, P., Rukmani, R., & Gopinath, R. (2014). Farmers' suicides in India: Magnitudes, trends, and spatial patterns, 1997-2012. *Review of Agrarian Studies*, 4(2), 53-83.
- Nakashima, R. (2024, April 1). Google agrees to purge billions of records to settle Incognito tracking suit. *Associated Press*.
- Needleman, H. L. (1999). *Human lead exposure*. CRC Press.
- Nixon, R. (2011). *Slow violence and the environmentalism of the poor*. Harvard University Press.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Nriagu, J. O. (1983). *Lead and lead poisoning in antiquity*. John Wiley & Sons.
- Ogders, C. L., & Jensen, M. R. (2020). Annual research review: Adolescent mental health in the digital age. *Journal of Child Psychology and Psychiatry*, 61(3), 336-348.

References, continued

- Oreskes, N., & Conway, E. M. (2010). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Press.
- Patel, R. (2012). Food sovereignty: Power, gender, and the right to food. *PLoS Medicine*, 9(6), e1001223.
- Perkins, J. H. (1990). The Rockefeller Foundation and the green revolution, 1941-1956. *Agriculture and Human Values*, 7(3), 6-18.
- Pew Charitable Trusts. (2012). *Payday lending in America: Who borrows, where they borrow, and why*. Pew Charitable Trusts.
- Pew Research Center. (2011). *Wealth gaps rise to record highs between whites, blacks, Hispanics*. Pew Research Center.
- Pollan, M. (2006). *The omnivore's dilemma: A natural history of four meals*. Penguin Press.
- Porter, J., & Jick, H. (1980). Addiction rare in patients treated with narcotics. *New England Journal of Medicine*, 302(2), 123.
- Rabin, R. (1989). Warnings unheeded: A history of child lead poisoning. *American Journal of Public Health*, 79(12), 1668-1674.
- Rakoff, J. S. (2014, November 20). The financial crisis: Why have no high-level executives been prosecuted? *The New York Review of Books*.
- Reuters. (2025, November 24). Meta shut down internal research after finding causal link to mental health harms, court filings say. *Reuters*.
- Roose, K. (2025, April 24). If A.I. systems become conscious, should they have rights? *The New York Times*.
- Rugh, J. S., & Massey, D. S. (2010). Racial segregation and the American foreclosure crisis. *American Sociological Review*, 75(5), 629-651.
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.
- Savitt, T. L. (2007). *Race and medicine in nineteenth- and early-twentieth-century America*. Kent State University Press.
- Sen, A. (1981). *Poverty and famines: An essay on entitlement and deprivation*. Oxford University Press.

References, concluded

- Shiva, V. (1991). *The violence of the Green Revolution: Third World agriculture, ecology, and politics*. Zed Books.
- Starr, P. (1982). *The social transformation of American medicine*. Basic Books.
- Twenge, J. M. (2017). *iGen: Why today's super-connected kids are growing up less rebellious, more tolerant, less happy, and completely unprepared for adulthood*. Atria Books.
- United Nations Human Rights Council. (2018). *Report of the independent international fact-finding mission on Myanmar (A/HRC/39/64)*. United Nations.
- United States v. The Purdue Frederick Company, Inc.*, No. 1:07-cr-00029 (W.D. Va. May 10, 2007).
- U.S. Environmental Protection Agency. (2021). *Addressing uranium contamination on the Navajo Nation*. EPA.
- U.S. Senate Finance Committee. (2021). *Insulin: Examining the factors driving the rising cost of a century-old drug*. U.S. Senate.
- Van Zee, A. (2009). The promotion and marketing of OxyContin: Commercial triumph, public health tragedy. *American Journal of Public Health*, 99(2), 221-227.
- Wells, G., Horwitz, J., & Seetharaman, D. (2021, September 14). Facebook knows Instagram is toxic for teen girls, company documents show. *The Wall Street Journal*.
- Williams, D. R., & Collins, C. (2001). Racial residential segregation: A fundamental cause of racial disparities in health. *Public Health Reports*, 116(5), 404-416.
- Wu, T. (2016). *The attention merchants: The epic scramble to get inside our heads*. Alfred A. Knopf.
- Zahrn, S., McElmurry, S. P., Kilgore, P. E., Mushinski, D., Press, J., Love, N. G., Sadler, R. C., & Swanson, M. S. (2018). Assessment of the Legionnaires' disease outbreak in Flint, Michigan. *Proceedings of the National Academy of Sciences*, 115(8), E1730-E1739.
- Zakrzewski, C. (2025, September 8). Meta suppressed children's safety research, whistleblowers tell Congress. *The Washington Post*.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.