Digital Twins in the Profit Machine: Deepfakes, Digital Personhood, and the Commercial Licensing of Identity

By Lílian de Macedo

A (Non-Human) Star Is Born

Consider the recent release of a motion picture that exhibits every hallmark of a blockbuster: an innovative screenplay, a compelling visual campaign, and a leading actor whose global box-office appeal remains unrivaled. In a traditional marketplace, those characteristics would virtually guarantee commercial success. However, this production presents a significant legal anomaly: the featured performer never appeared on set, never received directorial guidance, and never engaged with the cast. This phenomenon is not the result of archival editing, but rather the unauthorized deployment of artificial intelligence to synthesize an actor's likeness and persona, raising profound questions regarding the exhaustion of personality rights in the digital age.

This description may sound unrealistic, but is within the realm of existing possibilities—and it even has a name: deepfake. In Hollywood's surreal world, the ability to manipulate a person's face, voice, and even mannerisms to seemingly resemble a human being is already a reality. In early October 2025, the public was introduced to the self-proclaimed first AI-generated actress in the world, Tilly Norwood.[1] Her body, voice, and facial expressions—created by artificial intelligence developed by a group of Australian engineers and artists—resemble a human being so closely that it has become difficult to classify her as an artificial creation. According to the team behind the supposed actress, Tilly—the artificial intelligence—also assisted in writing the script for the science fiction film in which she stars and contributed to directing some scenes.

The virtual actress, however, sparked mixed reactions in certain sectors, particularly among actors. The American actors' union, SAG-AFTRA issued a statement to those responsible for the AI.[2] The document clarified that Tilly Norwood is not an actress, but rather a character generated by a computer program trained using the work of countless professionals. The statement emphasized that the training process used videos and images of numerous real human actors without their permission or compensation. The union further argued that the virtual being has no lived experience on which to base a performance, since she has no emotion.

---

[1] Accessible in: https://www.tillynorwood.com

[2] Accessible in Forbes: SAG-AFTRA Condemns AI 'Actress' Tilly Norwood—Joins Critics Emily Blunt, Whoopi Goldberg And More - https://www.forbes.com/sites/conormurray/2025/09/30/sag-aftra-condemns-ai-actress-tilly-norwood-joins-critics-emily-blunt-whoopi-goldberg-and-more/

The union's argument is compelling. Indeed, the AI was trained by humans with the purpose of eliciting emotional responses from other humans. And it is precisely this issue that has raised concern in another sector: the judiciary. Legal scholars and practitioners are now confronted with documents, videos, photographs, and other forms of evidence that may potentially be AI-generated. In many cases, identifying such use is nearly impossible without disclosure.

The Mirror Has Two Faces (and a Thousand Pixels)

The challenges posed by artificial intelligence in the legal field go far beyond evidentiary verification or even the drafting of legal pleadings and academic articles. This is because even the terminology and definitions currently used to describe AI's capabilities remain unsettled. A clear example is the term deepfake itself. Only a few years ago, it was largely unknown, and even today it lacks a consistent definition, varying depending on the source.

According to the U.S. government guide Increasing Threat of Deepfake Identities, available on the Department of Homeland Security website, "Deepfakes [are] an emergent type of threat falling under the greater and more pervasive umbrella of synthetic media [that] utilize a form of artificial intelligence/machine learning (AI/ML) to create believable, realistic videos, pictures, audio, and text of events which never happened."[3]

Wikipedia offers a similar definition, stating that "Deepfakes are images, videos, or audio that have been edited or generated using artificial intelligence, AI-based tools, or audio-video editing software."[4]

Both definitions largely converge, although they differ somewhat from Microsoft's approach. According to the tech company's website, "A deepfake is a fraudulent piece of content—typically audio or video—that has been manipulated or created using artificial intelligence. This content replaces a real person's voice, image, or both with similar-looking and sounding artificial likenesses."[5]

The Constant Gardener: Cultivating European AI Governance

Some nations have already established regulatory parameters to address the issue. The European Union, for example, adopted AI governance rules in mid-2024. The provisions—Regulation (EU) 2024/1689—classify AI systems according to risk level.[6] "Unacceptable risk" includes practices such as social scoring and cognitive behavioral manipulation targeting vulnerable

[3] Accessible in: https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf pg.3
[4] Accessible in: https://en.wikipedia.org/wiki/Deepfake
[5] Accessible in: https://www.microsoft.com/en-us/microsoft-365-life-hacks/privacy-and-safety/deepfakes
[6] Accessible in: https://eur-lex.europa.eu/eli/reg/2024/1689/oj

groups. “High risk” encompasses AI uses in critical infrastructure, including transportation and healthcare.

The penalties imposed by the regulation include fines of up to €35 million, approximately $40 million, or 7% of a company’s global annual turnover. In short, companies must ensure the provenance of their content—or face lawsuits and even operational bans within the EU.

But what about deepfakes specifically? Does the European Union propose any targeted measures? It does. European legislation currently categorizes deepfakes as a relatively low-risk issue compared to other AI-related concerns. Spam filters and video games, for example, are classified as minimal risk and are subject to little or no regulation.

The Verdict: Preempting the Deepfake Threat

This minimization of deepfake-related risks may be explained by the relative novelty of the topic. Yet the question remains: what happens when deepfakes cause irreparable harm to an individual within a judicial proceeding? For some time now, courts have questioned the reliability of videos and images introduced as evidence.

In an effort to mitigate these risks, the state of Tennessee enacted the ELVIS (Ensuring Likeness, Voice, and Image Security) Act.[7] The statute aims to protect individuals—whether or not famous—from the unauthorized commercial use of AI-generated content. Effective as of July 2024, the Act amended Tennessee’s 1984 Personal Rights Protection Act, which previously protected a person’s name and image, but not their voice. The amendment directly addresses the impact of AI on voice and likeness.

In the age of post-truth, an era in which emotional appeal and personal belief often outweigh objective facts, discourse becomes increasingly vulnerable to manipulation. To safeguard the integrity of the judicial process against such distortions, courts have relied on the Federal Rules of Evidence to address issues of authenticity in digital media.[8] The admissibility of audiovisual evidence is often central to case outcomes. When a party seeks to introduce non-testimonial evidence, it must satisfy the relevance requirement under Federal Rule of Evidence 401 and the authenticity requirement under Rule 901.

Rule 401 provides that evidence is relevant if “(a) it has any tendency to make a fact more or less probable than it would be without the evidence; and (b) the fact is of consequence in determining the action.”[9]

---

[7] https://www.tn.gov/governor/news/2024/1/10/tennessee-first-in-the-nation-to-address-ai-impact-on-music-industry.html

[8] Accessible in: https://www.uscourts.gov/forms-rules/current-rules-practice-procedure/federal-rules-evidence

[9] Fed. R. Evid. 401

Rule 403 allows courts to exclude relevant evidence when "its probative value is substantially outweighed by a danger of unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence." Two concerns are particularly salient in this context: 1) the risk of misleading the jury and 2) the danger of unfair prejudice. Because deepfakes are designed to produce highly realistic audiovisual representations, jurors may attribute an evidentiary reliability to such material that exceeds its actual probative value. This technological capacity to simulate authentic speech, gestures, and facial expressions increases the likelihood that jurors will treat fabricated images as documentary truth. In such circumstances, courts may reasonably conclude that the persuasive force of a deepfake stems less from its probative value than from its capacity to distort the fact-finding process. In essence, Rule 403 serves as a moderating mechanism in evidentiary admission.

Regarding authenticity, Rule 901 requires that "the proponent must produce evidence sufficient to support a finding that the item is what the proponent claims it is." Rule 901(b) lists ten non-exhaustive methods of authentication. For example, authenticity or audio recordings may be established through the opinion of a witness familiar with the voice. Video evidence may likewise be admitted, including recordings from automated systems such as CCTV, provided the jury concludes they are authentic.

It is worth noting that violations of evidentiary rules may lead to judicial sanctions. The use of synthetic media such as deepfakes in evidentiary contexts may also expose individuals to significant criminal liability under federal law. Courts possess the authority to sanction misconduct affecting judicial proceedings through findings of contempt, including under 18 U.S.C § 401, which empowers federal courts to punish acts that obstruct the administration of justice.

Moreover, the deliberate creation or introduction of manipulated audiovisual evidence may implicate statutes prohibiting witness or evidence tampering, particularly 18 U.S.C § 1512. In circumstances where individuals knowingly present false statements under oath in connection with such fabricated material, liability may also arise under the federal perjury statute, 18 U.S.C § 1621.

Together, these provisions illustrate that the legal system already possesses mechanisms capable of addressing certain forms of evidentiary manipulation associated with emerging artificial intelligence technologies.

In Time: The Digital Divide and the Price of Evidence

One possible solution to the deepfake problem would be to require judicial proof, prior to the admission of any evidence or document, that it was not generated by artificial intelligence.

However, this proposal raises a new question: Is there any tool capable of determining with absolute certainty whether content was created by AI?

Several detection methods have emerged. The first is human analysis. Certain errors remain common in AI-generated images, such as hands with more than five fingers, missing ears, or the absence of shadows in bright sunlight. In videos, lip movements may fail to synchronize with speech, and vocal tones often sound flat or unnatural.

It is clear that we are living through an embryonic stage of artificial intelligence—one marked by novelty and unpredictable consequences in both the legal sphere and everyday social life. Nevertheless, some companies and startups are already developing solutions, including tools specifically designed to detect deepfakes.

One application, Reality Defender, provides real-time detection tools for images, videos, and audio, including corporate impersonations and falsifications.[10] Another startup, Facia.ai offers solutions for detecting deepfake videos and images.[11] Meanwhile, X-PHY has developed a detector capable of identifying voice and face-swap forgeries without relying on external databases.[12]

Access to such technologies, however, remains limited due to high costs, typically only affordable by governments or large corporations. Facia.ai, for instance, employs a "price-per-call" model for its API integrations. The specific cost for deepfake detection is $0.40 per verification.[13] While this may appear negligible at first glance, for a celebrity or public figure to monitor the thousands of videos and images that surface daily across social media platforms, the cumulative cost becomes prohibitive.

In contrast, companies such as Reality Defender—winner of the RSA Conference Innovation Sandbox—typically do not disclose public pricing. Instead, they operate through enterprise licensing agreements, often reaching hundreds of thousands of dollars. These contracts encompass not only tool access but also large-scale SDK/API integration, 24/7 technical support, and the significant computational processing required to analyze multimodal content, including real-time audio, video, and imagery. Ultimately, 2026 market reports indicate that comprehensive detection solutions for large organizations can exceed $250,000 annually for full implementation, covering API integration and continuous communication channel monitoring.[14]

---

[10] Accessible in: https://www.realitydefender.com

[11] Accessible in: https://facia.ai
[12] Accessible in: https://x-phy.com
[13] Accessible in: *Pricing Plans*, FACIA.AI, https://facia.ai/pricing/
[14] Accessible*, e.g.*, Intel Market RSCH., AI Deepfake Detector Market Outlook 2026-2032 (2026); see also RSA Conference, *Reality Defender Wins RSAC Innovation Sandbox 2024*, rsaconference.com, https://www.rsaconference.com/about/press-releases/reality-defender-wins-rsac-innovation-sandbox-2024

However, what remains of the judiciary's accessibility to such tools? This issue entails both procedural and economic implications. While the adoption of real-time detection technologies by courts is technically feasible, these tools must navigate the rigorous standards of evidentiary authentication. Under Federal Rule of Evidence 901, the proponent of an item of evidence must "produce evidence sufficient to support a finding that the item is what the proponent claims it is."[15] Consequently, as deepfakes become increasingly indistinguishable from authentic media, courts may arguably demand expert testimony bolstered by cutting-edge detection software to satisfy this threshold of reliability.

Nevertheless, the current prohibitive cost of these tools creates a 'digital divide' within the courtroom: only high-net-worth litigants—such as celebrities or large corporations—possess the resources to afford the forensic analysis necessary to challenge or verify digital evidence.[16] Therefore, unless the judiciary expands the use of court-appointed experts to ensure access for lower-income litigants, or the cost of technology decreases drastically, the veracity of digital discourse in legal proceedings risks becoming a privilege of the few.[17] These efforts suggest that, despite the risks, technology itself may form part of the solution—provided it is accompanied by adequate regulation, transparency, and responsible use.

Digital Twins

But what about celebrities in this story? Do they have any means of protecting themselves against the traps of deepfakes? The answer remains uncertain, but there appears to be a light at the end of the tunnel. We are witnessing a normative migration: celebrities are transcending their initial digital vulnerability to embrace a framework of 'digital personhood'—a conceptual leap that necessitates a radical redefinition of privacy and property rights in the AI era.

For decades, the entertainment industry has tied its greatest successes to the human image. The figure of the celebrity has become an autonomous commercial asset, capable of generating revenue independently of any specific artistic work. Deepfakes, however, introduced the possibility of creating entire performances without the participation of the original artist. As a result, Hollywood faces the challenge of rethinking classical notions of authorship, consent, and authenticity. More than a technological problem, deepfakes undermine the very concept of individual control over one's own image, particularly when used for commercial or artistic purposes.

Ultimately, the primary challenge facing Hollywood is not the mere existence of generative technology, but rather the increasing sophistication required in talent agreements within the

---

[15] Fed. R. Evid. 901(a)
[16] *See* Danielle K. Citron & Robert Chesney, *Deepfakes and the New Disinformation Economy*, 105 VA. L. REV. 1753, 1789 (2019)
[17] *See* Fed. R. Evid. 706; *cf.*Herbert B. Dixon Jr., *Deepfakes: More Real Than Reality*, 58 JUDGES' J. 35, 37 (2019)

entertainment industry. Deepfake technology represents a paradigm shift: the transition of the actor from a physical service provider to a licensor of biometric data.

In this light, deepfakes allow cinematic icons to maintain their luster long after aging or death. This transcends mere nostalgia for classics like Gone with the Wind or Casablanca; it ensures the continuity of multi-billion dollar franchises by facilitating the use of an actor's likeness decades after they have left the screen. This utility extends beyond legacy content to ongoing productions. For instance, it allows for the completion of works should an actor pass away during filming—as seen in the rudimentary posthumous reconstruction of Paul Walker in the Fast & Furious franchise. Furthermore, unlike traditional CGI, deepfakes offer a more seamless "de-aging" process for actors returning to a sequel years later. While conventional digital rejuvenation often appears artificial, deepfakes utilize neural networks to project a youthful visage with unprecedented realism.

Indeed, the concept of de-aging has long haunted Hollywood, manifesting in the so-called "Uncanny Valley." Coined by Japanese roboticist Masahiro Mori in his seminal work *Bukimi No Tani*, the term describes the human brain's visceral rejection of opaque gazes, rigid skin textures, and the unnatural movements often found in digital effects.[18]

Beyond aesthetic naturalism, generative AI-driven deepfakes streamline complex production processes. Consider the use of stunt doubles: a protagonist's face can now be mapped onto a double's body with such precision that it enhances on-set safety without severing the audience's emotional connection to the lead actor. Moreover, if a director identifies the need for minor dialogue alterations during post-production, deepfakes can adjust the actor's lip movements, thereby eliminating the prohibitive costs of location returns and crew rescheduling.

Regarding dialogue, the technology facilitates automated lip-synchronization. AI can adjust facial musculature and mouth movements to correspond phonetically with foreign language audio. This preserves the integrity of the original performance while providing a more immersive experience for global audiences. Consequently, the legal landscape of entertainment should view deepfakes not as "identity theft," but as a novel financial asset. However, the transformation of deepfakes into a viable financial asset is contingent upon a robust framework of consent and fair compensation—a battleground recently defined by the 2023 SAG-AFTRA strike. The resulting agreement established groundbreaking protections by categorizing AI-generated content into "Employment-Based Digital Replicas" and "Independently Created Digital Replicas."[19]

Under these terms, studios must obtain "clear and conspicuous" consent and provide specific descriptions of the intended use of a digital replica.[20] Furthermore, the contract mandates that

[18] https://en.wikipedia.org/wiki/Uncanny_valley
[19] Acessible in: 2023 SAG-AFTRA TV/Theatrical Summary of Tentative Agreement, SAG-AFTRA, https://www.sagaftra.org/files/sa_documents/TV-Theatrical_2023_Summary_Agreement_Final.pdf
[20] *Id.* at 3

performers be compensated for the digital use of their likeness at rates equivalent to those they would have earned for the actual performance.[21] This collective bargaining success demonstrates that while technology facilitates a new economy of digital royalties, the legal barrier to its implementation remains the preservation of human agency and the mandatory compensation for the 'digital labor' performed by an actor's AI twin.

So, actors may now license their digital likeness, allowing studios to generate content while the performer is simultaneously engaged in other projects, thus birthing a new economy of digital royalties and intellectual property rights.

The financial implications of this technological shift are substantial. Modern so-called 'tentpole' film production budgets frequently exceed $200 million, with a significant percentage—often up to 20% to 30%—allocated to visual effects, international dubbing, and unforeseen reshoots due to scheduling conflicts.[22]

Deepfake-assisted production mitigates these logistical burdens by decoupling the performer's physical presence from the filming schedule. While high-end AI detection and synthesis tools can cost upwards of $250,000 for enterprise licensing, this capital expenditure is marginal compared to the 'burn rate' of a major production, which can reach $250,000 to $500,000 per day during active filming.[23] By utilizing AI-assisted likeness modeling to resolve scheduling constraints or health-related absences, studios can avoid catastrophic delays that often cost millions in unrecouped overhead, effectively transitioning production from a labor-intensive model to a scalable, software-driven asset.[24] In other words, the financial implications are substantial. Film production budgets often exceed $200 million, with large portions allocated to reshoots, scheduling conflicts, international dubbing, and visual effects. Deepfake-assisted production mitigates the need for costly logistics. In cases where actors are unavailable due to health or scheduling constraints, AI-assisted likeness modeling prevents catastrophic delays.

Beyond economic efficiency, deepfake technology has the potential to democratize entertainment access. Prohibitive production costs traditionally gatekeep the industry, favoring major studios. By lowering these financial barriers, independent filmmakers may gain access to tools previously reserved for the elite, diversifying storytelling and amplifying underrepresented voices. Deepfakes represent a turning point in Hollywood's evolution; when utilized ethically and regulated contractually, they expand creative horizons and democratize participation in the industry.

---

[21] *Id.* at 3
[22] Accessible in*:* Stephen Follows, *How Much Does a Blockbuster Movie Cost to Make?*, Sthephen Follows Film Data RSCH. (July 10, 2023), https://stephenfollows.com/how-much-does-a-blockbuster-movie-cost-to-make/
[23] *See* Brooks Barnes, *The Age of the $400 Million Movie*, N.Y. Times(May 22, 2024), https://www.nytimes.com/2024/05/22/business/media/blockbuster-movie-budgets.html
[24] *Cf.* Deloitte, *The Future of Content Creation: Generative AI in Media and Entertainment* 8-12 (2024)

Nevertheless, many performers remain skeptical of these purported benefits. In a recent dialogue on Episode #2440 of *The Joe Rogan Experience* podcast, Ben Affleck and Matt Damon explored the dichotomy between AI as a tool for efficiency versus a creative substitute. Affleck posited that while deepfakes may reduce logistical overhead, the technology lacks the capacity to mimic the human specificity that defines protectable intellectual property. For Damon, legal protection must remain anchored in the preservation of performance as a unique biometric and emotional act. This reinforces the thesis that deepfakes must be regulated through the lens of asset licensing rather than labor displacement.

In this context, Hollywood benefits from a legal doctrine that many jurisdictions around the world lack: the right of publicity. Over the course of the twentieth century, U.S. law developed the doctrine of the right of publicity, which protects an individual's economic interest in the commercial use of their identity. Unlike privacy rights, the right of publicity is not grounded in emotional distress, but in the unauthorized appropriation of economic value.

The U.S. Supreme Court expressly recognized this interest in *Zacchini v. Scripps-Howard Broadcasting Co*.[25] Hugo Zacchini was a human cannonball whose entire performance lasted approximately fifteen seconds. Despite his objections, a television station filmed and broadcast the full act during a news segment. The broadcaster argued that the First Amendment protected its conduct as news reporting. The Supreme Court disagreed, holding that broadcasting the entire performance eliminated the economic incentive for audiences to pay to see it live—an act described as the appropriation of the performer's entire economic value.

The logic of *Zacchini* has since been applied to contemporary disputes involving generative AI and deepfakes in Hollywood for three principal reasons. First, the Court recognized that a performer's act constitutes property. The right of publicity functions similarly to copyright or patent law by protecting the labor and investment required to create a professional identity. When AI uses an actor's likeness to perform in a film or advertisement, it appropriates the economic value of that performance in precisely the same way as the broadcast in *Zacchini*.

Second, *Zacchini* delineates the limits of First Amendment protection. Many deepfake creators argue that their works are protected as art, parody, or information. The *Zacchini* case established that free speech is not an absolute shield when the use substitutes the commercial value of the original performer. If a deepfake eliminates the need to hire the real actor, it violates the right of publicity.

Third, *Zacchini* implicitly protects digital identity. In Hollywood today, there is growing concern that studios may create or resurrect virtual actors based on real individuals. If broadcasting fifteen seconds of a performance exhausts an artist's economic value, using a deepfake to create an entire film exhausts that value even more profoundly.

---

[25] *Zacchini v. Scripps-Howard Broad. Co.*, 433 U.S. 562 (1977).

Thus, *Zacchini v. Scripps-Howard Broadcasting Co.* remains highly relevant in the deepfake era. When someone appropriates the spectacle—or even the persona—of an artist such that the public no longer needs the real individual, the violation is commercial in nature, not merely a matter of privacy.

AI may infringe not only commercial rights but also broader identity interests. The Ninth Circuit has held that evoking a celebrity's identity, even without literal reproduction, may violate the right of publicity. In *White v. Samsung Electronics America, Inc.*, Samsung produced a futuristic advertisement featuring a robot wearing a blonde wig, evening gown, and jewelry, standing beside a game-board reminiscent of *Wheel of Fortune*.[26] The show is, of course, hosted by Vanna White, who has blonde hair and wears long gowns with sparkly jewelry each night.

In *White v. Samsung Electronics America, Inc*., the plaintiff asserted four distinct causes of action: California Civil Code § 3344, regarding the unauthorized use of another's 'name, voice, signature, photograph, or likeness'; the common law right of publicity; the Lanham Act, 15 U.S.C. § 1125(a), for false endorsement; and intentional interference with a prospective economic advantage.

While the district court initially granted summary judgment for the defendants, the Ninth Circuit reversed the common law right of publicity and Lanham Act claims. Critically, the court held that the common law right of publicity is not limited to the specific categories listed in the California statute; rather, it protects a celebrity's entire identity against misappropriation, even when that identity is evoked through a symbolic representation—in this case, a robot in a blonde wig and evening gown.

So, although Samsung never used White's name, image, or voice, the court ruled in her favor, holding that the right of publicity protects identity itself. The court famously concluded that the "meanings" of identity are not limited to name or face, but extend to any "sufficiently specific" combination of attributes that lead the viewer to believe the celebrity is being depicted. As the court noted, the defendant's intent was to profit from the "value" of White's fame without paying for it, thereby constituting a misappropriation of her common law right of publicity.[27]

This reasoning may prove decisive in future deepfake cases. AI systems often generate likenesses without copying a specific photograph, instead reproducing style, facial features, and mannerisms. The defense frequently claims that no copyrighted image was used and that the AI created something "new." White rejects that argument: if the result unmistakably evokes a celebrity's identity to sell a product, it is unlawful.

---

[26] *White v. Samsung Elecs. Am., Inc.*, 971 F.2d 1512 (9th Cir. 1992).
[27] *Id.* at 1517.

The White decision, however, is not without controversy. In dissent, Judge Alex Kozinski warned that the ruling risked "stifling creativity," highlighting the tension between innovation and identity protection.

The Sound of an Intention

A recent example implicates the *White* holding. Actress Scarlett Johansson confronted OpenAI regarding the "Sky" voice of ChatGPT. According to Johansson in September 2023, OpenAI CEO Sam Altman invited her to voice the system, believing her voice could bridge the gap between humans and AI.[28] She declined. In May 2024, OpenAI released GPT-4 with a voice named "Sky," which users and Johansson herself found "eerily similar" to her own. Compounding the issue, Altman posted the word "her" on social media on launch day—a reference widely interpreted as invoking the 2013 film *Her*, in which Johansson voices an AI.[29]

Johansson's attorneys demanded information regarding the creation of the voice. OpenAI responded that it used a professional voice actor, not Johansson's voice. Nevertheless, under *White v. Samsung*, the deliberate evocation of Johansson's cinematic AI persona could constitute a violation of her right of publicity, even absent literal copying.

Johansson might also rely on another landmark case: *Midler v. Ford Motor Co*.[30] In that case, Ford sought to use Bette Midler's rendition of her song "Do You Want to Dance" in a commercial. After Midler refused, the advertiser hired a former backup singer and instructed her to imitate Midler's voice as closely as possible. The Ninth Circuit reversed the trial court and held that a distinctive voice is as personal as a face, noting that the voices at issue were almost indistinguishable and therefore causing actual consumer confusion. It follows that when someone is deliberately imitated to sell a product, it constitutes unlawful appropriation. The principles articulated in *Midler* are directly applicable to AI voice cloning, such as Johansson's case.

Another crucial doctrine is the Transformative Use Test, developed by the California Supreme Court to balance the right of publicity against the First Amendment. The test emerged in *Comedy III Productions, Inc. v. Gary Saderup, Inc.* (2001), where the court held that charcoal drawings of The Three Stooges were insufficiently transformative and therefore infringed publicity rights.[31] The test asks whether a work adds new expression, meaning, or message, or merely exploits a celebrity's fame. If the use is non-transformative, the right of publicity prevails. If it is transformative, free speech prevails.

---

[28] https://www.nbcnews.com/tech/tech-news/scarlett-johansson-shocked-angered-openai-voice-rcna153180
[29] Sam Altman (@sam_a), X (May 13, 2024, 10:11 AM), https://x.com/sam_a/status/1790080313466352031
[30] https://law.justia.com/cases/federal/appellate-courts/F2/849/460/37485/
[31] https://law.justia.com/cases/california/supreme-court/4th/25/387.html

California courts have applied the test with varying results. In *Kirby v. Sega,* a video game character was deemed to be transformative despite similarities to a real singer.[32] Conversely, in *No Doubt v. Activision*, the use of realistic band avatars performing their actual songs was held non-transformative.[33] Continuing application of the transformative use test allows courts to exclude most commercial deepfakes from First Amendment protection without creating new legal categories.

Deepfakes in computer-generated imagery also find important precedents in U.S. law. One such example is *Downing v. Abercrombie & Fitch* (2001).[34] Abercrombie & Fitch had published a clothing catalog with a surfing theme. To illustrate the catalog, the company purchased old photographs of surfers taken during a 1965 competition, thereby acquiring the copyright to those images. However, it did not obtain permission from the surfers depicted in the photographs for commercial use or advertising.

In court, Abercrombie argued that ownership of the copyright entitled it to use the images as it saw fit. The court disagreed and established two crucial principles. First, copyright law is distinct from the right of publicity. Accordingly, ownership of a digital file or its copyright does not confer the right to commercially exploit the likeness of the individual depicted without consent. Abercrombie also argued that the photographs were informational or historical in nature, portraying the culture of surfing. The court rejected this claim, finding that the use constituted disguised commercial exploitation, as the ultimate purpose was to sell clothing. Although the case predates the expansion of generative AI technologies, it remains frequently cited in Hollywood to adjudicate disputes involving deepfakes and the unauthorized use of digital images.

This form of image use is particularly controversial because many companies claim that they train artificial intelligence systems using images for which they hold licenses. *Downing v. Abercrombie* makes clear that this distinction is legally irrelevant. Even if an AI system holds copyright over the pixels it generates, if the final output evokes or replicates a real person for commercial purposes, that individual's right of publicity has been violated.

The *Downing* case is also frequently invoked in the fashion industry. Many brands have expressed interest in, or have already begun using, AI-generated models that closely resemble famous human models in order to save millions of dollars in appearance fees.

The standard for liability under *Downing* and its progeny does not require that a digital replica be 'confusingly similar' in the trademark sense; rather, the threshold is whether the celebrity is 'readily identifiable' from the content. Unlike the Lanham Act, which focuses on consumer

[32] https://law.justia.com/cases/california/court-of-appeal/2006/b183820.html
[33] https://law.justia.com/cases/california/court-of-appeal/2011/b223996/
[34] https://www.uniset.ca/other/cs2/265F3d994.html

confusion regarding endorsement, the California right of publicity serves to protect the individual's proprietary interest in their own persona. In *Downing*, the court clarified that even if a defendant uses an authentic photograph or a likeness that accurately depicts the subject, the lack of consent for commercial exploitation triggers liability if the public can perceive who is being depicted.

Thus, in the context of AI-generated models, a brand cannot escape liability by arguing that a 'digital twin' is merely 'inspired by' a celebrity if the specific attributes—such as facial geometry, distinctive style, and/or pose—allow a reasonable viewer to identify the specific public figure being imitated.The decision in *Downing v. Abercrombie* effectively forecloses this practice: if an AI system creates a model that the public identifies as confusingly similar a specific celebrity for the purpose of selling clothing, the brand may be held liable in the same manner as Abercrombie was.

This principle applies equally when social media is the storefront. Instagram, TikTok, X, and YouTube are saturated with deepfakes depicting celebrities endorsing products. Brands often attempt to characterize such content as parody or entertainment. Courts, however, routinely reject this defense, relying on the logic articulated in *Downing*: when the ultimate context is to induce a sale or generate profit for a platform, First Amendment protection gives way, and the right of publicity prevails.

This potential liability extends beyond individual accounts to the platforms that host and monetize such content. While Section 230 of the Communications Decency Act generally immunizes interactive computer services from liability for third-party content, this 'shield' explicitly excludes intellectual property claims.[35] Because the Ninth Circuit frequently characterizes the right of publicity as a form of intellectual property, social media giants—such as Instagram, TikTok, and X—may find themselves stripped of their statutory immunity when hosting deepfakes that misappropriate a celebrity's persona for commercial gain.[36]

Furthermore, as platforms shift from passive hosting to active content curation via AI-driven algorithms that prioritize high-engagement deepfake advertisements, they risk being viewed as co-developers of the infringing content rather than mere conduits. Consequently, under a contributory infringement framework, platforms that fail to implement robust notice-and-takedown procedures for deepfakes could face substantial exposure, as the commercial nature of these depictions often places them outside the protective ambit of the First Amendment.[37]

---

[35] 47 U.S.C. § 230(e)(2)

[36] *See Enigma Software Grp. USA, LLC v. Malwarebytes, Inc.*, 946 F.3d 1040, 1053 (9th Cir. 2019); *cf. Perfect 10, Inc. v. Visa Int'l Serv. Ass'n*, 494 F.3d 788, 794 (9th Cir. 2007)

[37] *See* Danielle K. Citron & Robert Chesney, *Deepfakes and the New Disinformation Economy*, 105 VA. L. REV. 1753, 1803 (2019)

At present, state courts appear to possess sufficient tools to address deepfake risks. However, as technology evolves, federal regulation may become necessary. Recognizing this need, the U.S. Congress has begun considering the NO FAKES (Nurture Originals, Foster Art, and Keep Entertainment Safe) Act—the first proposed federal law aimed at protecting individuals against AI-generated digital replicas.[38]

The bipartisan bill was introduced in the Senate in July 2024 and reintroduced in April 2025, with an identical version introduced in the House of Representatives in September 2024. It has garnered support from SAG-AFTRA, the Recording Industry Association of America (RIAA), and major technology and entertainment companies, including IBM and Disney.

The proposed legislation would create a new federal intellectual property right over voice and likeness, establishing that individuals own their digital identity. It prohibits the creation or distribution of highly realistic, readily identifiable digital replicas without consent and allows actions against creators, distributors, and platforms that fail to remove unauthorized content after notice. Notably, the right survives death for up to seventy years post-mortem, preventing the resurrection of deceased artists without family authorization. The Act also addresses deepfake pornography and requires explicit, informed, written consent for the creation of digital doubles. To avoid constitutional infirmity, the bill includes First Amendment carve-outs for news reporting, documentaries, historical works, satire, parody, and criticism.[39]

A federal statute would bring much-needed clarity. Currently, victims must navigate inconsistent state publicity laws, while technology companies lack clear regulatory guidance. The NO FAKES Act promises to unify standards and ensure that AI remains a tool for human creativity rather than a substitute for it.

Children of Men: Safeguarding the Human Legacy

Deepfakes pose a structural challenge to image rights and public trust in media authenticity. Yet they do not require a wholesale reinvention of legal doctrine. Existing principles—right of publicity, First Amendment limits, and evidentiary rules—already provide a robust foundation. What remains urgently necessary is a clear federal framework capable of defining rights and responsibilities in the AI era.

It is clear that the rights we now consider fundamental were the product of historical struggles and cannot be erased by algorithms. For this reason, the world is currently experiencing a form of regulatory race aimed at ensuring that technology serves humanity, rather than the opposite.

---

[38] https://www.congress.gov/bill/119th-congress/senate-bill/1367/text

[39] Nurture Originals, Foster Art, and Keep Entertainment Safe (NO FAKES) Act of 2024, S. 4875, 118th Cong. § 2(a) (2024)

Other major nations already have laws in force or regulatory proposals at advanced stages addressing this issue, even if still in an embryonic form. For one, the European Union definitively approved the AI Act in 2024 and it is subject to gradual implementation through 2026.[40] Among existing global regulatory frameworks, the Act is arguably the most comprehensive to date.

A brief overview of European legislation is necessary to illustrate its parameters and to compare it with those of other countries. In Europe, transparency has become mandatory. Any deepfake or AI-generated content that appears realistic must be clearly labeled as such. The rationale is that users must immediately know that what they see, hear, or perceive is artificial rather than human. European regulation also establishes categories of unacceptable risk, prohibiting AI systems that manipulate human behavior or employ any form of social scoring.

However, the European framework differs significantly from the American approach in its primary focus. While proposed U.S. legislation centers on the property interests of the artist, the European Union frames its regulation around the right of citizens not to be deceived. This shift in perspective constitutes the fundamental distinction between the two regulatory models.

France has likewise enacted specific legislation targeting digital fraud. Law No. 2024-449 of May 21, 2024**,** imposes criminal sanctions for such conduct and introduces prison sentences and fines of up to €60,000 (or about $70,000) for those who publish sexually explicit deepfakes or content intended to harm an individual's reputation without a clear disclosure that it was generated by artificial intelligence.[41] French law also places particular emphasis on the protection of minors, seeking to prevent the use of children's images in AI training models.

China has also enacted a specific legal framework addressing deepfakes through regulations designed to govern deep synthesis activities in internet information services. The regulation, entitled the Internet Information Service Deep Synthesis Management Provisions (commonly referred to as the Deep Synthesis Regulation) , strictly prohibits the use of AI to edit a person's face or voice without that individual's consent.[42] In addition, all deep synthesis services—that is, generative AI platforms—are required to embed invisible watermarks in their outputs to ensure traceability by authorities. The regulation further imposes liability on platforms that fail to act promptly, allowing the circulation of unlawful deepfake content.

Brazil has also been working toward the adoption of federal legislation on the subject, aiming to establish a Legal Framework for Artificial Intelligence**.** This initiative has been led by the Brazilian Federal Senate. Bill No. 2,338/2023 seeks to guarantee individuals the right to know when they are interacting with artificial intelligence and the right to challenge decisions made by algorithms.[43] The proposed legislation also aims to amend Brazil's Copyright Law by providing

[40] https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:32024R1689
[41] https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049563368
[42] https://digichina.stanford.edu/work/translation-internet-information-service-deep-synthesis-management-provisions-draft-for-comment-jan-2022/
[43] https://www25.senado.leg.br/web/atividade/materias/-/materia/157233

specific protection against AI-generated voices and images, closely aligning with the underlying principles of the NO FAKES Act**.**

Given the growing prevalence of these issues, we cannot deny that the evolution of humanity is intertwined with the evolution of Artificial Intelligence. The advancement of AI is not an isolated process; it feeds upon accumulated human labor, image, and creativity. Protecting these rights is not about opposing progress, but rather ensuring that progress rests upon an ethical foundation. Existing collective and individual rights must be secured to preserve the foundational liberties of Western society. These protections synthesize centuries of struggle for autonomy and privacy.

Artificial Intelligence is no longer a promise of the future. Many experts herald AI as the force structuring the architecture of our present and future. However, the fact that a technology is technically capable of replicating the human essence does not grant it the moral or legal right to do so. If the 20th century was marked by the struggle to decouple human identity from slavery and unrestrained commercial exploitation, the 21st century demands that the judiciary be more than a passive observer of innovation—it must be the ethical architect of what technology proposes. To ensure that technological evolution does not lead to the erosion of human dignity, our laws—such as the NO FAKES Act and the AI Legal Framework—must remain vigilant to one fact: Will we be a civilization that allows the machine to live at the expense of man's stolen identity, thereby renouncing our own freedom?

Therefore, the challenge is not to stop AI, but to empower ourselves to use it within legal frameworks. The role of legal scholars in this new context is to ensure that programming codes are never stronger than the rule of law.