
Regulating the “Unregulatable”: Why Policymakers Are in Paralysis When Regulating AI

Sarah Barrington and Hannah Bailey

Policymakers confronting generative AI have often accepted a striking premise: that AI is too complex, too fast-moving, and too “unprecedented” to be governed by existing frameworks. We argue that this premise is itself part of the problem. Many harms associated with generative AI, including fraud, impersonation, deceptive advertising, and exploitative synthetic media, are not wholly new categories of threat. They are familiar harms, now produced and distributed more cheaply, more quickly, and at greater scale. To move beyond the current policy paralysis, we propose decomposing AI into three layers: the model layer, the deployment layer, and the distribution layer. Once disaggregated in this way, AI no longer appears as an ungovernable monolith, but instead as a supply chain of software products and infrastructures that already sit within consumer protection, product liability, sectoral, and platform-governance regimes. Governing actors should stop treating the absence of legal authority as the obstacle. The real challenge, we argue, is the persistence of AI exceptionalism and the resulting under-enforcement of ordinary law.

1 Introduction

Generative AI has intensified policy anxiety with unusual speed. In only a few years, models capable of producing realistic text, audio, image, and video outputs have been integrated into search engines, office software, social media platforms, advertising tools, and consumer applications. Governments have responded with executive orders, regulatory principles, voluntary commitments, and a scattering of targeted statutory measures (Federal Communications Commission 2024; EU 2024b; Executive Office of the President 2023). Despite this activity, the overall response remains confused and inconsistent. Even where policymakers recognize serious harms, they still struggle over

whether existing law applies, which institutions should act, and whether entirely new governance frameworks are required.

Part of this confusion is conceptual. Industry actors and policymakers often describe AI as a qualitatively new category that sits outside ordinary legal and institutional governance tools. Terms such as “frontier models,” “artificial general intelligence,” “existential risk,” and “agentic AI” all reinforce the notion that AI is uniquely complex and difficult to govern. This framing has obvious political uses. For one, it casts AI as too novel and complex to be governed by the ordinary instruments of consumer protection, product liability, and competition law. It also helps sustain the inflated promise and valuation on which the contemporary AI market depends.

And yet, the language of “unprecedentedness” is itself highly precedented. Similar narratives accompanied earlier periods of deregulation; for example, during the rise of social media platforms, firms emphasized novelty, neutrality, and technical complexity while resisting classification under older media, advertising, and competition frameworks (Gillespie 2010; Napoli and Caplan 2017). Before the 2008 financial crisis, the financial derivatives industry made comparable arguments about complexity, speed, and the inadequacy of inherited rules (Financial Crisis Inquiry Commission 2011; Stout 2011; Awrey 2012). In both settings, firms and their allies warned that intervention would break innovation or exceed the capacity of existing law, and instead the firms offered voluntary best practices in place of binding obligations. As a result, in both instances there was a delay in ordinary oversight, which then shielded products and practices from liability until the harms were too widespread to ignore.

Seen in this light, many of the harms now attributed to AI are less novel than the surrounding narratives suggest. For example, deepfakes intensify long-standing problems of impersonation, fraud, and harassment. Automated content generation and recommendation systems similarly build on familiar advertising and engagement infrastructures. What is new is therefore not necessarily the type of harm enabled by AI, but rather the speed, scale, and opacity enabled by contemporary data pipelines, model architectures, and platform-enabled distribution. We propose that once AI is decomposed into three layers, namely model, deployment, and distribution, it comprises a supply chain of software products, services, and infrastructures that already sit within established legal regimes. Policymakers should instead be addressing the question of which actors at which layer should be held to which obligations, using which already available regulatory or legislative tools.

2 The Myth of Unprecedented AI: Familiar Harms at a New Scale

The harms currently associated with AI are concentrated in a particular subset of systems known as *generative AI*, meaning “computational techniques that are capable

of generating seemingly new, meaningful content such as text, images, or audio from training data” (Feuerriegel et al. 2024). Although these models are often framed as ushering in entirely new forms of digital harm, a more accurate interpretation is that they amplify long-standing problems. Fraud, harassment, propaganda, and deceptive advertising all predate the present AI boom.

We argue that the fundamental nature of these harms has not changed; rather, what has changed is the speed and scale with which specific technologies can produce these harms. In practice, “AI” in this context typically refers to a narrower set of technologies, comprising large language models (LLMs), vision-language models (VLMs), and text-, image-, and audio-to-video generation models, each with its own distinct architectures, training data, and points of failure, rather than a single monolithic technology. Referring to these systems by their technical names is important, since they are specific tools that produce the harms discussed below, not a single abstract “AI” category.

We recognize that a growing body of literature suggests generative AI may still introduce novel forms of harm. One such example of this is prominent work indicating that AI-generated content can be more persuasive than human-authored content. Even so, studies are inconsistent on this issue (for example, a study found that persuasiveness does not scale with language model size (Hackenburg et al. 2025)). So while future advances in AI could present new harms, this is not necessarily the case at present. In the case of AI-generated persuasive synthetic content, what AI currently offers is the easy generation of such synthetic media at low cost (Blakemore 2018; Barrington, Cooper, and Farid 2025).

This holds across many of the most prominent harms associated with AI. For example, mis- and disinformation predate the internet (Tucher 2022; Lasswell 1971; Rid 2020; Vosoughi, Roy, and Aral 2018), but a campaign that once required a team of operators can now generate text, visuals, or fabricated audio at scale with a single model and a few lines of input text (Bradshaw, Bailey, and Howard 2021; Goldstein et al. 2023). Generative AI has already enabled scams at both large and small scales (Chen and Magramo 2024; AARP 2024), but fraud and impersonation are centuries-old (Pappano 1937). Generative AI models can also produce non-consensual intimate imagery or fabricated material depicting minors (Ding and Suresh 2025; Farid 2022; Wong 2024; Kang 2025); however, it is worth noting that fabrication does not make this categorically different from authentic content, since synthetic content reproduces many of the same harms that authentic material inflicts on victims and viewers (Rini and Cohen 2022).

In all cases, generative AI is therefore better understood as a tool for automating and amplifying these harmful practices, rather than as the source of an entirely new category of harm. These continuities should change how the regulatory problem is framed. If the core harms extend earlier forms of digital and commercial wrongdoing, then the relevant question is why existing tools against fraud, deception, harmful products, exploitative content, and unfair practices are not being applied more consistently to a new set

of software products. Yet treating AI as an exceptional technology has encouraged policymakers to begin with “AI” as a freestanding object of regulation, rather than with the specific product, practice, or harm at issue. This tendency can distract from the more prosaic but more urgent task of enforcing existing law against familiar harms that now appear in automated form. We return to this distinction in Section 3.

This continuity of harms has not been matched by continuity in governance. Policymakers have often approached AI as a freestanding regulatory category, even though the substantive problems grouped beneath that label differ considerably. The National Conference of State Legislatures tracked over 1,000 state AI measures introduced in 2025, but those measures addressed widely differing matters, including government use, health care, employment, consumer disclosure, discrimination, children, chatbots, and synthetic media (Canada 2026). Their heterogeneity is precisely the point: a common “AI” label has become the organizing principle for a wide range of otherwise distinct policy problems. In short, policymakers too often begin by asking, “How should we regulate AI?” rather than, “What conduct or harm are we regulating, and does existing law already address it?” We should caveat this by noting that some AI-specific measures are warranted, particularly where existing law leaves notable gaps. Our concern, though, is with making the technology label the starting point of regulation, since doing so can obscure whether the underlying conduct is, for example, already governed as employment discrimination, deceptive advertising, unsafe medical practice, privacy intrusion, fraud, or another familiar category of harm. The example of Executive Order 14110 illustrates both approaches, as it directed agencies to consider AI within their existing sectoral authorities, but also established a novel compute-threshold reporting regime for certain dual-use foundation models (Executive Office of the President 2023).

Yet, the “Unprecedentedness” rhetoric continues to appear across the tech industry and media, and is repeatedly voiced by some of the industry’s most influential figures. In a single essay titled “The Adolescence of Technology,” Anthropic CEO Dario Amodei describes AI as posing “unprecedentedly large and rapid” labor-market disruption, “historically unprecedented levels of wealth concentration,” and a “serious and unprecedented threat” in cybersecurity (Amodei 2026a). Similarly, Google and DeepMind CEOs Pichai and Hassabis have likened the advent of AI to something as profoundly new as fire (Prakash 2023; Nelson 2026). Internal governance documentation at tech companies also encodes this framing: OpenAI’s Preparedness Framework formally defines a “critical capability” as one that could introduce “unprecedented new pathways to severe harm,” distinct from capabilities that merely amplify existing ones (OpenAI 2025). It is noteworthy that this language fits the classic playbook of “technological exceptionalism,” which justifies treating AI as falling outside the bounds of typical law and society (Kaminski and Jones 2024; Jones 2018).

A central strand of this persistent “Unprecedented” narrative is the conflation of current AI systems with artificial general intelligence (AGI). AGI is associated with a wide range

of definitions that generally converge on the concept of AI matching adult human-level cognition, but vary in the breadth of tasks a system can perform, the level of performance it must reach, the degree of autonomy it possesses, and whether it must equal or exceed human capabilities (Hendrycks et al. 2025). This variation, however, makes AGI a poor category for regulating present-day products, as the term does not identify a stable technical threshold, a particular deployment context, or a responsible actor. It instead refers to a hypothetical level or breadth of capability. Whatever one's view of whether AGI is attainable, it is not the technology currently embedded in most products and services. Present-day AI systems are better understood as narrower, probabilistic, and domain-variable tools deployed in particular contexts. They draft text, rank content, summarize documents, generate images, assist with coding, classify data, support decision-making, or power recommendation and advertising systems.

Treating AGI as the reference point for regulation distracts policymakers from immediate harms and distorts governance in three notable ways: first, by legitimizing a delay in regulation in anticipation of future AGI; second, by redirecting policy efforts toward speculative “alignment” and “existential risk” research at the expense of the sector-specific harms we discussed above; and finally, by privileging benchmarks of general capability over deployment-specific questions of safety, fairness, and market power (Narayanan and Kapoor 2025).

More broadly, the appeal to an unprecedented AI collapses a wide range of systems into a single notion of a monolithic “AI” category. Automation tools, recommender systems, chatbots, image generators, predictive models, and autonomous decision systems alike are all grouped together under the label of “AI,” despite operating differently and raising different regulatory questions. Furthermore, the framing of AI as monolithic is increasingly inaccurate on both technical and regulatory grounds. In most cases, contemporary AI products, such as agentic systems or coding assistants, comprise multiple models in the software pipeline, including not just an LLM but also retrieval systems, auxiliary classifiers, and API access (for example, Claude Code (Anthropic 2025)). Treating AI as a single object serves to obscure this internal structure and makes it even harder to understand which components are used in which deployment contexts. The result is that, at present, regulatory effort predominantly goes toward broad AI bills and advisory frameworks instead of using existing tools that target specific components, use cases, or actors, making ordinary tools like product safety, consumer protection, and competition law appear ill-suited by comparison.

3 The Path out of Policy Paralysis

Thus far we have discussed how AI can be strategically framed as an “unprecedented” technology. Here, we ask how that script then translates into policy inaction. We argue that the rhetoric of unprecedented AI operates through a recognizable playbook, and

importantly, several actors, from model developers to product deployers and platform distributors, are independently motivated to align with this playbook. Because these actors are all incentivized to depict AI as “unprecedented,” this can produce an effort that appears coordinated, even in the absence of explicit collusion. In this respect, AI governance resembles what platform scholars often describe as a “problem of many hands” (Thompson 1980), where responsibility is dispersed across multiple actors in ways that make accountability easier to evade. Within this “unprecedented” framing, we posit that three tenets are especially important.

Tenet 1: Complexity as a Shield against Ordinary Liability. The first tenet is that AI is too broad, too technical, and too general-purpose to fit within existing product liability and consumer protection frameworks. This framing often relies on umbrella terms such as “foundation models,” “agentic AI,” and the aforementioned “AGI.” These terms identify meaningful technical characteristics, but they operate at a level of generality that is often too broad to determine the obligations attached to a particular product or use. Bommasani et al. define a foundation model as “any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks” (Bommasani et al. 2021). This definition, while useful, is nevertheless broad relative to most product-level regulatory questions because it classifies models according to their training and potential adaptability while leaving unspecified the downstream application, user population, degree of human oversight, marketing claims, tool access, and relevant form of harm. Similarly, Shavit et al. define agentic AI systems as systems that can “pursue complex goals with limited direct supervision” (Shavit et al. 2023). In practice, such capacity often arises from a composite deployment in which a model is combined with planning or orchestration software, memory, and access to external tools or APIs. Agenticness is therefore better understood as a matter of degree and system configuration than as a bounded class of models. As such, the general label does not identify which component made the consequential decision or which actor controlled its deployment. These categories are thus useful for describing technical systems, but lack the detail necessary to understand how their deployment fits into existing product liability and consumer protection frameworks. This distinction is further developed in Section 3.2.

Tenet 2: Claims That Regulation Will Stifle Innovation. The second tenet underpinning the “unprecedented” narrative holds that binding regulation will stifle innovation, undermine competitiveness, or hand an advantage to geopolitical rivals. This is a claim that carries particular force given the growing number of ties linking AI to national security (NTIA 2024; Bureau of Industry and Security 2025). In practice, this logic steers policymakers toward hybrid arrangements that combine limited mandatory obligations with extensive voluntary commitments, as evidenced again by Executive Order 14110’s reporting requirements for dual-use foundation models, which nonetheless left firms substantial discretion (Executive Office of the President 2023; Guha et al. 2024; Bureau

of Industry and Security 2024). This appeal to competitiveness also rests on a selective reading of history: technologies regarded as strategically important, ranging from the Manhattan Project to ARPANET and the F-35, were typically heavily funded, directed, and constrained by the state, not left to open commercial competition. And yet modern-day AI is unusual by comparison, as it is framed as strategically vital while being pushed rapidly into mass consumer markets and, for frontier systems, sold or open-sourced in ways that extend their reach to potential adversaries (Bureau of Industry and Security 2025). Furthermore, strict regulatory regimes routinely coexist with innovation elsewhere, as in the pharmaceutical industry (U.S. Food and Drug Administration 2024; Stern 2017) (further explored in Section 3.2).

Tenet 3: Free Expression, Provenance, and Consent. The third tenet is perhaps the most effective at deflecting regulation. Because generative AI typically causes harm through the *content* a model produces when prompted by a user rather than through the model itself, industry actors can argue that any legislation targeting those outputs would infringe on the user’s freedom of speech. This argument is reinforced by recent First Amendment litigation over platform moderation and compelled disclosures (*Moody v. NetChoice, LLC* 2024; Lake 2024). Section 230 of the Communications Decency Act compounds this uncertainty: it gives platforms broad protection from liability for user-generated content, but it remains contested whether those protections should extend to content generated by AI systems rather than by third-party users (U.S. Congress 1996; Perault 2023; Volokh 2023). This argument lets developers direct attention toward speech protection, and so they can deter policymakers from regulating not only the outputs produced by generative AI, but also the broader AI development cycle.

Yet, within this debate there is still room for a regulatory response. Rather than banning synthetic content outright, policymakers could instead require disclosures when content is AI-generated, attach labels to it, or perhaps condition its algorithmic amplification and advertising on available provenance signals. Where such provenance information is recorded at the point of generation, it can reduce dependence on post-hoc detection, which is considered increasingly unreliable.

The C2PA Content Credentials standard, already adopted by Adobe, Microsoft, and other major platforms, provides one mechanism for attaching cryptographically signed information about the source and editing history of a digital asset (C2PA 2026a). However, it has limits: a valid credential establishes only that recorded provenance information remains unaltered, and does not confirm whether the underlying content is true or non-deceptive, and in some cases, credentials can be stripped through ordinary copying, editing, or platform re-processing (C2PA 2026a, 2026b, 2026c). The practical value of provenance also depends on coordinated adoption by model providers, editing applications, platforms, advertisers, and user interfaces. Some resistance to adoption is principled, as poorly designed implementations may disclose sensitive information or facilitate surveillance, risks that are particularly important for journalists, human-rights defenders, political

dissidents, and other users for whom anonymity is a condition of safety (C2PA 2026c, 2026b). The C2PA approach is also notably weaker for text than for image, audio, or video, in which ordinary copying, pasting, or paraphrasing readily separates generated text from any embedded manifest, so text requires other mechanisms such as account-level transparency or enforcement against deceptive representations. While imperfect, provenance still provides value and it should accordingly be treated as one evidentiary signal within a wider governance regime rather than as a complete authentication solution.

Yet, in practice, provenance systems, labels, and ranking adjustments have been implemented inconsistently (Mantzaris 2025; Mantzaris and Dutta 2025; Tiku and Verma 2025), and the association between provenance and speech regulation has given platforms cover to roll back voluntary safeguards, as in Meta’s January 2025 decision to end its third-party fact-checking program in the United States (McMahon, Kleinman, and Subramanian 2025). In short, while some may argue that interventions targeting AI-generated content may be constrained by speech regulation, even if this were the case, there is still room to regulate. Provenance measures, while themselves limited, provide one such starting point to protect users from much of this harm while remaining within the constraints of speech regulation.

3.1 Decomposing the AI Supply Chain

We argue that provenance measures are just one example of the broader approach that policymakers should take to AI regulation, that is, to regulate AI at the points where it already exists as familiar infrastructure, products, and distribution systems. Foundation model development is concentrated in a relatively small number of firms and operates on centralized infrastructure (Vipra and Korinek 2023; Korinek and Vipra 2024). At the same time, downstream products and distribution systems already sit within familiar bodies of law. The task at this point is therefore to allocate responsibility across the relevant layers rather than to search for a single master instrument capable of governing all of AI at once.

This is not the first call for a more disaggregated approach to AI governance. The EU AI Act, for example, distinguishes between providers, deployers, importers, distributors, product manufacturers, and providers of general-purpose AI models, and thus allocates obligations across the AI value chain (EU 2024b). Scholars such as Narayanan and Kapoor argue that AI should be understood as a “normal technology” governed through familiar institutional tools, while work on AI supply chains and the machine-learning life cycle has similarly located responsibility across specific actors and stages rather than treating AI systems as unitary objects (Narayanan and Kapoor 2025; Cobbe, Veale, and Singh 2023; Binns 2022). Our contribution is to connect that insight to the unprecedentedness playbook of tenets described above. We show that existing law can be mapped onto each layer, and why that law has so often gone unapplied: the rhetoric of AI exceptionalism has obscured where responsibility lies and encouraged regulators to treat governable systems as if they were beyond reach. Figure 1 sets out the contents of each layer, the

	Contents - What	Key Actors - Who	Governance Mechanisms - How
Model	Core AI Technologies Pre-trained AI models New model development Training data management Safeguards	National AI safety institutes Standards bodies (e.g., NIST) Frontier AI labs Cloud and compute providers	Joint industry-policy standards Auditing Red-teaming Compute thresholds Training data transparency
Deployment	Software Products & Wrappers Browser interfaces Mobile applications Marketing claims and use cases	Enterprise & government customers (procurement requirements) Courts and judicial systems App store operators Upstream infrastructure gatekeepers Industry-specific regulators	Product liability and class action litigation Deplatforming or service blocking (e.g., app stores, payment processing, cloud hosting)
Distribution	Distribution Platforms Social media feeds Search engines Messaging platforms Advertising systems	Social media companies FTC and FCC Industry consortia (e.g., C2PA) Congress and Senate judiciary committees	Advertising law Consumer protection law Platform terms of service Private rights of action (e.g., "TAKE IT DOWN" act) Antitrust law

Figure 1: An overview of the Model, Deployment, and Distribution layers. Reframing the AI monolith as three distinct layers shows where existing policy and regulatory frameworks overlap with technical components ("Contents - What"), stakeholders ("Key Actors - Who"), and applicable interventions ("Governance Mechanisms - How").

main actors shaping it, and the governance mechanisms already available. We then examine each in turn.

3.2 The Model Layer (Infrastructure)

The model layer comprises the core AI technologies themselves, including pre-trained models, training pipelines through which they are developed, safeguards, and compute infrastructure (see Figure 1, "Model" row). This is the layer at which much of the current policy debate is concentrated, particularly around "frontier models," compute thresholds, and existential risk, and it is the layer most often framed as beyond the reach of existing law. In practice, however, we argue that the model layer maps naturally onto supply chains we already know how to regulate. This layer maps onto Tenet 1 (i.e., the "complexity shield"), which breaks down AI into its constituent models and simplified functionality.

The physical infrastructure behind foundation models is already embedded in regulated systems. The training process itself runs on GPU clusters provided by a small number of cloud and compute providers, principally Amazon Web Services, Microsoft Azure, and Google Cloud, which together account for over 60% of the global cloud infrastructure market (Synergy Research Group 2025). The chips that power these clusters are manufactured by an even smaller number of firms, most notably NVIDIA and TSMC (TechInsights 2024), and are already governed, albeit imperfectly, by existing export control regimes (Gupta, Walker, and Reddie 2024). In other words, the physical and computational makeup of AI model development is already embedded in regulated supply

chains, and the concentration of these resources among a small number of providers (Vipra and Korinek 2023) means that the points of regulatory leverage are few and well-defined.

Training data transparency raises a further set of questions around intellectual property, privacy, and bias (Bommasani et al. 2023; Longpre et al. 2024). Yet each of these areas is already governed, at least in principle, by existing law. Copyright questions around foundation-model training, for example, rely on established doctrines such as fair use, market substitution, and substantial similarity (Henderson et al. 2023). In the *New York Times v. OpenAI* litigation, for instance, claims alleging copyright infringement, contributory and vicarious infringement, and DMCA violations all fall under existing intellectual property frameworks (*New York Times Co. v. Microsoft Corp.* 2023; *New York Times Co. v. Microsoft Corp.* 2025). Similarly, the EU GDPR imposes obligations on data processing regardless of whether that processing occurs in a model training pipeline or in a more conventional system (EU 2016).

The legal foundations for governing these issues therefore already exist. What regulators lack is the practical means to exercise that authority. Datasets are often undocumented even by their creators (Longpre et al. 2024), and meaningful audits require sustained human review. Rather than a problem of law, this is a resourcing and access problem. A substantial technical literature has already developed solutions to this problem, involving, among other things, encouraging standardized documentation practices such as datasheets for datasets (Gebru et al. 2021) and model cards (Mitchell et al. 2019), and audit methodologies for surfacing biased or unsubstantiated performance claims in deployed systems (Raji and Buolamwini 2019; Birhane et al. 2024). These contributions show that auditing AI systems is, in the narrow technical sense, a solved problem; what has not followed is the institutional mandate, funding, and legal obligation to apply them systematically. We argue that the inconsistency of their use is a governance failure rather than a technical one, and that responsibility for closing that gap lies at the model layer, with actors and mechanisms already identifiable in Figure 1.

Yet regulating inputs alone is insufficient, because the model produced from those inputs also requires direct oversight before release. Other industries have long followed this same logic. Pharmaceuticals must undergo premarket review and, in many cases, clinical trials before approval, while financial institutions are stress-tested under adverse scenarios to assess whether they can withstand systemic shocks (U.S. Food and Drug Administration 2024; U.S. Congress 2010; Federal Reserve Board 2024). In AI, the closest equivalents are auditing and red teaming, which are structured pre-deployment evaluations designed to identify failures before they occur at scale (Shevlane et al. 2023; Phuong et al. 2024; Feffer et al. 2024; Casper et al. 2024). Governments have already begun to move in this direction. The UK AI Security Institute has conducted evaluations of frontier AI systems since November 2023, including pre-release testing of some frontier models (AI Security Institute 2025; Perrigo 2025), and Executive Order 14110, before it

was rescinded, established reporting requirements for models above certain compute thresholds (Executive Office of the President 2023, 2025). The underlying principle is familiar from product regulation more broadly, namely that developers should be required to demonstrate safety before release.

The key actors at the model layer are AI safety institutes, standards bodies such as NIST, frontier AI laboratories, and cloud and compute providers (Figure 1). AI safety institutes now run pre-release and frontier-model evaluations, while NIST's AI Risk Management Framework provides a standards-based vocabulary for identifying, measuring, managing, and governing AI risk (Tabassi 2023; Bommasani et al. 2023; Henderson et al. 2023; EU 2016; Bureau of Industry and Security 2023). Governance at the model layer should focus on pre-deployment model safety, training data transparency, and compliance with existing data protection, intellectual property, and export obligations, without trying to anticipate every downstream use. That work belongs to the next two layers.

3.3 The Deployment Layer (Product)

The deployment layer is where the model infrastructure becomes a product (see Figure 1, "Deployment" row). It encompasses the software "wrappers" that expose model capabilities to end users, including browser-based interfaces such as ChatGPT and Google Gemini, mobile applications such as Grok, enterprise integrations such as Microsoft Copilot, and specialized tools in sectors such as health care diagnostics and financial credit scoring. As mentioned in Section 2, these products are composite systems rather than a single wrapped model. This composition is important for liability, as responsibility for a harmful output may fall to various actors within the product, such as the model provider, the agent-orchestration layer, or the specific tool integration that executed the action. In short, deployment-layer obligations need to be capable of reaching whichever component made the relevant design choice. At this layer, a firm has made specific design choices about which model to use, how to configure it, what guardrails to apply, what claims to make about its capabilities, and whom to market it to. These decisions collectively define the product, and with them come existing legal obligations. This targets our Tenet 2 (i.e., the "innovation shield") by showing where AI maps onto the existing product innovation cycle.

Two bodies of law are immediately relevant at this layer. The first is consumer protection law, which applies to all products regardless of the industry in which they are sold. In the United States, the Federal Trade Commission (FTC) has broad authority under Section 5 of the FTC Act to pursue "unfair or deceptive acts or practices," an authority that applies irrespective of whether the product in question uses AI or any other technology (U.S. Congress 1914). If a company markets an AI-powered legal-advice chatbot that provides inaccurate guidance, the relevant question in this particular case is whether the product's marketing claims match its actual performance. This is the same standard applied to any consumer product. The FTC has already signaled this intent through "Operation AI Comply," which has brought actions against companies marketing AI-powered products

with unsubstantiated claims (Federal Trade Commission 2024).

Indeed, legal scholars have noted that the FTC’s existing Section 5 authority, alongside other consumer protection statutes, is already sufficient to address many AI-related harms without the need for AI-specific legislation (Corrigan et al. 2024; Guha et al. 2024). The EU’s revised Product Liability Directive similarly extends strict liability to software, including AI-enabled products (EU 2024a).

The second body of law is sector-specific regulation, which imposes additional requirements when an AI product operates within a regulated industry. The principle is thus that, when an AI system is used in a regulated sector, it should be governed by the rules of that sector. This is already happening in many instances (Muehlematter, Daniore, and Vokinger 2021). The U.S. Food and Drug Administration, for example, has regulated AI-enabled medical devices under its existing premarket review authority, including granting the first De Novo authorization for an AI-based sepsis prediction tool in 2024 (Bhargava et al. 2024). Automated driving systems are regulated as components of vehicle safety under the National Highway Traffic Safety Administration (NHTSA) (NHTSA 2017). And, in financial services, AI-powered credit-scoring tools are subject to the Equal Credit Opportunity Act (U.S. Congress 1974) and Fair Credit Reporting Act (U.S. Congress 1970; Barocas and Selbst 2016). In each instance, again, the regulatory authority already exists. A medical AI product is ultimately a medical device, and a financial AI product is a financial instrument.

3.4 The Distribution Layer (Platform)

The distribution layer consists of the app stores, social media feeds, search systems, messaging platforms, advertising exchanges, and recommendation infrastructures through which AI products and their outputs reach users (see Figure 1, “Distribution” row). This is the layer that determines who sees what, in what context, with what degree of amplification, and with what commercial incentive behind that amplification. It is also the layer that is least developed in much current AI policy debate, even though many of the most visible harms associated with generative AI manifest here. This maps onto our Tenet 3 (i.e., “free expression”) by clearly drawing the boundary between the underlying technology and the content it produces.

We argue here that distribution deserves separate treatment because the downstream effects of AI are rarely determined by model capability alone. For example, the same synthetic image, chatbot output, or voice clip can be trivial in one setting and highly consequential in another depending on how it is ranked, recommended, targeted, monetized, or recirculated. Platform scholarship has long shown that recommendation, ranking, moderation, advertising systems, and visibility reduction are themselves forms of governance rather than neutral channels of transmission (Gillespie 2022; Gorwa 2024). The distribution layer should therefore be treated as an important part of the causal chain through which harm is produced.

While social media is an obvious example, we recognize that distribution is not always platform-controlled. AI outputs frequently reach audiences only after a user has generated content through a deployment-layer product and chosen to share it independently, via private messaging, screenshots, re-uploads, or peer-to-peer channels that sit outside any single platform's ranking or recommendation systems. This distinction is particularly important for regulatory design. Where distribution is platform-mediated, obligations can attach to ranking, recommendation, and advertising infrastructure, as discussed above. Where distribution is user-mediated, the more appropriate levers are notice-and-takedown obligations and rapid removal mechanisms rather than earlier detection methods, an approach already reflected in the federal TAKE IT DOWN Act, which requires platforms to remove non-consensual intimate imagery upon notice, regardless of how it reached the platform (U.S. Congress 2025). So while there are two pathways to regulation at the distribution layer, both are well-trodden and can be addressed in different ways.

Existing legal frameworks already reach many of these activities. Platform duties can arise through advertising law, consumer protection, intermediary regulation, privacy law, electoral law, and competition policy. In the European Union, the Digital Services Act imposes obligations around advertising transparency, recommender-system transparency, systemic-risk assessment, and risk mitigation by very large online platforms and search engines (EU 2022). EU enforcement against X and information requests to other large platforms have already focused in part on recommender systems and transparency obligations (European Commission 2023, 2024). In the United States, while the legal landscape is more fragmented, the underlying principles are nevertheless familiar. Platforms that profit from deceptive synthetic advertising, facilitate the spread of exploitative apps, or misrepresent their moderation and labeling practices remain exposed to consumer protection and related forms of scrutiny (Federal Trade Commission 2024; Lake 2024; Kaminski and Jones 2024).

Competition law also warrants a mention here, as distribution infrastructures are dominated by a small number of firms that control discovery, promotion, and monetization at scale. Search engines, app stores, and social platforms all determine which AI products gain visibility and which do not. They also influence which kinds of AI-generated outputs are commercially rewarded. Because of this, concentration at the distribution layer can therefore reinforce concentration at the model and deployment layers by privileging vertically integrated firms or by making access to audiences dependent on a small number of gatekeepers (Khan 2017; Crémer, Montjoye, and Schweitzer 2019; Gorwa 2024; Vipra and Korinek 2023; Korinek and Vipra 2024). In summary, a serious AI governance agenda therefore cannot avoid questions of platform power simply because they appear to belong to communications policy or competition law rather than to AI policy in the narrow sense.

4 Conclusions: New Technology, Same Old Problems

Generative AI has been treated as a governance crisis because it is so often described as unprecedented in form, scale, and consequence. This article argues that the more important problem lies in that description itself. The language of unprecedentedness has made ordinary regulation seem inapplicable at precisely the moment when existing law is most needed. We argue that the central obstacle to effective governance has therefore been a political and conceptual framing that casts governable products and infrastructures as though they sit beyond the reach of established institutions.

We argue that this framing draws attention away from the places where responsibility can actually be assigned, which can be made clear by instead breaking AI down into its model, deployment, and distribution layers, thereby making the regulatory problem much easier to assess. At each of these layers there are identifiable actors, existing obligations, and institutions with a basis to act. For that reason, we argue that claims that AI is fundamentally unregulatable should be treated with skepticism.

5 Addendum

In light of events in mid-2026, we add the following addendum. The Hugging Face incident in July 2026 (METR 2026) has been widely characterized as an unprecedented moment: agents built on OpenAI models circumvented sandbox controls and gained access to the public internet. While operating with reduced safeguards in a testing capacity, the agents obtained and used exposed credentials and previously unknown vulnerabilities to compromise parts of Hugging Face’s production infrastructure in a sophisticated attack comprising over a thousand agents and thousands of attacker actions (OpenAI 2026; Hugging Face 2026; METR 2026).

While the agentic coordination effort, along with the resulting intrusion into third-party production infrastructure, represents a significant advance in AI capabilities, we still maintain that the incident does not show that such systems are beyond ordinary governance, even in the cybersecurity domain. Ultimately, the harms posed by these models are familiar within cybersecurity law and practice. Conduct such as unauthorized access, the misuse of credentials, arbitrary code execution, and data exfiltration constitutes a well-established cybersecurity threat. Indeed, this incident resulted from a series of identifiable model actions and organizational choices made by the actors in question across the layers identified in this article. At the model layer, these were decisions about evaluation design and safeguards made by OpenAI. At the deployment layer, infrastructure and platform operators, here both OpenAI and Hugging Face, retained responsibility for access controls and the security of their systems. At the distribution layer, the account-integrity and content-moderation policies of open repositories and community platforms allowed malicious artifacts to be uploaded and disseminated.

This incident and its aftermath illustrate core tactics from the “unprecedentedness” playbook. The attack has generated substantial public discourse on existential AI risk, much of it advanced by the frontier developers themselves (Ramkumar 2026; Amodei 2026b). This discourse allows the recasting of near-term harms, such as security breaches, as novel and exceptional threats rather than familiar and immediate risks. This, in turn, creates space for AI leaders to suggest alternatives to regulation, such as reliance instead on “trust” (BBC News 2026) or voluntary commitments (Amodei 2026b). In short, while it is important to acknowledge that the agentic mechanism and autonomous scale of the Hugging Face incident were unusual, the harms that occurred, as well as the organizations, infrastructures, and design decisions through which they occurred, remain familiar and governable through existing means.

References

- AARP. 2024. "Chatbots and Voice-Cloning Fuel Rise in AI-Powered Scams." AARP Arizona. Accessed July 22, 2026. <https://states.aarp.org/arizona/chatbots-and-voice-cloning-fuel-rise-in-ai-powered-scams>.
- AI Security Institute. 2025. *Frontier AI Trends Report*. UK Department for Science, Innovation and Technology. <https://www.aisi.gov.uk/frontier-ai-trends-report>.
- Amodei, Dario. 2026a. "The Adolescence of Technology." Accessed July 13, 2026. <https://darioamodei.com/essay/the-adolescence-of-technology>.
- . 2026b. "We Must Pace the Frontier," September 12, 2026. Accessed September 25, 2026. <https://darioamodei.com/post/we-must-pace-the-frontier>.
- Anthropic. 2025. "Claude Code: Auto Mode." Accessed July 22, 2026. <https://www.anthropic.com/engineering/claude-code-auto-mode>.
- Awrey, Dan. 2012. "Complexity, Innovation, and the Regulation of Modern Financial Markets." *Harvard Business Law Review* 2 (2): 235–94. https://journals.law.harvard.edu/hblr/wp-content/uploads/sites/87/2012/11/HLB204_Complexity.pdf.
- Barocas, Solon, and Andrew D. Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104 (3): 671–732.
- Barrington, Sarah. 2025. "Generative AI Is Neither Too Unprecedented nor Too New to Regulate." Tech Policy Press, November 14, 2025. <https://www.techpolicy.press/generative-ai-is-neither-too-unprecedented-nor-too-new-to-regulate/>.
- Barrington, Sarah, Emily A. Cooper, and Hany Farid. 2025. "People Are Poorly Equipped to Detect AI-Powered Voice Clones." *Scientific Reports* 15 (1). <https://doi.org/10.1038/s41598-025-94170-3>.
- BBC News. 2026. "OpenAI CEO Sam Altman Says 'Right to Be Afraid of AI' — But 'World Should Trust' Tech Firms," September 14, 2026. Accessed September 25, 2026. <https://www.bbc.com/news/live/cm70d91rg7l4t>.
- Bhargava, Akhil, Carlos López-Espina, Lee Schmalz, Shah Khan, Gregory L. Watson, Dennys Urdiales, Lincoln Updike, et al. 2024. "FDA-Authorized AI/ML Tool for Sepsis Prediction: Development and Validation." *NEJM AI* 1 (12). <https://doi.org/10.1056/AIo2400867>.
- Binns, Reuben. 2022. "Human Judgment in Algorithmic Loops: Individual Justice and Automated Decision-Making." *Regulation & Governance* 16 (1): 197–211. <https://doi.org/10.1111/rego.12358>.
- Birhane, Abeba, Ryan Steed, Victor Ojewale, Briana Vecchione, and Inioluwa Deborah Raji. 2024. "AI Auditing: The Broken Bus on the Road to AI Accountability." In *2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, 612–43. IEEE. <https://doi.org/10.1109/satml59370.2024.00037>.

- Blakemore, Erin. 2018. "How Photos Became a Weapon in Stalin's Great Purge." A&E Television Networks, April 20, 2018. Accessed January 27, 2026. <https://www.history.com/articles/josef-stalin-great-purge-photo-retouching>.
- Board of Governors of the Federal Reserve System. 2024. "2024 Stress Test Scenarios." Federal Reserve Board, February 15, 2024. <https://www.federalreserve.gov/publications/2024-stress-test-scenarios.htm>.
- Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. 2021. "On the Opportunities and Risks of Foundation Models." *arXiv*, <https://doi.org/10.48550/arXiv.2108.07258>.
- Bommasani, Rishi, Kevin Klyman, Shayne Longpre, Sayash Kapoor, Nestor Maslej, Betty Xiong, Daniel Zhang, and Percy Liang. 2023. "The Foundation Model Transparency Index." *arXiv*, <https://doi.org/10.48550/arXiv.2310.12941>.
- Bradshaw, Samantha, Hannah Bailey, and Philip N. Howard. 2021. *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*. Working Paper 2021.1. Programme on Democracy and Technology, Oxford Internet Institute, University of Oxford. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/01/CyberTroop-Report-2020-v.2.pdf>.
- Bureau of Industry and Security, Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use; Updates and Corrections, 88 Fed. Reg. 73458 (Oct. 25, 2023), <https://www.federalregister.gov/documents/2023/10/25/2023-23055/implementation-of-additional-export-controls-certain-advanced-computing-items-supercomputer-and>.
- , Establishment of Reporting Requirements for the Development of Advanced Artificial Intelligence Models and Computing Clusters, 89 Fed. Reg. 73612, Proposed rule (Sept. 11, 2024), <https://www.federalregister.gov/documents/2024/09/11/2024-20529/establishment-of-reporting-requirements-for-the-development-of-advanced-artificial-intelligence>.
- , Framework for Artificial Intelligence Diffusion, 90 Fed. Reg. 4544 (Jan. 15, 2025), <https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion>.
- Canada, Chelsea. 2026. "New Trends Emerge as States Refine AI Legislation." National Conference of State Legislatures, January 22, 2026. Accessed August 1, 2026. <https://www.ncsl.org/news/details/new-trends-emerge-as-states-refine-ai-legislation>.
- Casper, Stephen, Carson Ezell, Charlotte Siegmann, Noam Kolt, Taylor Lynn Curtis, Benjamin Bucknall, Andreas Haupt, et al. 2024. "Black-Box Access Is Insufficient for Rigorous AI Audits." In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, 2254–72. <https://doi.org/10.1145/3630106.3659037>.

- Chen, Heather, and Kathleen Magramo. 2024. "Finance Worker Pays Out \$25 Million after Video Call with Deepfake 'Chief Financial Officer'." CNN, February 3, 2024. Accessed July 22, 2026. <https://www.cnn.com/2024/02/04/asia/deepfake-cfo-sc-am-hong-kong-intl-hnk>.
- Coalition for Content Provenance and Authenticity. 2026a. "C2PA and Content Credentials Explainer." C2PA Specifications, Version 2.4. Accessed August 1, 2026. <https://spec.c2pa.org/specifications/specifications/2.4/explainer/Explainer.html>.
- . 2026b. "C2PA Harms Modelling." C2PA Specifications, Version 2.4. Accessed August 1, 2026. https://spec.c2pa.org/specifications/specifications/2.4/security/Harms_Modelling.html.
- . 2026c. "C2PA Security Considerations." C2PA Specifications, Version 2.4. Accessed August 1, 2026. https://spec.c2pa.org/specifications/specifications/2.4/security/Security_Considerations.html.
- Cobbe, Jennifer, Michael Veale, and Jatinder Singh. 2023. "Understanding Accountability in Algorithmic Supply Chains." In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 1186–97. <https://doi.org/10.1145/3593013.3594073>.
- Complaint, The New York Times Company v. Microsoft Corporation and OpenAI, Inc., No. 1:23-cv-11195, S.D.N.Y., filed December 27, 2023 (2023), https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf.
- Corrigan, Jack, Owen Daniels, Lauren Kahn, and Danny Hague. 2024. *Governing AI with Existing Authorities: A Case Study in Commercial Aviation*. Center for Security and Emerging Technology (CSET), Georgetown University. <https://cset.georgetown.edu/publication/governing-ai-with-existing-authorities/>.
- Crémer, Jacques, Yves-Alexandre de Montjoye, and Heike Schweitzer. 2019. *Competition Policy for the Digital Era*. European Commission. <https://doi.org/10.2763/407537>.
- Ding, Michelle L., and Harini Suresh. 2025. "The Malicious Technical Ecosystem: Exposing Limitations in Technical Governance of AI-Generated Non-Consensual Intimate Images of Adults." *arXiv*, <https://doi.org/10.48550/arXiv.2504.17663>.
- European Commission. 2023. "Commission Opens Formal Proceedings against X under the Digital Services Act," December 18, 2023. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6709.
- . 2024. "Commission Sends Preliminary Findings to X for Breach of the Digital Services Act," July 12, 2024. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_3761.

European Parliament and Council of the European Union, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, Official Journal of the European Union, L 119, 4 May 2016, pp. 1–88 (2016), <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

———, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), Official Journal of the European Union, L 277, 27 October 2022, pp. 1–102 (2022), <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>.

———, Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on Liability for Defective Products and Repealing Council Directive 85/374/EEC, Official Journal of the European Union (2024), <https://eur-lex.europa.eu/eli/dir/2024/2853/oj>.

———, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence ... (Artificial Intelligence Act), Official Journal of the European Union (2024), <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

Executive Office of the President, Executive Order 14110 of October 30, 2023: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 88 Fed. Reg. 75191 (Nov. 1, 2023), <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence>.

———, Executive Order 14148 of January 20, 2025: Initial Rescissions of Harmful Executive Orders and Actions, 90 Fed. Reg. 8237 (Jan. 28, 2025), <https://www.federalregister.gov/documents/2025/01/28/2025-01901/initial-rescissions-of-harmful-executive-orders-and-actions>.

Farid, Hany. 2022. “Creating, Using, Misusing, and Detecting Deep Fakes.” *Journal of Online Trust and Safety* 1 (4). <https://doi.org/10.54501/jots.v1i4.56>.

Federal Communications Commission, Implications of Artificial Intelligence Technologies on Protecting Consumers from Unwanted Robocalls and Robotexts, Declaratory Ruling, FCC 24-17, CG Docket No. 23-362 (Feb. 8, 2024), <https://docs.fcc.gov/public/attachments/FCC-24-17A1.pdf>.

Federal Trade Commission. 2024. “FTC Announces Crackdown on Deceptive AI Claims and Schemes: Operation AI Comply,” September. Accessed July 22, 2026. <https://www.ftc.gov/news-events/news/press-releases/2024/09/ftc-announces-crackdown-deceptive-ai-claims-schemes>.

Feffer, Michael, Anusha Sinha, Zachary C. Lipton, and Hoda Heidari. 2024. “Red-Teaming for Generative AI: Silver Bullet or Security Theater?” *arXiv*, <https://doi.org/10.48550/arXiv.2401.15897>.

- Feuerriegel, Stefan, Jochen Hartmann, Christian Janiesch, and Patrick Zschech. 2024. "Generative AI." *Business & Information Systems Engineering* 66 (1): 111–26. <https://doi.org/10.1007/s12599-023-00834-7>.
- Financial Crisis Inquiry Commission. 2011. *The Financial Crisis Inquiry Report: Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. U.S. Government Printing Office. <https://www.govinfo.gov/content/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf>.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for datasets." *Communications of the ACM* 64 (12): 86–92. <https://doi.org/10.1145/3458723>.
- Gillespie, Tarleton. 2010. "The Politics of 'Platforms'." *New Media & Society* 12 (3): 347–64. <https://doi.org/10.1177/1461444809342738>.
- . 2022. "Do Not Recommend? Reduction as a Form of Content Moderation." *Social Media + Society* 8 (3). <https://doi.org/10.1177/20563051221117552>.
- Goldstein, Josh A., Girish Sastry, Micah Musser, Renee DiResta, Matthew Gentzel, and Katerina Sedova. 2023. "Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations." *arXiv*, <https://doi.org/10.48550/arXiv.2301.04246>.
- Gorwa, Robert. 2024. *The Politics of Platform Regulation: How Governments Shape Online Content Moderation*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780197692851.001.0001>.
- Guha, Neel, Christie Lawrence, Lindsey A. Gilmard, Kit Rodolfa, Faiz Surani, Rishi Bommasani, Inioluwa Deborah Raji, et al. 2024. "AI Regulation Has Its Own Alignment Problem: The Technical and Institutional Feasibility of Disclosure, Registration, Licensing, and Auditing." *George Washington Law Review* 92. <https://ssrn.com/abstract=4634443>.
- Gupta, Ritwik, Leah Walker, and Andrew W. Reddie. 2024. "Whack-a-Chip: The Futility of Hardware-Centric Export Controls." *arXiv*, <https://doi.org/10.48550/arXiv.2411.14425>.
- Hackenburg, Kobi, Ben M. Tappin, Paul Röttger, Scott A. Hale, Jonathan Bright, and Helen Margetts. 2025. "Scaling Language Model Size Yields Diminishing Returns for Single-Message Political Persuasion." *Proceedings of the National Academy of Sciences* 122 (10). <https://doi.org/10.1073/pnas.2413443122>.
- Henderson, Peter, Xuechen Li, Dan Jurafsky, Tatsunori Hashimoto, Mark A. Lemley, and Percy Liang. 2023. "Foundation Models and Fair Use." *Journal of Machine Learning Research* 24 (400): 1–79. <https://www.jmlr.org/papers/volume24/23-0569/23-0569.pdf>.

- Hendrycks, Dan, Dawn Song, Christian Szegedy, Honglak Lee, Yarin Gal, Erik Brynjolfsson, Sharon Li, Andy Zou, Lionel Levine, Bo Han, et al. 2025. "A definition of AGI." *arXiv*, <https://doi.org/10.48550/arXiv.2510.18212>.
- Hugging Face. 2026. "Security Incident Disclosure – July 2026," July 16, 2026. Accessed September 25, 2026. <https://huggingface.co/blog/security-incident-july-2026>.
- Jones, Meg Leta. 2018. "Does Technology Drive Law? The Dilemma of Technological Exceptionalism in Cyberlaw." *University of Illinois Journal of Law, Technology & Policy* 2018:249–84. <https://illinoisjltp.com/file/32/Jones.pdf>.
- Kaminski, Margot E., and Meg Leta Jones. 2024. "Constructing AI Speech." *Yale Law Journal Forum* 133:1212–58. <https://access.heinonline.com/HOL/P?h=hein.journals/yjlfor133&i=1212>.
- Kang, Cecilia. 2025. "A.I.-Generated Images of Child Sexual Abuse Are Flooding the Internet." *New York Times*, July 10, 2025. <https://www.nytimes.com/2025/07/10/technology/ai-csam-child-sexual-abuse.html>.
- Khan, Lina M. 2017. "Amazon's Antitrust Paradox." *Yale Law Journal* 126 (3): 710–805. <https://www.yalelawjournal.org/note/amazons-antitrust-paradox>.
- Korinek, Anton, and Jai Vipra. 2024. *Concentrating Intelligence: Scaling and Market Structure in Artificial Intelligence*. 33139. National Bureau of Economic Research. <https://doi.org/10.3386/w33139>.
- Lake, James B. 2024. "Labels, Social Media, Consumer Protection, and the First Amendment." *Florida Law Review Forum* 75:35–45. <https://www.floridalawreview.com/article/92518-labels-social-media-consumer-protection-and-the-first-amendment>.
- Lasswell, Harold D. 1971. *Propaganda Technique in World War I*. Originally published in 1927 as *Propaganda Technique in the World War*. Cambridge, MA: MIT Press.
- Longpre, Shayne, Robert Mahari, Naana Obeng-Marnu, William Brannon, Tobin South, Katy Ilonka Gero, Alex Pentland, and Jad Kabbara. 2024. "Position: Data Authenticity, Consent, & Provenance for AI Are All Broken: What Will It Take to Fix Them?" In *Proceedings of the 41st International Conference on Machine Learning*, 235:32711–25. <https://proceedings.mlr.press/v235/longpre24b.html>.
- Mantzaris, Alexios. 2025. "The Indicator Guide to AI Labels." *Indicator*, October 2, 2025. <https://indicator.media/p/the-indicator-guide-to-ai-labels>.
- Mantzaris, Alexios, and Nasha Dutta. 2025. "Tech Platforms Promised to Label AI Content. They're Not Delivering." *The Indicator*, October 23, 2025. <https://indicator.media/p/tech-platforms-fail-to-label-ai-content-c2pa-metadata>.
- McMahon, Liv, Zoe Kleinman, and Courtney Subramanian. 2025. "Facebook and Instagram Get Rid of Fact Checkers." *BBC News*, January 7, 2025. <https://www.bbc.com/news/articles/cly74mpy8klo>.

- METR. 2026. *Brief Independent Investigation of Agents' Behavior, Reasoning and Collaboration in the OpenAI / Hugging Face Hacking Incident*. <https://metr.org/blog/2026-08-26-openai-hugging-face-incident-investigation/>.
- Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. "Model Cards for Model Reporting." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–29. <https://doi.org/10.1145/3287560.3287596>.
- Moody v. NetChoice, LLC, 603 U.S. 707 (2024), https://www.supremecourt.gov/opinions/23pdf/22-277_d18f.pdf.
- Muehlematter, Urs J., Paola Daniore, and Kerstin N. Vokinger. 2021. "Approval of Artificial Intelligence and Machine Learning-Based Medical Devices in the USA and Europe (2015–2020): A Comparative Analysis." *The Lancet Digital Health* 3 (3): e195–e203. [https://doi.org/10.1016/S2589-7500\(20\)30292-2](https://doi.org/10.1016/S2589-7500(20)30292-2).
- Napoli, Philip M., and Robyn Caplan. 2017. "Why Media Companies Insist They're Not Media Companies, Why They're Wrong, and Why It Matters." *First Monday* 22 (5). <https://doi.org/10.5210/fm.v22i5.7051>.
- Narayanan, Arvind, and Sayash Kapoor. 2025. "AI as Normal Technology: An Alternative to the Vision of AI as a Potential Superintelligence." Knight First Amendment Institute, April 15, 2025. <https://knightcolumbia.org/content/ai-as-normal-technology>.
- National Highway Traffic Safety Administration. 2017. "Automated Driving Systems: A Vision for Safety 2.0." Accessed July 22, 2026. <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>.
- National Telecommunications and Information Administration. 2024. *Dual-Use Foundation Models with Widely Available Model Weights*. U.S. Department of Commerce. <https://www.ntia.gov/federal-register-notice/2024/dual-use-foundation-artificial-intelligence-models-widely-available-model-weights>.
- Nelson, Jason. 2026. "DeepMind CEO Says AGI Will Be Bigger Than Electricity or Fire." Decrypt, July 14, 2026. Accessed July 22, 2026. <https://decrypt.co/373511/deep-mind-ceo-agi-bigger-than-electricity-fire>.
- OpenAI. 2025. "Our Updated Preparedness Framework," April 15, 2025. Accessed July 22, 2026. <https://openai.com/index/updating-our-preparedness-framework/>.
- . 2026. "The Hugging Face Incident and the Road Ahead," August 26, 2026. Accessed September 25, 2026. <https://openai.com/index/hugging-face-incident-and-the-road-ahead/>.
- Opinion and Order, The New York Times Company v. Microsoft Corporation and OpenAI, Inc., No. 1:23-cv-11195, S.D.N.Y., April 4, 2025 (2025), <https://www.nysd.uscourts.gov/sites/default/files/2025-04/yf%2023cv11195%20OpenAI%20MTD%20opinion%20april%204%202025.pdf>.

- Pappano, Albert Earl. 1937. "The False Neros." *The Classical Journal* 32 (7): 385–92. <https://www.jstor.org/stable/3291534>.
- Perault, Matt. 2023. "Section 230 Won't Protect ChatGPT." *Journal of Free Speech Law* 3:363–74. <https://www.journaloffreespeechlaw.org/perault.pdf>.
- Perrigo, Billy. 2025. "Inside the U.K.'s Bold Experiment in AI Safety." *Time* (January 16, 2025). <https://time.com/collections/davos-2025/7204670/uk-ai-safety-institute/>.
- Phuong, Mary, Matthew Aitchison, Elliot Catt, Sarah Cogan, Alexandre Kaskasoli, Victoria Krakovna, David Lindner, et al. 2024. "Evaluating Frontier Models for Dangerous Capabilities." *arXiv*, <https://doi.org/10.48550/arXiv.2403.13793>.
- Prakash, Prarthana. 2023. "Alphabet CEO Sundar Pichai Says That A.I. Could Be 'More Profound' Than Both Fire and Electricity—But He's Been Saying the Same Thing for Years." *Fortune*, April 17, 2023. Accessed July 22, 2026. <https://fortune.com/2023/04/17/sundar-pichai-a-i-more-profound-than-fire-electricity/>.
- Raji, Inioluwa Deborah, and Joy Buolamwini. 2019. "Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products." In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429–35. <https://doi.org/10.1145/3306618.3314244>.
- Ramkumar, Amrith. 2026. "Anthropic Researcher Quits Over 'Out-of-Control' AI Fears." *Wall Street Journal*, September 9, 2026. <https://www.wsj.com/tech/ai/anthropic-researcher-quits-over-out-of-control-ai-fears-707b7628>.
- Rid, Thomas. 2020. *Active Measures: The Secret History of Disinformation and Political Warfare*. New York: Farrar, Straus / Giroux.
- Rini, Regina, and Leah Cohen. 2022. "Deepfakes, Deep Harms." *Journal of Ethics and Social Philosophy* 22 (2): 143–61. <https://doi.org/10.26556/jesp.v22i2.1628>.
- Shavit, Yonadav, Sandhini Agarwal, Miles Brundage, Steven Adler, Cullen O'Keefe, Rosie Campbell, Teddy Lee, et al. 2023. *Practices for Governing Agentic AI Systems*. White Paper. OpenAI. <https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf>.
- Shevlane, Toby, Sebastian Farquhar, Ben Garfinkel, Mary Phuong, Jess Whittlestone, Jade Leung, Daniel Kokotajlo, et al. 2023. "Model Evaluation for Extreme Risks." *arXiv*, <https://doi.org/10.48550/arXiv.2305.15324>.
- Stern, Ariel Dora. 2017. "Innovation under Regulatory Uncertainty: Evidence from Medical Technology." *Journal of Public Economics* 145:181–200. <https://doi.org/10.1016/j.jpubeco.2016.11.010>.
- Stout, Lynn A. 2011. "Derivatives and the Legal Origin of the 2008 Credit Crisis." *Harvard Business Law Review* 1 (1): 1–38. <https://access.heinonline.com/HOL/P?h=hein.journals/hbusrew1&i=15>.

- Synergy Research Group. 2025. "Cloud Market Share Trends: Big Three Together Hold 63% While Oracle and the Neoclouds Inch Higher." Accessed July 22, 2026. <https://www.srgresearch.com/articles/cloud-market-share-trends-big-three-together-hold-63-while-oracle-and-the-neoclouds-inch-higher>.
- Tabassi, Elham. 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.100-1>.
- TechInsights. 2024. "Data-Center AI Chip Market – Q1 2024 Update." Accessed July 22, 2026. <https://www.techinsights.com/blog/data-center-ai-chip-market-q1-2024-update>.
- Thompson, Dennis F. 1980. "Moral Responsibility of Public Officials: The Problem of Many Hands." *The American Political Science Review* 74 (4): 905–16. <https://doi.org/10.2307/1954312>.
- Tiku, Nitasha, and Pranshu Verma. 2025. "We Uploaded a Fake Video to 8 Social Apps. Only One Told Users It Wasn't Real." *Washington Post*, October 22, 2025. <https://www.washingtonpost.com/technology/2025/10/22/ai-deepfake-sora-platforms-c2pa/>.
- Tucher, Andie. 2022. *Not Exactly Lying: Fake News and Fake Journalism in American History*. New York: Columbia University Press.
- U.S. Congress, Federal Trade Commission Act, 15 U.S.C. Section 45(a) (1914), <https://www.law.cornell.edu/uscode/text/15/45>.
- , Fair Credit Reporting Act, 15 U.S.C. Sections 1681–1681x (1970), <https://www.law.cornell.edu/uscode/text/15/1681>.
- , Equal Credit Opportunity Act, 15 U.S.C. Sections 1691–1691f (1974), <https://www.law.cornell.edu/uscode/text/15/1691>.
- , 47 U.S.C. Section 230: Protection for Private Blocking and Screening of Offensive Material, Communications Decency Act of 1996 (1996), <https://www.law.cornell.edu/uscode/text/47/230>.
- , Dodd-Frank Wall Street Reform and Consumer Protection Act, Public Law 111-203, 124 Stat. 1376 (2010), <https://www.govinfo.gov/app/details/PLAW-111publ203>.
- , Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act, Public Law 119-12, Known as the TAKE IT DOWN Act (2025), <https://www.congress.gov/119/plaws/publ12/PLAW-119publ12.pdf>.
- U.S. Food and Drug Administration. 2024. "Development and Approval Process: Drugs." U.S. Food and Drug Administration. Accessed April 27, 2026. <https://www.fda.gov/drugs/development-approval-process-drugs>.

- Vipra, Jai, and Anton Korinek. 2023. "Market Concentration Implications of Foundation Models." *arXiv*, <https://doi.org/10.48550/arXiv.2311.01550>.
- Volokh, Eugene. 2023. "Large Libel Models? Liability for AI Output." *Journal of Free Speech Law* 3:489–558. <https://www.journaloffreespeechlaw.org/volokh4.pdf>.
- Vosoughi, Soroush, Deb Roy, and Sinan Aral. 2018. "The Spread of True and False News Online." *Science* 359 (6380): 1146–51. <https://doi.org/10.1126/science.aap9559>.
- Wong, Matteo. 2024. "High School Is Becoming a Cesspool of Sexually Explicit Deep-fakes." *The Atlantic*, September 26, 2024. <https://www.theatlantic.com/technology/archive/2024/09/ai-generated-csam-crisis/680034/>.

Authors

Sarah Barrington is a PhD candidate at the School of Information, University of California, Berkeley. Email: sbarrington@berkeley.edu.

Hannah Bailey is an Assistant Professor at the Institute for Strategy and Technology, Carnegie Mellon University. Email: hbailey@andrew.cmu.edu.

Acknowledgments

We are grateful to Justin Hendrix, Dean Jackson, Sam Woolley, and Hany Farid for their helpful insights and early thoughts shaping the piece. This commentary builds on a provocation one of the authors wrote for Tech Policy Press (Barrington 2025).

Funding Statement

The authors received no specific funding for this work.

Disclosure Statement

The authors declare no competing interests. A large language model was used to cross-check bibliography entries and for spelling and grammar guidance.

Keywords

Generative AI; AI governance; technological exceptionalism; regulatory enforcement; platform governance