Preservation vs. Fabrication: An Ethical Framework of Consent, Transparency, and Integrity for Posthumous AI Art

Xiantao Zhang
Beihang University
zhangxiantao@buaa.edu.cn

Abstract

Generative AI's evolution from stylistic mimicry to the digital resurrection of deceased artists poses a profound ethical challenge to artistic integrity and cultural heritage. This position paper argues that posthumous creation by AI becomes a fundamental transgression when it moves beyond restoration to fabricate new works. We draw a sharp distinction between the legitimate use of AI as a ***tool*** *to realize an artist's verifiable intent and its ethically impermissible use as an* ***agent*** *of resurrection that fabricates new works. Our argument draws on a philosophical critique of authenticity and Walter Benjamin's "Aura," a review of inadequate legal frameworks, and a case study typology that illustrates a spectrum of ethical acceptability. In response, we propose a normative ethical framework grounded in three core principles: (I) the primacy of the artist's explicit, prior consent; (II) mandatory and permanent transparency in labeling; and (III) a Principle of Integrity, enforced by a Non-Transcendence Rule which dictates that AI may be used to fulfill intent but never to invent it. This framework offers clear, actionable guidance for stakeholders to ensure that technology serves to preserve, rather than corrupt, the authentic human essence of our artistic legacies.*

1. Introduction

1.1. From Stylistic Mimicry to Digital Resurrection

Generative Artificial Intelligence has reached an ethically fraught crossroads in the art world: the digital resurrection of deceased artists to create new works. This practice has evolved rapidly from the comparatively simple "in the style of..." prompts of image generators like DALL-E [8] to the complex, disquieting reality of digital resurrections. Recent, widely discussed examples, such as the AI-driven reanimation of the late rapper Tupac Shakur [53] and a posthumous comedy special in the voice of George Carlin [31], highlight the urgency of this issue. This is not merely an incremental technological advance; it is a paradigm shift

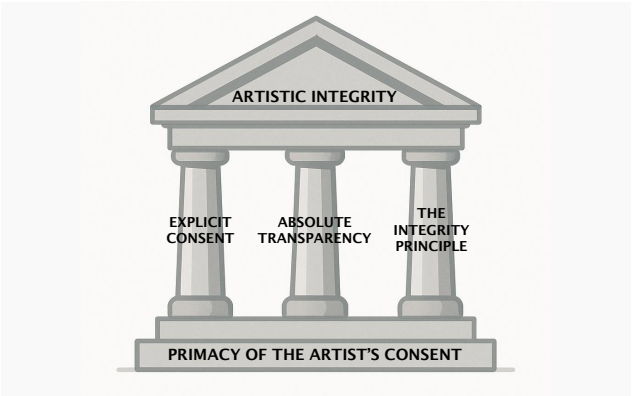

Figure 1. The proposed framework rests on the Primacy of the Artist's Consent, upheld by three principles: (I) Explicit Consent from the artist, (II) Mandatory Transparency in labeling, and (III) the Principle of Integrity, which dictates that AI may serve but never fabricate artistic intent.

that compels us to confront a stark question: **Do we have the right to create on behalf of the dead?**

The broader societal phenomenon of creating AI representations of the deceased has been explored under the term "Generative Ghosts" by Morris and Brubaker [39]. Their work outlines a comprehensive design space for these "AI afterlives" and systematically anticipates a wide range of benefits and risks for individuals and society. While their analysis covers the entire landscape, this paper narrows the focus to a specific, uniquely contentious domain: the use of AI to create new works of *art* in a deceased artist's name. We argue that this application poses distinct challenges to authenticity and cultural heritage that demand a specific *normative* ethical framework, rather than a descriptive one.

Our central contention is that this practice is not a benign tribute but a transgression that blurs the line between authentic legacy and artificial fabrication. We therefore propose a robust ethical framework to counter this threat. We argue that to preserve the core value of human creation, any posthumous AI generation must be governed by three inviolable pillars: (I) it must be grounded in the artist's explicit Consent; (II) it must be accompanied by absolute Trans-

parency for the public; (III) and most critically, it must adhere to a Principle of Integrity that forbids technology from transcending an artist's life's work to fabricate new intent—a line we call the *Non-Transcendence Rule*. This three-pillared framework is visualized in Figure 1.

1.2. Posthumous Generative Agents: A Shift from Tool to Agent

To establish a rigorous foundation for our argument, we must first define our terms. The practice of using artificial intelligence to simulate the deceased is sometimes referred to as "Digital Necromancy" [27] or "Digital Doppelgängers" [60], and Demetrius Floudas [20] identifies them as "Thanabots." Morris and Brubaker [39] describe this broader phenomenon as creating "AI afterlives" and specifically define an advanced AI agent known as "Generative Ghosts." More crucially, such *ghosts* can simulate a person's personality based on their digital footprint and generate entirely new content.

Within this broader category, this paper is specifically concerned with what we term "**Posthumous Generative Agents**": AI systems that do not merely parrot a deceased artist's existing data but generate novel, seemingly original artistic content. The potential applications for general AI afterlives are vast. They range from providing emotional solace to grieving relatives, a practice which carries its own significant mental health risks [39], to powering educational simulations. For our purposes, we focus on the creation of new art. This fusion of technology and remembrance gives rise to profound ethical problems: unauthorized identity appropriation, the absence of consent, and the potential for narrative distortion.

The core of the ethical dilemma is a paradigm shift in AI's role from a passive tool to an active agent, a distinction central to the challenges of AI afterlives [39]. Early applications, such as artistic style transfer [13, 14, 22, 24, 29, 32, 49, 50, 57], clearly positioned AI as a **tool** in the hands of a human creator. The emergence of *Posthumous Generative Agents*, however, transforms AI from a passive tool into an active **agent**.

When AI is a tool, it extends human creativity [44, 59]. But when AI becomes an agent that simulates a deceased artist and *creates* in their name, it usurps the unique **agency** that belongs to the artist alone. It is this usurpation that triggers a foundational crisis of authorship, creative intent, and artistic authenticity. The problem, therefore, is not the technology itself, but the qualitative change in its role in the creative process. This demands that our ethical and legal frameworks must evolve beyond evaluating the use of a tool and move towards regulating the impersonation by an agent. It is precisely this paradigm shift that necessitates the comprehensive framework of **Consent, Transparency, and Integrity** we advocate for throughout this paper.

2. The Philosophical Foundation

Understanding the ethical challenge of *Posthumous Generative Agents* requires a philosophical analysis of an artwork's intrinsic value. This section argues that the essence of art lies not in its replicable style, but in a trinity of concepts that AI cannot access: the artist's lived experience, the historical uniqueness Walter Benjamin called the "**Aura**" [7], and the expressive intent of the creator.

2.1. The Nature of Authenticity

The core tenet of our position is this: the value of an artwork is profoundly more than the sum of its aesthetic or stylistic properties, which an algorithm can learn and replicate. Its true **authenticity** is inextricably rooted in its origin within the artist's lived experience, emotions, intellect, and creative intent at a specific moment in time. A generative model can flawlessly reproduce a painter's brushstroke or a singer's timbre, but it cannot replicate the lived experience that informed them. An AI has never fallen in love, witnessed a sunset with a sense of wonder, endured a hangover, or confronted its own mortality. It possesses neither human frailty nor human struggle.

An AI's output, however virtuosic, is therefore a mimicry lacking the intrinsic link to human experience that gives art its soul. Research has suggested that even when AI-generated art is preferred on certain aesthetic criteria, human participants can still distinguish AI from human creations at a rate significantly higher than chance [26], implying that some ineffable quality of humanity or authenticity is perceptible to us. This perceived gap is often tied to the narrative behind the work; when confronted with AI-generated art, viewers seek the artist's emotional journey and backstory to grant it a value commensurate with traditional art, a connection that is inherently absent in posthumous fabrications [30].

2.2. The Total Collapse of Benjamin's *Aura*

To deepen this critique, we must revisit Walter Benjamin's landmark essay, *The Work of Art in the Age of Mechanical Reproduction* [7]. Benjamin defined the *Aura* of an artwork as its unique presence in time and space, its singular existence at the place where it happens to be. This *Aura* is tied to its history, its tradition, and the ritualistic context from which it emerged.

Nearly a century ago, Benjamin argued that mechanical reproduction initiated the "decay of the Aura" by detaching art from its original domain [7]. Generative AI, however, pushes this process to a radical conclusion, prompting a thorough re-examination of authenticity in the digital age [51]. The challenge AI poses is fundamentally different from that of mechanical reproduction. While Benjamin's focus was on works with clear human intentionality, AI introduces a "distributed agency" spread across the human

prompter, the algorithm, and the vast training data, creating novel interpretations rather than mere copies [6].

This leads not just to a decay of the Aura, but to its total collapse. Benjamin's theory described how a mechanical copy causes the original's Aura to wither. Generative AI, however, goes a step further: it creates not a copy that lacks Aura, but an "**Aura Vacuum**." An AI-generated artwork never possessed an Aura to begin with. Some scholars even argue that AI art operates as pure inauthenticity, deriving a new form of creativity from its very elusiveness and lack of a singular origin story [10]. While this may open new creative pathways for living artists, in the posthumous context, it results in a forgery born without history, context, or soul. Others contend that a new aura of information can emerge from the wide digital dissemination of AI works [23], but we argue this digital popularity cannot substitute for the authentic, embodied presence central to an artist's legacy. The challenge AI poses is therefore more fundamental than that of Benjamin's era. The debate, therefore, is no longer between an original and its copy, but about the very status of a cultural product born inherently inauthentic.

2.3. Expressive Authenticity

To further refine our analysis, we introduce the concept of "**Expressive Authenticity**" to distinguish true artistic creation from AI's pattern-matching. Genuine art, regardless of its form, originates from the creator's internal motivations, emotional drives, and the desire to communicate a specific message. It is the embodiment of the artist's agency—the culmination of their personal choices, intentions, and worldview. These very concepts of autonomy, authenticity, and authorship become fraught when applied to generative systems, which lack genuine intent and are instead driven by algorithms and training data [36].

AI-generated art, by contrast, is often criticized as soulless precisely because its creative impulse stems not from a deep expressive need, but from a fascination with technological novelty, the pursuit of commercial gain, or a rebellious stance against the art establishment. It mimics traditional forms but, as Marshall McLuhan might have argued, it fails to "massage" our senses in a new way because it pretends to be another medium rather than creating a new form of interaction [19, 37]. This fundamental disregard for creative intent renders the AI's output incapable of possessing true artistic value, even if it is technically flawless.

This valuation of authenticity is not merely elitist or academic; it is a deeply internalized cultural value. This is powerfully demonstrated by the public reaction to a specific case: the AI completion of Keith Haring's *Unfinished Painting* [41, 48, 56]. The public backlash was not simple anti-AI sentiment; it was an intuitive and visceral response to the gross violation of expressive authenticity and authorial intent. The public instinctively understood that the immense power of Haring's work lay precisely in its unfinished state—a heart-wrenching statement on the void left by the AIDS epidemic and his own impending death. The AI completion did not add value; it obliterated it. This case proves that any technology that bypasses or ignores the sanctity of artistic intent will face immense social and ethical resistance. This violation of artistic legacy is precisely the harm that our proposed Principle of Integrity, with its core *Non-Transcendence Rule*, is designed to prevent, providing a powerful, real-world justification for its necessity.

3. Legal Gaps and Limitations

A *Posthumous Generative Agent* operates in a legal landscape profoundly ill-equipped to govern it. Existing legal doctrines, designed in an era of human authors and tangible copies, reveal significant gaps and inadequacies when faced with the novel challenges of posthumous AI generation. This section argues that these frameworks suffer from a fundamental *tool mismatch*, requiring a solution beyond their scope.

3.1. A Legal Tool Mismatch

Applying current laws to posthumous AI creation is like trying to solve a 21st-century problem with 20th-century tools. The core legal tools at our disposal, including Copyright Law, Moral Rights, and the Right of Publicity, each fail in critical ways.

Copyright Law This framework falters at the most basic level. The United States Copyright Office, for example, has explicitly stated that works generated entirely by AI are not eligible for copyright protection because they lack the requisite **human authorship** [42, 43]. While this stance rightly defends the centrality of human creativity, it also leaves AI-generated content in a legal vacuum, creating uncertainty around ownership, usage rights, and infringement liability. Copyright law was designed to incentivize human creators; it is fundamentally mismatched for a world with non-human creators. The legal ambiguity extends to the very inputs of the process—whether training on copyrighted data constitutes fair use—and the outputs, which lack clear ownership, creating a difficult tension between copyright law and computer science practices [18, 21].

Moral Rights (Droit Moral) Originating in the French tradition, moral rights are perhaps the most relevant existing tool [2, 3]. They protect an author's personal, non-economic connection to their work. Chief among these is **the right of integrity**—the right to object to any distortion, mutilation, or other modification that would be prejudicial to the author's honor or reputation [25, 46]. Using an AI to generate new works in a deceased artist's name could certainly be

Aspect	United States	Europe (France)
Key Legislation/Principle	Visual Artists Rights Act (VARA) / Right of Publicity	Droit Moral (Moral Rights)
Duration of Moral Rights	Artist's lifetime only; Inconsistent post-mortem	Perpetual
Scope of Protection	Narrow; limited to specific visual arts	Broad; covers all work types
Implications for AI Generation	**Weak Protection.** Rights terminate at death, creating a major loophole. Publicity rights are inconsistent and commercially focused, offering unstable defense.	**Strong Protection.** Perpetual rights provide a permanent legal basis for heirs to block AI works that harm the artist's reputation or oeuvre integrity.

Table 1. A Comparative Analysis of Posthumous Moral Rights Frameworks.

argued as a violation of this right, an issue that sits at the heart of the debate over AI-created art [16]. However, as legal analysts point out, the strength and duration of these rights vary dramatically across jurisdictions, making them an unreliable shield against posthumous fabrication [34].

The Right of Publicity This right protects an individual's name, image, and likeness—their persona—from unauthorized commercial use [35]. It has been successfully invoked by estates, such as that of Tupac Shakur, to demand the removal of songs using AI-generated vocal impersonations [9]. However, its power is inconsistent. The right of publicity varies state by state in the U.S. and country by country elsewhere. Furthermore, its protection is often limited to commercial exploitation, leaving a grey area for works that claim to be expressive or non-commercial [1].

This tool mismatch demonstrates that a piecemeal application of existing laws is insufficient. These doctrines were built on the philosophical assumption of a human author and their original expression. AI agents shatter both premises, necessitating a new legal or ethical paradigm. This legal vacuum demonstrates that existing law is insufficient on its own, reinforcing the urgent need for the clear, proactive ethical guidelines provided by our proposed framework.

3.2. Divergent Rights: The U.S. and Europe

The inadequacy of the legal framework is thrown into sharp relief when comparing the dominant legal traditions of the United States and continental Europe, particularly France.

In the United States, the federal Visual Artists Rights Act of 1990 (VARA) [54] grants moral rights to creators, but its protections are notably narrow. First, its scope is limited to a small subset of works of visual art, excluding films, posters, and most digital media. Critically, VARA's moral rights terminate at the artist's death. This creates a vast legal loophole, effectively giving a green light to the posthumous manipulation and imitation of an artist's style as soon as they are deceased, a critical weakness highlighted in recent legal analyses [34].

In stark contrast, the French legal tradition treats moral rights (*droit moral*) [2, 3] as an extension of the author's personality. Consequently, these rights are considered "perpetual, inalienable, and non-transferable." Even after a work's copyright has expired and it has entered the public domain, the artist's heirs or a designated representative can perpetually defend the work's integrity and the artist's name. This robust, personality-centric protection provides a much stronger legal foundation for resisting the improper resurrection of an artist by AI, a point of contrast often noted when examining the potential for AI to infringe on moral rights [16].

This divergence stems from a fundamental difference in philosophy. The U.S. system tends to view an artist's rights as a form of **property right**—economic, alienable, and finite. The European model, however, emphasizes them as a **personal right**—an inalienable right tied to human dignity. A summary of these key legal differences is presented in Table 1. Since the core harm of AI resurrection is an affront to an artist's legacy, reputation, and dignity, the European focus on personality rights offers a far more appropriate and powerful model for global governance. Indeed, this rights-based philosophy provides a strong legal precedent that aligns directly with our framework's emphasis on the artist's personal agency, as embodied in the Principle of Consent and Principle of Integrity.

3.3. The Dawn of AI-Specific Legislation

Recognizing a growing legal void, some legislatures are beginning to act. A landmark example is California's Assembly Bill 1836 [11], which specifically addresses the use of AI to create digital replicas of deceased individuals.

The law aims to prohibit the unauthorized use of a *digital replica* of a deceased personality in an expressive work, such as a film or sound recording. It defines a *digital replica* as a realistic, computer-generated representation of an individual's voice or likeness where the individual did not perform. However, the legislation also reflects the difficult balancing act between protecting legacies and safeguarding free speech. It includes significant exemptions for works of news, public affairs, commentary, criticism, and parody.

The precise line between a prohibited fake creation and a permissible parody or criticism is sure to become a legal battleground. Nonetheless, the emergence of such laws

marks a critical legislative trend: an acknowledgment that old laws are not fit for purpose and that rules must be tailored for the age of the *Posthumous Generative Agent*.

4. Case Studies: A Typology of Posthumous AI Applications

To ground our theoretical discussion, this section examines three distinct categories of posthumous AI application. These examples are not merely anecdotal; they form a clear typology that reveals a spectrum of ethical acceptability, ranging from respectful restoration to profound violation. This typology is outlined in Table 2. By analyzing the AI's role and its relationship to the artist's original intent in each case, we can draw sharp, defensible ethical lines.

4.1. Restoration—The Ethically Acceptable

A positive and clear-cut example of AI's potential is the "Operation Night Watch" project at Amsterdam's Rijksmuseum [17, 47]. One of the project's goals was to reconstruct the sections of Rembrandt's 1642 masterpiece that were cut off in 1715 to fit the painting into a smaller space. The key to this project's ethical success is that the AI was not creating new content from scratch. Its reconstruction was strictly guided by two crucial data sources: first, an ultra-high-resolution scan of the existing original, which allowed the AI to deeply learn Rembrandt's unique style, colour, and lighting; and second, a complete, contemporary copy of the full painting created by Gerrit Lundens before the original was trimmed.

In this context, the AI's role was that of a highly sophisticated **Restoration Tool**. It translated the composition from the Lundens copy into the style of Rembrandt, even correcting for minor perspectival distortions present in the copy. Its objective was to *restore*, not to *create*. The process was rigorously constrained by historical evidence with the clear goal of recovering the artist's verifiable, original intent. From an ethical standpoint, this application is laudable, as it deepens our understanding and appreciation of a cultural treasure's original state. This case serves as a perfect example of our Principle of Integrity in action: the AI serves, but does not transcend, the artist's verifiable intent, making it an ethically sound application.

4.2. Completion—The Ethically Unacceptable

In stark contrast to the Rembrandt restoration, the AI completion of Keith Haring's *Unfinished Painting* serves as a prime example of ethical transgression [41, 48, 56]. The artistic power of this piece resides precisely in its unfinished state. Created in the final stages of his life, Haring deliberately left most of the canvas blank, with his iconic figures confined to one corner, dripping paint at their edges. This vast emptiness is not an absence of content; it *is* the content—a powerful and heartbreaking statement on the void left by lives, including his own, cut short by the AIDS epidemic. The work's meaning is embodied in this conceptual gesture.

The AI-driven completion, which filled the blank space with more Haring-style figures, completely ignored and thereby nullified this core artistic concept. In this instance, the AI functioned as a **Concept Destroyer**. It demoted a profound work of conceptual art, rich with pain, protest, and historical context, into a purely decorative image. The fierce public and critical backlash, denouncing the act as a desecration and disrespectful, powerfully demonstrates that this application was ethically indefensible. It shows that when AI is applied without any understanding of context or intent, it causes profound harm. This type of application, where AI is used to complete or even uncover works, directly engages with complex questions of moral rights and authorial integrity [16]. This violation is a textbook example of why our *Non-Transcendence Rule* is an essential ethical line that must not be crossed.

4.3. Resurrection—The Most Egregious Violation

This final category represents the most extreme ethical breach, where AI is used to resurrect a deceased artist to create entirely new content. Recent examples are rife, and have been noted as key points of concern in the academic discourse [39] on this topic: a posthumous comedy special released in the voice of George Carlin [5, 15, 31], a Drake song featuring an AI-generated Tupac Shakur vocal [4, 45], and videos of the late singer Coco Lee "singing" new songs on social media [28, 58].

In these cases, the AI's role is that of an **Identity Impersonator**. It does not restore or complete; it fabricates intent where none existed, creating a false artistic history and a deceptive performance. While a minority of fans may see this as a novel form of tribute, the dominant public reaction and the explicit condemnation from family members reveal a widespread ethical consensus against this practice. It is viewed not only as a deep disrespect to the deceased but also as a devaluation of the creative process itself. This form of digital necromancy, which goes far beyond restoration or evidence-based completion, constitutes a clear and untouchable ethical red line. It represents a flagrant violation of all three of our proposed principles: it proceeds without specific Consent, it is inherently deceptive and lacks Transparency, and it completely shatters the Principle of Integrity by fabricating intent wholesale.

5. Position and Framework

Based on the philosophical, legal, and practical harms identified, we now propose a constructive ethical framework. A reactive posture, relying on outdated legal tools and case-by-case public outcry, is insufficient to address this systemic challenge. What is required is a proactive, clear, and robust

Aspect	Restoration	Completion	Resurrection
Case Study	Rembrandt, *The Night Watch*	Keith Haring, *Unfinished Painting*	AI Tupac / Coco Lee
AI's Role	Restorative Tool	Conceptual Sabotage	Identity Impersonator
Relation to Artist's Intent	**Fulfills** a verifiable original intent.	**Violates** the work's established conceptual intent.	**Fabricates** new, un-authored intent.
Core Ethical Issue	Fidelity and historical accuracy.	Violation of conceptual integrity and authorial intent.	Identity appropriation, familial distress, and legacy distortion.
Ethical Acceptability	**High:** Recovers history based on verifiable data.	**Very Low:** Destroys the work's core conceptual meaning.	**Very Low:** Violates artist's agency and forges cultural history.

Table 2. A Typology of Posthumous AI Art Applications.

ethical framework that can guide artists, their estates, technology developers, and cultural institutions alike.

This need for proactive governance is not unique to artistic applications but reflects a global consensus on the principles required for trustworthy AI. Major technology companies, academic institutions, and governmental bodies have established ethical frameworks built on core values like transparency, fairness, and accountability [12, 38]. National standards bodies, such as NIST in the United States, have developed risk management frameworks to embed these principles into the entire AI lifecycle [40, 52, 55]. Our proposed framework builds on this global foundation, translating these high-level principles into specific, actionable rules tailored to the unique ethical challenges of posthumous artistic creation.

We propose a framework built on three interdependent and inviolable principles: **Explicit Consent**, **Absolute Transparency**, and **Unyielding Integrity**. This framework is designed not to stifle technology, but to channel it—to ensure AI serves as a tool for preservation and understanding, rather than as an agent of forgery and distortion.

5.1. The Principle of Consent

Core Requirement: Any posthumous use of generative AI to restore, complete, or create in an artist's name must be authorized by the artist's explicit, informed, and specific prior consent.

This principle is not merely a theoretical construct; it reflects the expressed desires of working artists. A recent large-scale survey of over 450 artists found that the vast majority believe in the importance of transparency and informed consent regarding the use of their work in training AI models [33]. Their views underscore a professional consensus: artists demand a say in how their creative identity is used, reinforcing the ethical necessity of making their consent the primary requirement.

This principle establishes the artist's intent as the absolute, paramount ethical consideration. General or vague permissions are inadequate to navigate the complexities of generative technology. Instead, authorization must be treated as a form of **digital creative will**—an advance directive for one's artistic legacy that clearly delineates the scope, nature, and limits of any permissible AI intervention.

Crucially, while legal consent from an estate or rightsholder is a necessary procedural step, it is **ethically insufficient** as a standalone justification. The artist's legacy is not merely property to be managed; it is the extension of a unique creative consciousness. Heirs and estate managers, while often well-intentioned, face inherent conflicts of interest, including financial incentives, public pressure, and the simple impossibility of truly divining an artist's wishes regarding technologies they never encountered. Their role is one of stewardship, not of posthumous co-authorship. Therefore, in the absence of explicit and specific directives from the artist, the only ethically defensible position is a **presumption of prohibition**. To create on behalf of the dead without their express permission is an act of profound ethical overreach.

5.2. The Principle of Transparency

Core Requirement: Any legitimately authorized work involving posthumous AI generation must be labeled clearly, prominently, and permanently, identifying the nature and extent of the AI's contribution.

This principle is a non-negotiable safeguard for cultural and historical integrity. Its purpose is twofold. First, it is an act of intellectual honesty that protects the public, art historians, and the market. It prevents AI-generated artifacts from being mistaken for, or devaluing, an artist's authentic oeuvre, thereby preventing the **corruption of the historical record**. An art historian a century from now must be able to distinguish a true Rembrandt sketch from a sophisticated fabrication.

Second, transparency is the mechanism that makes accountability possible. It ensures that future generations understand the distinction between a human cultural legacy and a technologically-derived product. Labels such as "AI-assisted restoration," "AI-completed based on existing plans," or "AI-generated in the style of [Artist's Name]" are not disclaimers but vital pieces of provenance. Without

this bright-line disclosure, we risk a slow, digital poisoning of our cultural well, where the authentic and the artificial become indistinguishable. Transparency does not diminish the work; it fortifies the truth.

5.3. The Principle of Integrity

Core Requirement: AI intervention must be strictly limited to serving and realizing an artist's established and verifiable intent. It must never be used to transcend their life's work by fabricating new, un-authored intent.

This principle is the cornerstone of the entire framework, creating a clear operational boundary that directly addresses the philosophical harms of inauthenticity and the practical violations seen in our case studies. It codifies the distinction between respectful service and unethical fabrication. This boundary is defined by the **Non-Transcendence Rule**:

- **Permissible (Serving Intent):** AI may be used as a high-fidelity instrument for restoration where the original state is known (as with Rembrandt's *The Night Watch*). It may also be used to complete an unfinished work, but only where there is an exceptionally high degree of certainty derived from the artist's own detailed sketches, notes, or architectural plans. In these cases, AI is subordinate to and serves a pre-existing, verifiable intent.
- **Impermissible (Fabricating Intent):** AI must be forbidden from functioning as a proxy creator to generate entirely new works from scratch. This includes composing "new" songs for a deceased singer, painting canvases in the name of a deceased artist, or writing new material for a deceased comedian. Such acts constitute a form of digital forgery and the fabrication of cultural history. They violate the fundamental principle that art is tethered to an artist's lived experience, agency, and singular vision. The AI-driven completion of Keith Haring's *Unfinished Painting* stands as the quintessential cautionary tale, demonstrating the profound conceptual damage inflicted when this Non-Transcendence Rule is broken.

By enforcing this rule, we protect the very essence of artistic legacy: the irreplaceable link between a creator and their creation.

6. Limitations and Future Directions

A rigorous inquiry demands acknowledging the limitations of our proposed framework. These boundaries do not invalidate our core position; rather, they highlight the complexities of the issue and map out essential avenues for future research, debate, and policy development. This section identifies four key areas where our proposed framework invites further refinement and discussion.

6.1. The Historical Consent Paradox

The central pillar of our framework is the Principle of Explicit Consent, which functions effectively for contemporary and future artists who can create a digital creative will. However, this principle becomes anachronistic and creates a paradox when applied to historical figures. Artists such as Rembrandt, Mozart, or Shakespeare could not have provided explicit consent for technologies they could never have conceived.

A rigid application of our framework's presumption of prohibition would forbid even beneficial, historically-grounded applications like the Rembrandt restoration. This is not our intent. For pre-digital era artists, the ethical weight must therefore shift almost entirely to the **Principle of Integrity** and the **Principle of Transparency**. Any intervention would require an exceptionally high burden of proof, grounded in rigorous academic scholarship and historical evidence, to demonstrate that the AI is serving a verifiable, original intent. This limitation reveals that a one-size-fits-all application of the framework is insufficient. A critical future direction is for cultural heritage institutions, historians, and ethicists to develop a specialized protocol for historical figures that honors the spirit of our framework while adapting to the reality of the historical record.

6.2. The Ambiguity of the Creative Boundary

This paper draws a bright line—the *Non-Transcendence Rule*—between permissible completion and impermissible creation. In landmark cases like Keith Haring's painting, the distinction is stark. However, in practice, this boundary can be far more ambiguous.

Consider an author who dies leaving behind detailed notes, character outlines, and a half-finished manuscript for a novel, or a composer who leaves sketches and leitmotifs for a symphony but no full orchestration. At what point does an AI-assisted completion cross the line from fulfilling a detailed plan to fabricating new creative choices? Our framework provides the essential guiding principle, but it does not resolve every marginal case.

This limitation highlights the need for **governance structures and expert arbitration**. Who decides when the available source material is sufficient to justify a completion? A panel of peer experts? The artist's designated literary executor? This points to a future where artists' estates may need to appoint not just legal executors, but creative integrity committees to interpret an artist's wishes and adjudicate these complex boundary cases.

6.3. From Ethics to Enforcement

This paper deliberately articulates an *ethical* framework, not a detailed *legislative roadmap*. While we argue this ethical foundation should inform future law, a significant gap remains between proposing principles and achieving global,

enforceable regulation. As we have shown, the legal traditions in the United States and Europe are already divergent, and achieving international consensus on issues of copyright and moral rights is notoriously difficult.

Furthermore, generative AI models are often developed and deployed by transnational corporations, operating in a space that can outpace national legislation. The practical enforcement of principles like Absolute Transparency depends on the willingness of these technology companies to build in the necessary safeguards and on international bodies to establish shared standards. A key limitation is the monumental challenge of translating these ethical imperatives into harmonized, enforceable regulations. This is a critical task for policymakers, international legal scholars, and technology ethicists in the years to come.

6.4. AI Afterlives Beyond Art

Our analysis is primarily focused on the domain of artistic and cultural creation, where concepts of authenticity, intent, and legacy are paramount. However, the phenomenon of *AI afterlives* [39] extends far beyond this scope into other complex areas:

- **Education:** AI simulations of historical figures like Albert Einstein or Marie Curie for educational purposes.
- **Therapy:** "Grief-bots" that simulate a deceased loved one to provide solace, an application area with its own host of potential benefits and mental health risks [39].
- **Commercial Exploitation:** The use of a deceased celebrity's digital replica to endorse new products.

While our three principles—Consent, Transparency, and Integrity—remain highly relevant in these contexts, their specific application and weighting would need to be recalibrated. For an educational bot, integrity might mean unwavering historical accuracy. For a therapeutic tool, integrity could involve psychological safety and preventing emotional dependency. For commercial use, the right of publicity and prevention of fraudulent endorsement become the central concerns. This paper's scope is intentionally focused, but a crucial next step is for researchers to adapt and apply this ethical framework to these other, equally challenging, applications of what Morris and Brubaker call *Generative Ghosts* [39].

7. Conclusion

Generative AI holds immense potential as a tool for cultural heritage, aiding in restoration and analysis. However, when it oversteps an artist's life and intent, it transforms from a tool into an agent of erosion, undermining the very integrity it has the potential to serve. Technological innovation must not come at the cost of artistic integrity, the seriousness of our cultural heritage, or the dignity of human creation.

We stand at a critical juncture, forced to choose between embracing technological novelty and defending the irreplaceable human core of art. While prior work has mapped the broad design space and socio-technical risks of what have been termed *Generative Ghosts* [39], this paper has argued for a specific, normative path forward for the domain of art. We have proposed a clear ethical framework built on the foundational principles of **Artist's Consent, Absolute Transparency, and a strict Non-Transcendence Rule of Integrity** to serve as a navigational guide. This framework is not anti-technology; rather, it is pro-humanity. It advocates for a paradigm where AI is confined to serving and realizing an artist's established, authentic intent, and is strictly prohibited from fabricating it.

The responsibility for navigating this future does not fall to any single group. We issue a call to action for a broad coalition of stakeholders to engage in this urgent dialogue. Artists must consider how they wish their legacies to be managed in this new era. Their estates and cultural institutions must act as responsible stewards. Technology developers bear a profound duty to design systems with ethical guardrails. And legislators must work to create legal protections that are fit for purpose. A collective commitment is required to establish responsible norms and safeguards before irreversible damage is done to our cultural record.

Ultimately, the goal is to ensure that AI technology serves as a responsible force that genuinely aids our understanding and preservation of art, not one that dilutes or supersedes its most precious quality. In the roar of the machine, we must listen for the authentic voice of human experience and expression. The *Posthumous Generative Agent* is a marvel of engineering, but it is an empty phantom. We must instead dedicate ourselves to protecting the true ghost in the machine: the indelible, irreplicable spirit of a human life, complete with its struggle, joy, intention, and complexity, which constitutes the very soul of art.

References

[1] Comedy III Productions, Inc. v. Gary Saderup, Inc. 4

[2] Berne convention for the protection of literary and artistic works, 1928. Rome Act, Article 6bis. 3, 4

[3] French intellectual property code, 2023. 3, 4

[4] Daniel Atherton. Incident number 792: Unauthorized use of ai to replicate tupac shakur's and snoop dogg's voices in drake's 'taylor made freestyle', 2024. 5

[5] Daniel Atherton. Incident number 627: Unauthorized ai impersonation of george carlin used in comedy special, 2024. 5

[6] Oshri Bar-Gil. The transformation of artistic creation: from benjamin's reproduction to ai generation. *AI & SOCIETY*, 2025. 3

[7] Walter Benjamin. *The Work of Art in the Age of Mechanical Reproduction*. Penguin UK, 2008. 2

[8] James Betker, Gabriel Goh, Li Jing, Tim Brooks, Jianfeng Wang, Linjie Li, Long Ouyang, Juntang Zhuang, Joyce Lee, Yufei Guo, Wesam Manassra, Prafulla Dhariwal, Casey Chu,

Yunxin Jiao, and Aditya Ramesh. Improving image generation with better captions. *OpenAI*, 2023. 1

[9] Jon Blistein. Tupac estate demands drake take down 'taylor made freestyle' over ai vocals. *Rolling Stone*, 2024. 4

[10] James Brusseau and Luca Turchet. How ai generates creativity from inauthenticity, 2025. 3

[11] California State Assembly. Assembly bill no. 1836: Use of likeness: digital replica, 2024. 2023-2024 Regular Session. Chaptered by Secretary of State, September 17, 2024. 4

[12] Center for Security and Emerging Technology. Translation white paper on trustworthy artificial intelligence. Technical report, Georgetown University, CSET, 2020. 6

[13] Haibo Chen, lei zhao, Zhizhong Wang, Huiming Zhang, Zhiwen Zuo, Ailin Li, Wei Xing, and Dongming Lu. Artistic style transfer with internal-external learning and contrastive learning. In *Advances in Neural Information Processing Systems*, pages 26561–26573. Curran Associates, Inc., 2021. 2

[14] Haibo Chen, Lei Zhao, Zhizhong Wang, Huiming Zhang, Zhiwen Zuo, Ailin Li, Wei Xing, and Dongming Lu. Dualast: Dual style-learning networks for artistic style transfer. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 872–881. Computer Vision Foundation / IEEE, 2021. 2

[15] Samantha Chery. George carlin's estate settles lawsuit over ai comedy special, 2024. 5

[16] Helen Catherine Darby. Ai-created art and the moral rights of authors. Columbia Journal of Transnational Law Bulletin, 2021. 4, 5

[17] Isis Davis-Marks. Lost edges of rembrandt's 'night watch' are restored using artificial intelligence. *Smithsonian Magazine*, 2021. 5

[18] Deven R. Desai and Mark Riedl. Between copyright and computer science: The law and ethics of generative ai, 2024. 3

[19] Quentin Fiore and Marshall McLuhan. *The medium is the massage*. New York: Random House, 1967. 3

[20] Demetrius Floudas. Digital necrolatry: Thanabots and the prohibition of post-mortem ai simulations. Submitted to EU AI Office's Plenary: Drafting the Code of Practice for General-Purpose Artificial Intelligence, 2025. 2

[21] Giorgio Franceschelli and Mirco Musolesi. Copyright in generative deep learning. *Data & Policy*, 4:e17, 2022. 3

[22] Tsu-Jui Fu, Xin Eric Wang, and William Yang Wang. Language-driven artistic style transfer. In *Computer Vision - ECCV 2022 - 17th European Conference, Tel Aviv, Israel, October 23-27, 2022, Proceedings, Part XXXVI*, pages 717–734. Springer, 2022. 2

[23] Nettrice Gaskins. The aura of ai-generated art. *Medium*, 2023. 3

[24] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. A neural algorithm of artistic style, 2015. 2

[25] Jane C. Ginsburg. Tale of two copyrights: Literary property in revolutionary france and america. *Tulane Law Review*, 64: 991, 1989. 3

[26] Anna Yoo Jeong Ha, Josephine Passananti, Ronik Bhaskar, Shawn Shan, Reid Southen, Haitao Zheng, and Ben Y. Zhao. Organic or diffused: Can we distinguish human art from ai-generated images? In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*, page 4822–4836, New York, NY, USA, 2024. Association for Computing Machinery. 2

[27] James Hutson and Jay Ratican. Life, death, and ai: Exploring digital necromancy in popular culture—ethical considerations, technological limitations, and the pet cemetery conundrum. *Metaverse*, 4(1), 2023. 2

[28] Bryan Ke. Chinese vlogger sparks outrage for using ai to 'resurrect' coco lee, godfrey gao. *Yahoo! News*, 2024. 5

[29] Dmytro Kotovenko, Artsiom Sanakoyeu, Sabine Lang, and Björn Ommer. Content and style disentanglement for artistic style transfer. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pages 4421–4430. IEEE, 2019. 2

[30] Peter Kun, Matthias Freiberger, Anders Sundnes Løvlie, and Sebastian Risi. Ai art perceptions with genframe – an image generating picture frame. In *DRS2024: Boston*. Design Research Society, 2024. 2

[31] Christopher Kuo. George carlin's estate sues podcasters over a.i. episode. *The New York Times*, 2024. 1, 5

[32] Yanghao Li, Naiyan Wang, Jiaying Liu, and Xiaodi Hou. Demystifying neural style transfer. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, page 2230–2236. AAAI Press, 2017. 2

[33] Juniper Lovato, Julia Zimmerman, Isabelle Smith, Peter Dodds, and Jennifer Karson. Foregrounding artist opinions: A survey study on transparency, ownership, and fairness in ai generative art, 2024. 6

[34] Carolyn Wimbly Martin and Isabel Jonson. Legal and ethical issues in posthumous art and artificial intelligence. Lutzker & Lutzker Insight, 2024. 4

[35] J.T. McCarthy. *The Rights of Publicity and Privacy*. Number 1. Thomson/West, 2005. 4

[36] Jon McCormack, Toby Gifford, and Patrick Hutchings. Autonomy, authenticity, authorship and intention in computer generated art, 2019. 3

[37] Marshall McLuhan. *Understanding media: The extensions of man*. MIT press, 1994. 3

[38] Microsoft Azure. What is responsible ai. https://learn.microsoft.com/en-us/azure/machine-learning/concept-responsible-ai, 2025. 6

[39] Meredith Ringel Morris and Jed R. Brubaker. Generative ghosts: Anticipating benefits and risks of ai afterlives, 2024. 1, 2, 5, 8

[40] National Institute of Standards and Technology (NIST). Ai risk management framework. https://www.nist.gov/itl/ai-risk-management-framework, 2023. 6

[41] Artnet News. A keith haring painting 'completed' using a.i. generates backlash. *Artnet News*, 2024. 3, 5

[42] U.S. Copyright Office. Copyright registration guidance: Works containing material generated by artificial intelligence. Federal Register, 88 FR 16190, 2023. 3

[43] U.S. Copyright Office. Letter re: Zarya of the dawn. Official letter from the U.S. Copyright Office, 2023. Registration # VAu001480196. 3

[44] Katherine O'Toole and Emőke Ágnes Horvát. Extending human creativity with ai. *Journal of Creativity*, 34(2):100080, 2024. 2

[45] Isaiah Richard. Drake diss track with AI tupac shakur, 'taylor made,' has been taken down, 2024. 5

[46] Sam Ricketson and Jane Ginsburg. *International Copyright and Neighbouring Rights: The Berne Convention and Beyond*. Oxford University Press, Oxford, UK, 2022. 3

[47] Rijksmuseum. For the first time in 300 years the night watch is complete again. Rijksmuseum Press Release, 2021. 5

[48] Kalhan Rosenblatt. Ai altered a keith haring painting about the aids crisis — and, for some, ruined its meaning. *NBC News*, 2024. 3, 5

[49] Manuel Ruder, Alexey Dosovitskiy, and Thomas Brox. Artistic style transfer for videos. In *Pattern Recognition - 38th German Conference, GCPR 2016, Hannover, Germany, September 12-15, 2016, Proceedings*, pages 26–36. Springer, 2016. 2

[50] Manuel Ruder, Alexey Dosovitskiy, and Thomas Brox. Artistic style transfer for videos and spherical images. *Int. J. Comput. Vis.*, 126(11):1199–1219, 2018. 2

[51] David Salas Espasa and Mar Camacho. From aura to semi-aura: reframing authenticity in ai-generated art—a systematic literature review. *AI & SOCIETY*, 2025. 2

[52] Ron Schmelzer and Kathleen Walch. A framework for trustworthy ai. Blog post, Project Management Institute, 2025. 6

[53] Hayley Tsukayama. How the tupac 'hologram' works. *The Washington Post*, 2012. 1

[54] United States Congress. Visual artists rights act of 1990, 1990. Codified at 17 U.S.C. § 106A. 4

[55] University of California, San Francisco. Trustworthy ai. https://ai.ucsf.edu/trustworthy, 2025. 6

[56] Elaine Velie and Rhea Nayyar. Ethical questions arise after ai "completes" keith haring painting. *Hyperallergic*, 2024. 3, 5

[57] Xin Wang, Geoffrey Oxholm, Da Zhang, and Yuan-Fang Wang. Multimodal transfer: A hierarchical deep convolutional neural network for fast artistic style transfer. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017. 2

[58] Bryan Wong. Chinese content creator uses ai to 'resurrect' coco lee, godfrey gao; calls it an 'expression of love'. *CNA Lifestyle*, 2024. 5

[59] Eric Zhou and Dokyun Lee. Generative artificial intelligence, human creativity, and art. *PNAS Nexus*, 3(3):pgae052, 2024. 2

[60] Hazem Zohny. The specter of corporate necromancy: Who controls the dead in the age of digital doppelgängers? *The American Journal of Bioethics*, 25(2):113–115, 2025. 2