Mimetic AI Systems: Understanding and Regulating the Use of Generative Models for Impersonation

Norman Bukingolts
Department of Engineering Education
University of Florida
nbukingolts@ufl.edu

Abstract

Generative artificial intelligence models are being used to imitate the words, voices, bodies, and artistic styles of private and public figures with unprecedentedly high accuracy and scalability. Despite their usage offering cost efficiencies over employing the human counterpart, associated drawbacks – scams and fraud, baseless social death or defamation, and an erosion of trust in online information environments – are growing and reaching criticality. *Mimetic AI systems* use generative models which leverage knowledge extracted from data provided during training or inference time to capture and reproduce the actions, decisions, and preferences of specific individuals in novel contexts. In this paper, I explain how such systems power the creation and distribution pipelines of deepfakes, digital doubles, voice clones, and other impersonations. I then conduct a normative ethics assessment of these systems and discuss their benefits and risks to key stakeholders: system operators, targets, and audiences, as well as their creators, intermediaries, and regulators. Finally, I propose several regulatory solutions and outline their possible implementation challenges to support initiatives in AI governance aimed at addressing the multifaceted obstacles which mimetic AI systems pose to the integrity, value, and endurance of authentic human expression.

Introduction

The ability to digitally impersonate oneself and others carries clear and profound ethical implications for personal identity, information integrity, and questions of authenticity, ownership, and consent (McIlroy-Young et al. 2022a; Paris and Donovan 2019). Indeed, the misuse of generative artificial intelligence (GenAI) to efficiently create deepfakes which misinform, harm, or otherwise deceive people has been identified, discussed, and evaluated at length in recent literature (Fallis 2021; Hancock and Bailenson 2021; de Rancourt-Raymond and Smaili 2023; Vatreš 2021; U.S. Department of Homeland Security 2021).

The ”symptoms” of the ”illness” and a common ”cause” (i.e. the harmful outcomes of unsanctioned impersonations enabled by GenAI systems) are known – and ”treatments” (i.e. AI governance solutions) are being tested. However, the ”prognosis” (i.e. the success of GenAI regulation) stands to benefit from the revival of research which examines, clarifies, and contextualizes the mime behind the mimicry: the GenAI models and systems behind the deepfakes.

McIlroy-Young et al. introduced the *mimetic model* in 2022 as:

> ”...an algorithm that is trained on data from a specific individual in a given domain, and which is designed to accurately predict and simulate the behavior of this individual in new situations from the domain” which ”Crucially. . . is generative in the sense that it does not exist simply to predict a specific person’s behavior, but to produce this behavior in context” (McIlroy-Young et al. 2022a).

Three years later, GenAI systems are capable of inexpensively and efficiently counterfeiting human expression across text, audio, image, and video modalities with remarkable precision (Landymore 2025; Scott 2024; Newsroom 2025; Thompson 2024; Wiggers 2025a). Correspondingly, a rapidly expanding market – projected to be worth billions of U.S. dollars by 2030 – exists for deepfake technologies which replicate individual styles and characteristics (Grand View Research 2024a,b).

Notwithstanding, potential losses exceed expected gains: estimates from January 2025 project the cumulative cost of fraud from deepfakes to reach $40 billion USD by 2027 (Jennings-Trace 2025). Major governing bodies have codified provisions to mitigate the risks of deepfakes into law, emphasizing transparency measures such as explicit disclosure of content origin, such as in Article 50 of the European Union’s AI Act (The European Parliament and the Council of the European Union 2024), and requiring consent from targets of voice cloning, such as in Tennessee’s ELVIS Act (Madarang 2024).

These new developments call for renewed discussion. I begin by posing and examining the following questions:

- Which GenAI models and systems are used to impersonate individuals, and in what capacity?
- Who are the users, targets, and other stakeholders of the GenAI models and systems used for impersonation?
- What innovations enabled certain GenAI models to develop the ability to impersonate people?
- How are the generative models which are capable of impersonation integrated into GenAI systems?

Establishing Relevant Terms

McIlroy-Young et al. outlined several definitions concerning the machine learning models used for imitating human behavior and their stakeholders (McIlroy-Young et al. 2022a). For this discussion, I choose to redefine and introduce several terms – beginning with the technology itself.

- A *mimetic model* is a machine learning algorithm which uses analytical and generative capabilities to produce an accurate simulation of the behavior of an individual or their proxy in a novel context within a particular modality by drawing on training knowledge, runtime adaptation to examples, or both.
- A *mimetic AI system* is a software platform (e.g., web application) which deploys one or more mimetic models across one or more modalities, allowing its users to interact with the model(s) through prompts and retrieve their output for downstream distribution.

Three roles are involved in the usage of mimetic AI systems: the *operator* who uses the system and distributes outputs, the *target* whose characteristics are being imitated, and the *audience* who perceives the outputs. Figure 1 shows a directed graph of the relationships between each role, outlining a simple framework for visualizing the possible routes to impersonation.

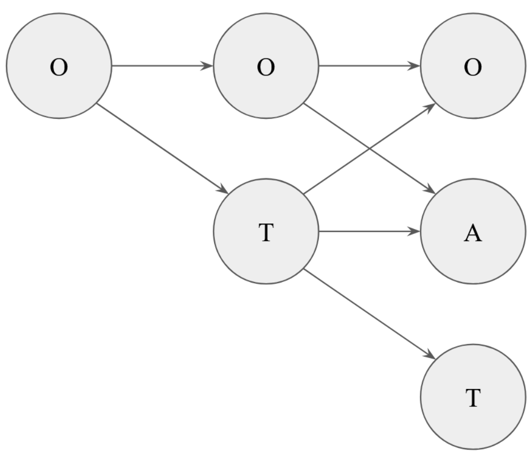

Figure 1: A graph demonstrating the dynamics of impersonation and intentionality between operator (O), target (T), and audience (A), answering the question: "who impersonates whom, and for whom to witness?" Possible routes include: operator impersonates operator for operator (O-O-O), operator impersonates operator for audience (O-O-A), operator impersonates target for operator (O-T-O), operator impersonates target for target (O-T-T), and operator impersonates target for audience (O-T-A).

The operator can be anyone with access to a mimetic AI system and a distribution network of any scale, from an individual experimenter sharing outputs with friends via direct message to a business distributing outputs at-scale to many customers on multiple social media platforms. In contrast,

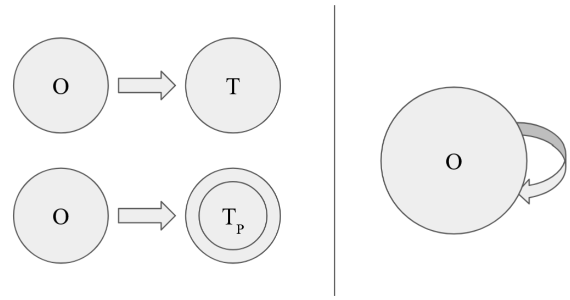

Figure 2: Private and public figures (top left, T) as well as proxy figures (bottom left, T_P) are targets of interpersonal impersonation, while operators target themselves in intrapersonal impersonation (right, O).

the target must be a nameable entity with specific characteristics that can be captured and reproduced. As illustrated in Figure 2, mimetic AI systems support both *intrapersonal impersonation*, where an operator targets themselves, and *interpersonal impersonation*, where an operator targets another individual or entity. Three potential targets exist, all ultimately corresponding to one or more nameable persons:

- A *public figure* is an individual who has deliberately placed themselves in the public eye and gained fame or prominence in societal affairs (Legal Information Institute, Cornell University 2020), such as a celebrity, politician, or artist, whose style and imitable characteristics are publicly known and accessible.
- A *private figure* is an ordinary individual who has not explicitly sought public attention (Lubin Austermuehle DiTommaso, PC 2025), whose style and imitable characteristics are privately known yet remain accessible via descriptions and samples of writing, voice, appearance, or behavior.
- A *proxy figure* is a singular or collective unit which represents the imitable characteristics of one or more individuals. Existing at a layer of abstraction above a legal full name, this non-fictional or fictional entity aggregates the style(s) of one or more public or private figures who contributed to its creation, as exemplified in Figure 3.

Whereas there is always one human target in the impersonation of a private or public figure, the imitation of a proxy figure can constitute multi-target impersonation. For example, constructing a unimodal output – such as a voice clone of a character from a television show – is an impersonation by proxy of one person: the voice actor. However, constructing a multimodal output – such as imitating a character's appearance and words in conjunction with its voice – is an impersonation by proxy of multiple people: the writer(s), artist(s), animator(s), and others. As such, impersonations of private and public figures may be derived from imitations of proxies and multiply across modalities.

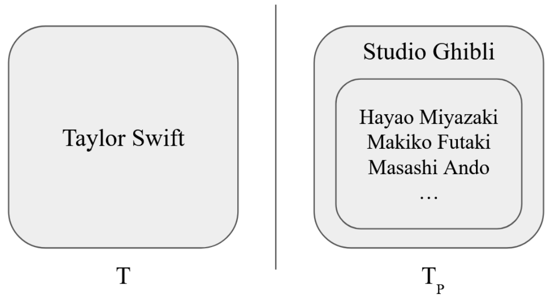

Figure 3: While Taylor Swift is a public figure target (T), Studio Ghibli is a proxy figure target (T_P) because its artistic style represents the collective contributions of its human stakeholders.

The audience encompasses anyone who observes the system's output, which necessarily includes the operator but often may extend to the target and external individuals or groups (e.g., the general public), depending on the output's distribution. Referencing the graph in Figure 1, I specify:

- *Contained interpersonal impersonation* (i.e. O-T-O) occurs when the operator impersonates others for themself to witness, whereas *uncontained interpersonal imitation* (i.e. O-T-T and O-T-A) occurs when the operator impersonates others for others to witness, possibly including but not limited to the target.
- *Contained* and *uncontained intrapersonal impersonation* (i.e. O-O-O and O-O-A, respectively) – or *delegation to digital self* – occurs when the operator impersonates themself for themself or others to witness, respectively.

Three stakeholders design mechanisms and rules which influence the usage, distribution, and risk mitigation of mimetic models, mimetic AI systems, and their outputs:

- The *creator* develops the model(s) or system(s), or both.
- The *intermediary* creates the content-sharing platform(s) on which model outputs are shared.
- The *regulator* establishes enforceable rules, often but not necessarily legal frameworks, which govern models, systems, and content-sharing platforms.

Stakeholders are not necessarily distinct entities. For example, one company can be a creator and an intermediary if it owns or develops a mimetic AI system and a social media network which integrates that system or allows its (or others') outputs. Creators and intermediaries may take on internal regulatory responsibilities by setting and self-applying AI safety measures in the absence of requirements from external regulators (e.g., governmental bodies).

Technical History and Status Quo

The origins of mimetic models can be contextualized by examining recent developments where deep learning architectures were used to solve tasks in computer vision and natural language processing.

Image classification was famously accomplished with high accuracy in 2012 using a convolutional neural network (CNN) named ImageNet (Krizhevsky, Sutskever, and Hinton 2012). In the same year and shortly after, speech recognition was advanced by deep feed-forward neural networks (Hinton et al. 2012; Deng, Hinton, and Kingsbury 2013). Named-entity recognition (NER) in unstructured text became a primary use case for bidirectional encoder representations from transformers after their introduction in 2018 (Devlin et al. 2018).

Image synthesis – particularly style transfer, such as changing photographs into impressionist paintings – was demonstrated with CNNs in 2016 (Gatys, Ecker, and Bethge 2016) and further documented in 2018 (Jing et al. 2018), evolving in 2020 to employ generative adversarial networks, notably the CycleGAN (Zhu et al. 2020), and diffusion probabilistic models (Ho, Jain, and Abbeel 2020). Speech synthesis followed a similar timeline, being proven feasible with deep neural networks in 2016 with WaveNet (van den Oord et al. 2016) and advanced in 2018 (Shen et al. 2018). NER, among other tasks (e.g., sentiment analysis, question answering, etc.), was proven extensible to language models by using generative pre-trained transformers the following year with GPT-2 (Radford et al. 2019).

Further innovations multiplied analytical capabilities and enabled knowledge gained from multiple modalities to be shared in a unified architecture, augmenting generative model efficiency and output quality. By 2020, text-to-speech models were able to clone voices from just fifteen seconds of audio; the model deployed on the *15.ai* website notably interwove the DeepMoji sentiment analysis model (Felbo et al. 2017) to inject emotional context into synthesized speech (Temitope 2024). Contrastive language-image pre-training was introduced in 2021 (Radford et al. 2021) and enabled textual semantics to influence the visual style of outputs in latent text-to-image diffusion models, such as Stable Diffusion (Rombach et al. 2022; Ramesh et al. 2022).

Similar developments for audio and video followed; contrastive language-audio pre-training was introduced (Elizalde et al. 2022) alongside innovations in text-to-video models (Hong et al. 2022). Since then, advancement has continued in understanding and guiding style transfer across modalities (Wang et al. 2023; Demerlé et al. 2024; Li et al. 2025), demonstrating a major trend towards multimodal architectures in the state of the art for mimetic model development. For example, OpenAI introduced the multimodal GPT-4o model in 2024 as:

> ”...a single new model [trained] end-to-end across text, vision, and audio, meaning that all inputs and outputs are processed by the same neural network” (OpenAI 2024a), whose 2025 implementation of ”image generation excels at...precisely following prompts, and leveraging 4o's inherent knowledge base and chat context – including transforming uploaded images or using them as visual inspiration” (OpenAI 2025b).

Explanation and Integration

The ability of mimetic models to recognize and impersonate targets arises from a combination of training (i.e. pre-training or fine-tuning) knowledge and runtime (i.e. inference time) adaptation to reference materials. Purpose-built mimetic models exist, such as chess player models trained from scratch (McIlroy-Young et al. 2022b), and effective approaches have been demonstrated for fine-tuning large language models (LLMs) for impersonation (Shipper 2023; Ressmeyer, Masling, and Liao 2019). Indeed, as in many cases for machine learning systems, representation volume in the training dataset directly impacts learning quality (Appian Corporation 2025), resulting in more accurate recognition and impersonation of frequently encountered human subjects (Center for News, Technology & Innovation 2024).

However, mimetic models can also be prompted to impersonate individuals entirely absent from training data. As training datasets grow more comprehensive, models learn to interpret and reproduce fundamental patterns of human expression – such as physiological characteristics in vision models or grammar and narrative voice in language models (Riemer and Peter 2024) – through emergent abilities based on latent representations, regardless of explicit design intention (Wei et al. 2022; Bereczki and Ádám Liber 2023). Domain generalization can allow understanding aspects of a provided sample well enough to be able to generate a stylistically similar output in an unseen context (Thomas and Ghadiyaram 2025).

Mimetic AI systems, as implementations of mimetic models, function through structured user interaction patterns which highlight particular input-output transformations that preserve or translate stylistic elements across contexts. Common examples include standalone applications or integrated services with user interfaces ranging from simple text prompts to sophisticated fields accepting multimodal inputs, such as images or other media, for the mimetic model to reference (Franzen 2024; Google Cloud 2025c).

The core interaction typically involves system operators specifying a target and desired content parameters, depending on the modalities in which the system functions. For example, operators can identify targets by name through natural language, such as a prompt to "write in Hemingway's style", or by providing reference samples, such as uploading voice recordings for cloning (OthersideAI 2025; ElevenLabs Inc. 2025a). Many platforms offer adjustable parameters which control attributes of generation, such as stylistic adherence or "creativity levels" via model temperature (Adobe Inc. 2025; Google Cloud 2025b).

The type of a mimetic model or mimetic AI system is expressed by the relationship between its input and output modalities (i.e. "x-to-y", where x is input modality and y is output modality). Notably, type designations are not mutually exclusive, meaning multiple types can be assigned simultaneously. Table 1 outlines ten system types, separating what would be the audio modality into speech and music to highlight the distinctness of their usage and implementation.

Outputs are displayed to operators within the platform such that they may be extracted from the immediate environment and distributed to an outside audience (e.g., copying-and-pasting on-device; downloading and sharing across devices), or even reused as input for further generation before distribution. Crucially, outputs can include content generated by the mimetic model as well as the model itself. For example, web application-based systems exist which allow users to not just generate and download synthetic speech but also to create text-to-speech models from sample files and share them (101soundboards.com 2025).

The Normative Ethics of Mimetic AI Systems

Normative ethical frameworks present reliable materials for constructing a stance on mimetic models and mimetic AI systems. Ethical considerations of their benefits and risks can inform both private moral judgements and public policy.

I present multiple frameworks herein to avoid the pitfalls of unilaterally discussing and defining the magnitude of impacts on stakeholders, which vary widely in scope and severity. For example, some outcomes may be proportional to a numerical factor, such as the number of targets or the size of the audience. Others shrink or grow based on system type (e.g., image-to-image vs. speech-to-speech) or target type (e.g., private vs. public figures). No one framework is most correct; interpreting outcomes with multiple ethical theories accounts for their variety by offering a consistent, enduring, and interoperable set of evaluative lenses.

Deontology

The deontological (i.e. rule-based) approach combines the Kantian principle of moral duty to respect other people, *observantia*, with a juridical duty to respect laws (Campos 2019; Dillon 2022). For example, the assertion that mimetic AI systems should be avoided because using them is illegal presents a seemingly straightforward ethical position. However, its application remains problematic in the presence of a *non liquet* (i.e. a gap in the law) which carries unclear implications on respect for the systems' stakeholders.

For example, recent reports from the U.S. Copyright Office concerning GenAI content argue for the protection of human likeness (U.S. Copyright Office 2023, 2024, 2025a,b). While sentiments promoting such protection may be interpreted as respectful, there are still inadequate guardrails to prevent impersonation via mimetic AI system as evidenced by the systems' unregulated proliferation under current laws constraining GenAI usage. Recent lawsuits against AI companies have been initiated by targets of impersonation who claim that training on their data – especially on copyrighted content – without their consent is disrespectful (Knibbs 2024; Gilbertson and Reisner 2024).

Contemporary legal uncertainty creates practical challenges for all potential targets seeking protection by enabling a permissive environment where the counterargument – mimetic AI system usage is permissible because no laws explicitly prohibit it – gains traction despite constituting a normative fallacy known as the *is-ought gap* (Hume 1739; Black 1964; Cohon 2018). Recognizing the absence of prohibition does not constitute justification is, therefore, essential to a fair deontological stance.

System type	System description	Usage example	System examples
Text-to-Text	Transforms textual inputs into new textual outputs while preserving stylistic elements	Drafting a paragraph describing a character's thoughts in the style of Virginia Woolf for a novel	ChatGPT (OpenAI 2025a); Claude (Anthropic PBC 2025)
Text-to-Image	Generates visual content from textual descriptions	Creating new artwork in the style of Frida Kahlo for a museum exhibit	ChatGPT via GPT-4o Image Generation (OpenAI 2025b); Vertex AI via Imagen 4 (Google Cloud 2025a)
Image-to-Image	Transforms existing images according to specified parameters, styles, or target characteristics	Converting a photograph into an image in the style of Studio Ghibli to share with a friend	ChatGPT via GPT-4o Image Generation (OpenAI 2025b); Midjourney Omni-Reference (Midjourney, Inc. 2025b,a)
Text-to-Speech	Converts written text into natural-sounding speech	Using a celebrity's voice clone to generate narration for a documentary	AI Voice Lab (Respeecher 2025); Dia (Nari Labs 2025b,a); Speechify Studio (Speechify Inc. 2025)
Text-to-Music	Generates musical compositions from textual descriptions	Creating a soundtrack in the style of Hans Zimmer for a film	Suno (Suno, Inc. 2024); Loudly (Loudly GmbH 2025); Mubert Render (Mubert Inc 2025)
Speech-to-Speech	Transforms speech input into new output which embodies a speaker's vocal characteristics	Changing the voice recordings for a character in a television show to a different language for dub localization	ElevenLabs Voice Changer and Dubbing (ElevenLabs Inc. 2023, 2025e,d); Replica Studios STS (Replica Studios, Inc. 2025)
Music-to-Music	Translates musical elements from one piece to another, retaining a particular style	Transforming a base composition to match the style of a specific artist for a new song	AIVA (Aiva Technologies SARL 2025); SoundGen (Audio Design Desk, Inc. 2023); Stable Audio (Stability AI Ltd 2025)
Text-to-Video	Generates dynamic visual content from textual descriptions or narratives	Creating short clips mimicking Wes Anderson's cinematographic style to use in a film	Sora (OpenAI 2024c, 2025c); Vertex AI via Veo 3 (Google Cloud 2025a)
Image-to-Video	Creates video sequences by animating or otherwise extending static image inputs	Creating a clip with a digital twin for personalized content marketing	Creative Reality™ Studio 3.0 (D-ID 2025); Invideo AI (Invideo 2025); Gen-4 References (Runway AI, Inc. 2025)
Video-to-Video	Extends video inputs into new outputs while preserving visual style	Creating a deepfake where an actor's performance is transformed to mimic another person's expressions	DeepFaceLab (Liu et al. 2023); Pollo.ai style transfer (Pollo.ai 2025); DeeVid (DeeVid AI 2025)

Table 1: Types of mimetic AI systems, their description, usage examples, and real-world implementation examples.

Virtue Ethics

Impersonating oneself and others using mimetic AI systems challenges the understanding of what constitutes a virtuous character, particularly regarding integrity in human expression. The style of an individual's work embodies their commitment to genuineness – a core virtue that philosopher Charles Taylor examines in his work on authenticity (Taylor 1991). Being true to oneself has moral significance beyond just self-expression; it represents a promise to be genuine when communicating with others, which extends to the information and resources that people share.

Impersonation is inherently inauthentic, even when performed without GenAI. As such, the argument follows that mimetic AI systems should be avoided because using them disenfranchises the human author who would have otherwise assumed authentic authorship of the output. However, if full consent for digital impersonation were able to be granted to the operator by a consenting target – and if that target were adequately given control of the model's usage equivalent to that of its operator and creator – and if commensurate compensation for the model's usage were granted to the target – then the mimetic model's outputs could begin to be interpreted as those of an authorized, albeit incomplete, representative of the human target.

Whether or not permission is granted, generating and sharing the outputs of mimetic models forsakes commitments to authenticity because generation relies on a latent space – a representation of data that captures its underlying structure and features but by definition cannot wholly encapsulate its entire nature (Bergmann 2025). Debate and research is ongoing to determine the inner representational complexity within mimetic models such as LLMs (Bender et al. 2021; Lindsey et al. 2025). Regardless of any evidence that may emerge from such investigations which suggests that mimetic models exhibit human-level understanding of human expression and its motivators, their reliance on a latent space to produce outputs necessarily makes them inauthentic to a human target.

Care Ethics

Impersonation without consent can represent a failure of moral attention to the target's vulnerability and autonomy. This perspective is particularly concerned with power asymmetries (Abma et al. 2020): mimetic AI system creators and operators can appropriate a target's voice or style with minimal effort, and the target surrenders to that appropriation without agency in the process. The argument follows that mimetic AI systems should be avoided because using them is careless in relation to the target.

The assignment of carelessness to system usage depends heavily on context. For example, impersonation can attempt to honor and extend relationships (Sander-Staudt 2025) (e.g., a student learning from a mentor's style with their permission), or impersonation can exploit without reciprocity or recognition (e.g., operators who replicate artists' styles without their permission). However, proving care is widely shown in the creation and operation of contemporary mimetic AI systems is more ambiguous and difficult than finding examples which suggest the opposite. For instance, despite substantial commitments from companies to prevent unauthorized voice cloning (ElevenLabs Inc. 2025c), cases persist where individuals assert that voice clones of them were created without their consent or compensation (Pappas 2024). These circumstances illustrate enduring relational harm – not merely intellectual property violation but a breach of the implicit social contract between creative professionals and their audiences.

Proponents of the affirmative position that impersonation can be done with care would defend mimetic model usage and proliferation for benevolent intentions (e.g., for educational purposes). Opponents would argue that no matter the intent, carelessness is demonstrated by the absence of rigid, enforced standards for the compensation, consent, control, and other supporting mechanisms for impersonation targets to experience the same benefits from mimetic AI system deployment as their operators and creators.

Consequentialism

Mimetic AI systems are perhaps most directly evaluated by presenting arguments based on their outcomes, yielding the perspective that using them is justified when it produces positive consequences and unjustified otherwise. This approach, though simple, encompasses significant complexity as different consequentialist theories offer distinct and competing evaluative lenses.

Classical utilitarianism examines aggregate happiness, suggesting that widespread access to mimetic AI systems might maximize overall pleasure – a view closely aligned with hedonistic consequentialism, which positions pleasure as the primary good (Driver 2022). Under this framework, using mimetic AI systems is justified when it brings pleasure to its stakeholders, acknowledging impersonation as both entertainment (e.g., parody) and as a means to desired ends, such as generating new music in the style of a dead musician when no more could or would otherwise be created. Importantly, pleasure in this case is not exclusively derived from satisfying creative whims. It also results from saving resources (e.g., time, money, and energy) by outsourcing a given task to a mimetic model instead of a human – often at the expense of the latter.

For example, pleasure derived by operators engaging in interpersonal impersonation often directly results in displeasure for the target(s), such as artists who experience both dignitary harm and potential economic loss when their styles are replicated without consent or compensation (Derico 2024). In the case of intrapersonal impersonation, the delegation of a task to a digital double may result in temporary pleasure (e.g., convenience and relief from avoiding work) but long-term displeasure from addiction to using the mimetic model, as has been documented with general-purpose GenAI systems (Zhou and Zhang 2024).

Preference utilitarianism considers stakeholders' competing desires (Simões 2013). For example, targets of impersonation may prefer restricting access to their likeness, whereas operators and creators may prefer expanding it to improve and encourage usage of their models and systems. Creators and intermediaries who stand to benefit financially from mimetic AI system proliferation may exhibit preferences at odds with regulators and audiences who could benefit from efforts opposing it.

A necessary condition for consequentialist justifications is identifying and reconciling any unbalanced distributions of pleasure and preferences. Jeremy Bentham's *hedonic calculus* presents a method for quantifying the pleasure and pain which result from an action based on seven circumstances: intensity, duration, certainty, remoteness, fecundity, purity, and extent (Bentham 1789; LibreTexts Humanities 2024). Such strategies have precedent for application in industry-standard harms modeling (Coalition for Content Provenance and Authenticity 2025b), acting in service of reaching balanced consequentialist perspectives and solutions.

Assessing Benefits and Risks

Understanding the benefits and risks of mimetic AI systems to their stakeholders helps to explain value judgements supporting or opposing their usage and regulation. I assume benefits arise predominantly from successful, convincing imitation of targets' characteristics, whereas risks are present regardless of whether systems succeed or fail. When these systems achieve high-fidelity imitation, they enable unauthorized impersonation and deception; when they do not, they can incur costs beyond misrepresentation.

Saving Resources

Mimetic AI systems offer immediate practical advantages over human labor through dramatic resource optimization in bypassing the biggest practical limitations of accessing targets: unavailability, unwillingness, and expensiveness. For example, having a renowned physicist explain complex concepts in an educational video would typically require their direct participation – an often unrealistic proposition given their limited time. Platforms such as Cameo (Baron App, Inc. 2025) offer limited engagement with public figures, but outputs can be brief, expensive, non-interactive, and restricted to a range of requests which ultimately rely on external discretion and schedules.

In contrast, content creation that would typically require days or weeks of human effort can be accomplished in hours or minutes with a mimetic AI system, reducing production costs and timelines by orders of magnitude. For example, a human novelist might require five to eight months, or much longer, to complete a 100,000-word manuscript (Chesson 2021), while a text-to-text system could generate a draft in the style of a successful author in a single day, or even less (Kokoski 2023). The financial implications are correspondingly significant. For instance, employing professional human voice actors for a commercial project might cost hundreds or thousands of U.S. dollars in studio time, talent and usage fees, and other production expenses (Global Voice Acting Academy, Inc. 2025), while a text-to-speech system using high-quality voice clones could generate commercially viable output at a fraction of the cost (ElevenLabs Inc. 2025b).

Similarly, operators can extend their own communication capacity beyond human limitations in a qualitatively different form of automation than generic GenAI assistance. A busy professional might leverage a mimetic AI system to draft contextually appropriate emails to colleagues in their narrative voice and style, maintaining consistent representation across communications that would be impossible to manage manually given time constraints. In each of these scenarios, mimetic AI systems offer their underlying model(s) to operators as a *force multiplier* to generate more content in the target's style than the target could feasibly produce through human means (McIlroy-Young et al. 2022a).

In the long-term, time and cost savings surge alongside the volume of generated output. These results radically alter the perceived productivity, influence, and legacy of operators and targets by audiences. In other words, using mimetic AI systems can facilitate consistent exposure to specific styles, voices, or visual aesthetics across contexts, cultivating audience familiarity and preference at scale. For example, an advertiser using such systems to enforce a brand's distinct creative style across thousands of personalized content pieces (Portman 2023) may be able to establish stronger preference formation than would be possible without them.

With cognitive and perceptual conditioning via the *mere exposure* (Zajonc 1968; Pilat and Krastev 2021b) and *illusory truth* (Henderson, Simons, and Barr 2021; Pilat and Krastev 2021a; Dreyfuss 2017) effects, the capacity of mimetic AI systems to generate high volumes of stylistically consistent content enables strategic reinforcement of specific messages, associations, or held beliefs. Further, anthropomorphism in chatbots – a possible design paradigm for a text-to-text system employed in conversational applications – enables the simulation of human personal connection at scale, which potentially strengthens parasocial relationships (Maeda and Quan-Haase 2024; Akbulut et al. 2024) between audiences and targets, particularly public and proxy figures.

Losing Resources

Mimetic AI systems enable highly personalized individual-level fraud via scams targeting vulnerable individuals. Examples include voice cloning to impersonate family members in emergency scenarios, such as ”grandparent scams”, and more broad applications of synthetic identity fraud for extorting money (U.S. Federal Communications Commission 2025; Timoney 2025). Businesses and organizations also face threats of institutional-level fraud via mimetic models, such as impersonation of managers to authorize large wire transfers (Magramo 2024) and impersonation of customers to bypass identity verification protocols in end-user applications (Vahl 2024).

The corresponding economic impact is accelerating at an alarming rate and shows potential for long-term damage. Current projections indicate financial losses attributable to fraud from deepfakes alone reach approximately $200 million USD in the first quarter of 2025 (Giardina 2025). More concerning is the growth trajectory, with projected cumulative losses potentially reaching $40 billion USD by 2027 (Jennings-Trace 2025).

As mimetic AI systems make the simulation of artistic styles increasingly accessible, they also alter the financial and cultural incentives which have traditionally sustained creative professions. Emerging artists face a daunting landscape where their styles can be rapidly replicated and mass-produced, potentially undermining the scarcity value that has historically supported their livelihoods (Hoover 2023; Smith 2024; Derico 2024). This changing dynamic is discouraging those who would become artists from investing in the development of their creativity (Nguyen 2023), representing a potential long-term diminishment of cultural innovation and authentic human expression.

A deeper form of dispossession arises as a consequence, where individuals lose not just economic control but a sense of ownership over their creative identity. Artists report experiencing a form of creative paralysis when their styles become widely imitated by GenAI systems, leading some to abandon their signature approaches to avoid accusations of using GenAI themselves (Watson 2025). Online creative communities have witnessed divisive conflicts over suspected AI-generated content, resulting in false accusations, member expulsions, and harrowing effects on legitimate creators (Stokel-Walker 2023).

Erosion of Trust

Mimetic AI systems inherently compress and simplify the complex patterns they attempt to imitate, creating unavoidable distortions in their outputs. In reducing multidimensional human expression into statistical approximations reliant on a latent space (Bergmann 2025), their usage involves sacrificing nuance and context. This loss becomes particularly apparent and problematic when mimetic models are used to extrapolate beyond knowledge granted by training or runtime adaptation to examples.

When systems are implemented, for instance, which let museum-goers converse with a voice clone and language model impersonating the renowned and deceased Spanish artist Salvador Dalí (Eaton 2024), they receive plausible but speculative outputs. These extrapolations lack the authentic developmental trajectory, lived experiences, and contextual grounding that would inform genuine creative expression. The authority to represent how an individual would respond to novel circumstances remains with that individual alone – a boundary which mimetic AI systems inherently cross. This creates a dangerous illusion of accessibility to inaccessible perspectives, potentially distorting the collective understanding of historical or contemporary individuals, artistic movements, and intellectual traditions through convincing but ultimately inauthentic representations.

Correspondingly, the reputational damage to targets can be devastating and immediate because mimetic model outputs can be generated at scale and distributed rapidly across multiple platforms (McIlroy-Young et al. 2022a). The social consequences vary in severity but disproportionately affect public and private figures who engage in reputation-dependent professions or social circles, particularly women and those belonging to marginalized or minority populations (Paris and Donovan 2019; U.S. Department of Homeland Security 2021). Clear examples include a synthetic video showing a political figure performing illicit acts (Hunter 2024), fabricated audio of a principal spouting racial bigotry (Banerji 2024), or AI-generated pornography featuring individuals without their consent (Chesney and Citron 2019). Beyond external social consequences, targets experience significant internal psychological distress, including identity violation and consequent, persistent anxiety (Wei 2024).

More broadly, the popularization of mimetic AI system usage contributes to an erosion of trust in digital information ecosystems. Fueled by the rapidly increasing volume of AI-generated content in misinformation claims since 2023 (Dufour et al. 2024), this erosion occurs thanks to the *liar's dividend*, a term coined by law professors Bobby Chesney and Danielle Citron: a phenomenon where authentic content can be dismissed as synthetic, allowing genuine misconduct to be plausibly denied as a fabrication by GenAI (Chesney and Citron 2019). Persistent exposure to impersonations of specific individuals or entities enabled by mimetic AI systems gradually reshapes public perception of their authentic characteristics. As a result, when synthetic approximations of a public figure's speaking style, political opinions, or creative output circulate widely, these simulations begin to influence public understanding of the actual person (Goldstein and Lohn 2024).

This influence represents the long-term effects of cognitive and perceptual conditioning. As GenAI content becomes commonplace, audiences develop *informational learned helplessness*, surrendering to the idea that distinguishing authentic from synthetic content is impractical or impossible (Nolan and Kimball 2021). This creates a dangerous feedback loop where mimetic model outputs shape public expectations, which then influence future outputs, potentially diverging further from authentic sources with each iteration.

Potential Solutions and Challenges

Mimetic AI systems can generate disproportionate risks relative to their benefits, particularly through their potential to undermine autonomy and privacy, erode trust, and exacerbate existing power imbalances. While they offer temporal and financial efficiency advantages for creative applications, these benefits accrue inequitably to technology companies and privileged users, while vulnerable targets (e.g., individuals susceptible to extortion or misinformation, and established or emerging artists) bear crushing financial and social burdens which disincentivize genuine expression and manifest at scale through rights and privacy violations.

Recognizing the inevitable preservation of interest and investment in the continued development of mimetic models and mimetic AI systems, I argue that targeted mitigation strategies, rather than prohibition, offer the most practical path forward to their governance. I hereby list possible solutions combining top-down governance mechanisms targeting creators, intermediaries, and regulators with bottom-up strategies for operators, targets, and audiences. I also document and discuss various obstacles to their implementation.

Top-Down Solutions

I propose the following as promising avenues for scalable, systemic regulatory change:

- Implementing friction and containment mechanisms – such as the guardrails implemented by OpenAI and Google to circumvent deepfakes ahead of 2024 global elections (OpenAI 2024b; Deslandes 2024) – that detect when users attempt to generate or share content impersonating specific individuals, triggering graduated responses including consent verification, ethical warnings, multi-step confirmations, deference to human sources, and usage limitations for high-risk scenarios. Automatic content recognition can be applied not just to text prompts but also to the provision of reference materials, particularly when processing file uploads by users to block the use of unauthorized materials, such as copyrighted recordings in music-to-music systems (Audible Magic Corporation 2024).
- Integrating compensation mechanisms for targets of interpersonal impersonation, such as royalties or usage fees, as well as the ability to monitor and restrict access to mimetic models. Mechanistic interpretabilty tools such as a "microscope" for LLMs (Lindsey et al. 2025) can be used to identify internal concepts within mimetic models related to a certain impersonation target (e.g., the name of a public figure, or the name of a studio) and set observable thresholds and triggers for financial disbursement or an interruption in system functionality.
- Including domain-specific authorities and specialists in the design, deployment, and regulation lifecycles for mimetic AI systems, such as nationally or internationally-recognized professional associations. Precedent exists for such organizations representing the lived experience of potential impersonation targets to meaningfully contribute to advancing AI ethics and regulation alongside technology companies (U.S. Copyright

Office 2025b). They also influence industry movements, such as the *fAIr Voices* initiative by the National Association of Voice Actors which calls for companies offering synthetic speech products to provide commitments to ethical training, consent, control, compensation, security, and data sourcing (National Association of Voice Actors 2025).

- Expanding assertive provenance mechanisms and inferred context tools to encourage the transparent and accessible attribution of synthetic media throughout its lifecycle. Software solutions, such as Adobe's *Content Credentials* (Adobe, Inc. 2025) and Google's *SynthID* or "About this result" (Google DeepMind 2025; Hebbar and Wolf 2024), are part of meaningful commitments by industry leaders to collaborative initiatives (Coalition for Content Provenance and Authenticity 2025a). Hardware solutions such as *OriginStory* – using sensors implanted in microphones to detect biosignals such as vocal cord vibrations for creating a watermark to be embedded at the point of audio creation (Triolo 2024) – also show promise as provenance mechanisms.
- Continuing efforts which establish enforceable frameworks to regulate the training and deployment of mimetic AI systems as well as the distribution of their outputs. For example, in the United States, the ELVIS Act (Madarang 2024) passed in Tennessee in 2024 and the proposed federal-level NO FAKES Act (Sen. Coons 2024) seek to protect the rights of private and public figures to their individual likeness. Similarly, the U.S. Copyright Office reports that viable licensing and compensation solutions other than blanket fair use determinations exist for using copyrighted data to train GenAI models (U.S. Copyright Office 2025b), which can help to bolster data privacy protections for human creative professionals while still fostering technical innovation.

Circumvention or Exploitation Operators determined to generate and share unauthorized impersonations can employ various techniques to bypass both assertive provenance and friction measures, as well as exploit compensation or control mechanisms.

Simple actions such as screen recording, cross-platform posting, or format conversion can strip embedded watermarks and attribution signals (Hebbar and Wolf 2024). More sophisticated circumvention might involve using GenAI technologies, as demonstrated by the use of an image-to-image model to remove watermarks from protected content (Wiggers 2025b). Similarly, interface friction can be bypassed through prompt engineering techniques which disguise malicious intentions via jailbreaking methods (Xu et al. 2024) or by migrating to alternative platforms with fewer restrictions. Thresholds for compensation disbursement or system interruption can be artificially reached, for example by spamming application programming interfaces (APIs) with an exorbitant volume of malicious requests. These conditions create a race between protection mechanisms and subterfuge, where defensive measures must continuously evolve across mimetic AI systems and output distribution channels to address new evasion strategies.

Direct Opposition Creators and intermediaries of mimetic AI systems naturally stand to benefit from their continued, uninterrupted development and deployment. As such, they may oppose certain regulatory proposals.

For example, as of early 2025, some companies whose GenAI products may be characterized as mimetic AI systems are actively opposing proposals to limit their access to training data, particularly to copyrighted data (Katzke 2025; Merica 2024). They argue that draconian barriers to accessing data would stifle innovation, thereby weakening the viability of the nation in which they are headquartered to be a global leader in AI development (Berger 2025). Continuation of the status quo follows, within which mitigable harms persist. Open dialogue and collaboration between stakeholders regarding frameworks of self-regulation, co-regulation, and traditional government regulation can help acknowledge legitimate innovation needs while establishing boundaries which address avoidable risks (G'sell 2024).

Bottom-Up Solutions

To provide immediate protection while top-down measures develop, I advocate for the following:

- Using and supporting the development of tools which identify and shield human-made content, such as digital watermarking and data cloaking. Examples include *Glaze* (Shan et al. 2025) and *Nightshade* (Shan et al. 2024) to protect human art and photographs, *AudioSeal* (Roman et al. 2024) and *AntiFake* (Yu, Zhai, and Zhang 2023) to protect human voices, *HarmonyCloak* (Meerza, Sun, and Liu 2025) to protect human music, and *SynthID* (Google DeepMind 2025) and similar systems to protect human writing and video.
- Engaging with the arts and humanities to cultivate a view which does not depict human creativity as a means to be automated nor as an end to be outsourced to technology. Prosocial behavior and the fostering of interpersonal relationships result from arts engagement (Rugg et al. 2021). It follows that operators who have a personal stake in the arts may be discouraged from misusing mimetic AI systems for unauthorized impersonation by virtue of empathizing with targets, particularly with the private and public figures behind nonhuman proxies.
- Attempting to employ human means to perform a task, or reconsidering the task's necessity entirely, before deciding to outsource it to a mimetic AI system. Whether conducting interpersonal impersonation or delegation to digital self, operators have the chance to avoid the pitfalls of both neo-luddism and technological addiction by embracing avenues where human lived experience that cannot be encapsulated by GenAI is prized, such as in the creative struggle behind artistic endeavors (Cave 2023).
- Applying media literacy concepts which encourage clarifying content authenticity to the usage of mimetic AI systems and the distribution or perception of their outputs. Audiences and targets equipped with assertive provenance and inferred context tools can more directly avoid or diminish the negative impacts of deepfakes as a form of misinformation (Goldstein and Lohn 2024; Hebbar

and Wolf 2024), reducing the incentives for malicious operators to produce or share deceptive outputs.

- Acting in recognition of mimetic AI system complexity to avoid misuse of their capabilities. An appropriate allegory is the use of heavy machinery. It is desirable, for example, to make the use of a forklift simple and accessible. However, because the machinery is so powerful and potentially dangerous, the Occupational Safety and Health Administration requires safety instruction and certification to legally use it in the United States (U.S. Department of Labor 2025). Mimetic AI systems also possess superhuman might and risk, but their usage mandates no official training nor certification and their user-facing documentation may lack meaningful explanations of harms or instructions to avoid them. In the absence of such requirements and materials, operators should still exercise caution, seeking and heeding guidance on best practices for data privacy in GenAI applications.

Lack of Accessibility While assertive provenance and data cloaking tools are emerging across modalities, their accessibility remains a significant barrier to widespread adoption, particularly among non-technical users who may be most vulnerable to the risks of mimetic AI systems.

Tools defending visual likeness are the most accessible as of early 2025. *Glaze* and *Nightshade* are available through direct download (SAND Lab, University of Chicago 2024), *Glaze* is accessible through a web platform called *WebGlaze* (SAND Lab, University of Chicago 2025), and a similar data cloaking method called *AI Disturbance* has been integrated into the ibisPaint drawing software (ibis inc. 2025). However, comparable technologies for other modalities – such as voice cloning protection via *AntiFake* (Yu, Zhai, and Zhang 2023) – require technical expertise for implementation, including command-line operations, manual file processing, or custom code execution. Investment in developing more user-friendly implementations, particularly with commonly used creative platforms, can make data protection accessible to those without technical expertise.

Lack of Motivation Even when defensive technologies and educational resources are available, barriers exist to impede their adoption. A 2023 study by Maier et al. of digital natives and their concern with data privacy claims:

> “...few actually take steps to self-manage privacy e.g. by adjusting settings or encrypting their data. Instead most people appear resigned and simply surrender to (potential) violations to their privacy” (Maier et al. 2023).

With mimetic AI systems, these barriers could be amplified by limited public understanding of AI capabilities, difficulty in conceptualizing abstract future harms, and the absence of immediate negative feedback when protective measures are neglected. Potential targets may adopt an ”it won’t happen to me” mindset until directly affected (Louwersmedia Group 2024), at which point remediation becomes significantly more difficult than prevention would have been. Awareness campaigns highlighting risks and data protection pathways may help mitigate such attitudes and foster a shift in public opinion favoring engagement with data privacy.

Conclusion

In 1881, a columnist for the Michigan Medical News journal wrote about one doctor copying another, stating:

> ”...is not plagiarism the most subtle form of flattery? It is the tribute that mediocrity pays to genius” (Leonard 1881).

Mimetic AI systems are far from mediocre plagiarists of human genius; they are that genius’ best and most hubristic attempt yet at impersonating itself. This paper’s investigation into their mechanisms, implementations, and outcomes – as well as the challenging yet promising paths to their oversight and regulation – shows how one of the most dangerous contemporary misues of GenAI might be understood and mitigated. I believe many avenues exist for future research evaluating and addressing mimetic AI models, systems, and their impacts. For example, exploring novel mechanisms to integrate during the moments at which content is created or consumed by both people and mimetic AI systems could enable the timely prevention, detection, and remediation of unethical impersonation attempts while safeguarding and incentivizing authentic human expression. Stakeholders can expect incoming improvements in multimodal GenAI models and the integration of autonomous agentic behavior into GenAI systems, as well as new laws restricting or expanding access to copyrighted materials for GenAI training, to shape upcoming mimetic AI system designs, integrations, and worldwide influence.

Ethical Statement

I present the following ethical considerations, researcher positionality, and adverse impact statements as further context to the reader about various factors and considerations which influenced my authorship of this paper.

Ethical Considerations

This paper focuses on the misuse of GenAI models to promote deception and misinformation, as well as how to understand and mitigate it. Ethical stances on deception itself – and modern manifestations in harmful deepfakes – are clear and well-established, arising from moral views and corresponding legal codifications introduced and championed far before the advent of GenAI technologies. However, perspectives shared by technical and non-technical communities discussing which measures ought to be taken to mitigate risks from the misuse of GenAI models are often at odds for a variety of legitimate reasons. My key ethical consideration, as such, is to treat such perspectives and their stakeholders fairly in the process of finding sources and synthesizing my own understanding and claims from them. As a result, my literature review includes a wide variety of primary and secondary sources in an attempt to ensure representativeness.

Researcher Positionality

The perspectives and frameworks I present as the author of this paper are influenced by my interdisciplinary background and lived experience. Specifically, I write with empathy for targets of unauthorized impersonation and for technical innovators who strive to prevent the misuse of their work.

Adverse Impact

The unintended negative consequences of this paper revolve around enabling the risks associated with impersonation via mimetic AI systems. For example, several examples of systems that could be employed for unauthorized impersonation were presented with direct citations; these could plausibly be found and used by a reader with malicious intent. To account for that, this paper was authored without including specific instructions to generate such impersonations and highlights commonly integrated safety mechanisms which trace and disincentivize misconduct.

Acknowledgments

I extend my sincerest gratitude to Sonja Schmer-Galunder for her guidance, support, and feedback during the writing process. I also thank Michael Link for his thoughtful remarks on various topics.

References

101soundboards.com. 2025. AI Text to Speech AI Computer Voices List - 130,583+ Voices. https://www.101soundboards.com/tts. Accessed: 2025-05-20.

Abma, T. A.; Visse, M.; Hanberger, A.; Simons, H.; and Greene, J. C. 2020. Enriching evaluation practice through care ethics. *Evaluation*, 26(2): 131–146.

Adobe, Inc. 2025. Content Credentials overview. https://helpx.adobe.com/content/help/en/creative-cloud/help/content-credentials.html. Accessed: 2025-05-20.

Adobe Inc. 2025. Set styles for image generation. https://helpx.adobe.com/content/help/en/firefly/generate-images-with-text-to-image/customize-generated-images/set-styles-for-image-generation.html. Accessed: 2025-05-19.

Aiva Technologies SARL. 2025. AIVA, the AI Music Generation Assistant. https://www.aiva.ai/. Accessed: 2025-04-20.

Akbulut, C.; Weidinger, L.; Manzini, A.; Gabriel, I.; and Rieser, V. 2024. All Too Human? Mapping and Mitigating the Risk from Anthropomorphic AI. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 7(1): 13–26.

Anthropic PBC. 2025. Meet Claude. https://www.anthropic.com/claude. Accessed: 2025-05-23.

Appian Corporation. 2025. What Makes a Good Dataset? https://docs.appian.com/suite/help/25.1/collect-data.html. Accessed: 2025-04-20.

Audible Magic Corporation. 2024. Ensuring Content Integrity: Suno Partners with Audible Magic for User Uploads. https://www.audiblemagic.com/2024/10/18/ensuring-content-integrity-suno-partners-with-audible-magic-for-user-uploads/. Accessed: 2025-05-22.

Audio Design Desk, Inc. 2023. SoundGen. https://soundgen.io/guide. Accessed: 2025-04-20.

Banerji, O. 2024. Why Schools Need to Wake Up to the Threat of AI 'Deepfakes' and Bullying. https://www.edweek.org/technology/why-schools-need-to-wake-up-to-the-threat-of-ai-deepfakes-and-bullying/2024/12. Accessed: 2025-04-21.

Baron App, Inc. 2025. Personalized videos from your favorite stars. https://www.cameo.com. Accessed: 2025-04-20.

Bender, E. M.; Gebru, T.; McMillan-Major, A.; and Shmitchell, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, 610–623. New York, NY, USA: Association for Computing Machinery. ISBN 9781450383097.

Bentham, J. 1789. *An Introduction to the Principles of Morals and Legislation*. London: Printed for T. Payne and Son.

Bereczki, T.; and Ádám Liber. 2023. AI's emergent abilities a 'double-edged sword' | IAPP. https://iapp.org/news/a/ais-emergent-abilities-a-double-edged-sword. Accessed: 2025-04-20.

Berger, V. 2025. The AI Copyright Battle: Why OpenAI And Google Are Pushing For Fair Use. https://www.forbes.com/sites/virginieberger/2025/03/15/the-ai-copyright-battle-why-openai-and-google-are-pushing-for-fair-use/. Accessed: 2025-04-21.

Bergmann, D. 2025. What is latent space? https://www.ibm.com/think/topics/latent-space. Accessed: 2025-04-20.

Black, M. 1964. The Gap Between "Is" and "Should". *The Philosophical Review*, 73(2): 165–181.

Campos, A. S. 2019. Kant on Acting from Juridical Duty. *International Journal of Philosophical Studies*, 27(4): 498–514. Publisher: Routledge.

Cave, N. 2023. The Red Hand Files - Issue #218. https://www.theredhandfiles.com/chat-gpt-what-do-you-think/. Accessed: 2025-05-22.

Center for News, Technology & Innovation. 2024. Synthetic Media & Deepfakes. https://innovating.news/article/synthetic-media-deepfakes/. Accessed: 2025-04-20.

Chesney, B.; and Citron, D. 2019. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107.

Chesson, D. 2021. How Long Does it Take to Write a Book? https://kindlepreneur.com/how-long-to-write-a-book/. Accessed: 2025-04-20.

Coalition for Content Provenance and Authenticity. 2025a. About - C2PA. https://c2pa.org/about/about/. Accessed: 2025-05-22.

Coalition for Content Provenance and Authenticity. 2025b. C2PA Harms Modeling. https://c2pa.org/specifications/specifications/1.4/security/Harms_Modelling.html. Accessed: 2025-05-22.

Cohon, R. 2018. Hume's Moral Philosophy. In Zalta, E. N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2018 edition.

D-ID. 2025. D-ID's Creative Reality™ Studio | Generative AI Video Creator. https://www.d-id.com/photo-to-video/. Accessed: 2025-04-20.

de Rancourt-Raymond, A.; and Smaili, N. 2023. The unethical use of deepfakes. *Journal of Financial Crime*, 30(4): 1066–1077. Copyright - © Emerald Publishing Limited; Last updated - 2024-12-16; SubjectsTermNotLitGenreText - United States–US.

DeeVid AI. 2025. Video to Video Generator. https://deevid.ai/video-to-video. Accessed: 2025-04-20.

Demerlé, N.; Esling, P.; Doras, G.; and Genova, D. 2024. Combining audio control and style transfer using latent diffusion. https://arxiv.org/abs/2408.00196. arXiv:2408.00196.

Deng, L.; Hinton, G.; and Kingsbury, B. 2013. New types of deep neural network learning for speech recognition and related applications: an overview. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 8599–8603.

Derico, B. 2024. 'A tech firm stole our voices - then cloned and sold them'. https://www.bbc.com/news/articles/c3d9zv50955o. Accessed: 2025-04-20.

Deslandes, N. 2024. Google to require disclosures for digitally altered election ads. https://techinformed.com/google-to-require-disclosures-for-digitally-altered-election-ads/. Accessed: 2025-05-22.

Devlin, J.; Chang, M.; Lee, K.; and Toutanova, K. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *CoRR*, abs/1810.04805.

Dillon, R. S. 2022. Respect. In Zalta, E. N.; and Nodelman, U., eds., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2022 edition.

Dreyfuss, E. 2017. Want to Make a Lie Seem True? Say It Again. And Again. And Again. https://www.wired.com/2017/02/dont-believe-lies-just-people-repeat/. Accessed: 2025-04-20.

Driver, J. 2022. The History of Utilitarianism. In Zalta, E. N.; and Nodelman, U., eds., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2022 edition.

Dufour, N.; Pathak, A.; Samangouei, P.; Hariri, N.; Deshetti, S.; Dudfield, A.; Guess, C.; Escayola, P. H.; Tran, B.; Babakar, M.; and Bregler, C. 2024. AMMeBa: A Large-Scale Survey and Dataset of Media-Based Misinformation In-The-Wild. https://arxiv.org/abs/2405.11697. arXiv:2405.11697.

Eaton, K. 2024. Hello Dalí: Museum Exhibit Uses AI Tech for a Surreal Twist That Lets Visitors 'Talk' to the Artist. https://www.inc.com/kit-eaton/hello-dal-museum-exhibit-uses-ai-tech-for-a-surreal-twist-that-lets-visitors-talk-to-artist.html. Accessed: 2025-04-20.

ElevenLabs Inc. 2023. Introducing speech to speech. https://elevenlabs.io/blog/speech-to-speech. Accessed: 2025-04-20.

ElevenLabs Inc. 2025a. Create a replica of your voice that sounds just like you. https://elevenlabs.io/voice-cloning. Accessed: 2025-05-19.

ElevenLabs Inc. 2025b. ElevenLabs Pricing for Creators & Businesses of All Sizes. https://elevenlabs.io/pricing. Accessed: 2025-05-20.

ElevenLabs Inc. 2025c. ElevenLabs Voices: A comprehensive guide. https://elevenlabs.io/voice-guide. Accessed: 2025-05-23.

ElevenLabs Inc. 2025d. Localize content across 29 languages with AI dubbing. https://elevenlabs.io/dubbing-studio. Accessed: 2025-04-20.

ElevenLabs Inc. 2025e. Transform your voice into another with our AI voice changer. https://elevenlabs.io/voice-changer. Accessed: 2025-04-20.

Elizalde, B.; Deshmukh, S.; Ismail, M. A.; and Wang, H. 2022. CLAP: Learning Audio Concepts From Natural Language Supervision. https://arxiv.org/abs/2206.04769. arXiv:2206.04769.

Fallis, D. 2021. The Epistemic Threat of Deepfakes. *Philosophy & Technology*, 34(4): 623–643.

Felbo, B.; Mislove, A.; Søgaard, A.; Rahwan, I.; and Lehmann, S. 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Franzen, C. 2024. What's the best interface for gen AI? It all depends on the use case. https://venturebeat.com/ai/whats-the-best-interface-for-gen-ai-it-all-depends-on-the-use-case/. Accessed: 2025-04-19.

Gatys, L. A.; Ecker, A. S.; and Bethge, M. 2016. Image Style Transfer Using Convolutional Neural Networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2414–2423.

Giardina, C. 2025. Deepfake-Enabled Fraud Has Already Caused $200 Million in Financial Losses in 2025, New Report Finds. https://variety.com/2025/digital/news/deepfake-fraud-caused-200-million-losses-1236372068/. Accessed: 2025-04-20.

Gilbertson, A.; and Reisner, A. 2024. Apple, Nvidia, Anthropic Used Thousands of Swiped YouTube Videos to Train AI. https://www.proofnews.org/apple-nvidia-anthropic-used-thousands-of-swiped-youtube-videos-to-train-ai/. Accessed: 2025-04-20.

Global Voice Acting Academy, Inc. 2025. Rate Guide - GVAA. https://globalvoiceacademy.com/gvaa-rate-guide-2/. Accessed: 2025-04-20.

Goldstein, J. A.; and Lohn, A. 2024. Deepfakes, Elections, and Shrinking the Liar's Dividend. https://www.brennancenter.org/our-work/research-reports/deepfakes-elections-and-shrinking-liars-dividend. Accessed: 2025-04-20.

Google Cloud. 2025a. Expanding Vertex AI with the next wave of generative AI media models. https://cloud.google.com/blog/products/ai-machine-learning/announcing-veo-3-imagen-4-and-lyria-2-on-vertex-ai. Accessed: 2025-05-23.

Google Cloud. 2025b. Experiment with parameter values | Generative AI on Vertex AI. https://cloud.google.com/vertex-ai/generative-ai/docs/learn/prompts/adjust-parameter-values. Accessed: 2025-04-20.

Google Cloud. 2025c. Generative AI beginner's guide | Generative AI on Vertex AI | Google Cloud. https://cloud.google.com/vertex-ai/generative-ai/docs/learn/overview. Accessed: 2025-05-19.

Google DeepMind. 2025. SynthID. https://deepmind.google/technologies/synthid/. Accessed: 2025-04-20.

Grand View Research. 2024a. AI Voice Cloning Market Size, Share & Trends Report, 2030. https://www.grandviewresearch.com/industry-analysis/ai-voice-cloning-market-report. Accessed: 2025-04-19.

Grand View Research. 2024b. Deepfake AI Market Size, Share And Growth Report, 2030. https://www.grandviewresearch.com/industry-analysis/deepfake-ai-market-report. Accessed: 2025-04-19.

G'sell, F. 2024. Regulating Under Uncertainty: Governance Options for Generative AI. https://cyber.fsi.stanford.edu/content/regulating-under-uncertainty-governance-options-generative-ai. Accessed: 2025-05-23.

Hancock, J. T.; and Bailenson, J. N. 2021. The Social Impact of Deepfakes. *Cyberpsychology, Behavior, and Social Networking*, 24(3): 149–152. PMID: 33760669.

Hebbar, N.; and Wolf, C. 2024. Determining trustworthiness through provenance and context. https://static.googleusercontent.com/media/publicpolicy.google/en//resources/determining_trustworthiness_en.pdf. Accessed: 2025-04-20.

Henderson, E. L.; Simons, D. J.; and Barr, D. J. 2021. The Trajectory of Truth: A Longitudinal Study of the Illusory Truth Effect. *Journal of Cognition*.

Hinton, G.; Deng, L.; Yu, D.; Dahl, G. E.; Mohamed, A.-r.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T. N.; and Kingsbury, B. 2012. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, 29(6): 82–97.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising Diffusion Probabilistic Models. https://arxiv.org/abs/2006.11239. arXiv:2006.11239.

Hong, W.; Ding, M.; Zheng, W.; Liu, X.; and Tang, J. 2022. CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers. https://arxiv.org/abs/2205.15868. arXiv:2205.15868.

Hoover, A. 2023. Voice Actors Are Bracing to Compete With Talking AI. https://www.wired.com/story/ai-voice-actors-jobs-threat/. Accessed: 2025-04-20.

Hume, D. 1739. *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. London: Printed for John Noon.

Hunter, C. 2024. How a deepfake almost ruined my political career. https://www.ted.com/talks/cara_hunter_how_a_deepfake_almost_ruined_my_political_career. Accessed: 2025-04-21.

ibis inc. 2025. 67. AI Disturbance - How to use ibisPaint. https://ibispaint.com/lecture/index.jsp?no=192&lang=en. Accessed: 2025-04-20.

Invideo. 2025. Free Image to Video AI Generator Online | Invideo AI. https://invideo.io/make/image-to-videos/. Accessed: 2025-04-20.

Jennings-Trace, E. 2025. AI deepfakes estimated to cause $40 billion in losses by 2027. https://www.techradar.com/pro/security/ai-deepfakes-estimated-to-cause-usd40-billion-in-losses-by-2027. Accessed: 2025-04-20.

Jing, Y.; Yang, Y.; Feng, Z.; Ye, J.; Yu, Y.; and Song, M. 2018. Neural Style Transfer: A Review. https://arxiv.org/abs/1705.04058. arXiv:1705.04058.

Katzke, C. 2025. AI Safety Newsletter #35: Lobbying on AI Regulation. https://newsletter.safe.ai/p/ai-safety-newsletter-35-lobbying. Accessed: 2025-04-21.

Knibbs, K. 2024. Every AI Copyright Lawsuit in the US, Visualized. https://www.wired.com/story/ai-copyright-case-tracker/. Accessed: 05-20-2025.

Kokoski, C. 2023. I Wrote a Novel With ChatGPT In 30 Minutes. https://medium.com/the-bald-writer/i-wrote-a-novel-with-chatgpt-in-30-minutes-c15602f271b5. Accessed: 2025-04-20.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. ImageNet classification with deep convolutional neural networks. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, 1097–1105. Red Hook, NY, USA: Curran Associates Inc.

Landymore, F. 2025. An AI Model Has Officially Passed the Turing Test. https://futurism.com/ai-model-turing-test. Accessed: 2025-04-19.

Legal Information Institute, Cornell University. 2020. public figure | Wex | US Law | LII / Legal Information Institute. https://www.law.cornell.edu/wex/public_figure. Accessed: 2025-04-20.

Leonard, C. H. 1881. A Plagiarism in "High Places.". *Michigan Medical News*, 4(1): 295–296.

Li, C.; Park, M.; Rossi, C.; and Li, Z. 2025. Text-Driven Video Style Transfer with State-Space Models: Extending StyleMamba for Temporal Coherence. https://arxiv.org/abs/2503.12291. arXiv:2503.12291.

LibreTexts Humanities. 2024. 7.1.6: Hedonic Calculus. https://human.libretexts.org/Courses/Folsom_Lake_College/PHIL_300%3A_Introduction_to_Philosophy_(Bauer)/07%3A_Ethics/7.01%3A_Utilitarianism/7.1.06%3A_Hedonic_Calculus. Accessed: 2025-04-20.

Lindsey, J.; Gurnee, W.; Ameisen, E.; Chen, B.; Pearce, A.; Turner, N. L.; Citro, C.; Abrahams, D.; Carter, S.; Hosmer, B.; Marcus, J.; Sklar, M.; Templeton, A.; Bricken, T.; McDougall, C.; Cunningham, H.; Henighan, T.; Jermyn, A.;

Jones, A.; Persic, A.; Qi, Z.; Thompson, T. B.; Zimmerman, S.; Rivoire, K.; Conerly, T.; Olah, C.; and Batson, J. 2025. On the Biology of a Large Language Model. *Transformer Circuits Thread*.

Liu, K.; Perov, I.; Gao, D.; Chervoniy, N.; Zhou, W.; and Zhang, W. 2023. Deepfacelab: Integrated, flexible and extensible face-swapping framework. *Pattern Recognition*, 141: 109628.

Loudly GmbH. 2025. AI Text to Music. https://www.loudly.com/text-to-music. Accessed: 2025-04-20.

Louwersmedia Group. 2024. 'Risks posed by AI are increasing,awareness not yet' - Hotelvak. https://hotelvak.eu/en/hoteltech/risks-due-to-ai-increase-awareness-not-yet/. Accessed: 2025-04-21.

Lubin Austermuehle DiTommaso, PC. 2025. Public Figures vs. Private Figures: Which one are you? https://www.chicagobusinesslawfirm.com/services/defamation/public-figures-vs-private-figures-which-one-are-you/. Accessed: 2025-04-20.

Madarang, C. 2024. Elvis Act Signed Into Tennessee Law to Protect Musicians From AI Deepfakes. https://www.rollingstone.com/music/music-news/elvis-act-tennessee-signed-ai-impersonation-1234992395/. Accessed: 2025-04-15.

Maeda, T.; and Quan-Haase, A. 2024. When Human-AI Interactions Become Parasocial: Agency and Anthropomorphism in Affective Design. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, 1068–1077. New York, NY, USA: Association for Computing Machinery. ISBN 9798400704505.

Magramo, H. C., Kathleen. 2024. Finance worker pays out $25 million after video call with deepfake 'chief financial officer'. https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html. Accessed: 2025-04-19.

Maier, E.; Doerk, M.; Reimer, U.; and Baldauf, M. 2023. Digital natives aren't concerned much about privacy, or are they? *i-com*, 22(1): 83–98.

McIlroy-Young, R.; Kleinberg, J.; Sen, S.; Barocas, S.; and Anderson, A. 2022a. Mimetic Models: Ethical Implications of AI that Acts Like You. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '22, 479–490. ACM.

McIlroy-Young, R.; Wang, R.; Sen, S.; Kleinberg, J.; and Anderson, A. 2022b. Learning Models of Individual Behavior in Chess. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '22, 1253–1263. ACM.

Meerza, S. I. A.; Sun, L.; and Liu, J. 2025. HARMONYCLOAK: Making Music Unlearnable for Generative AI . In *2025 IEEE Symposium on Security and Privacy (SP)*, 430–448. Los Alamitos, CA, USA: IEEE Computer Society.

Merica, D. 2024. The AI industry uses a light lobbying touch to educate Congress from a corporate perspective. https://apnews.com/article/artificial-intelligence-lobbying-congress-technology-regulation-d54fe64f3fd135fdca41bdaf528282b8. Accessed: 2025-04-21.

Midjourney, Inc. 2025a. Omni-Reference --oref. https://www.midjourney.com/updates/omni-reference-oref. Accessed: 2025-05-15.

Midjourney, Inc. 2025b. V7 Alpha. https://www.midjourney.com/updates/v7-alpha. Accessed: 2025-05-15.

Mubert Inc. 2025. Royalty-free text to music AI generator by Mubert. https://mubert.com/render. Accessed: 2025-04-20.

Nari Labs. 2025a. GitHub - nari-labs/dia: A TTS model capable of generating ultra-realistic dialogue in one pass. https://github.com/nari-labs/dia/. Accessed: 2025-05-19.

Nari Labs. 2025b. Nari Labs' Dia: The Cool AI That Makes Words Talk. https://narilabs.org/article/nari-labs-dia-ai-speech. Accessed: 2025-05-19.

National Association of Voice Actors. 2025. fAIr Voices: Consent, Control, Compensation. https://navavoices.org/fair-voices/. Accessed: 2025-05-22.

Newsroom. 2025. The Secret Behind Today's Most Realistic Text-to-Speech Voices. https://moderndiplomacy.eu/2025/04/12/the-secret-behind-todays-most-realistic-text-to-speech-voices/. Accessed: 2025-04-19.

Nguyen, K. 2023. AI Is Causing Student Artists to Rethink Their Creative Career Plans | KQED. https://www.kqed.org/arts/13928253/ai-art-artificial-intelligence-student-artists-midjourney. Accessed: 2025-04-20.

Nolan, S. A.; and Kimball, M. 2021. Giving Up: Informational Learned Helplessness. https://www.psychologytoday.com/us/blog/misinformation-desk/202112/giving-informational-learned-helplessness. Accessed: 2025-04-20.

OpenAI. 2024a. Hello GPT-4o. https://openai.com/index/hello-gpt-4o/. Accessed: 2025-05-16.

OpenAI. 2024b. How OpenAI is approaching 2024 worldwide elections. https://openai.com/index/how-openai-is-approaching-2024-worldwide-elections/. Accessed: 2025-04-20.

OpenAI. 2024c. Sora System Card. https://openai.com/index/sora-system-card/. Accessed: 2025-05-19.

OpenAI. 2025a. ChatGPT | OpenAI. https://openai.com/chatgpt/overview/. Accessed: 2025-05-22.

OpenAI. 2025b. Introducing 4o Image Generation. https://openai.com/index/introducing-4o-image-generation/. Accessed: 2025-04-15.

OpenAI. 2025c. Sora | OpenAI. https://openai.com/sora/. Accessed: 2025-05-19.

OthersideAI. 2025. Famous Writer Style Mimic | Emulate the style of your favorite writers | HyperWrite AI Writing Assistant. https://hyperwriteai.com/aitools/famous-writer-style-mimic. Accessed: 2025-04-20.

Pappas, L. A. 2024. Actors Say AI Co. ElevenLabs Cloned Their Voices. https://www.law360.com/articles/1874870/

actors-say-ai-co-elevenlabs-cloned-their-voices. Accessed: 2025-05-23.

Paris, B.; and Donovan, J. 2019. Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence. https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf. Accessed: 2025-05-19.

Pilat, D.; and Krastev, S. 2021a. Illusory Truth Effect - The Decision Lab. https://thedecisionlab.com/biases/illusory-truth-effect. Accessed: 2025-04-15.

Pilat, D.; and Krastev, S. 2021b. Mere Exposure Effect - The Decision Lab. https://thedecisionlab.com/biases/mere-exposure-effect. Accessed: 2025-04-15.

Pollo.ai. 2025. Video to Video AI - Free Video Style Transfer | Pollo AI. https://pollo.ai/video-to-video. Accessed: 2025-04-20.

Portman, P. 2023. Generative AI And The Art Of Personalization. https://www.forbes.com/councils/forbestechcouncil/2023/11/13/generative-ai-and-the-art-of-personalization/. Accessed: 05-21-2025.

Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; Krueger, G.; and Sutskever, I. 2021. Learning Transferable Visual Models From Natural Language Supervision. https://arxiv.org/abs/2103.00020. arXiv:2103.00020.

Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; and Sutskever, I. 2019. Language Models are Unsupervised Multitask Learners. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf. Accessed: 2025-05-20.

Ramesh, A.; Dhariwal, P.; Nichol, A.; Chu, C.; and Chen, M. 2022. Hierarchical Text-Conditional Image Generation with CLIP Latents. https://arxiv.org/abs/2204.06125. arXiv:2204.06125.

Replica Studios, Inc. 2025. Speech to Speech - AI Voice Changer. https://www.replicastudios.com/products/speech-to-speech. Accessed: 2025-04-20.

Respeecher. 2025. AI Voice Lab. https://www.respeecher.com/ai-voice-lab. Accessed: 2025-05-23.

Ressmeyer, R.; Masling, S.; and Liao, M. 2019. "Deep Faking" Political Twitter using Transfer learning and GPT-2. https://cs229.stanford.edu/proj2019aut/data/assignment_308832_raw/26647402.pdf. Accessed: 2025-04-19.

Riemer, K.; and Peter, S. 2024. Conceptualizing generative AI as style engines: Application archetypes and implications. *International Journal of Information Management*, 79: 102824.

Roman, R. S.; Fernandez, P.; Défossez, A.; Furon, T.; Tran, T.; and Elsahar, H. 2024. Proactive Detection of Voice Cloning with Localized Watermarking. https://arxiv.org/abs/2401.17264. arXiv:2401.17264.

Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; and Ommer, B. 2022. High-Resolution Image Synthesis with Latent Diffusion Models. https://arxiv.org/abs/2112.10752. arXiv:2112.10752.

Rugg, G.; Novak-Leonard, J.; Reynolds, M.; Tani, E.; Bertsche, L.; and Eife, E. 2021. The Outcomes of Arts Engagement for Individuals and Communities. https://www.norc.org/content/dam/norc-org/pdfs/NORC%20Outcomes%20of%20Arts%20Engagement%20-%20Full%20Report.pdf. Accessed: 2025-05-22.

Runway AI, Inc. 2025. Gen-4 References. https://academy.runwayml.com/gen4/gen4-references. Accessed: 2025-05-23.

SAND Lab, University of Chicago. 2024. Nightshade - User's Guide. https://nightshade.cs.uchicago.edu/userguide.html. Accessed: 2025-04-21.

SAND Lab, University of Chicago. 2025. WebGlaze - What is WebGlaze. https://glaze.cs.uchicago.edu/webglaze.html. Accessed: 2025-04-21.

Sander-Staudt, M. 2025. Care Ethics | Internet Encyclopedia of Philosophy. https://iep.utm.edu/care-ethics/. Accessed: 2025-04-20.

Scott, C. 2024. Study finds ChatGPT's latest bot behaves like humans, only better | Stanford School of Humanities and Sciences. https://humsci.stanford.edu/feature/study-finds-chatgpts-latest-bot-behaves-humans-only-better. Accessed: 2025-04-19.

Sen. Coons, C. A. D.-D. 2024. S.4875 - 118th Congress (2023-2024): NO FAKES Act of 2024. https://www.congress.gov/bill/118th-congress/senate-bill/4875. Accessed: 2025-04-15.

Shan, S.; Cryan, J.; Wenger, E.; Zheng, H.; Hanocka, R.; and Zhao, B. Y. 2025. Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models. https://arxiv.org/abs/2302.04222. arXiv:2302.04222.

Shan, S.; Ding, W.; Passananti, J.; Wu, S.; Zheng, H.; and Zhao, B. Y. 2024. Nightshade: Prompt-Specific Poisoning Attacks on Text-to-Image Generative Models. https://arxiv.org/abs/2310.13828. arXiv:2310.13828.

Shen, J.; Pang, R.; Weiss, R. J.; Schuster, M.; Jaitly, N.; Yang, Z.; Chen, Z.; Zhang, Y.; Wang, Y.; Skerry-Ryan, R.; Saurous, R. A.; Agiomyrgiannakis, Y.; and Wu, Y. 2018. Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions. https://arxiv.org/abs/1712.05884. arXiv:1712.05884.

Shipper, D. 2023. How to Make AI Write Like Your Favorite Author. https://every.to/chain-of-thought/how-to-make-ai-write-like-your-favorite-author. Accessed: 2025-04-20.

Simões, M. C. 2013. Hare's preference utilitarianism: an overview and critique. *Trans/Form/Ação*, 36(2): 123–134.

Smith, D. 2024. 'It's very easy to steal someone's voice': how AI is affecting video game actors. https://www.theguardian.com/technology/2024/mar/29/how-ai-is-affecting-video-game-actors. Accessed: 2025-04-20.

Speechify Inc. 2025. AI Voice Cloning: Clone Your Voice Instantly. https://speechify.com/voice-cloning/. Accessed: 2025-05-19.

Stability AI Ltd. 2025. Stable Audio - User Guide. https://stableaudio.com/user-guide/audio-to-audio. Accessed: 2025-05-19.

Stokel-Walker, C. 2023. A Huge Subreddit Suspended A User For Posting AI Art, But The Work Is 100% Human-Made. https://www.buzzfeednews.com/article/chrisstokelwalker/art-subreddit-illustrator-ai-art-controversy. Accessed: 2025-04-15.

Suno, Inc. 2024. Create Music with Text. https://help.suno.com/en/articles/3197377. Accessed: 2025-04-20.

Taylor, C. 1991. *The Ethics of Authenticity*. Harvard University Press. ISBN 978-0-674-26863-0.

Temitope, Y. 2024. 15.ai Creator reveals journey from MIT Project to internet phenomenon. https://guardian.ng/technology/15-ai-creator-reveals-journey-from-mit-project-to-internet-phenomenon/. Accessed: 2025-05-16.

The European Parliament and the Council of the European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). https://eur-lex.europa.eu/eli/reg/2024/1689. Accessed: 2025-05-20.

Thomas, X.; and Ghadiyaram, D. 2025. What's in a Latent? Leveraging Diffusion Latent Space for Domain Generalization. https://arxiv.org/abs/2503.06698. arXiv:2503.06698.

Thompson, S. A. 2024. A.I. Can Now Create Lifelike Videos. Can You Tell What's Real? https://www.nytimes.com/interactive/2024/09/09/technology/ai-video-deepfake-runway-kling-quiz.html. Accessed: 2025-04-19.

Timoney, M. 2025. Gen AI is ramping up the threat of synthetic identity fraud. https://www.bostonfed.org/news-and-events/news/2025/04/synthetic-identity-fraud-financial-fraud-expanding-because-of-generative-artificial-intelligence.aspx. Accessed: 2025-04-20.

Triolo, T. 2024. ASU researchers develop special microphone to verify human speech. https://news.asu.edu/20240503-science-and-technology-asu-researchers-develop-special-microphone-verify-human-speech. Accessed: 2025-05-22.

U.S. Copyright Office. 2023. Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence. https://www.copyright.gov/ai/ai_policy_guidance.pdf. Accessed: 2025-04-19.

U.S. Copyright Office. 2024. Copyright and Artificial Intelligence, Part 1: Digital Replicas. https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf. Accessed: 2025-04-19.

U.S. Copyright Office. 2025a. Copyright and Artificial Intelligence, Part 2: Copyrightability. https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf. Accessed: 2025-04-19.

U.S. Copyright Office. 2025b. Copyright and Artificial Intelligence, Part 3: Generative AI Training (Pre-Publication Version). https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf. Accessed: 2025-05-20.

U.S. Department of Homeland Security. 2021. Increasing Threat of Deepfake Identities. https://www.dhs.gov/sites/default/files/publications/increasing/threats_of_deepfake_identities_0.pdf. Accessed: 2025-04-15.

U.S. Department of Labor. 2025. Powered Industrial Trucks (Forklift) eTool. https://www.osha.gov/etools/powered-industrial-trucks/training. Accessed: 2025-05-22.

U.S. Federal Communications Commission. 2025. 'Grandparent' Scams Get More Sophisticated | Federal Communications Commission. https://www.fcc.gov/consumers/scam-alert/grandparent-scams-get-more-sophisticated. Accessed: 2025-04-20.

Vahl, S. 2024. Cloned customer voice beats bank security checks. https://www.bbc.com/news/articles/c1lg3ded6j9o. Accessed: 2025-04-20.

van den Oord, A.; Dieleman, S.; Zen, H.; Simonyan, K.; Vinyals, O.; Graves, A.; Kalchbrenner, N.; Senior, A.; and Kavukcuoglu, K. 2016. WaveNet: A Generative Model for Raw Audio. https://arxiv.org/abs/1609.03499. arXiv:1609.03499.

Vatreš, A. 2021. Deepfake Phenomenon: An Advanced Form of Fake News and Its Implications on Reliable Journalism. *Društvene i humanističke studije*, 6(3(16)): 561–576. Number: 3(16).

Wang, H.; Wu, P.; Rosa, K. D.; Wang, C.; and Shrivastava, A. 2023. Multimodality-guided Image Style Transfer using Cross-modal GAN Inversion. https://arxiv.org/abs/2312.01671. arXiv:2312.01671.

Watson, A. 2025. The uncomfortable tension caused by AI in artistic non-realism. https://technews.iit.edu/2025/03/06/the-uncomfortable-tension-caused-by-ai-in-artistic-non-realism/. Accessed: 2025-04-15.

Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D.; Chi, E. H.; Hashimoto, T.; Vinyals, O.; Liang, P.; Dean, J.; and Fedus, W. 2022. Emergent Abilities of Large Language Models. https://arxiv.org/abs/2206.07682. arXiv:2206.07682.

Wei, M. 2024. The Psychological Effects of AI Clones and Deepfakes. https://www.psychologytoday.com/us/blog/urban-survival/202401/the-psychological-effects-of-ai-clones-and-deepfakes. Accessed: 2025-04-20.

Wiggers, K. 2025a. Deepfake videos are getting shockingly good. https://techcrunch.com/2025/02/04/deepfake-videos-are-getting-shockingly-good/. Accessed: 2025-04-19.

Wiggers, K. 2025b. People are using Google's new AI model to remove watermarks from images. https://techcrunch.com/2025/03/17/people-are-using-googles-new-ai-model-to-remove-watermarks-from-images/. Accessed: 2025-04-21.

Xu, Z.; Liu, Y.; Deng, G.; Li, Y.; and Picek, S. 2024. A Comprehensive Study of Jailbreak Attack versus Defense for Large Language Models. https://arxiv.org/abs/2402.13457. arXiv:2402.13457.

Yu, Z.; Zhai, S.; and Zhang, N. 2023. AntiFake: Using Adversarial Audio to Prevent Unauthorized Speech Synthesis. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, CCS '23, 460–474. New York, NY, USA: Association for Computing Machinery. ISBN 979-8-4007-0050-7.

Zajonc, R. B. 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2, Pt.2): 1–27. Place: US Publisher: American Psychological Association.

Zhou, T.; and Zhang, C. 2024. Examining generative AI user addiction from a C-A-C perspective. *Technology in Society*, 78: 102653.

Zhu, J.-Y.; Park, T.; Isola, P.; and Efros, A. A. 2020. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. https://arxiv.org/abs/1703.10593. arXiv:1703.10593.