# Deferred Failure, Verification Cost, and Organizational Competence Under Automation

Flyxion
Independent Researcher

November 2026

**Abstract**

An interview with Cory Doctorow that takes the Microsoft 365 outage of November 12, 2026 and the Federal Aviation Administration ground stop of November 13, 2026 as its starting point advances a structural argument about automation and organizational fragility that is stronger than its framing suggests. The analysis reads the interview by timestamp and separates five layers of claim: factual premises, causal mechanisms, explanations of motive, forecasts, and prescriptions. The most durable contribution is the observation that running systems outlive the people who understand them, so that reductions in maintenance capacity are rewarded at once and penalized late, with a conception of process knowledge as a distributed capability that documentation does not capture. The argument about plausible machine errors is sound when stated as a conditional claim about verification cost with an explicit quality constraint. The attribution of the outages to layoffs is not established, since reporting identifies technical triggers, a latent networking defect and a router installation, with no demonstrated staffing link. The financial forecasts rest on unmeasured chains of conditions, and the political program is not derived from the diagnosis. A restricted formulation of the central claim is proposed, supported by a partial-internalization model of reserve elimination and a quality constraint on the verification-cost comparison. It is testable, independent of any financial collapse, and general beyond information technology.

**Keywords:** automation; process knowledge; verification cost; resilience; labor

# 1 Source and Method

The interview under analysis is a video conversation between Cory Doctorow and an interviewer who is not named in the captions, with a running time of about one hour and six minutes [1]. It refers to the Microsoft 365 outage and the nationwide ground stop as recent events and uses them as its opening motivation, although the captions do not state when the conversation was recorded. The transcript consists of automatically generated captions without speaker labels, so attribution of a passage to Doctorow or to the interviewer rests on content and conversational structure, and a few proper names are visibly mis-transcribed. Locations in the text below are given as caption start times in hours, minutes and seconds, which is the only stable locator the material offers.

The program follows a recognizable sequence. A cold excerpt of the finance discussion plays first (00:00:00 to 00:01:25), followed by a narrated introduction that reviews the outages, the industry's public promises, and the government's branding of the technology as superintelligence (00:01:30 to 00:06:48). Doctorow's account of the outages and their institutional background occupies roughly (00:08:58 to 00:21:05). The interviewer then turns to the industry's promises (00:21:08), and Doctorow explains why employers are buying the technology and why he judges its economics unworkable (00:22:41 to 00:34:30). A discussion of who drives automation, with insurance claims and air traffic control as cases, runs to (00:39:35). The longest sustained section concerns financial fragility and initial public offerings (00:39:39 to 00:52:00), and it is followed by Doctorow's treatment of process knowledge and the training pipeline (00:53:49 to 00:59:30) and his closing argument about political opportunity (01:00:02 to 01:04:10). The final two minutes are promotional.

Where the interview asserts facts that can be checked, the analysis compares them with reporting available through late November 2026 and states where such reporting confirms, contradicts, or does not reach the claim. The analysis distinguishes five layers: factual premises, causal mechanisms, explanations of motive, forecasts, and prescriptions. Explanations of motive deserve a layer of their own because they are neither mechanisms nor forecasts. When Doctorow says that executives want the technology because it lets them dismiss subordinates who refuse unreasonable instructions, he attributes a purpose to a class of actors, and such an attribution can be plausible without being established, since it requires evidence about procurement decisions, incentives and internal communications that an interview rarely supplies. The distinction separates the defensible claim that deployment can weaken professional refusal from the much stronger

claim that adoption is driven by the wish to weaken it. The interview moves among these layers quickly, and much of its persuasive force comes from the way a well-supported mechanism lends credibility to a motive, a forecast or a program that rests on further premises of its own.

## 2 Outage Reporting and the Inference From Failure to Understaffing

This section addresses factual premises and the causal mechanism that the interview infers from them. The narration concedes at (00:02:50) that artificial intelligence probably had little direct connection to most of the outages and locates the connection instead in layoffs justified by claimed efficiencies from the technology: tens of thousands of Microsoft employees (00:03:03) and roughly four hundred Federal Aviation Administration employees dismissed under the Department of Government Efficiency effort (00:03:30). Neither figure is sourced in the program. Doctorow's own opening is more cautious than the framing around it. He allows that there is some aspect of coincidence (00:08:58), names the shared features as early computerization, layered legacy systems, mergers that eliminate redundant staff, and an ideology of austerity (00:09:26 to 00:11:21), and describes the replacement of staff with automation as a final blow that may or may not have been the last straw (00:11:53 to 00:12:03). The proposition actually advanced is therefore about accumulated technical debt and depleted reserves, with AI-motivated layoffs as one contributing stressor among several.

Reporting on both incidents supports the narrower description. Microsoft's preliminary post-incident report places the Microsoft 365 incident between 4:20 PM UTC on November 12 and 4:10 PM UTC on November 13 [2]. Microsoft attributed it to a problem introduced during an earlier network capacity upgrade that surfaced after a more recent network change [3], and one account reports a Microsoft executive tracing the defect to a patch introduced in July 2026 [4]. The cause was a latent defect in the networking gateway fronting many services, and the reports describe no AI involvement in the proximate failure, which does not exclude every indirect contribution. The ground stop of November 13 followed the installation of a new router at an FAA facility in El Dorado, Arkansas, during work to replace copper wiring with fiber. The router flooded the agency's network with data and overwhelmed an older system, officials could not initially locate the cause, and the FAA had not determined a root cause when ABC

News reported on November 17, having paused further router installations pending investigation [5, 6].

This reporting is compatible with Doctorow's picture of layered systems whose seams can rupture (00:09:48 to 00:10:53), since in both cases an interaction between a change and an older configuration produced the failure. It supplies no information about staffing in the teams responsible for the changes, and the FAA trigger was new equipment, not aging equipment. The reports establish technical mechanisms and leave the organizational explanation open.

The interview nevertheless treats the outages as evidence for the staffing thesis. Doctorow says there is a prima facie case that the cuts went too deep (00:15:49), anticipates the reply that the dismissed workers were unrelated to the failed services, and answers that managers who knew which workers to retain would not have had servers fail (00:15:57 to 00:16:13). The interviewer reads a line from Doctorow's blog asserting that Microsoft's servers were up while the laid-off workers were on the payroll (00:13:22 to 00:13:40). This reasoning infers inadequate capacity from the occurrence of failure. It treats the occurrence of an event whose base rate is unknown as evidence of a particular cause. As a general rule it cannot be disconfirmed, because every large organization eventually suffers incidents whatever its staffing, and a contrast between before and after requires a baseline of incident frequency, severity and recovery time under the earlier staffing level. The program also describes an outage lasting two days (00:01:35 to 00:01:47), which Doctorow affirms (00:16:20). Microsoft's own record gives about twenty-four hours of impact spanning two calendar days [2]; the interviewer's recollection of colleagues unable to work across both days (00:16:13 to 00:16:39) is consistent with a long tail of recovery, but the report is the better guide to duration.

The defensible form of the claim is that reductions in operational expertise raise the probability, duration or severity of failures even when the immediate effects are difficult to observe. A test would compare incident frequency and recovery time across organizations or across periods, controlling for the volume of change and the age of the systems. The interview offers a cluster of incidents in its place, and a cluster of incidents has several possible explanations, one of which is chance.

# 3 Delayed Consequences and the Persistence of Running Systems

This section addresses a mechanism. The most important argument in the interview appears in a passage prompted by a book on organizing technology workers (00:16:48). Doctorow reports its observation that a strike on an automobile assembly line stops production at once, whereas software continues to run for a long time after its authors stop working (00:17:28), and he adds the corollary that dismissing the authors has consequences that arrive much later because their code keeps running (00:17:46). He then generalizes to a class of policy problems in which cause and effect are separated in time, naming privacy violations, monopoly, and the delayed harms of smoking (00:17:55 to 00:19:40), and observes that such separation serves scammers, confounds policymakers, and shields managers (00:20:16). The Twitter acquisition supplies a concrete case (00:19:43 to 00:20:01). Doctorow says the company's site reliability engineers were good at their jobs, so that whole server racks could be unplugged without service degrading, and that the removal of redundancy meant the first failed power supply would take the service down. The anecdote exhibits the mechanism cleanly: reserve capacity is invisible while it is intact, it is consumed by cuts without any visible loss, and its absence is revealed only by the first fault it would have absorbed.

The argument makes few empirical commitments. It does not require that any particular outage resulted from layoffs, that AI works or fails, or that any financial collapse occurs. It requires only that systems keep operating after the loss of their maintainers and that the organization's accounting records the saving immediately and the loss late. The continued operation of a system is then not sufficient evidence that the organization retains the capacity to repair, modify, or reconstruct it, and the crucial distinction is between continued operation and continued maintainability. This is close in structure to the classical observation that automation leaves human operators with the residual tasks that require the most skill while eroding the practice by which that skill is maintained [17], and to the analysis of tightly coupled and interactively complex systems in which small faults propagate in ways that no one anticipated [18].

Doctorow's contrast between efficiency and resilience (00:10:55 to 00:11:21) states the same point at the level of management philosophy: a person who is not working at this moment looks like waste, and keeping that person in reserve in case something breaks looks like a cost without a product. His debt metaphors, which compare accumulated technical debt to paying one credit card with another and to a maturity mismatch

(00:12:03 to 00:12:26), are serviceable but slightly misleading. Debt accrues interest on a known schedule, whereas the loss of reserve capacity resembles the cancellation of insurance. The analogy is not exact, since an insurance premium is known and the implicit premium here, the salary of a worker who might not be needed, is itself uncertain in value. The saving is the premium not paid, the loss is realized only if the insured event occurs, and because the events are infrequent most periods contain no visible loss at all. That is why survival is weak evidence of adequacy, and it explains why ordinary performance measurement can reward decisions that progressively reduce the organization's ability to remain functional.

The same structure appears in a second form in the interview, where the cut is justified not by an existing substitute but by a promised one. Doctorow cites a report, which he attributes to a magazine article, of an official at a research institute pleading against the termination of long-running cancer studies and being told that AI would cure cancer (00:14:36 to 00:14:55). The claim is unverified here, but it states the pattern that makes the interview's thesis more than a restatement of austerity critique: reserves are eliminated in anticipation of capabilities that have not been demonstrated, so that the decision does not depend on the capability arriving but only on the belief that it will.

# 4 Process Knowledge and the Limits of Documentation

This section addresses a mechanism and the factual status of one illustrative anecdote. Near the end Doctorow introduces process knowledge (00:53:49) as everything that keeps a business running without being intellectual property, and distinguishes it from tacit knowledge, which he takes to refer to what a single worker knows (00:54:04). Process knowledge is distributed across workers within a division, across divisions within a firm, and across firms within a sector. His example is a jam-prone machine that one person knows how to repair, with a retired former repairer who will still do it for fifty dollars if the current repairer is on vacation (00:54:24 to 00:54:36), and he says this is the difference between a factory that runs and one that stops. The concept has clear antecedents in the study of tacit knowing [20] and in ethnographic work on how field technicians maintain diagnostic competence through shared stories and informal exchange [19], but Doctorow's emphasis is distinctive in locating the capability in a network of relationships

and not in any individual.

The work-to-rule discussion (00:54:53 to 00:55:15) is the strongest evidence offered for the concept. If employees follow only documented procedures and withhold informal judgment, an organization can cease to function even though everyone is formally doing the assigned work. This undermines the assumption that an organization is fully represented by its official procedures, and it explains why replacing a worker with a system that reproduces the worker's visible output may not reproduce the worker's contribution. That contribution may include recognizing exceptional circumstances, deciding that a procedure is inappropriate, and knowing whom to call, and these capabilities are exercised rarely enough that productivity measurement scarcely registers them.

There is an internal tension in the way the claim is stated. Doctorow says process knowledge cannot be recovered and must be regrown over years (00:54:40 to 00:54:50), yet his own example is one of recovery, in which a retired worker is reachable by telephone and willing to return for a fee (00:54:30 to 00:54:36). The more accurate claim is that process knowledge can be recovered when its carriers remain reachable and willing, and otherwise must be reconstituted at a cost in time that may exceed what a failure allows. Some of it can also be reconstructed from documentation, logs, source code and external experts. What cannot be guaranteed is recovery at an acceptable cost within the time available during an incident. The Chips Act illustration (00:55:21), in which a four-year horizon was too short to yield a return because manufacturing lore had to be built, is plausible as an instance but multi-causal, and it serves better as an illustration than as a test.

An important distinction that the interview does not draw is between preserving records and preserving operational capability. An organization might retain every source file, incident report and internal message and still lose the ability to operate its systems, because the capability consists of relationships among information, authority, equipment, available personnel and opportunities for corrective feedback. A documented recovery procedure may be accurate and yet depend on equipment that has been replaced, permissions that have lapsed, or a department that no longer exists. Conversely, a sufficiently complete archive can make some lost capabilities cheaper to reconstruct. Doctorow describes the loss of this capacity but offers no operational definition or measurement of it, which is the principal gap between the concept and a research program. A workable definition treats organizational competence as the set of interventions that the organization can successfully execute under specified conditions, within a specified time and budget, and not as the amount of information it holds. On that definition recovering knowledge

and recovering capability come apart in both directions. An organization may identify the person who knows how to repair a machine and still fail to repair it, because parts are unavailable, authorization has lapsed or the equipment has changed, and it may restore the capability by a different procedure without ever reconstructing the original knowledge. The retired repairer then illustrates recovery of a capability under favorable conditions and not recovery of process knowledge as such.

The interview extends process knowledge forward in time through the training pipeline. Knowledge is acquired by working next to those who have it, absorbing it by observation and by asking questions that no one thought to ask (00:55:38 to 00:55:50), which means that a halt in hiring interrupts transmission as well as retaining. Doctorow argues that the economy risks losing both ends, the experienced workers and the entrants who would replace them (00:57:22), and he cites the example of Greek medical students during the debt crisis, concluding that an oncologist cannot be produced on demand because training takes twelve to fifteen years (00:56:50 to 00:57:20). The Greek claim is asserted without a source, but the structural point about cohort gaps is sound regardless of the example. The interviewer and Doctorow then apply it to radiologists who retire without successors, left to review a hundred times more images (00:58:51 to 00:59:22), and to mathematicians, where the announcement that AI can do the work discourages entry and so becomes self-fulfilling (00:59:24 to 00:59:30).

One anecdote in this section deserves separate treatment because it functions as the main worked example of process knowledge. Doctorow describes a chip foundry with a batch that failed quality assurance and could run only at half its nominal clock speed, whose manager knew of a customer division that had been unable to get a low-end computer approved, offered the chips cheaply, and thereby produced a product that he calls the most successful launch Apple has had in ten years, after which the defective stock was exhausted and new chips were manufactured (00:55:52 to 00:56:35). Published reporting confirms the outline of the ending. The MacBook Neo uses binned versions of the A18 Pro processor, chips with one graphics core disabled that would otherwise have been discarded or sold as lower grades, and demand was strong enough that Apple was reported to be running out of the binned stock [10, 11]. The details differ from the interview's account. The reports describe binning as Apple's use of its own inventory with a disabled graphics core and not a reduced clock speed, they do not describe a foundry manager initiating a sale to an internal division, and the company's own public characterization concerned its best launch week for first-time Mac buyers and not a ten-year record [10]. As told, it is not confirmed and carries less evidential weight than its

position in the argument suggests. It is used here only to illustrate process knowledge, not as evidence for the concept, and its literal accuracy does not affect that use.

# 5 Verification Cost and the Economics of Replacing Skilled Workers

This section addresses a mechanism and states the verification-cost condition formally. Doctorow's central technical argument concerns the economics of checking machine output (00:32:23 to 00:34:00). For automation to allow a firm to dismiss skilled workers without sacrificing quality, the remaining workers must act as humans in the loop who catch the system's errors, and those errors must be detectable by them. He imagines dismissing nine tenths of radiologists and having the remainder review ten times as many reports (00:33:01 to 00:33:10), and he argues that the errors a statistical system makes are the ones statistically closest to correct while still being wrong (00:32:40). If a chatbot reports a spacecraft in a patient's chest, the reviewer will notice, but a lesion that resembles a tumor in most cases and is not one in this case will be difficult to catch at scale (00:33:21 to 00:33:40). This is persuasive as a statement about cost. Let $c_h$ denote the cost of producing a result by independent human work, $c_g$ the cost of generating a candidate by machine, $c_v$ the cost of verifying a candidate to the required standard, and $\epsilon$ the expected cost of defects that survive verification, including corrective work and liability. Replacement is economical only if

$$c_g + c_v + \epsilon < c_h \quad \text{subject to} \quad P(\text{undetected consequential error}) \leq \delta, \tag{1}$$

where $\delta$ is the maximum admissible risk appropriate to the application. A different requirement would be noninferiority to the human process, that the undetected error probability not exceed that of the human baseline by more than a stated margin, and the two should not be confused, since automated and human processes with error probabilities of, say, 0.5 and 0.1 percent can both satisfy a threshold of one percent without being equivalent in quality. The constraint is necessary because the cost comparison is meaningful only against some stated quality standard. Without it a process can satisfy the inequality by reducing verification effort and allowing unacceptable errors to escape, and the expected-loss term $\epsilon$ does not repair this on its own, since two processes with the same error rate can differ greatly in the severity of the failures that remain. The

quantities are also not constants, and the unit of analysis matters. Let $n$ be the number of items each reviewer must inspect, $t$ the verification effort spent on each item, and $q(n, t)$ the residual probability of a consequential error. Higher volume does not automatically raise the per-item cost of verification. An organization under time pressure can reduce $t$, so that measured verification cost per item falls while $q$ rises, or it can hold the standard fixed, in which case the time needed to clear the workload grows. These are different operational responses, and the feasible configurations are those with $q(n, t) \leq \delta$, with $c_v$ understood as the minimum cost of achieving that constraint at the required throughput. This is the form that turns the radiology example into a testable economic claim: dismissing nine tenths of the radiologists multiplies $n$ by ten, and the replacement strategy fails when the organization cannot satisfy its throughput, cost and reliability requirements together. Fluent output reduces $c_g$ sharply without necessarily reducing $c_v$, and when verification requires the same expertise as production, $c_v$ approaches $c_h$ and the saving disappears. The interviewer's description of insurance adjusters who are accountable for plausible but wrong machine output (00:37:13) makes the same point from the side of accountability.

Two qualifications refine the argument. First, the near-correctness of surviving errors is partly a selection effect. Any reviewer filters out the errors that look wrong, so the errors that remain will by construction be those that look right, and the relevant empirical quantity is the detection rate at the review volumes that the replacement plan imposes, not the character of the generator's errors in isolation. The selection is also endogenous to the verification effort $t$ introduced below, since a reviewer who spends less effort per item filters out fewer errors, so that the plausibility of surviving errors is partly a property of the reviewer's threshold and not only of the generator's error distribution. The interview does not present measurements of detection rates, and it acknowledges nothing about how detection might decay with volume, although the literature on supervision of automated systems suggests it does [17]. Second, domains differ in whether cheap automatic verifiers exist. In diagnostic imaging or claims adjudication, ground truth is expensive and the verification argument is strong. In domains where a program, a solver or a rule checker can test a candidate, $c_v$ can be moved into software and the economics change, although residual error does not vanish. Doctorow's account of the technology as theory-free inference, in which a system is given tokens and asked to guess without being told what the world is (00:25:53 to 00:29:00, 00:32:31), describes an unconstrained sequence predictor and explains his chess example, in which such a system sometimes places a piece on an occupied square (00:28:41). It does not address

systems that combine generation with external constraints, and the properties of the unconstrained case do not establish the limits of every assisted workflow. The brief remark that the originating paper on attention [15] won a Turing Prize and a Nobel (00:27:07 to 00:27:15) is inaccurate as stated. The paper itself received neither. The 2018 Turing Award went to Bengio, Hinton and LeCun for work on deep learning, and the 2024 Nobel Prize in Physics went to Hopfield and Hinton for foundational work on neural networks, so the remark conflates awards to related research and attaches them to the paper, which is a different error from misattributing a result to the wrong person.

The economic conclusion is stated more carefully than its summary suggests. Doctorow does not say that augmentation is useless. He affirms that a tool prompting a radiologist to look again at a few cases a day could save a life a week (00:33:36 to 00:33:50), and his claim is that the cost of developing theory-free inference cannot be repaid by making radiology more expensive and slightly better (00:33:50 to 00:34:00). This is a conditional claim about the relationship between development cost and the revenue available from augmentation, and it is plausible given the scale of investment he assumes. It does not follow that no augmentation business can be profitable, since viability depends on deployment costs, market size, error rates and the value of improvements, and the interview presents no revenue data. The narration's report that firms are reducing token spend and rehiring laid-off employees (00:03:42 to 00:03:52) would be relevant evidence but is not sourced. Doctorow also leaves a door open: after a bankruptcy, assets bought cheaply might support modest uses such as a confirmation prompt for professionals (00:34:14 to 00:34:30).

# 6 Motive Claims and the Distinction Between a Scheme and a Scam

This section addresses explanations of motive. Doctorow gives two explanations of why employers and investors want the technology. The first he calls materialist (00:23:59). A sales call that offers the chance to dismiss most of the subordinates who tell a manager that an idea is illegal, immoral or impossible, and to replace them with a system that will produce whatever is requested, is attractive in itself (00:24:18 to 00:24:50), and it is attractive especially to managers whose reports are professionally obliged to refuse, such as lawyers, physicians and accountants (00:24:49 to 00:25:16). The pitch need not be true, only plausible long enough for the seller to exit (00:25:19). The second explanation

is ideological: a capital allocator who is continually thwarted by experts finds in the technology the first mechanism for converting money into expertise without persuading any expert to cooperate (00:29:15 to 00:30:04). The interviewer presses the point that the sums involved reflect belief in a transformative technology and not only a sales opportunity (00:23:29), and Doctorow's reply folds belief into the account by describing the venture as a mix of true believers and swindlers who go farther together than either would alone (00:30:24 to 00:31:30).

These are explanations of motive and demand, a layer distinct from mechanism, and the interview gives no data that would discriminate among them. The licensing observation is analytically valuable, however, because it identifies a function, refusal under professional duty, that is simultaneously a verification function and a safety function. Replacing the people who refuse with a system that complies removes the check at the point where it is most valuable, and this connects the economics of verification to the loss of maintenance capacity discussed earlier. Both can be described as a conversion of capacity for refusal and repair into capacity for output. That unification is not stated in the interview, but it follows from its premises, and it is a more robust formulation than the claim that the technology is simply a scam.

The scheme and scam distinction (00:30:41 to 00:32:20, 00:34:08) is introduced through the story of Ponzi, whose postal reply coupon arbitrage might have worked if scaled, but who found it easier to collect new money. A scheme is an enterprise whose underlying economics could pay off, and a scam is one that cannot and depends on new money. Doctorow classifies the current investment as a scam because replacing skilled labor with unreliable output is not feasible. The classification depends on the verification argument and inherits its limits. It also compresses a spectrum. An enterprise can have useful products and a speculative valuation at once, and the word scam implies intent in a way the economic argument does not require. Doctorow himself separates the technology from its financing, insisting that technology does not have cooties and that the correct response to useful tools made by bad actors is to clone them as free software (00:22:41 to 00:23:29). The distinction he actually defends is therefore between usefulness and valuation, which is more modest and more defensible than the rhetorical contrast of scheme and scam.

# 7 Directions of Automation and Evidence From Workers

This section addresses mechanism and motive, and extends the mechanism beyond the interview. Doctorow proposes that labor-driven automation historically serves quality, because a craftsperson picks up a tool to do better work, whereas capital-driven automation serves throughput, and that when capital has market power it serves throughput at the expense of quality (00:35:10 to 00:35:40). His evidence is a contrast between business journals of the late 1990s and 2000s, which discussed how executives could cope with young workers demanding access to the web, and current journals, which discuss how to cope with workers who refuse to use AI (00:36:01 to 00:36:22). The contrast is suggestive about who is pushing adoption, though it rests on impressionistic reading of a genre and not on adoption data. The interviewer's reference to survey after survey of workplaces (00:36:46 to 00:36:52) would supply such data if cited.

The historical generalization is too broad as stated, since workers also adopt tools to increase throughput and employers sometimes invest in quality, safety and reliability. The more defensible distinction concerns who controls the objective of the automation and what is measured. A tool introduced to assist a professional can be judged by whether it improves the professional's judgment, whereas a tool introduced to eliminate positions is judged by how many employees can be removed while measured output stays above a threshold. Those are different optimization problems, and the second is damaging when the organization measures throughput but not the maintenance capacity, judgment and institutional knowledge being removed. Doctorow's own qualifier about market power matters here. Where customers cannot leave, quality degradation carries no penalty, which connects the automation argument to his earlier remark that the firms in question are monopolies that can afford not to care (00:15:25 to 00:15:49), and to his definition of the degradation of services as what organizations do when they expect no consequences (00:14:05). The interview provides no evidence about whether the November outages produced consequences for the firms concerned, which would be the direct test of that expectation.

The insurance example (00:36:54 to 00:38:47) illustrates the interaction with verification cost. Claims adjusters are pressed by their employers to deny as many claims as possible (00:37:30), claimants respond by using AI to write claims that are long and heavily documented and sometimes cite insurance regulations and policy clauses that do not exist (00:38:21 to 00:38:33), and adjusters are then told to use AI to cope. The re-

duction in the cost of producing a claim raises the volume that must be reviewed, which is a form of externalized verification cost, and the adjuster, whose bargaining power is low, bears it. Doctorow's remark that an observer wanting to reach this destination would not start from here (00:38:47) extends the point to air traffic control, where a shortage of controllers produced by decades of hostile policy is now being answered with proposals for automation (00:38:54 to 00:39:22). The link between the training shortage and the November ground stop is, as discussed above, not shown by the reporting on the stop itself.

The insurance case generalizes in a way that the interview only gestures at. The verification inequality compares costs borne by a single party, but in correspondence between a sender and a reviewer the two sides of the inequality fall on different parties. For a programmer, sending a million messages was never expensive, and what restrained the practice was a cultural convention that it is rarely a good idea, together with the skill gap that kept most people from trying. Nothing in this requires an exploit, which is why the situation differs from hacking: the ordinary, documented operation of mail systems, application portals and submission forms is being used as designed, so there is no defect to patch and the only barrier left is the convention. When assistants make routine programming knowledge available to everyone, the skill gap closes and the convention is left to carry the load alone. A claimant, applicant, student or author can then produce a long, heavily documented submission, or thousands of variants of one, at almost no cost, and the receiving institution must absorb the resulting rise in review volume, so that its total review cost, which is the volume multiplied by the per-item effort, and its residual error probability both grow with $n$. Per-item cost need not rise, since reviewers may spend less on each item, and the relevant quantity here is the total cost borne by the receiving institution. The sender can then satisfy a version of the inequality while the reviewer bears the whole increase, which is the sense in which the cost is externalized. Admissions offices, journals, employers and instructors occupy the reviewer's position in the same structure. This mechanism differs from the loss of an organization's own reserves, since there the capacity falls against a stable workload and here the workload expands against unchanged staffing, although both produce an unsustainable ratio of demands to verification capacity. The interview supplies only the insurance example, and the extension to coursework, journal submissions and job applications is an inference from the same mechanism, not something it establishes.

The extension also has a signaling consequence. A simple congestion externality would explain the overload, but it would not explain why screening cannot rely on the

content of what arrives. A signal is informative when its cost differs across types of sender, so that it pays some types to produce it and not others [14]. A lengthy and carefully documented submission may have had that property to the extent that its cost was correlated with the quality the reviewer sought, and the defensible claim is that cheap generation destroys the correlation between the cost of production and the quality sought, and with it whatever discriminatory value such a production-cost signal possessed. The claim does not require that the signal was ever a good measure, and cheap generation may simply reveal that institutions were relying on weak proxies, a pattern summarized as Goodhart's law, in which a measure that becomes a target ceases to be a good measure. Cheaper generation also need not increase the number of distinct senders. It may raise the number of submissions per sender, the length of each, or the number of variants of one submission, and these have different implications for screening and can be measured separately. Reviewers are then pushed toward automated screening, which raises the sender's incentive to generate further variations, so that the arrangement tends toward an arms race in which the cost of attention is the scarce resource. The testable quantities are submissions per reviewer, time per review, and the acceptance rate and error rate at each volume, compared across periods before and after generation became cheap.

# 8 Financial Fragility and the Limits of Mechanism

This section addresses factual premises and forecasts. The avalanche image that opens and recurs in the program (00:00:00, 00:41:47) is offered as an account of how the financial structure supporting the AI industry could fail: a long accumulation of small instabilities, then a sudden collapse triggered by one more. Doctorow frames the question as lying between two poles, Stein's Law, which says that what cannot go on forever eventually stops (00:41:31), and the aphorism that markets can stay irrational longer than an investor can stay solvent (00:50:00), which he attributes to Galbraith and which is more usually credited to Keynes, with doubtful provenance in either case. Stein's Law is true and does not yield a forecast. It says nothing about the time scale or about whether the particular trajectory in question is one that cannot continue. Doctorow acknowledges that the combination makes timing impossible (00:49:53 to 00:50:06), and the interviewer concedes that earlier predictions of imminent collapse have failed (00:49:37). The consequence is that the central forecast is weakly constrained. If the bubble bursts, the analysis is vindicated, and if it continues, the continuation can be read as further

accumulation of instability. Without a specified threshold or a discriminating prediction, the avalanche analogy explains possible mechanisms of collapse better than it predicts one.

The interview proposes several mechanisms. Data center financing is described as stacked, so that one party withdrawing from one arrangement would cause dozens to fail (00:42:53 to 00:43:10), with Gulf investors who now need to rebuild damaged terminals named as a possible source of withdrawal. These are claims about exposures, leverage and contractual dependencies, and none is quantified. A financial system can contain fragile arrangements without being near a particular failure threshold, and establishing that collapse is likely would require evidence about leverage, liquidity, dependencies and available buffers that the interview does not supply.

The most specific factual claims concern the SpaceX offering. Doctorow states that Nasdaq gave the listing an arrangement allowing entry to its main index without the usual seasoning (00:43:22 to 00:43:55), that S&P declined after fund managers objected (00:43:40), that index funds are therefore compelled to buy shares of which almost none were offered to the public (00:45:57 to 00:46:24), and that insiders could sell on the first day (00:44:49). Published reporting supports the core of the mechanism. Nasdaq adopted a fast-entry rule, approved on March 30 and effective on May 1, 2026, that admits a new listing ranked within the top forty Nasdaq-100 members by total market capitalization after fifteen trading days, and the offering had a float of roughly four percent [7]. After the fifteen days, a new listing's weight is the lesser of its total market capitalization and three times its float-adjusted capitalization [7], so a small float is scaled up in the index weight, with one analysis describing a 4.3 percent float treated as 12.9 percent for weighting, and passive funds benchmarked to the index must buy regardless of price [8]. SpaceX listed on June 12 and joined the index on July 7 [9]. Two details are only partly confirmed. S&P did not follow Nasdaq: its consultation on mega-cap companies produced no change to the methodology for the family that includes the S&P 500, so that new listings still wait at least twelve months [12, 13]. That matches the outcome Doctorow describes, but the reporting consulted does not say that S&P acted after fund managers objected, which is the causal claim. The analytical weight rests on the mechanism, fast entry combined with a small float and float-adjustment scaling, which is documented, and the contrast with S&P is rhetorically important but evidentially weaker. On insider sales, a report on the index entry refers to a lockup expiring on August 6 [9], which sits uneasily with insiders selling on the first day unless certain holders were exempt, and the interview cites no source for the claim, so that it

is not confirmed by the reporting consulted.

That the mechanism is real shows that index-driven demand can be induced mechanically by rule changes. It does not show that the arrangement damages savers or produces a systemic event, because the exposure of an individual saver depends on the index tracked. Doctorow's statement that everyone's savings are in such funds (00:44:20 to 00:44:45) is accurate in spirit, but a broad total-market fund holds a small position in a low-float company, with one fund expected to hold under two tenths of one percent [7], whereas a fund tracking the Nasdaq-100 holds more. The question of who bears losses is therefore a question about index methodology that the interview does not examine.

From this base the interview builds a conditional chain about the prospective public offerings of AI developers (00:47:50 to 00:49:30). If Nasdaq were denied a similar concession by political pressure on regulators (00:48:20), and if a failed offering frightened the market, then margin calls might follow and lenders might refuse to roll over loans, and the lenders to data center builders might withdraw in a cascade, so that a single puncture could burst the whole structure (00:48:57 to 00:49:30). Each link is conditional and none is assigned a likelihood, which is the usual condition of such narratives and the reason they cannot be assessed as forecasts. Doctorow supports the political consequences with the claim that statisticians found foreclosures to be the strongest predictor of a precinct switching to Donald Trump (00:47:12 to 00:47:40). A precinct-level association does not by itself identify individual-level behavior, and the point is presented as an analogy and not as a measured effect, as is the earlier anecdote about a Senate race and the price of hamburgers (00:46:39 to 00:47:05).

# 9 The Political Conclusion

This section addresses prescriptions. Doctorow closes by invoking Milton Friedman's observation that in a crisis ideas move from the periphery to the center (01:00:52 to 01:01:10). In Friedman's own formulation the actions taken in a crisis depend on the ideas that are lying around [16]. Doctorow argues that the job of those who hold different views is to keep ideas available, and he lists labor power, wealth redistribution, prevention of the accumulation of wealth, and the seizure of the means of computation (01:01:39 to 01:02:00). The interviewer adds that the fragility of current systems opens space for radical reform of how technological systems are built, maintained and governed (01:02:36 to 01:03:30), and Doctorow compares the situation to rising seas, where the outcome depends on whether societies welcome refugees or turn to what he calls eco-

fascism (01:03:33 to 01:03:59).

These are political prescriptions and they have a different evidentiary status from the operational diagnosis. The diagnosis of institutional fragility could support several responses: stronger worker protections, improved engineering governance, mandatory operational reserves, better documentation, competition policy, revised liability for outages in critical infrastructure, or changes to how change management is audited. The interview does not compare these alternatives. It also compresses the historical analogy. Accounts of the earlier neoliberal turn emphasize an organized intellectual and institutional infrastructure assembled over decades [21], so that the condition for ideas to move from the periphery was not merely that they existed but that organizations were ready to carry them, a condition that Doctorow's phrasing of keeping ideas lying around leaves implicit. The remark about a midterm election as an instance of ideas moving to the center (01:02:02) is a partisan expectation outside the scope of the analysis and is not evaluated here.

The closing line may be the most useful for the technical argument. Utopia, Doctorow says, is not a world in which things never go wrong but one in which failures are addressed gracefully (01:03:59). This is a statement of resilience as a design value, and it is consistent with the earlier question he poses about what the cushion is and how a system fails gracefully (00:20:57 to 00:21:05). It offers a standard by which the interview's own prescriptions could be compared: whether each measure increases the capacity of institutions to detect deviations, absorb faults, and recover.

# 10 A Restricted Statement of the Central Claim

This section restates the mechanisms as a testable proposition. The preceding sections suggest a restatement that preserves what is most defensible in the interview and discards what depends on contested forecasts. An organization evaluated on current output and current cost will tend to eliminate resources whose value consists in preventing, diagnosing or repairing infrequent failures, because the value of such resources appears only when a failure occurs, and a failure may be remote in time from the decision that caused it and difficult to attribute to it. The tendency is stronger where the organization faces little competitive or regulatory consequence for failure, and where a prospective capability promises to substitute for the eliminated resource. The incentive to eliminate reserves can arise before the substitute has shown that it performs as promised, which is why the proposition does not depend on any assessment of the technology's quality and

does not depend on a financial collapse. It applies to hospitals, public administration, manufacturing and software infrastructure alike.

The tendency is not automatic, because it depends on the decision rules, incentives and information of the organization. A manager who anticipates future losses and bears their cost may rationally retain substantial reserves even when current measurement does not reward them, so the proposition becomes sharper once the conditions under which elimination is privately advantageous are stated. Let $r$ denote retained operational capacity, $C(r)$ its present cost, $p(r)$ the probability of a consequential failure and $L(r)$ the expected loss conditional on failure. An organization that bears its full expected cost chooses $r$ to minimize

$$J(r) = C(r) + p(r)L(r). \tag{2}$$

A decision-maker who internalizes only a fraction $\alpha$ of the loss faces instead

$$J_\alpha(r) = C(r) + \alpha\, p(r)L(r), \qquad 0 \le \alpha \le 1. \tag{3}$$

Suppose $C$ is increasing, the product $pL$ is decreasing in $r$, and $J_\alpha$ is twice differentiable and convex with $J''_\alpha(r^*) > 0$ at a nondegenerate interior optimum $r^*(\alpha)$, which therefore satisfies $C'(r) + \alpha(pL)'(r) = 0$. Implicit differentiation gives, for $\alpha \in (0, 1]$ at an interior optimum,

$$\frac{dr^*}{d\alpha} = -\frac{(pL)'(r^*)}{J''_\alpha(r^*)} > 0, \tag{4}$$

so that reducing $\alpha$ lowers the privately optimal reserve. The formula applies only at an interior optimum. At $\alpha = 0$ the objective reduces to $C$, which is increasing, and the optimum is the lowest admissible reserve, so the derivative need not describe that endpoint. Two different phenomena should be kept apart. In the first, the expected failure loss $pL$ is convex and decreasing in $r$, so that when reserves are cut from a sufficiently high level early cuts barely raise it and later cuts raise it at an increasing rate, as with $Ae^{-kr}$ for positive constants $A$ and $k$. Whether a particular organization begins in that regime is an empirical question that the functional form does not settle. This produces a long period of apparently harmless reductions followed by severe consequences, but it is compatible with the convex case above, and the optimum then changes continuously with $\alpha$. In the second, the convexity assumption fails, which is plausible where reserve capacity consists of indivisible personnel, redundant facilities or discrete recovery arrangements. Nonconvex costs, threshold effects or multiple local minima then allow the globally optimal reserve to change discontinuously as $\alpha$ falls. Both can resemble the avalanche image

that the financial section examines, but only the second makes the optimal capacity itself discontinuous, although a convex loss of steep enough curvature can resemble a threshold in observed failure data even when the optimum moves continuously. The two do not have a simple empirical signature, because a discontinuity can sit in three different objects: the organization's capacity decision, the underlying failure-loss function, and the realized consequences of a particular incident. A discrete staffing decision can produce a jump in the chosen reserve while observed failures remain stochastic and their severity varies continuously, and a continuously declining reserve can cross a physical or organizational threshold and produce an abrupt loss of function. Staffing and incident-severity data alone therefore may not distinguish the cases, and an empirical analysis would have to identify which of the three objects is discontinuous. The problem is then not simply that organizations measure current costs but that the decision-maker's objective excludes part of the expected cost of eliminating resilience. The parameter $\alpha$ is small where managers are evaluated over horizons shorter than the lag between a decision and its failure, where the costs of failure fall on customers, employees or the public, and where market power or regulatory insulation removes the penalty for failing, which restates Doctorow's observation about monopolies in a form that can be tested. These mechanisms are not equivalent. A manager may bear the full monetary consequences of failure and still discount losses that arrive after a departure, whereas another may care about long-term outcomes and face incentives that exclude harm to customers or the public, so a fuller model would separate the time at which failure costs arise, with a discount rate for the decision-maker, from the fraction of those costs that the decision-maker expects to bear. The sign of the comparative static does not depend on the decomposition, since $dr^*/d\alpha > 0$ holds for any $\alpha$ so defined, whether it is small because the decision-maker discounts heavily or because the costs are externalized. The decomposition matters for policy, because it determines which instrument targets which component. The static model is sufficient for the limited proposition here. It contains no substitute, so its result is independent of substitute quality by construction. In an actual organization a successful replacement could lower the expected loss or the cost of capacity and change the optimum, and the defensible claim is only that the incentive to eliminate reserves can arise before the substitute's performance has been demonstrated.

The restated proposition is testable. Its implications include a lag between reductions in operational staffing and increases in the severity or duration of incidents, which could be measured in organizations that report incident data; a relationship between the age and interdependence of systems and the effect of staffing reductions, as the discussion

of layered legacy systems implies; a dependence on market power, so that organizations facing real exit by customers should show weaker effects; and, for the claim about replacing skilled workers, a measured relationship between review volume per reviewer and the rate at which plausible errors are detected. For process knowledge, the testable quantity is the time and cost required to restore a capability after its carriers are lost, as a function of whether those carriers remain reachable and whether documentation, logs and code survive.

Across the five layers the interview is strongest on mechanism and weakest on forecast. The cited reports identify technical causes of the November incidents and no AI involvement in the proximate failure mechanisms, and the index mechanism around the SpaceX offering is partly confirmed. The explanations of motive are plausible but attribute purposes without evidence, and the defensible form is that deployment can weaken professional refusal. The central danger the interview identifies is therefore not that AI is breaking the internet, which the reporting on the named incidents does not support, but that institutions may be using the promise of automation to justify the removal of capabilities whose value becomes visible only after they are gone.

# References


[1] Video interview with Cory Doctorow, "How AI Is Breaking the Internet (and Everything Else)," English auto-generated captions, running time approximately 1:06:33, discussing the November 12 and 13, 2026 outages; recording date not stated in the captions.

[2] Microsoft, *Preliminary Post Incident Report for Microsoft 365*, report dated November 15, 2026 (copy hosted by Next.ink).

[3] Lawrence Abrams, "Microsoft Says Months-Old Bug Caused Global Microsoft 365 Outage," BleepingComputer, November 16, 2026. https://www.bleepingcomputer.com/news/microsoft/microsoft-says-months-old-bug-caused-global-microsoft-365-outage

[4] "Microsoft Outage Explained: A Three-Month-Old Bug's Global Impact," MobileSyrup, November 16, 2026. https://mobilesyrup.com/2026/11/16/microsoft-365-outage-explainer/

[5] "What Forced the US to Ground All Planes? A Router Switch, We Now Know," Al Jazeera, November 15, 2026. https://www.aljazeera.com/news/2026/11/15/what-forced-the-us-to-ground-all-planes-a-router-switch-we-now-know

[6] Ayesha Ali and Meredith Deliso, "FAA Halts Router Installation Following Nationwide Ground Stop," ABC News, November 17, 2026. https://6abc.com/story/faa-halts-router-installation-following-nationwide-ground-stop/20048610/

[7] "The SpaceX IPO: How Index Funds Are Adapting," Morningstar, 2026. https://www.morningstar.com/funds/spacex-ipo-how-index-funds-are-adapting

[8] "The SpaceX Mega IPO: Why Index Choice Matters," CME Group, 2026. https://www.cmegroup.com/articles/2026/the-spacex-mega-ipo-why-index-choice-matters.html

[9] "Nasdaq Rewrote Its Rules for SpaceX; Now $4.3B in Passive Buying Forces Index Funds' Hands," International Business Times, 2026.

[10] "Apple Is Reportedly Facing a Massive Dilemma With the MacBook Neo," MacRumors, April 7, 2026. https://www.macrumors.com/2026/04/07/macbook-neo-massive-dilemma/

[11] "Why the MacBook Neo 2 May Arrive Sooner Than Expected," Macworld, 2026. https://www.macworld.com/article/3108728/why-the-macbook-neo-2-may-arrive-sooner-than-expected.html

[12] Kyle Anthony, "What the SpaceX IPO Reveals About Index Methodology," Cboe, June 25, 2026. https://etfmarket.cboe.com/canada/en/news/what-the-spacex-ipo-reveals-about-index-methodology

[13] "Upcoming Mega-Cap IPOs: Implications for Index Investors," Northern Trust, June 12, 2026. https://www.northerntrust.com/united-states/insights-research/2026/point-of-view/upcoming-mega-cap-ipos-implications-for-index-investors

[14] Michael Spence, "Job Market Signaling," *Quarterly Journal of Economics* 87(3), 355–374, 1973.

[15] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention Is All You Need," *Advances in Neural Information Processing Systems* 30, 2017.

[16] Milton Friedman, *Capitalism and Freedom*, University of Chicago Press, 1962; preface to the 1982 edition.

[17] Lisanne Bainbridge, "Ironies of Automation," *Automatica* 19(6), 775–779, 1983.

[18] Charles Perrow, *Normal Accidents: Living with High-Risk Technologies*, Basic Books, 1984.

[19] Julian E. Orr, *Talking About Machines: An Ethnography of a Modern Job*, Cornell University Press, 1996.

[20] Michael Polanyi, *The Tacit Dimension*, Doubleday, 1966.

[21] Philip Mirowski and Dieter Plehwe, editors, *The Road from Mont Pèlerin: The Making of the Neoliberal Thought Collective*, Harvard University Press, 2009.