Cambridge Forum on AI: Law and Governance (2025), **1**, e30, 1–22
doi:10.1017/cfl.2025.10012

Cambridge Forum

RESEARCH ARTICLE

Audio deepfakes and the regulation of the landlords of creativity

Bao Kham Chau[1,2] and George He[3]

[1]Cornell University Cornell Tech, New York, New York, USA; [2]Harvard University Berkman Klein Center, Cambridge, Massachusetts, USA and [3]Harvard Law School Library Innovation Lab, Cambridge, Massachusetts, USA
Corresponding author: Bao Kham Chau; Email: baokham.chau@gmail.com

(Received: 1 November 2024; revised: 28 April 2025; accepted: 4 May 2025)

Abstract

This paper begins with a brief technical explanation of generative AI and how only a small subset of entities – the landlords of creativity – have access to the computational resources and expertise needed to create foundation models that enable audio deepfakes. It then examines how regulatory regimes in America, the European Union (EU), and China address the misuse of generative AI. Although each framework seeks to regulate generative AI in different ways, the paper argues that none effectively assigns liability to the landlords of creativity. Finally, the paper proposes holding these landlords responsible for their renters' malicious usage. This proposal not only is technically feasible but also is conceptually aligned with established legal doctrines in the American, EU, and Chinese frameworks.

Keywords: generative AI regulation; generative AI liability; landlords of creativity; AI providers regulation; AI providers liability; comparative generative AI regulations

1. Introduction

In March 2019, the CEO of a UK-based energy firm received a call from his superior, the chief executive of the parent company, urgently instructing him to transfer €220,000 to a Hungarian supplier (Damiani, 2019). Trusting the familiar voice, he complied, only to later discover that fraudsters had used Artificial Intelligence (AI)-based audio cloning technology to mimic his boss's voice, making this one of the first known instances where AI was employed in a voice phishing scam. This incident serves as an example of how audio deepfakes present a distinct and increasingly pernicious risk out of all deepfake formats. While visual deepfakes such as a widely circulated video of the Ukrainian President urging troops to surrender have already demonstrated their disruptive potential (Williamson, 2022), audio deepfakes carry even greater challenges due to their lack of discernible visual cues. Audio deepfakes, also known as voice cloning, can produce realistic synthetic audio clips mimicking an individual's voice (Mcuba, Singh, Ikuesan & Venter, 2023). Because they are harder to detect without expert intervention, the misuse of audio deepfakes in propaganda campaigns could erode trust and mislead the public more insidiously and effectively than visual deepfakes, potentially causing significant social and political discord.

Deepfakes' nefarious consequences are more pronounced as we move into the age of generative AI, marked by the release of ChatGPT in April 2022. Developed by OpenAI based on the neural network architecture described in the transformer paper (Vaswani et al., 2017), ChatGPT has redefined the boundaries of conversational computing. Alongside other Large Language Models

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

(LLMs) (Wei et al., 2022), ChatGPT's ability to engage in human-like conversations has spurred advances across diverse fields, especially in facilitating breakthroughs in complex research areas (Chau & Livermore, 2024; Zhang & Kamel Boulos, 2023). Pharmaceutical companies, for example, are using generative AI to uncover insights from vast datasets, aiming to reduce the 12–15 years required to bring drugs to market (Hughes, Rees, Kalindjian & Philpott, 2011). Similarly, physicists leverage AI for more efficient simulations of complex electron interactions, potentially leading to novel materials and catalysts (Kirkpatrick et al., 2021; Menon & Ranganathan, 2022).

Despite these advances, the integration of generative AI into various applications also highlights substantial risks.[1] In addition to posing threats to national security, generative AI raises new challenges to civil regulatory regimes such as copyright laws (Lee, Cooper & Grimmelmann, 2024). Generative AI's capability to create art, music, literature, and other forms of creative work has raised many questions about originality and authorship (Appel, Neelbauer & Schweidel, 2023).[2] AI models that allow the generation of sound-alike audio are becoming more accessible (Magowan, 2023), creating an existential threat to the monetization models of existing players in the music industry (Reynolds, 2024). Due to the complicated data supply chain, it is difficult to determine who is liable for a particular act of copyright infringement.

These challenges have prompted significant investments into detection methods and the ethical implications of their use, but detection alone is insufficient (Federal Trade Commission, n.d.). Given these real-world repercussions, courts, policymakers, and scholars are exploring how existing liability regimes can interact with the deployment of AI to incentivize safe development and use (Henderson, Hashimoto & Lemley, 2023). However, these liability regimes do not adequately take into account the complexity of training foundation generative AI models and the current generative AI ecosystem (Bommasani et al., 2021).

In particular, these liability regimes do not consider who owns the infrastructure to process the data, train the models, and deploy said models. This ownership oversight could be fatal to apportioning liability. As it stands, only a handful of organizations have the hardware infrastructure to support the acquisition of data for and the actual training of generative AI models (IoT Analytics, 2023). Instead of selling the models to the end-users, these businesses primarily lease the models out to consumers through Application Programming Interfaces (APIs). Under this Software as a Service (SaaS) framework, consumers can indirectly use the creative outputs of generative AI, but they do not directly own the models that produce such outputs. Nor are they responsible for the deployment of said generative AI models. This relationship is akin to a landlord-tenant dynamic: the 'landlords of creativity' own the creative infrastructure (i.e., the generative AI models and the data centers where the models are deployed), while the tenants merely rent access through the landlords' APIs.

Because these landlords own all the critical infrastructure and have the technical power to control its usage, this paper proposes placing these landlords of creativity at the center of the liability analysis. Specifically, we suggest placing the default liability onto the landlords because they not only are the cheapest cost avoiders, but also are obligated to maintain their infrastructure in a 'habitable' condition. We distinguish between criminal/civil liabilities and commercial/non-commercial usage of the generative AI models, and recommend that the landlords can reduce liability if they (1) limit the capacities provided to unidentified customers and/or (2) keep an audit trail of generated outputs

[1] Self-driving cars, an earlier example of AI-powered technology, have been involved in numerous crashes, some resulting in fatalities (United States Department of Transportation, National Highway Traffic Safety Administration, 2022). Similarly, researchers have shown that AI-powered diagnostic systems can be manipulated to produce incorrect medical diagnoses, raising significant ethical and safety concerns (Zhou et al., 2021). Relevant to this paper, the use of LLMs and generative AI in deepfake technology poses particularly insidious threats when leveraged to spread misinformation and propaganda.

[2] Traditionally, copyright laws protected human-created works, but AI-generated works reside in a gray area, especially in the music and art industries where attributability is critical for royalty payments. Some singers in the music industry are embracing this new wave of voice technology by making their voices available for public use (Vanessa Romo, 2023), while many more cautious artists are facing unauthorized use of their voices (Coscarelli, 2023; Reed, 2023).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

for registered users, which include the attribution and identification of user inputs for generated content.

We begin with a brief technical explanation of generative AI and how only a small subset of entities – the landlords of creativity – have access to the computational resources and expertise needed to create foundation models that enable deepfakes. We then examine how regulatory regimes in America, the European Union (EU), and China address the misuse of generative AI. Although each framework seeks to regulate generative AI in different ways, we argue that none effectively assigns liability to the landlords of creativity, who own the infrastructures needed to process data for, train, and deploy LLMs. Finally, we propose holding these landlords responsible for their renters' malicious usage. This proposal not only is technically feasible but also is conceptually aligned with established legal doctrines in the American, EU, and Chinese frameworks.

2. The liability problem of generative AI

Generative AI is a subset of AI focusing on producing new and meaningful content – such as text, images, or audio – based on training data (Feuerriegel, Hartmann, Janiesch & Zschech, 2024). Unlike traditional AI, which centers on classification, prediction, or decision-making, generative AI models aim to create content that closely mimics real data (Chen, Morris, Stevens Aubrey & Wang, 2019). A prime example is Generative Adversarial Networks, which use a generator and discriminator in tandem: the generator creates realistic data while the discriminator attempts to distinguish it from real samples (Goodfellow et al., 2014; Wiggers, 2019). Over time, this adversarial setup enhances both components until the discriminator can no longer reliably tell generated content from real data, making the output inherently difficult to detect (Kim, Ren & Yang, 2021; Kong, Lee, Kim & Jaekyoung, 2020). While not all generative AI architectures are adversarial, the outputs across architectures are increasingly indistinguishable from real content.

Detecting AI-generated audio remains significantly harder than identifying visual deepfakes, largely because audio anomalies are less perceptible. Visual flaws – like odd lighting or anatomical inconsistencies – are generally easier to spot than subtle tonal irregularities in voice recordings. As a result, adversarial audio attacks are both easier to carry out and harder to detect (Carlini & Wagner, 2018). This challenge is heightened by the pace of generative audio development, which outstrips research on detection (Leffer, 2024). In particular, voice synthesis has rapidly advanced due to improvements in encoding and transformer models (Wang et al., 2023). Yet despite these technical strides, many synthesis models lack guarantees that their outputs can be uniquely identified – raising concerns for regulators seeking to curb abuses of these models (Khosrowi, forthcoming).

A key unresolved issue is determining liability for misuses. In copyright cases, for example, it is unclear where infringement occurs within the generative AI supply chain or who should bear responsibility (Lee et al., 2024, p. 7). While many regulatory frameworks target both users and providers, the reality is more complex. Foundation models are often built by one party (the provider) but fine-tuned and deployed by others (users) before reaching the end consumers. For instance, over 67% of AI startups build on OpenAI's models like ChatGPT (Jones, 2024). If copyrighted content is generated by ChatGPT and shared by an end-user, who is liable – the end-user, the fine-tuner, or OpenAI? We argue that those who operate the foundational infrastructure – the 'landlords of creativity' – must be included in the liability calculus.

This is because, among the various stakeholders involved in the generative AI supply chain, only the entities operating at the data, compute, and developer layers are intimately involved in developing foundation models (Amazon Web Services, Inc, n.d.; Jones, 2023).[3] These layers are responsible for

[3] Foundation models are models trained on massive datasets. Foundation Models (FMs) are large deep learning neural networks that have changed the way data scientists approach machine learning (ML). Rather than develop artificial intelligence (AI) from scratch, data scientists use a foundation model as a starting point to develop ML models that power new applications more quickly and cost-effectively. The term foundation model was coined by researchers to describe ML models trained on a

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

data ingestion and the initial training of foundation models. They manage data ingestion, model training, and technical scaling, placing them in a strong position to address challenges related to model size and complexity. As models grow, it becomes nearly impossible for many entities to train or run state-of-the-art systems on their own hardware.

Take, for example, generative pre-trained transformer (GPT) models. These models expanded from 117 million parameters in 2017 to 1.5 billion in 2019 (OpenAI, 2019), reaching 175 billion with GPT-3 and an estimated 1.76 trillion with GPT-4 (Brown et al., 2020). Other generative AI models have similarly grown, as illustrated in Table 1 below:

As model sizes grow, so too does the volume of training data required. Researchers have shown that scaling model parameters must be matched by dataset expansion to sustain performance gains (Hoffmann et al., 2022; Kaplan et al., 2020). For instance, Hoffmann et al. (2022) found that MT-NLG 530B underperformed when trained on just 300 billion tokens – despite its 530 billion parameters. In audio generation, this scaling translates to over 60,000 hours of voice recordings from more than 7,000 speakers, a sharp increase from historical norms of under 1,000 hours (Kahn et al., 2020; Wang et al., 2023).

This exponential growth has driven a massive surge in computational demands. OpenAI used 1,024 NVIDIA A100 GPUs over 34 days to train GPT-3, costing more than $4.6 million (Narayanan et al., 2021). Training newer models like Gemini Ultra may cost as much as $191 million (Stanford University Human-Centered Artificial Intelligence, 2024). At the same time, access to training data is becoming more limited. A 2024 audit showed roughly 25% of critical domain content and 5% of data across three major datasets – C4, RefinedWeb, and Dolma – were restricted (Stanford University Human-Centered Artificial Intelligence, 2024, pp. 2, 63). As freely available data shrinks, developers must either use biased, outdated datasets or pay for better alternatives. These costs effectively bar all but the largest organizations from building foundation models (Bommasani et al., 2021). Yet such firms often restrict public access to their models, limiting usage to approved partners (OpenAI Help Center, n.d., Section 2). This centralization raises concerns about competitive fairness and safety (Federal Trade Commission, 2023).

The concentration of AI providers is even more pronounced in the field of generative voice synthesis, or text-to-speech (TTS). Unlike generative text-to-text and text-to-image models (e.g., GPT, XLNet, and T5), which process discrete tokens as input, TTS models require continuous audio data. Using token-based models for TTS tasks introduces additional errors and noise into the training process (Mehrish, Majumder, Bhardwaj, Mihalcea & Poria, 2023). To mitigate these issues, a large audio dataset is necessary. However, obtaining vast amounts of high-quality audio

Table 1. Examples of the sizes of generative AI models developed by Google, OpenAI, and Meta (Giattino, Mathieu, Samborska and Roser, n.d.).

Model	Number of Parameters	Year
GPT-1	117 million	2018
GPT-2	1.5 billion	2019
GPT-3	175 billion	2020
GPT-4	Estimated 1.76 trillion	2023
Llama Guard	7 billion	2023
Llama 3-405B	405 billion	2024
Deepseek-V3	671 billion	2025

broad spectrum of generalized and unlabeled data and capable of performing a wide variety of general tasks such as understanding language, generating text and images, and conversing in natural language (Amazon Web Services, Inc, n.d.; Jones, 2023).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

data is significantly more challenging than sourcing extensive text and image datasets from the internet. Of the more than 6,000 languages worldwide, adequate data exists to train TTS foundation models for only about a dozen (Xu et al., 2020). Although many researchers have attempted to develop TTS systems with low data resources, their proposals have yet to be operationalized and integrated into open-source, commercial-grade TTS systems (Elneima & Bińkowski, 2022; Gabryś et al., 2022). This limits the number of successful TTS models to a few organizations, such as Microsoft's VALL-E.

Despite this concentration, current liability frameworks focus almost exclusively on end-users. While this may be adequate for visual deepfakes, it is less so for audio, where users lack visual cues to assess legality. If end-users unknowingly create unlawful content, it may be appropriate to hold the infrastructure providers accountable – especially if the providers have been repeatedly alerted to illicit uses of their models. Much like a landlord being made aware of illegal tenant activities, these providers should at least be partly liable when misuse persists on their platforms.[4]

3. Existing liability frameworks

As of December 2024, there are over 1,600 policy initiatives aimed at regulating AI worldwide (Mariani, Eggers & Kishnani, 2024). This paper examines three major regulatory frameworks – the American, EU, and Chinese – not to endorse them, but to recognize their global influence on AI development and governance (Bradford, 2020; Kardon, 2023; Zick, 2013). It surveys these frameworks (see Figure 1) and identifies shortcomings, particularly in how they apportion liability for generative AI misuse.

We analyze how each framework handles the 'model ownership' problem across three stages of the generative AI supply chain: data ingestion, model training, and deployment. We find that none of these regimes sufficiently consider how the foundation model creators – the landlords of creativity

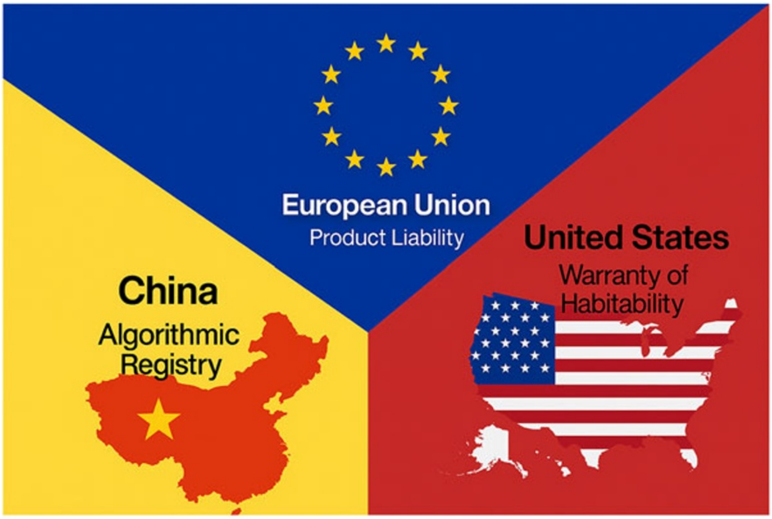

Figure 1. Overview of three AI frameworks.

[4] See, for example, Hemmings v. Pelham Wood Ltd. Liability Ltd. Partnership (2003), which holds the landlord liable when a tenant was shot by an intruder in his apartment because of the landlord's negligence. Similarly, see Pichardo v. Big Diamond, Inc., 215 S.W.3d 497 (Tex. App. Fort Worth 2007), which holds that one who controls a premise has a duty to use ordinary care to protect invitees from foreseeable criminal acts of third parties.

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

– should be held liable at different stages. We conclude with a discussion of an industry-led initiative as a potential model for addressing liability.

3.1. American AI regulatory frameworks

In the United States, generative AI development remains largely unregulated. Although frameworks like First Amendment jurisprudence and Section 230 of the Communications Decency Act may apply, they do not provide comprehensive regulatory coverage (Comp, 2023; Volokh, Lemley & Henderson, 2023). Moreover, Justice Gorsuch has even questioned Section 230's applicability to AI-generated content (Gonzalez v. Google LLC, 2023):

> You've got to do something beyond that. As I take your argument, you think that the Ninth Circuit's 'neutral tools' rule is wrong because, in a post-algorithm world, artificial intelligence can generate some forms of content, even according to neutral rules. I mean, artificial intelligence generates poetry, it generates polemics today. That – that would be content that goes beyond picking, choosing, analyzing, or digesting content. And that is not protected. Let's – let's assume that's right, okay? Then I guess the question becomes, what do we do about YouTube's recommendations? And – and as I see it, we have a few options.

The result is a patchwork of copyright, patent, data protection, and liability statutes that create uncertainty in regulation. This section surveys U.S. efforts to regulate generative AI, focusing on the lack of attention to foundation model ownership at each stage of the supply chain.

3.1.1 Proposed legislation

There is no comprehensive federal AI framework. Instead, several states – including California, New York, and Illinois – have enacted laws addressing various aspects of AI development and usage. For example, California passed a law requiring its Department of Technology to conduct a comprehensive inventory of all high-risk 'computational process[es] derived from machine learning, statistical modeling, data analytics, or artificial intelligence' used or considered for use in state agencies. This includes an inventory of '[t]he measures in place, if any, to mitigate the risks … of inaccurate, unfairly, discriminatory, or biased decisions' rendered by these computational processes (Cal. Gov't Code § 11,546.45.5, West). Similarly, Michigan enacted a law requiring the disclosure of whether a political advertisement was generated in whole or substantially by AI (Mich. Comp. Laws Ann. § 169.247, West). In addition to these state statutes, over 40 states have introduced more than 400 AI-related bills. As of February 7, 2024, New York leads with 65 bills under consideration, followed by California with 29, Tennessee with 28, Illinois with 27, and New Jersey with 25 (Heath, 2024; National Conference of State Legislatures, 2024).

At the federal level, Congress has introduced 141 AI-related pieces of legislation since the start of the 118th session (Brennan Center for Justice, 2024). Three of the twelve pending legislations relate to the regulation of deepfakes. The Protect Elections from Deceptive AI Act, introduced by Senator Klobuchar, prohibits the distribution of materially deceptive AI-generated audio or visual media related to federal candidates (Protect Elections from Deceptive AI Act, 2023). It allows for 'a covered individual whose voice or likeness appears in, or who is the subject of, a materially deceptive AI-generated audio or visual media' to seek injunctive or other equitable relief. The Deepfakes Accountability Act similarly requires producers of 'false personation record[s]' to disclose that the record has been created through generative AI or similar technologies (H.R. 5586, 2023, §2). It establishes civil and criminal liabilities for failing to disclose such content and creates a private right of action. Additionally, Representative Brittany Pettersen introduced a bill to establish a Task Force on AI in the Financial Services Sector, which would report to Congress on the use of AI in this sector, including how bad actors could utilize audio deepfakes to compromise and access consumers' financial accounts (H.R. 5808, 2023, §1). These pending legislations, however, have slim chances of being enacted into law (GovTrack.US, 2024a, estimating an 11% chance for the Protection Elections

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

from Deceptive AI Act, 2024b, estimating a 2% chance for the Deepfakes Accountability Act, 2024c, estimating a 1% chance for the Preventing Deep Fake Scams Act).

3.1.2 Executive orders

To address this regulatory gap and guide the design and use of AI, the Biden Administration published the Blueprint for an AI Bill of Rights. This document outlines five non-binding principles and associated practices aimed at protecting the American public from harms associated with the design, use, and deployment of AI systems in both public and private organizations.[5] Building on the AI Bill of Rights, President Biden issued the Executive Order on Safe, Secure, and Trustworthy AI on October 30, 2023 (The White House, 2023 (hereinafter 'AI Executive Order')).[6] While most of this Executive Order applies only to the federal government, private sector organizations that contract with the federal government for the use of AI may also have to comply with various requirements.

Specifically, the Order directs federal agencies and their contractors to set new standards for AI safety and security. It also requires AI developers to share their safety test results with regulators and orders the National Institute of Standards and Technology to establish standards, tools, and tests to evaluate the safety, security, and trustworthiness of AI and AI-generated content (Fed. Reg. 715196-7, 2023). Second, the AI Executive Order calls on Congress to pass comprehensive privacy legislation and requires agencies such as the National Science Foundation (NSF) to devote resources to research on privacy-preserving techniques compatible with the use of AI-powered systems (88 Fed. Reg. 71,596, at 75,206). Third, the Order seeks to ensure that AI advances equity and civil rights by mandating that agencies and federal contractors address unlawful discrimination and other harms resulting from the use of AI in governmental programs (88 Fed. Reg. 71,596, at 75,212). Fourth, the AI Executive Order directs federal health agencies to advance the responsible use of AI in the development of affordable and life-saving drugs (88 Fed. Reg. 71,596, at 75,214). Additionally, the Order requires the Secretary of Education to create resources to support educators deploying AI-enabled educational tools with appropriate guardrails for use in classrooms (88 Fed. Reg. 71,596, at 75,216).

Fifth, the AI Executive Order directs federal agencies to develop principles and best practices to mitigate harms while maximizing the benefits of AI for workers (88 Fed. Reg. 71,596, at 75,210). This includes addressing job displacement, labor standards, workplace equity, and data collection concerning workers. Sixth, the Order expects the Director of the NSF to help promote AI innovation by implementing the National AI Research Resource and providing AI researchers and students access to key AI resources and data (88 Fed. Reg. 71,596, at 75,206). Seventh, the Biden Administration seeks to promote multilateral cooperation in the regulation of responsible AI by tasking the State Department and other agencies to establish international frameworks for managing AI's risks while harnessing its benefits (88 Fed. Reg. 71,596, at 75,223). Finally, the Order endeavors to ensure the responsible deployment of AI in the government by requiring federal agencies to issue guidance for

[5] The first principle declares that individuals should be protected from unsafe or ineffective AI systems. Thus, these systems should be developed 'with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risk, and potential impacts of the system' (The White House, 2022 (hereinafter 'Biden AI Bill of Rights')). The second principle seeks to protect individuals from algorithmic discrimination, such as when AI-powered systems contribute to unjustified differential treatment or impacts based on race, color, ethnicity, sex, religion, age, national origin, disability, veteran status, genetic information, or any other classification protected by law. This principle encourages AI developers to take 'proactive and continuous measures' to ensure systems are designed fairly (Biden AI Bill of Rights). The third principle highlights the need to protect individuals from abusive data practices and asserts the right to have control over one's own data. It suggests that designers and developers should ask for permission in plain language understandable by the average individual. If obtaining consent in advance is not possible, it is necessary to design and implement privacy safeguards. The fourth principle advises that individuals should be informed when an automated system is being used to make decisions impacting them. All information should be provided in clear, comprehensible language and updated whenever significant changes occur in the system. Finally, the fifth principle states that individuals should have the opportunity to opt out of being subjected to decisions made by automated systems, where appropriate. Instead of an automated system, these individuals should have access to a human alternative.

[6] This executive order has since been revoked by another executive order on January 23, 2025 (The White House, 2025).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

their use of AI and accelerate their adoption of AI products, services, and personnel (88 Fed. Reg. 71,596, at 75,218).

Although these broad guidelines and enumerated rights do not specifically target deepfakes, generally, and audio deepfakes, specifically, the guidelines and rights do broadly cover deepfakes. Regulating the trustworthiness of AI and AI-generated content, for example, implicate the use of deepfakes. Promoting the use of responsible AI similarly covers deepfakes' use and dissemination. As discussed below, however, these broad guidelines and regulations fall short in establishing a clear liability regime. Misuses of generative AI are still being adjudicated under existing liability regimes such as product liability and vicarious liability.

3.1.3 American regulatory gaps

In the absence of clear legislative or executive liability standards, the American judiciary is left to resolve disputes with no unified theory. Recent lawsuits (e.g., Kadrey v. Meta, 2023; Andersen v Stability AI Ltd, 2023; Dubus et al v. NVIDIA Corporation, 2024) reflect this ambiguity.

Federal legislative proposals indiscriminately lumped the three stages of the generative AI supply chain (i.e., processing, training, deploying) into one liability bucket. They do not take into account the fact that the entity that they seek to hold liable – 'any person' (H.R. 5586, 2023, § 1041(a)); '[b]ad actors' (H.R. 5808, 2023, § 2(6)); and 'a person, political committee, or other entity' (Protect Elections from Deceptive AI Act, 2023, § 325(b)) – seems to be the end-users of the generative AI model. This fails to distinguish between malicious end users and well-meaning actors misled by upstream model owners – who currently face no liability. This issue is especially acute for audio deepfakes, where intent is harder to prove. Indeed, unlike visual deepfakes, it is harder for the end users to know that the generated audio content is created based on a particular personality. It is therefore harder to prove malicious intent for downstream users of audio deepfakes when determining criminal liability. Proposals regulating generative AI should accordingly take into account model ownerships in apportioning liability.

If the executive branch is inclined to take on the lead role in regulating generative AI, it must then tackle this issue of liability. As it stands, however, not only do the aforementioned executive orders not consider model ownership at the three stages of the generative AI supply chain, they do not even tackle the issue of liability itself. The Blueprint for an AI Bill of Rights, for example, has no legal force and does not create new legal rights. In particular, the Blueprint does not attach any liability to the violation of the enumerated rights, nor does it specify who is liable for such violations.[7] Even though a number of federal agencies have modified their policies to conform to the suggestions in the Blueprint, their interpretations and implementations are not standardized.[8] The diverse responses highlight a related and more important weakness of the Blueprint – it is sector-specific.[9] The Blueprint contains guidelines for some industries (e.g., health, labor, and education)

[7] The Office of Science and Technology Policy conceded that the Blueprint is 'a white paper published by the White House Office of Science and Technology Policy … [It] is non-binding and does not constitute U.S. government policy' (Biden AI Bill of Rights, 2022, at 2).

[8] In response to the Blueprint, the Department of Defense continues to operate under its AI Ethical Principles and the associated Responsible AI Strategy and Implementation Pathway (U.S. Department of Defense, 2022a, 2022b). The Equal Employment Opportunity Commission, on the other hand, took more concrete steps by releasing technical guidance explaining how employers' misuse of AI systems could violate Title VII and other anti-discrimination laws (U.S. Equal Employment Opportunity Commission, 2023). These actions pale in comparison to the Federal Trade Commission, which has started exploring regulations to address issues of AI discrimination and fraud (Federal Trade Commission, 2024).

[9] While there are benefits to a flexible, industry-specific approach (Laffont & Tirole, 2000), generative AI is a technology not limited to one specific sector, and 'if this technology goes wrong, it can go quite wrong' (Oversight of A.I.: Rules for Artificial Intelligence, 2023, Statement of Sam Altman, CEO of OpenAI). Sam Altman testified before Congress that '[o]ne concern of particular importance to OpenAI is the risk of racing dynamics leading to a decline in safety standards, the diffusion of bad norms, and accelerated AI timelines, each of which heighten societal risks associated with AI' (OpenAI, 2023). Because this decline in safety standards is not industry-specific, Altman urged Congress to create uniform safety standards for advanced AI.

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

while remaining virtually silent on others (e.g., agriculture, robotics, and ridesharing). Although the AI Executive Order attempts to standardize AI governance across industries, it similarly does not impose any liability on the private sector. Nor does the Order impose liability on federal agencies or their contractors for failing to adhere to its guidelines.[10] This limited scope is further narrowed in the wake of the Supreme Court's decision to overturn the *Chevron* doctrine in Loper Bright Enterprises et al. v. Raimondo (2024).[11]

3.2. The EU AI regulatory frameworks

Unlike America, the EU has enacted and is in the process of finalizing major legislations controlling the use of AI and AI-powered services. Specifically, EU regulations and directives can be divided into two groups – those relevant to AI governance, and those explicitly overseeing the development and use of AI. Given the EU's prolific foray into this emerging field, it is not possible to examine all regulations and directives. We instead focus our analysis on what we consider to be representative legislations that deal with the three stages of the generative AI supply chain. In particular, we examine the General Data Protection Regulation (GDPR), the Digital Services Act (DSA), and the EU AI Act and how they regulate the processing, training, and deployment of generative AI models that are used to create deepfakes. While laudable, we believe that incorporating the landlords of creativity's ownership of foundation models into these EU regulations and directives would help them better achieve their regulatory goals.

3.2.1 GDPR

Although the GDPR does not explicitly mention AI or machine learning, many provisions in the law are relevant to the use and development of AI. As an initial matter, the operation of AI systems is in tension with the principle of data minimization of Article 5(1)(c) of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, 2016/679, 2016 (hereinafter 'GDPR'). This principle requires that personal data be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed' (GDPR, art. 5(1)(c)). Yet, the very idea of machine learning involves using algorithms to discover unexpected correlations and insights. Similarly, AI is affected by Article 22 of the GDPR, which provides for a general right 'not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her' (GDPR, art. 22). As AI systems are increasingly replacing humans in recruitment, lending, and other services, their usage makes it more likely that a decision will be based 'solely' on automated processing, which would then bring the system under the purview of Article 22 of the GDPR.

Relevant to deepfakes, generally, and audio deepfakes, specifically, are the provisions concerning the definition and use of personal data (GDPR, arts. 4–6, 9, 13–15, 17, 19, 21–22). The collection of de-identified data to train LLMs raises issues with the core tenets of the data protection principles underlying the GDPR – purpose limitation, data minimization, and the limitation on automated decisions. Article 5(1)(b) set forth the principle of purpose limitation, where personal data should be 'collected for specific, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes' (GDPR, art. 5(1)(b)). This restriction is relevant to AI systems as these systems usually would collect and process data for one purpose but then use it for another

[10] For example, it requires the Secretary of Labor to 'issue guidance to make clear that employers that deploy AI to monitor or augment employees' work must continue to comply with protections that ensure that workers are compensated for their hours worked, as defined under the Fair Labor Standards Act of 1938, 29 U.S.C. 201 et seq., and other legal requirements,' but does not place liability on employers that violate such guidance (88 Fed. Reg. 71,596, at 75,210). It merely directs the administrative state to study and issue reports, guidelines, and plans on the regulation of AI.

[11] The *Chevron* doctrine, which was overturned in *Loper Bright Enterprises v. Raimondo*, 144 S. Ct. 2244, 219L. Ed. 2d 832 (2024), required courts to defer to an agency's interpretation of a statute, unless it is not reasonable or unless Congress had clearly expressed a different intent.

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

(GDPR, art. 14(4)). For example, audio data collected to clone celebrity voices could be used to train generative AI models that could then later be used to foment discords.[12] Under the GDPR, then, it seems that the landlords of creativity might be liable for the illicit use of the data collected. As discussed below, however, the connection between the data collected and the generated deepfakes might be too attenuated to trigger the obligations imposed by GDPR. Further research is perhaps needed to 'carry[] out a data protection impact statement' as required under Article 35 of the GDPR (GDPR, art. 35).

3.2.2. DSA

The DSA (Council Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), 2022/2065, henceforth DSA) does not specifically target generative AI, but covers the deployment of LLMs that are used to generate deepfakes. In particular, the DSA requires platforms to manage and mitigate systemic risks arising from their services. (DSA, arts. 26, 34). It requires large online platforms to cooperate with authorities and independent auditors to show that the platforms have enacted effective mitigation strategies against content manipulation (DSA, arts. 28, 31). Although it exempts large online platforms from being liable for the underlying content, these platforms must implement systems for content detection, flagging, and removal, along with clear transparency requirements to alert users about the manipulated nature of the content (DSA, arts. 14–16).

In addition to targeting platforms where deepfake content is published, the DSA also imposes obligations – and potential liabilities – on hosting services that store information at the request of third parties (DSA, art. 6). This provision could plausibly extend to the digital infrastructure providers that deploy and maintain generative AI systems – the landlords of creativity. While these landlords may not directly publish or moderate content, their infrastructure power the services that generate and disseminate synthetic media. Under Article 6, if such services knowingly host manifestly illegal content, they are obligated to act. This raises an important regulatory question: should these foundational providers bear responsibility for monitoring or reporting harmful uses of the models they host, even if they lack direct control over the end-user interactions? A narrow interpretation of the DSA might exempt them – but such an exemption risks creating blind spots in accountability. To ensure robust safeguards against harmful generative AI content, the DSA's framework may need some revisions to clarify the duties of these infrastructural intermediaries.

3.2.3 EU AI act

To specifically address AI governance, the EU passed an AI-specific regulatory framework titled Council Regulation (EU) 2024/1689 of 13 June 2024 (Artificial Intelligence Act), 2024/1689 of the European Parliament and of the Council of 13 June 2024 (hereinafter 'EU AI Act'). The EU AI Act seeks to regulate AI systems to 'ensure that AI is trustworthy and safe and is developed and used in accordance with fundamental rights obligations' (EU AI Act, preamble(3)). To accomplish this goal, it takes a risk-based approach. The Act defines four different types of risk categories. First, AI systems that create an unacceptable risk (e.g., social scoring systems and manipulative AI), contravening EU values and considered to be a clear threat to fundamental rights, will generally be banned in the EU (EU AI Act, art. 50). Second, AI systems that are classified as high-risk AI (e.g., credit scoring system) will be subjected to comprehensive mandatory compliance obligations. High-risk AI systems will also be subjected to conformity assessments to evaluate their compliance with the Act (EU AI Act, art. 6). Third, AI systems classified as limited-risk (e.g., Chatbots, emotion recognition and biometric categorization systems, and systems generating deepfakes) will be subject to more minimal transparency obligations (EU AI Act, art. 5). Finally, all other AI systems not falling under one of the three main risk classes (e.g., AI-enabled recommender systems or spam filters) are classified

[12] In the UK, an audio deepfake of the London mayor, Sadiq Khan, purportedly making inflammatory remarks was uploaded online, which almost caused 'serious disorder' (Spring, 2024).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

as minimal/no-risk. The EU AI Act allows the free use of minimal-risk AI systems, while voluntary codes of conduct are encouraged (European Commission, n.d.).

The EU AI Act's risk based approach is a more comprehensive regulatory framework that attempts to lay the groundwork for imposing liability on the use of AI. The Act, however, does not explicitly address liability for AI-related harms. As outlined above, it merely creates a typology of AI risks and sets out obligations for entities at different levels of risk. In not considering the technical complexities and the problem of ownerships of generative AI, however, 'the AI Act is unlikely to achieve what it aims to do, namely the creation of conditions for trustworthy AI' (Kusche, 2024).[13] Indeed, 'the lack of civil society expertise in standard-setting' raises questions surrounding the need for a nuanced, inclusive, and context-specific approach to risk identification and mitigation (Gamito & Marsden, 2024). As the EU is working towards a liability framework, it is necessary to consider the ownership of foundation models throughout the generative AI supply chain.

3.2.4 EU regulatory gaps

The ways EU regulatory frameworks apportion liability do not neatly track with the roles of actors in the generative AI supply chain discussed in Section II. Neither the GDPR nor the DSA adequately regulate the landlords of creativity, despite these actors owning the LLMs used for deepfakes. In theory, these entities fall under the purview of GDPR and DSA, as they handle data ingestion, LLM training, and AI deployment – with GDPR governing data use and DSA covering AI deployment.

However, the current wording of both GDPR and DSA primarily targets downstream users and platforms. Take, for example, an AI-generated audio clip in Slovakia that was circulated widely just before the 2023 general election, purportedly showing Progressive Slovakia's chairman discussing election fraud strategies (An Incident Database, 2023). To comply with DSA requirements, Meta labeled the deepfake as satire. Meanwhile, Slovak authorities targeted the unknown creator of the clip, without addressing liability for the landlord responsible for the foundational AI model (Women Press Freedom, 2024).

This oversight arises because GDPR and DSA do not reflect the structure of the generative AI market. Although fine-tuned models produce most deepfakes (Mubarak, Alsboui & Alshaikh et al., 2023), their outputs heavily depend on foundation models owned by landlords (Ohm, 2024). Users rarely interact directly with foundation models, instead using intermediary fine-tuned models derived from landlords' foundation models. Consequently, even if deepfakes result primarily from foundational model training, landlords could evade liability under GDPR Article 14(4) and DSA Articles 14–16 since their models aren't directly implicated in downstream illicit use.

The EU AI Act better targets the landlords by differentiating between 'providers' (developers) and 'deployers' (users) of AI systems (EU AI Act, arts. 3(3)-(4)). However, its liability provisions still largely focus on downstream users. Article 50(4), for instance, mandates disclosure of AI-generated deepfakes by deployers (EU AI Act, art. 50(4)). This works clearly if end-users directly access the landlords' models but becomes ambiguous with intermediaries involved. In this case, it is unclear if the landlords are the one required to label the generated content or if this obligation falls on one of the intermediaries. Additionally, the Act's risk classification does not currently label landlords as high-risk, despite their foundational role in high-risk applications (EU AI Act, art. 6). Equally important, the AI Act imposes obligations rather than direct liability.

To address these limitations, the EU has proposed the Product Liability Directive (PLD) and AI Liability Directive (AILD), aiming to clarify civil liability frameworks for AI damages (European Commission, 2022a, 2022b).[14] While these proposals take the right steps forward, they must more effectively address the complexity of assigning liability specifically to the landlords of creativity.

[13] For example, Kusche criticizes that '[t]he category of high risk is apparently not defined by some type of technical risk assessment but the result of political judgments that are implicitly linked to values' (Kusche, 2024).

[14] The European Commission, however, has decided to withdraw the proposal on February 11, 2025 due to the lack of agreement on a final text.

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

3.3. Chinese AI regulatory frameworks

The Chinese state has also been prolific in developing AI regulations (Sheehan, 2023). Similar to the American and European approaches, some Chinese regulations do not explicitly target AI but cover related, tangential fields. Three are notably relevant to generative AI and deepfakes.

In 2021, the Chinese Communist Party (CCP) issued the Provisions on the Management of Algorithmic Recommendations in Internet Information Services (互联网信息服务算法推荐管理规定) (hereinafter 'Chinese Algorithm Recommendation Regulation'), applying broadly to internet services employing algorithmic recommendations, such as social media and e-commerce. It grants users rights including turning off algorithmic recommendations, deleting personalization tags, and receiving explanations of algorithm impacts (art. 17). Additionally, it introduced an algorithm registry, requiring providers to submit details like the provider's name, algorithm type, self-assessment reports, and displayed content (art. 24). Violations incur fines between CNY 10,000 to 100,000 (art. 31).

At around the same period, the CCP promulgated the Provisions on the Administration of Deep Synthesis Internet Information Services (互联网信息服务深度合成管理规定) (hereinafter 'Chinese Deepfakes Regulation'). This regulation not only covers service providers, but also users and any other entities involved in the use of deep fake (which it calls 'deep synthesis') technologies (Chinese Deepfakes Regulation, arts. 2, 6). It requires that deepfakes be labeled as such so that these contents do not 'cause confusion or mislead the public' (Chinese Deepfakes Regulation, art. 17). While the Chinese Deepfakes Regulation does not explicitly set out penalties for noncompliance, it does state that the violators would 'be punished in accordance with relevant laws and administrative regulations' (Chinese Deepfakes Regulation, art. 22).

Finally, the CCP promulgated the Interim Measures for the Management of Generative Artificial Intelligence Services (生成式人工智能服务管理暂行办法) (hereinafter 'Chinese Generative AI Regulation') on July 10, 2023. This regulation applies to the use of all generative AI technologies that are used to provide services to the public. It notably excludes development and application of generative AI technologies that have not been used to provide services to the public (Chinese Generative AI Regulation, art. 2). The Chinese Generative AI Regulation imposes highly onerous obligations on the providers of generative AI, requiring providers ensure that intellectual property rights are not violated, and that the providers '[e]mploy effective measures to increase the quality of training data, and increase the truth, accuracy, objectivity, and diversity of training data' (Chinese Generative AI Regulation, art. 7). If generative AI providers are found to have violated the Chinese Generative AI Regulation, they could be subjected to penalties 'in accordance with the provisions of the PRC Cybersecurity Law, The PRC Data Security Law, the PRC Law on the Protection of Personal Information, The PRC Law on Scientific and Technological Progress, and other such laws and administrative regulations' (Chinese Generative AI Regulation, art. 21).

Similar to the EU, the Chinese regulatory frameworks do not optimally impose liability on the entities that are responsible for creating the outputs of generative AI. For example, while the Chinese Deepfakes Regulation requires 'deep synthesis service providers' to watermark the outputs, this requirement can be imposed on downstream consumers of foundation models (Chinese Deepfakes Regulation, arts. 17, 23). Under the Chinese definition of 'deep synthesis service provider,' an entity that tweaks (i.e., fine-tunes) a pre-trained foundation model which then generates offending outputs would be held liable even though they have no knowledge of such offending outputs. Chinese courts have already imposed liability on downstream generative AI consumers in copyright cases. In February 2024, for example, the Guangzhou Internet Court held an undisclosed AI Company operating a website liable for copyright infringement (Yiu & Bond, 2024). The AI Company, however, did not create the AI model that generated the infringing output. Instead, the model was 'provided using an unnamed third-party provider's AI model which was connected via a programmable interface … to the [defendant AI Company's] website' (Yiu & Bond, 2024).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

3.4. Bigtech and the coalition for content provenance and authenticity

In addition to national regulatory frameworks, leading technology and media organizations (e.g., Adobe, Microsoft, Google) joined together in the Coalition for Content Provenance and Authenticity (C2PA) to regulate AI-generated content by developing open technical standards that certify the source and history (provenance) of digital media.[15] The C2PA develops open technical standards to verify and communicate the origin, authenticity, and provenance of digital content. Specifically, the C2PA has defined an interoperable specification for embedding and verifying metadata identifying the creator and history of content called Content Credentials (Coalition for Content Provenance and Authenticity, 2025) – a framework which is media-agnostic, supporting images, video, audio, and documents.

This standard would work across all stages of the generative AI supply chain, reaching even the landlords of creativity. Metadata can indicate which foundation model was used in generating content. Third-party tools and platforms could use the embedded metadata to verify the authenticity of a content. The C2PA was designed with the intention of promoting transparency and trust by providing audit trails for contents and helping to combat misinformation, disinformation, and manipulated media (e.g., deepfakes).

Despite its comprehensive framework, the C2PA faces several challenges in effectively labeling and authenticating voice recordings. For example, embedding detailed provenance information could raise privacy concerns that would trigger privacy laws such as the United States Health Insurance Portability and Accountability Act or the EU GDPR. Additionally, even when the provenance of the data is embedded, nothing prevents it from being stripped away during subsequent edits or conversions. To be effective, industry initiatives such as the C2PA need to be coupled with an actual liability regime, as proposed below.

4. Proposed liability framework

The current generative AI landscape is dominated by a few companies with the computational power to train foundation models, which we identify as 'landlords.'[16] These companies include OpenAI and Microsoft, with 39% and 30% share of the market, respectively (IoT Analytics, 2023). Unlike regular vendors, these landlords' business model revolves around a subscription-based SaaS model where users (i.e., lessees) pay to use pre-trained generative AI models via the landlords' APIs. This SaaS model essentially means that the users do not own the models, but merely lease them from the landlords of creativity (OpenAI, n.d.; Microsoft, n.d.).

The 'landlord' terminology provides a fitting analogy to conceptualize liability in the generative AI and audio deepfake domain. Just as traditional landlords own properties and lease them out to tenants, generative AI companies like OpenAI and Microsoft own and control the underlying foundation models and lease access to these powerful tools to users via subscription-based APIs. This terminology emphasizes the nature of the relationship – ownership, control, and conditional access – making clear the central position these 'landlords of creativity' hold within the generative AI supply chain (Jones, 2023; Lee et al., 2024).

Adopting the landlord terminology also appropriately highlights the asymmetrical control landlords hold over generative AI systems, similar to how property landlords have exclusive authority over maintaining and ensuring habitability of leased premises (Restatement (Second) of Property, Land. & Ten, 1977). Since these generative AI landlords alone have control over critical components such as data selection, training methods, model weights, and safety protocols, they are uniquely positioned

[15] This type of self-regulation might be preferred, as the engineering profession might have more expertise to allocate liability (Chau, 2024).

[16] According to IoT Analytics (2023) OpenAI and Microsoft take 79% of the generative AI foundation models and platforms market while Google take 7%, AWS take 8%, and the rest take 16% of the market.

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

to mitigate risks and prevent harms more efficiently and cost-effectively than their lessees (Calabresi, 1975).

Moreover, the landlord analogy aligns with historical legal principles related to liability. For instance, the landmark case MacPherson v. Buick Motor Co. establishes that manufacturers cannot evade liability simply because their products reach consumers through intermediaries (MacPherson v. Buick Motor Co, 1916). Similarly, AI landlords cannot evade responsibility simply because they distribute generative models through APIs or other intermediaries. Further supporting this analogy, current indemnification practices by major landlords like OpenAI and Microsoft parallel the implicit duties of landlords under common law, reinforcing their role in ensuring their 'property' – the generative AI models – remains safe and free from defect (Hawk, 2023; OpenAI, 2024b). Thus, the landlord framework is both practically and legally appropriate as it clearly delineates roles and responsibilities, facilitating a fair and effective liability regime that matches the realities of the generative AI landscape.

The current AI market complicates the effort to regulate generative AI because the aforementioned frameworks do not distinguish between the landlords of creativity and the lessees. The American frameworks, for example, lump both together as entities using 'automated systems … [that are] derived from machine learning, statistics, or other data processing or artificial intelligence techniques, and exclude passive computing infrastructure' (Biden AI Bill of Rights, 2022, at 10; AI Executive Order 75,193-4). Similarly, the EU AI Act does not clearly distinguish between the two parties, merely seeking to regulate 'a natural or legal person, public authority, agency, or other body that develops an AI system,' which is defined as any 'software that is developed with' machine learning techniques defined in the Annex (EU AI Act, art. 3). In the same way, the Chinese frameworks impose liabilities on any '[o]rganizations or individuals that use generative AI' (Chinese Generative AI Regulation, art. 5; Song & Mo, 2024).

4.1. The liability dilemma

Regulating the wrong party, or allowing for any increase in regulation, can stifle innovation. On the other hand, regulation is needed to protect consumers as well as businesses from the increasingly sophisticated attack vectors offered by audio deepfakes. In this section, we examine which parties can most appropriately bear the burdens of liability in a way which maximizes innovation while minimizing harm.

We argue that placing liability on the lessees is suboptimal because they are not the cheapest cost avoiders. They do not control the copious amount of data that were used to train the foundation model. Nor do they control the weights of the generative AI algorithms. Everything in the pre-trained foundation models is controlled by the landlords. Thus, the landlords of creativity should 'have better knowledge of the risks involved and of ways of avoiding them than alternative bearers … [the landlords of creativity are] in a better position to use that knowledge efficiently to choose the cheaper alternative … [and are] better placed to induced modifications in the behavior of others where such modifications is the cheapest way to reduce the sum of accident and safety costs' (Calabresi, 1975).

However, placing liabilities on the landlords of creativity is also problematic because their models could be put to unpredictable uses. In the simplest case, a foundation model is trained by a landlord for malicious uses. In this instance, the landlord would be fully responsible for said uses. Things are, however, more complicated because most foundation models are not trained for a particular malicious or high-risk use. Furthermore, if the lessees (or sublessees) repurpose (i.e., fine-tune) the foundation models for another, legal usage, it becomes more difficult to determine liability for when the models produce harmful outputs.

An illustrative example of this complexity can be seen in the landmark *Betamax Case*, where the U.S. Supreme Court ruled that Sony was not liable for copyright infringement by users, as the Betamax video recording device had substantial non-infringing uses (Sony Corp. of Am. v. Universal City Studios, Inc, 1984). This precedent highlights how tools that serve a wide array of purposes, not all

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

of which can be predicted by their creators, complicate the attribution of liability. Compounding this problem is the fact that the landlords cannot exercise the requisite control over their lessees/sublessees, which is necessary to hold the landlords liable under the principles of *respondeat superior* or franchisor/franchisee liability.

Due to this muddled picture, no US courts have determined that AI models are products for purposes of product liability. It is, however, undeniable that the landlords of creativity stand atop the complex generative AI supply chain (Jones, 2023). Similar to a complex manufacturing chain, the landlord's 'obligation to the consumer must [therefore] keep pace with the changing relationship between them; it cannot be escaped because the marketing of a product has become so complicated as to require one or more intermediaries' (MacPherson v. Buick Motor Co, 1916).

The EU has recently recognized this product liability dimension of generative AI and proposed a revision of the PLD and the AILD. Although the directives do not solely focus on the landlords of creativity, they at least recognize AI as a product and significantly expand the liability of AI providers (European Commission, 2022; European Parliament, 2024b). Scholars studying the American generative AI regulatory frameworks agree with this recategorization and have begun to explore placing liability on the landlords of creativity and other online platforms (Henderson et al., 2023; Janger & Twerski, 2023; Sharkey, 2022; Volokh et al., 2023). This imposition of liability makes sense because the landlords are the cheapest cost avoiders. Similar to online platforms such as Amazon, OpenAI's and Microsoft's 'position[s] in the distribution chain allow [them] to take cost-effective steps to reduce accidents … [and] it is not a close call … [to impose liability on them because] the benefits of the actions [OpenAI and Microsoft] can take to minimize accidents vastly outweigh the costs of these actions to' the landlords (Loomis v Amazon.com, Inc.).

4.2. Reducing liability with unknown users

As providers of generative AI model interfaces, landlords control the services accessible to users. Some landlords limit access frequency by tiers, while others offer granular control based on GPT models and training types (Microsoft, 2024b; OpenAI, 2024b). To mitigate malicious use, landlords can restrict capacities available to unidentified or untrusted users to extensively 'red-teamed' functionalities,[17] thus reducing potential liability for illicit actions.

When unregistered users generate and disseminate disinformation or propaganda that threatens national security, restricting their access should reduce the liability of the platform providers. This is because the whitelisted services offered by the landlords would have been thoroughly tested to meet the standards set by applicable AI regulatory frameworks.[18] In such cases, the landlords can be seen as having fulfilled their due diligence and should not face criminal liability for the unpredictable actions of users. However, as the cheapest cost avoiders, they may still bear civil liability. Similarly, if unregistered users produce infringing copyrighted content, the landlords should be held civilly liable to the rightful copyright holders.

This assumption of responsibility is not new, as it mirrors the implied warranty of habitability of landlord-tenant relationships. Under American common law, the landlord has a duty to maintain the property in a habitable condition. '[W]hen the landlord violates this duty, he becomes subject to liability for physical harm resulting from such violation' (Restatement (Second) of Property, Land. & Ten, 1977 § 17.6). Here, the property of the landlords of creativity is the generative AI model, and the

[17] 'Red-teamed' models are those that have been extensively tested by researchers to find flaws and vulnerabilities.

[18] Biden AI Bill of Rights, at 18 (requiring that automated systems to 'undergo extensive testing before deployment' that are domain-specific to ensure that the technology will work in its real-world context); AI Executive Order at 75,219 (recommending to agencies to adopt 'external testing for AI, including AI red-teaming for generative AI, to be developed in coordination with the Cybersecurity and Infrastructure Security Agency'); EU AI Act at 70–71 (defining the AI regulatory sandbox where data shall be processed for the purposes of 'developing and testing certain innovative AI systems'); Chinese Deepfakes Regulation art. 7 (mandating that service providers review their algorithm mechanism).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

Table 2. Overview of liability proposal (self-work)

	Description	Liability Proposal
Landlord/Lessee	The landlord creates and hosts the generative AI model, leasing it out to the renter via APIs.	Default liability for the landlord. Landlord can shift liability with an audit trail.
Joint Landlords	The landlord creates the generative AI model, but then sells or licenses it to another organization to host. That organization then leases the model (with or without modification) to the renter via APIs.	Default liability for landlord and commercial host. Liability could be divided via contract. Liability can be shifted with an audit trail.
Open Source	The landlord creates the generative AI model, but then open-sources it, without commercializing the model.	No liability for landlords.

landlords must make sure that said model does not harm the tenants. Various landlords of creativity have already implicitly accepted this principle. OpenAI, for example, has already agreed to indemnify its 'API customers … [for] any third party claim that Customer's use or distribution of Output infringes a third party's intellectual property right' (OpenAI, 2024b). Similarly, Microsoft also committed to 'defend and indemnify commercial customers from lawsuits for copyright infringement' arising from the usage of its Azure OpenAI service (Hawk, 2023). Our proposal will merely expand this indemnification coverage to any output provided to any users.

4.3. Indemnification policies of AI landlords

When dealing with registered users, the majority of the liability cases are expected to fall into two categories – commercial versus non-commercial distribution. The commercial distribution is further divided into two cases, one where the landlord creates and hosts the generative AI models and the other where the one company creates the models but another hosts them (see Table 2).

For registered, paying users, landlords may shift liability for harmful or infringing content back to the user if they maintain a clear audit trail linking specific user inputs to generated outputs. This record should show, for instance, that the user prompted the model to 'create an audio of the U.S. president praising religious extremists' or 'create a song similar to the current most listened-to track.' In such cases, the user – not the landlord – is primarily liable.

However, landlords are not fully absolved. If the prompt involves a celebrity or public figure, landlords should: (1) notify the affected party, (2) retain the generated content for an extended period, and (3) warn the user before delivering the content. After notification, existing frameworks such as the DMCA govern further action. If the prompt concerns a private individual, liability shifts to the user upon distribution under existing laws (e.g., right of publicity, privacy). Upon notice of misuse, landlords are responsible for ensuring such content cannot be reproduced. This aligns with landlord-tenant principles: liability arises where the landlord knew or should have known of the harm and failed to take reasonable steps to prevent it (Restatement (Second) of Property, Land. & Ten (1977)).

The proposal is also complementary to various reporting requirements of the American, EU, and Chinese AI frameworks. The Biden AI Bill of Rights recommends an '[i]ndependent evaluation and reporting that confirms that the system is safe and effective, including reporting of steps taken to mitigate potential harms, should be performed and the results made public whenever possible.' Similarly, the EU AI Act mandates that '[p]roviders of high-risk AI systems placed on the Union market shall report any serious incident or any malfunctioning of those systems which constitutes a breach of obligations under Union law.' Finally, the Chinese Algorithm Recommendation Regulation requires providers to submit to the algorithm registry 'the service provider's name, service form, field of application, type of algorithm, algorithm self-assessment report, content to be displayed.'

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

In cases where the organization that trains the models does not monetize said models, this proposal imposes no liability for the landlords. While this case is far and few in between, it is not unusual. One of the more prominent examples is Meta's open-source of Llama generative AI model (see, e.g., https://llama.meta.com, Meta's open-source of Llama generative AI model). Open source model makers include in their licenses provisions that shift the risk of liability from the developers of the models to the end-users. This proposal recognizes this disclaimer of liability, which tracks with how the EU deals with open-source generative AI models (European Parliament, 2024b). For example, the scope of the revised PLD does not include 'free and open-source software that is developed or supplied outside the course of a commercial activity' (European Commission, 2022a).

4.4 A more complete framework

Our proposed framework addresses the regulatory gaps in existing American, European, and Chinese frameworks by clearly delineating responsibilities in the generative AI ecosystem through the landlord-lessee analogy. Specifically, our proposal assigns default liability to the landlords of creativity, recognizing them as the entities best positioned to prevent and mitigate harms associated with generative AI systems. The current American regulatory proposals, such as the Biden AI Bill of Rights and the AI Executive Order, do not sufficiently distinguish between developers (landlords) and end-users (lessees). They broadly define AI and AI systems in a manner that lumps together all entities involved in the use or development of these systems, from infrastructure providers to end-users. In doing so, these frameworks dilute accountability and overlook the crucial distinction between those who control the underlying technology and those who merely lease access to it.

This gap can be explicitly addressed by assigning primary liability to the landlords – companies like OpenAI and Microsoft – that control foundation models, datasets, model weights, and safety protocols. This contrasts with the ambiguous U.S. framework, ensuring clearer accountability and incentivizing landlords to proactively manage risks. This liability framework also aligns more closely with historical U.S. product liability legal precedents (MacPherson v. Buick Motor Co.) by reinforcing accountability up the supply chain rather than dispersing it.

The EU AI Act and the revised PLD, while more advanced in recognizing generative AI as a product, still fail to adequately distinguish the unique positions occupied by landlords versus lessees. The EU's approach tends to spread liability across various actors without sufficient granularity, potentially causing confusion and diluting incentives for risk mitigation.

The proposed liability framework enhances clarity and innovation velocity by assigning default liability specifically to landlords, thereby incentivizing entities with the greatest control to implement robust safety and mitigation measures. Unlike the broad EU framework, which holds AI system developers liable without distinction, our framework provides landlords with clear mechanisms – such as audit trails and red-teaming – to demonstrate responsible conduct and shift liability appropriately when warranted. Furthermore, our proposal extends current indemnification practices in the industry, reinforcing implicit duties akin to the implied warranty of habitability. By doing so, it leverages established legal principles to foster safer AI practices without imposing undue burdens on innovation.

5. Conclusion

The rapidly evolving generative AI and audio deepfakes landscape demands a regulatory framework that is both flexible and technically feasible. While laudable, the regulatory frameworks of the United States, the EU, and China miss the tree for the forest and do not fully address the landlord-lessee relationships of the generative AI ecosystems. Our recommendation to impose default liability on the 'landlords of creativity' – those entities with the resources and technical capabilities to develop foundation generative AI models – offers a pragmatic approach to mitigating potential harms while fostering innovation.

The landlords of creativity are not only the cheapest cost avoiders, but also are obligated to provide an infrastructure free of defects. Therefore, they should be liable in most cases of generative AI misuse. To reduce liability, the landlords could limit the features available to unregistered users, allowing them to access only features that have been extensively tested in accordance with the requirements of the American, EU, and Chinese regulatory frameworks. In the case of registered renters, the landlords could reduce liability by having an audit trail that keeps track of registered users' prompts and outputs. Additionally, the landlords should have a notification system to notify affected entities of potentially malicious usage of their models.

Conforming to these requirements would not be financially or technically onerous to the landlords. Many landlords have already agreed to indemnify their lessees if they are sued by a third party (OpenAI Service Terms). Additionally, it is industry practice to keep logs for debugging purposes (Menn, 2021). Our proposal would not incur significant technical debt, as the primary logging infrastructure is already in place at many of these landlords' premises. Indeed, our proposal is also in line with the C2PA standard that the software industry is creating. Adopting this proposal would therefore allow the landlords space to innovate while still addressing the potential harms of such innovation.

Acknowledgement. Bao Kham Chau thanks for their comments to James Grimmelman, Jordi Weinstock, Chang Cai, Anthony Adamthwaite, the CornellTech DLI Seminar, and participants in the Yale Law School's Propaganda and Emerging Technologies Conference, the Stanford Law School's Law and Society in an Age of Crisis, the Public Governance and Emerging Technologies Conference, and the JUST-AI Jean Monnet Summer Colloquium. George He thanks for their comments to Christopher Bavitz, Jack Cushman, the Harvard Berkman Klein Center for Internet and Society, and participants in the Stanford Law School's Law and Society in an Age of Crisis, and the JUST-AI Jean Monnet Summer Colloquium.

Funding statement. We received no funding from any source for this article.

Competing interest statement. The authors declare none.

References

Amazon Web Services, Inc. (n.d.). *Foundation models*. https://aws.amazon.com/what-is/foundation-models/

Analytics, I. (2023, December). *The leading generative AI companies*. https://iot-analytics.com/wp/wp-content/uploads/2023/12/INSIGHTS-RELEASE-The-leading-generative-AI-companies.pdf

Andersen v Stability AI Ltd, 700 F. Supp. 3d 853 (N.D. Cal. 2023)

Appel, G., Neelbauer, J., & Schweidel, D. A. (2023, April 7). Generative AI has an intellectual property problem. *Harvard Business Review*. https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., … Wei, F. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*. https://crfm.stanford.edu/assets/report.pdf

Bradford, A. (2020). *The Brussels effect: how the European Union rules the world*. Oxford: Oxford University Press.

Brennan Center for Justice. (2024, October 1). Artificial intelligence legislation tracker. https://www.brennancenter.org/our-work/research-reports/artificial-intelligence-legislation-tracker

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Amodei, D. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*. https://arxiv.org/abs/2005.14165

Calabresi, G. (1975). Concerning cause and the law of torts: an essay for Harry Kalven, Jr. *University of Chicago Law Review*, *43*(1), 69–84.

Carlini, N., & Wagner, D. (2018). Audio adversarial examples: Targeted attacks on speech-to-text. In *2018 IEEE Security and Privacy Workshops (SPW)* (1–7). IEEE. https://doi.org/10.1109/SPW.2018.00009

Chau, B. K. (2024). Engineering a fiduciary: Expanding the regulatory scope of algorithmic bias. https://jolt.law.harvard.edu/digest/engineering-a-fiduciary-expanding-the-regulatory-scope-of-algorithmic-bias.

Chau, B. K., & Livermore, M. (2024). Computational legal studies comes of age. *European Journal of Empirical Legal*, *1*(1), 89–104. https://publicera.kb.se/ejels/article/view/19684

Chen, G., Morris, P. P., Stevens Aubrey, J., & Wang, T. (2019). Understanding programmatic creative: The role of AI. *Journal of Advertising*, *48*(4), 347–355. https://doi.org/10.1080/00913367.2019.1652128

China Law Translate. (2023). Interim measures for the management of generative artificial intelligence services [Translation]. https://www.chinalawtranslate.com/en/generative-ai-interim/ (Accessed 19 February 2024).

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

Coalition for Content Provenance and Authenticity, C2PA, https://c2pa.org/.

Coalition for Content Provenance and Authenticity. (2025). *Specifications*. https://c2pa.org/specifications/specifications/1.4/index.html

Comp, L. A. § 169.247 (West).Congressional Research Service. (2023, December 28). *Section 230 immunity and generative artificial intelligence*. https://crsreports.congress.gov/product/pdf/LSB/LSB11097

Coscarelli, J. (2023, April 19). An A.I. hit of fake 'Drake' and 'The Weeknd' rattles the music world. *The New York Times*. https://www.nytimes.com/2023/04/19/arts/music/ai-drake-the-weeknd-fake.html

Council Regulation (EU) 2022/2065 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. http://data.europa.eu/eli/reg/2022/2065/oj

Council Regulation (EU) 2024/1689 of 13 June 2024 (Artificial Intelligence Act), 2024 O.J. L, 12 July 2024. http://data.europa.eu/eli/reg/2024/1689/oj

Damiani, J. (2019). A Voice Deepfake Was Used to Scam a CEO Out of $243,000. *Forbes*, https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/.

Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2023, 118th Cong.(2023).

Dubus et al v. NVIDIA Corporation, 4:2024cv02655 (N.D. Cal. 2024)

Elneima, A., & Bińkowski, M. (2022). Adversarial text-to-speech for low-resource languages. In *Proceedings of the Seventh Arabic Natural Language Processing Workshop*. https://aclanthology.org/2022.wanlp-1.8/

European Commission. (2022, September 28). Proposal for a directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive - AILD), COM(2022) 496 final. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496

European Commission. (2022, September 28). Proposal for a directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive - AILD), COM(2022) 496 final. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496

European Commission. (2022a). Proposal for a directive on liability for defective products (COM(2022) 495 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0495

European Commission. (2022b). Proposal for a directive on adapting non-contractual civil liability rules to artificial intelligence (COM(2022) 496 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0496

European Commission. (n.d.). *AI Act: Shaping Europe's Future*. Retrieved October 20, 2024, from https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

European Parliament. (2024a, March 12). Legislative resolution P9_TA(2024)0132 on the proposal for a directive on liability for defective products (COM(2022)0495 – C9-0322/2022 – 2022/0302(COD)). *New Product Liability Directive*. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0132_EN.html

European Parliament. (2024b, March 12). Legislative resolution P9_TA(2024)0130 on the proposal for a regulation on horizontal cybersecurity requirements for products with digital elements (Cyber Resilience Act). https://www.europarl.europa.eu/doceo/document/TA-9-2024-0130_EN.pdf

FBI, Criminals Use Generative Artificial Intelligence to Facilitate Financial Fraud, Alert Number: I-120324-PSA (December 3, 2024), https://www.ic3.gov/PSA/2024/PSA241203.

Fed. Reg. 715196-7 (Nov. 1, 2023) (hereinafter "AI Executive Order").

Federal Trade Commission. (2023, June 29). Generative AI raises competition concerns. https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns

Federal Trade Commission. (2024). Trade regulation rule on commercial surveillance. https://www.reginfo.gov/public/do/eAgendaViewRule?pubId=202110&RIN=3084-AB69

Federal Trade Commission. (n.d.). The FTC Voice Cloning Detection Challenge. Retrieved February 18, 2024, from https://www.ftc.gov/news-events/contests/ftc-voice-cloning-detection-challenge.

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business and Information Systems Engineering*, *66*(1), 111–126. https://doi.org/10.1007/s12599-023-00834-7

Gabryś, A., Huybrechts, G., Ribeiro, M.S., Chien, C.M., Roth, J., Comini, G., Barra-Chicote, R., Perz, B., Lorenzo-Trueba, J. *et al* (2022). Voice Filter: Few-shot text-to-speech speaker adaptation using voice conversion as a post-processing module https://arxiv.org/abs/2202.08164

Gamito, M. C., & Marsden, C. T. (2024). Artificial intelligence co-regulation? The role of standards in the EU AI Act. *International Journal of Law and Information Technology*, *32*(1), eaae011.

George, H., & Roberts Kingman, K. (2022). AI - Is it Art, yet? *Harvard Journal of Law and Technology Digest*, https://jolt.law.harvard.edu/digest/ai-is-it-art-yet.

Giattino, C., Mathieu, E., Samborska, V., & Roser, M. (n.d.). Data page: Parameters in notable artificial intelligence systems. *Our World in Data*. Retrieved October 14, 2024, from https://ourworldindata.org/grapher/artificial-intelligence-parameter-count.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., and Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (vol 27, 2672–2680). New York: Curran Associates, Inc.

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

GovTrack.US. (2024a). Protection Elections from Deceptive AI Act. https://www.govtrack.us/congress/bills/118/s2770

GovTrack.US. (2024b). Deepfakes Accountability Act. https://www.govtrack.us/congress/bills/118/hr5586

GovTrack.US. (2024c). Preventing Deep Fake Scams Act. https://www.govtrack.us/congress/bills/118/hr5508

Hawk, J. (2023, November 15). Microsoft Azure AI, data, and application innovations help turn your AI ambitions into reality. *Microsoft Azure*. https://azure.microsoft.com/en-us/blog/microsoft-azure-ai-data-and-application-innovations-help-turn-your-ai-ambitions-into-reality/

Heath, R. (2024). Exclusive: States are introducing 50 AI-related bills per week. *Axios*. https://www.axios.com/2024/02/14/ai-bills-state-legislatures-deepfakes-bias-discrimination

Hemmings v. Pelham Wood Ltd. Liability Ltd. Partnership, 375 Md. 522, 826 A.2d 443 (2003)

Henderson, P., Hashimoto, T., & Lemley, M. (2023). Where's the liability in harmful AI speech? *Journal of Free Speech Law*, *3*, 589.

Hoffmann, J., Borgeaud, S., Cai, T., Millican, K., Kyriakidis, P., Wang, M., Karp, A. R. (2022, March 29). Training compute-optimal large language models. *arXiv preprint*ar *Xiv:2203.15556*. https://arxiv.org/abs/2203.15556v1

Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, *162*(6), 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

An Incident Database. (2023). *Incident 573: Deepfake Recordings Allegedly Influence Slovakian Election*. https://incidentdatabase.ai/cite/573/

Janger, E. J., & Twerski, A. D. (2023). Functional tort principles for internet platforms: Duty, relationship, and control. *Yale Journal of Law & Technology*, *26*(1), 1.

Jones, E. (2023, July 17). *Explainer: What is a foundation model?* Ada Lovelace Institute. https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/

Jones, H. (2024, June 14). How startups are using AI. *Kruze Consulting*. https://kruzeconsulting.com/blog/how-startups-using-ai/.

Kadrey v. Meta Platforms, Inc., No. 23-CV-03417-VC, 2023 WL 8039640 (N.D. Cal. Nov. 20, 2023).

Kahn, J., Rivière, M., Zheng, W., Kharitonov, E., Xu, Q., Mazaré, P.-E., … Fuegen, C. et al. (2020). Libri-Light: A benchmark for ASR with limited or no supervision. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 7669–7673.

Kaplan, J., McCandlish, S., Henighan, T., Brown, D. M., Chess, B., Child, R., … Sutskever, I. (2020, January 23). Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*. https://arxiv.org/abs/2001.08361

Kardon, I. (2023). Hearing on "Rule by Law: China's Increasingly Global Legal Reach". https://www.uscc.gov/sites/default/files/2023-05/Isaac_Kardon_Testimony.pdf.

Khosrowi, D., Finn, F., & Clark, E. (forthcoming). Engaging the many-hands problem of generative-AI outputs: A framework for attributing credit. *AI & Ethics*.

Kim, C., Ren, Y., & Yang, Y. (2021). Decentralized attribution of generative models. In *Proceedings of the International Conference on Learning Representations (ICLR)*. https://openreview.net/forum?id=_kxlwvhOodK

Kirkpatrick, J., McMorrow, B., Turban, D. H. P., Gaunt, A. L., Spencer, J. S., Matthews, A. G. D. G., … Cohen, A. J. (2021). Pushing the frontiers of density functionals by solving the fractional electron problem. *Science*, *374*(6573), 1385–1389. https://doi.org/10.1126/science.abj6511

Kong, J., Lee, H., Kim, J., & Jaekyoung, B. (2020). HiFi-GAN: Generative adversarial networks for efficient and high fidelity speech synthesis. In *Advances in Neural Information Processing Systems* (vol 34, 1–12). New York: Curran Associates, Inc.

Kusche, I. (2024). Possible harms of artificial intelligence and the EU AI act: Fundamental rights and risk. *Journal of Risk Research*, 1–14. https://doi.org/10.1080/13669877.2024.2350720.

Laffont, -J.-J., & Tirole, J. (2000). *Competition in telecommunications*. Cambridge: The MIT Press.

Lee, K., Cooper, A. F., & Grimmelmann, J. (2024). Talkin' 'Bout AI generation: Copyright and the generative-AI supply chain. *Journal of the Copyright Society*. https://dx.doi.org/10.2139/ssrn.4523551.

Leffer, L. (2024, January 26). AI audio deepfakes are quickly outpacing detection. *Scientific American*. https://www.scientificamerican.com/article/ai-audio-deepfakes-are-quickly-outpacing-detection/

Loomis v Amazon.com, Inc, 277 Cal. Rptr. 3d 769, 789 (Cal. Ct. App. 2021).

MacPherson v. Buick Motor Co, 217 N.Y. 382, 111 N.E. 1050 443 (1916) (Traynor, J. concurring).

Magowan, J. (2023). It's like I've got this music in my mind: Protecting human authorship in the age of generative artificial intelligence. *UC Law Journal*, *75*, 233.

Mariani, J., Eggers, W. D., & Kishnani, P. K. (2024). *The AI regulations that aren't being talked about*. Deloitte Center for Government Insights. https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-regulations-around-the-world.html

Mcuba, M., Singh, A., Ikuesan, R. A., & Venter, H. (2023). The effect of deep learning methods on deepfake audio detection for digital investigation. *Procedia Computer Science*, *219*, 211–218. https://doi.org/10.1016/j.procs.2023.01.080

Mehrish, A., Majumder, N., Bhardwaj, R., Mihalcea, R., & Poria, S. (2023, May 30). A review of deep learning techniques for speech processing. *arXiv preprint arXiv:2305.00359*. https://arxiv.org/pdf/2305.00359

Menn, J. (2021, December 16). Major tech companies struggle to plug holes in logging software. Reuters. https://www.reuters.com/technology/major-tech-companies-struggle-plug-holes-logging-software-2021-12-16/

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

Menon, D., & Ranganathan, R. (2022). A generative approach to materials discovery, design, and optimization. *ACS Omega*, *7*(29), 25958–25971. https://doi.org/10.1021/acsomega.2c02311

Microsoft. (2024a). *Azure OpenAI Service pricing overview*. Retrieved February 24, 2024, from https://azure.microsoft.com/en-us/pricing/details/cognitive-services/openai-service/.

Microsoft. (2024b). *VALL-E 2*. https://www.microsoft.com/en-us/research/project/vall-e-x/vall-e-2/

Microsoft. (n.d.). *Azure OpenAI service pricing*. Retrieved February 23, 2024, from https://azure.microsoft.com/en-us/pricing/details/cognitive-services/openai-service/.

Ministry of Education. (2023, July 10). *Interim measures for the management of generative artificial intelligence services* [Translation].Translated at https://www.chinalawtranslate.com/en/generative-ai-interim/

Ministry of Public Security. (2022, November 25). *Provisions on the Administration of Deep Synthesis Internet Information Services* [Translation]. Translated at https://www.chinalawtranslate.com/en/deep-synthesis/

Mubarak, R., Alsboui, T., Alshaikh, O. et al. (2023). A survey on the detection and impacts of deepfakes in visual, audio, and textual formats. *IEEE Access*, *11*, 144497–14452. https://doi.org/10.1109/ACCESS.2023.3344653

Narayanan, D., Chen, S., Wu, Y., Zhang, Z., Zhou, L., Liu, S., … Wei, F. (2021, August 23). Efficient large-scale language model training on GPU clusters using Megatron-LM. *arXiv preprint arXiv:2104.04473*. https://arxiv.org/pdf/2104.04473

National Conference of State Legislatures. (2024, January 12). Artificial intelligence 2023 legislation. https://www.ncsl.org/technology-and-communication/artificial-intelligence-2023-legislation

National Internet Information Office, Ministry of Industry and Information Technology of the People's Republic of China, Ministry of Public Security of the People's Republic of China, & State Administration for Market Regulation. (2021, December 31). *Provisions on the Management of Algorithmic Recommendations in Internet Information Services*. Translated at https://www.chinalawtranslate.com/en/algorithms/

Ohm, P. (2024). Focusing On Fine-Tuning: Understanding The Four Pathways For Shaping Generative AI. *Science and Technology Law Review*, *25*(2). https://doi.org/10.52214/stlr.v25i2.12762

OpenAI. (2024a). *OpenAI service terms*. Retrieved February 26, 2024, from https://openai.com/policies/service-terms.

OpenAI. (2024b). *Rate limits*. Retrieved February 24, 2024, from https://platform.openai.com/docs/guides/rate-limits?context=tier-free.

OpenAI. (2019, February 14). *Better language models and their implications*. https://openai.com/index/better-language-models/

OpenAI. (2023, March 23). *GPT-4 system card*. https://www.judiciary.senate.gov/imo/media/doc/2023-05-16_-_qfr_responses_-_altman_addendum.pdf

OpenAI. (n.d.). *ChatGPT pricing*. Retrieved February 23, 2024, from https://openai.com/chatgpt/pricing.

OpenAI Help Center. (n.d.). *How to access Sora?* OpenAI. Retrieved February 21, 2024, from https://help.openai.com/en/articles/8958981-how-to-access-sora.

OpenAI Service Terms. (2025, April 4) https://openai.com/policies/service-terms/.

Oral Argument at 49, *Gonzalez v. Google LLC*, 143 S. Ct. 1191 (2023) (Gorsuch, J).

Oversight of A.I.: Rules for Artificial Intelligence, 118th Cong. (2023) (statement of Sam Altman, CEO of OpenAI).

Pichardo v. Big Diamond, Inc, 215 S.W.3d 497 (Tex. App. Fort Worth 2007)

Preventing Deep Fake Scams Act, H.R. 5808, 118th Cong. (2023).

Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 1 (2023).

Reed, R. (2023, May 8). AI created a song mimicking the work of Drake and The Weeknd: What does that mean for copyright law? *Harvard Law School Today*. https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law/

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016. (2016). General Data Protection Regulation. *Official Journal of the European Union*, L 119, 1–88. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC

Restatement (Second) of Property, Land. & Ten. § 17.6 (1977)

Reynolds, M. (2024, January 18). AI-generated music is everywhere; is any of it legal? *ABA Journal*. https://www.abajournal.com/web/article/ai-generated-music-is-everywhere-is-any-of-it-legal

Sastry, G., Heim, L., Belfield, H., Anderljung, M., Brundage, M., Hazell, J., Coyle, D. (2024, February 13). Computing power and the governance of artificial intelligence. *arXiv preprint arXiv:2402.08797*. https://arxiv.org/pdf/2402.08797

Sharkey, C. M. (2022). Products liability in the digital age: Online platforms as "cheapest cost avoiders". *Hastings Law Journal*, *73*, 1327.

Sheehan, M. (2023). *China's AI Regulations and How They Get Made*. Carnegie Endowment for International Peace.

Song, S., & Mo, W. (2024, April 26). China's first case regarding AI-generated voice infringement. https://www.kwm.com/cn/en/insights/latest-thinking/china-s-first-case-regarding-ai-generated-voice-infringement.html

Sony Corp. of Am. v. Universal City Studios, Inc, 464 U.S. 417 (1984).

Spring, M. (2024, February 13). Sadiq Khan says fake AI audio of him nearly led to serious disorder. https://www.bbc.com/news/uk-68146053

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press

Stanford University Human-Centered Artificial Intelligence. (2024). *Artificial Intelligence Index Report 2024*. https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf

U.S. Department of Defense. (2022a, February 24). DOD adopts ethical principles for artificial intelligence. https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/

U.S. Department of Defense. (2022b, June). Responsible artificial intelligence strategy and implementation pathway. https://www.ai.mil/docs/RAI_Strategy_and_Implementation_Pathway_6-21-22.pdf

U.S. Equal Employment Opportunity Commission. (2023, May 18). *Select issues: Assessing adverse impact in software, algorithms, and artificial intelligence used in employment selection procedures under Title VII of the Civil Rights Act of 1964*. https://www.eeoc.gov/laws/guidance/select-issues-assessing-adverse-impact-software-algorithms-and-artificial

United States Department of Transportation, National Highway Traffic Safety Administration. (2022). Summary report: Standing general order on crash reporting for automated driving systems (DOT HS 813 324). https://www.nhtsa.gov/sites/nhtsa.gov/files/2022-06/ADS_SGO_Summary_Report.pdf

Vanessa Romo, Grimes invites fans to make songs with an AI-generated version of her voice, NPR (April 24, 2023, 7:21 PM ET), https://www.npr.org/2023/04/24/1171738670/grimes-ai-songs-voice.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, 5998–6008.

Volokh, E., Lemley, M. A., & Henderson, P. (2023). Freedom of speech and AI output. *Journal of Free Speech Law*, *3*, 651.

Wang, C., Chen, S., Wu, Y., Zhang, Z., Zhou, L., Liu, S., … Wei, F. (2023). Neural codec language models are zero-shot text to speech synthesizers. *Microsoft*, https://www.microsoft.com/en-us/research/publication/neural-codec-language-models-are-zero-shot-text-to-speech-synthesizers/.

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Fedus, W. (2022). Emergent abilities of large language models. *Transactions on Machine Learning Research*, *22*, https://openreview.net/forum?id=yzkSU5zdwD.

The White House. (2022). *Blueprint for an AI Bill of Rights: Making automated systems work for the American people*. Author. https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf

The White House. (2023). *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*. Author. https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

The White House. (2025). *Removing Barriers to American Leadership in Artificial Intelligence*. Washington, DC: Author. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/

Wiggers, K. (2019, December 26). Generative adversarial networks: What GANs are and how they've evolved. *VentureBeat*. https://venturebeat.com/2019/12/26/gan-generative-adversarial-network-explainer-ai-machine-learning/

Williamson, E. (2022, March 17). Q&A: With Zelenskyy surrender hoax, the feared future of deepfakes is here. *UVA Today*

Women Press Freedom. (2024). Slovakia: Deepfake audio clip aims to manipulate voters and discredit journalist Monika Tódová ahead of election. https://www.womeninjournalism.org/threats-all/slovakia-deepfake-audio-clip-aims-to-manipulate-voters-and-discredit-journalist-monika-tdov-ahead-of-election. (Accessed 14 April 2025).

Xu, J., Tan, X., Ren, Y., Qin, T., Li, J., Zhao, S., & Liu, T.-Y. (2020, August 9). *LRSpeech: Extremely low-resource speech synthesis and recognition*. arXiv preprint arXiv:2008.03687. https://arxiv.org/pdf/2008.03687

Yiu, C., & Bond, T. (2024, April 10). Liability of AI service providers for copyright infringement: Guangzhou Internet Court reaches world's first decision. Bird & Bird. https://www.twobirds.com/en/insights/2024/china/liability-of-ai-service-providers-for-copyright-infringement

Zhang, P., & Kamel Boulos, M. N. Article 286. (2023). Generative AI in medicine and healthcare: Promises, opportunities, and challenges. *Future Internet*, *15*(9), 286. https://doi.org/10.3390/fi15090286.

Zhou, Q., Ma, K., Xia, Y., Zheng, Y., Nair, B., Vargo, A., … Wu, S. (2021). A machine and human reader study on AI diagnosis model safety under attacks of adversarial images. *Nature Communications*, *12*. https://doi.org/10.1038/s41467-021-27441-7

Zick, T. (2013). *The cosmopolitan First Amendment*. Cambridge: Cambridge University Press.

Bao Kham Chau is a visiting fellow at the Cornell University, Cornell Tech NYC, and an affiliate at the Harvard Berkman Klein Center for Internet and Society. His research interest focuses on critical intersections between law, technology, and governance. Prior to entering the legal profession, he worked as a senior software engineer at several BigTech companies, where he built patented machine learning and artificial intelligence features. Bao also helped co-found the Cornell Law Xenophobia Meter Project, which developed a proprietary AI model to detect xenophobic hate speech from ingested online content.
He received his Juris Doctorate at the University of Virgina School of Law, his Master of Arts at Harvard University, and his Bachelor of Arts in Computer Science, History, and Political Science at the University of California, Berkeley.
George He is a research fellow at Harvard Library Innovation Lab.

Cite this article: Chau B.K. and He G. (2025). Audio deepfakes and the regulation of the landlords of creativity. *Cambridge Forum on AI: Law and Governance* **1**, e30, 1–22. https://doi.org/10.1017/cfl.2025.10012

https://doi.org/10.1017/cfl.2025.10012 Published online by Cambridge University Press