Towards a New Copyright Paradigm for Generative AI: Bridging Human-Machine Creativity Through Legal and Policy Reform

Daiwon Hyun
Graduate School of Metaverse
Sogang University
dhyun@sogang.ac.kr

Sunho Park
Graduate School of Metaverse
Sogang University
pshmeta@sogang.ac.kr

Eutteum Kim
Graduate School of Metaverse
Sogang University
3kjacob@gmail.com

Abstract

Generative AI challenges copyright frameworks for cultural heritage preservation. We analyze U.S. and EU approaches to AI-generated content through comparative legal analysis, revealing fundamental inadequacies in addressing machine learning's distributed agency. Recent 2025 rulings including Bartz v. Anthropic establish that AI training's legal status depends on transformativeness, data provenance, and market harm rather than categorical rules.

We propose a five-principle harmonization framework: transparency through content provenance standards, protected research activities, tiered compensation for training and outputs, evidentiary requirements for creative control, and risk-based compliance. Analysis of recent policy developments including the U.S. Copyright Office's January 2025 guidance and EU's GPAI Code of Practice reveals emerging consensus around human authorship requirements. Our framework provides practical guidance for navigating divergent jurisdictional approaches while enabling international interoperability in AI-driven cultural preservation.

1. Introduction

The emergence of generative artificial intelligence marks a transformative moment in the relationship between technology and creativity. Systems like Stable Diffusion, DALL-E, and Midjourney have demonstrated unprecedented capabilities in generating artistic works that blur the distinction between human and machine authorship. This technological revolution arrives at a critical juncture for cultural heritage preservation, where institutions worldwide seek to digitize, analyze, and extend artistic legacies using computational methods. Yet the legal frameworks governing these activities remain rooted in pre-digital concepts of authorship, originality, and creative control.

The year 2025 has witnessed a dramatic evolution in the legal landscape surrounding AI and copyright. In February, the Thomson Reuters v. Ross Intelligence ruling suggested narrow protections for AI training on copyrighted materials. By June, however, the Bartz v. Anthropic and Kadrey v. Meta decisions introduced nuanced distinctions based on transformativeness and data provenance. Meanwhile, regulatory bodies have accelerated their responses: the U.S. Copyright Office released comprehensive guidance on AI copyrightability in January, the European Commission published its GPAI Code of Practice in July, and South Korea issued detailed copyright registration guidelines in June. These rapid developments underscore both the urgency and complexity of establishing coherent legal frameworks for AI-generated content.

This paper addresses a fundamental question: how can international copyright law evolve to accommodate AI-driven cultural preservation while maintaining appropriate protections for human creativity? We examine this challenge through the lens of the Workshop on Cultural Continuity of Artists (WCCA) and its André Kim fashion legacy project, which exemplifies the practical tensions between technological innovation and legal constraints in cultural heritage preservation.

1.1. Research Context and Motivation

The intersection of generative AI and cultural heritage preservation represents both immense opportunity and significant legal risk. Modern diffusion models and transformer architectures can learn from artistic corpora to generate novel works that capture stylistic essence while producing original compositions [26]. This capability has profound implications for cultural institutions seeking to preserve and extend artistic legacies in the digital age.

Consider the André Kim fashion legacy project undertaken by the Workshop on Cultural Continuity of Artists (WCCA). The project encompasses over 10,000 high-resolution garment scans, thousands of design sketches, and extensive archival footage documenting five decades of Korean haute couture innovation. Training generative models on this dataset could enable unprecedented exploration of Kim's design philosophy, potentially generating new designs that honor his aesthetic while addressing contemporary fashion challenges. Yet each step in this process, from data ingestion to model training to output generation, raises unresolved questions about copyright ownership, fair use boundaries, and the legal status of AI-mediated creativity [13].

The urgency of addressing these questions extends beyond individual projects. Museums, archives, and cultural institutions worldwide are investing billions in digitization initiatives. Without clear legal frameworks, these investments risk creating vast digital collections that cannot be fully utilized for AI-driven research, education, or creative exploration. The resulting legal uncertainty chills innovation and potentially undermines the fundamental mission of cultural preservation: ensuring that artistic heritage remains accessible and relevant for future generations.

1.2. Problem Statement

The fundamental challenge lies in reconciling computational creativity with legal frameworks designed for human authorship. This incompatibility creates cascading uncertainties that affect every stakeholder in the AI ecosystem.

The training phase presents immediate legal risks. When a diffusion model ingests thousands of copyrighted artworks to learn stylistic patterns, does this constitute fair use or infringement? The 2025 U.S. court decisions offer contradictory guidance: Ross Intelligence prohibits competitive use, while Bartz v. Anthropic permits transformative learning from legally acquired data. European regulations provide clearer exceptions for research but impose opt-out requirements that create compliance complexity. South Korean law demands explicit licensing for commercial applications. This jurisdictional fragmentation forces developers to navigate a legal minefield where identical technical operations carry vastly different legal consequences across borders.

The output phase introduces additional complexity. If an AI system trained on André Kim's designs generates a novel garment, who owns the copyright? The human prompter who provided creative direction? The AI company whose model produced the output? The estate of André Kim whose works informed the model's understanding? Or perhaps no one, if the work lacks sufficient human authorship? The January 2025 U.S. Copyright Office guidance requires "substantive human creative control," but provides limited clarity on what constitutes sufficient human involvement in iterative AI workflows.

Platform liability compounds these uncertainties. A museum hosting an interactive "design in André Kim's style" exhibit faces reactive notice-and-takedown obligations in the United States but proactive filtering requirements under EU Article 17. The same technical system thus requires fundamentally different architectures to achieve compliance across jurisdictions, dramatically increasing development costs and operational complexity.

1.3. Research Contributions

This paper makes three primary contributions to the intersection of AI technology and copyright law:

First, we provide a systematic comparison of U.S. and EU approaches to AI-generated content, analyzing key differences in fair use doctrine, transformative use standards, and platform liability frameworks. This analysis extends beyond abstract legal principles to examine practical implications for AI system design and deployment.

Second, we develop a structured framework for evaluating legal risks associated with AI-driven cultural heritage projects. This framework considers factors including the commercial nature of use, jurisdictional scope, and technical architecture to provide actionable guidance for project planning.

Third, we propose five principles for international interoperability that balance innovation incentives with creator rights protection. These principles are designed to serve as minimum common standards that can be implemented across jurisdictions while respecting local legal traditions.

1.4. Paper Organization

The remainder of this paper is structured as follows: Section 2 analyzes the specific intellectual property challenges arising from AI-generated content in cultural heritage contexts. Section 3 provides a detailed comparison of U.S. and EU regulatory approaches with emphasis on practical implications. Section 4 presents our proposed framework for international harmonization. Section 5 discusses implications and future research directions.

2. Intellectual Property Challenges in AI-Generated Cultural Heritage

The collision between generative AI and copyright law exposes fundamental incompatibilities in how we conceptualize creativity, ownership, and cultural preservation. Each phase of the AI pipeline (data acquisition, model training, and content generation) challenges established legal principles while creating new categories of risk that existing frameworks cannot adequately address.

2.1. Legal Status of Training Data

Neural network training involves creating multiple copies of protected works through digitization, preprocessing, and storage [19]. Each operation potentially constitutes reproduction that may infringe copyright absent applicable exceptions. U.S. courts consider such uses potentially transformative when serving different purposes than original expressive functions, while European jurisprudence maintains stricter reproduction rights boundaries requiring explicit statutory exceptions.

The distinction between commercial and non-commercial use significantly impacts legal analysis. Academic research may benefit from broader exceptions under U.S. fair use and EU text/data mining provisions (Article 3, CDSM). However, commercial AI development faces stricter scrutiny. Article 4 CDSM permits rights holder opt-out from commercial mining, potentially requiring negotiated licenses. This creates uncertainty for hybrid research-commercial projects.

2.1.1. Impact of Model Architecture on Copyright Risk

Architecture choice carries significant copyright implications [27]. GANs present advantages for copyright-conscious applications: effective results with smaller curated datasets facilitating provenance tracking, adversarial training encouraging novel generation reducing substantial similarity risk, and suitability for style-specific applications [24].

Diffusion models require massive datasets precluding individual rights clearance. Studies demonstrate higher memorization rates, particularly for repeated training images [3]. The iterative denoising process creates direct mathematical relationships between training data and outputs, strengthening derivation claims. Economic analyses suggest differentiating "small" models with licensed datasets from "large" models requiring collective licensing [12]. For the André Kim project, choosing between a licensed GAN versus large diffusion model represents fundamental legal strategy.

2.2. Derivative Works and Transformative Use

When models train predominantly on single artists' corpora, outputs may exhibit substantial similarity to protected elements. Courts assess total concept, sequence, arrangement, and distinctive creative choices. Application to AI content remains untested.

Recent research demonstrates measurable memorization and verbatim reproduction [3]. This "regurgitation" correlates with model size and data repetition frequency. Extraction attacks recover exact copyrighted passages, code snippets, and visual elements. This evidence undermines AI-human learning analogies: the Copyright Office notes humans retain "imperfect impressions" while AI creates "perfect copies" [28]. Rights holders can now demonstrate protected content presence through extraction methods, strengthening substantial similarity claims as in The New York Times v. OpenAI case.

Memorization necessitates technical safeguards including differential privacy, deduplication, and output filtering, creating tension between compliance and capability. Whether trained models constitute derivative works affects sharing and commercialization.

2.3. Style Appropriation and Artistic Identity

While copyright doesn't protect style per se, AI systems' high-fidelity stylistic reproduction challenges this boundary [11]. Contemporary models capture underlying design principles, compositional strategies, and creative decision-making processes constituting distinctive artistic approaches, raising questions about creator protection adequacy.

2.4. Liability Distribution in AI Systems

Multi-actor development complicates liability frameworks. Direct infringement requires volitional conduct, but determining sufficient volition becomes complex across users providing prompts, platforms providing access, developers creating systems, and data providers. Secondary liability applies based on knowledge, material contribution, and control ability. Platform terms often include broad licenses and indemnification clauses shifting risk to users while retaining commercial benefits, with enforceability largely untested for AI-generated works.

These challenges demonstrate current frameworks' inadequacy for AI-generated cultural heritage content. The following section examines jurisdictional attempts to resolve these issues through divergent regulatory approaches.

3. Comparative Analysis of U.S. and EU Regulatory Frameworks

The regulatory treatment of AI-generated content diverges significantly between the United States and European Union, reflecting fundamental differences in legal philosophy. The U.S. maintains a utilitarian framework emphasizing innovation through flexible fair use doctrine and ex post judicial review, while the EU adopts a prescriptive authors' rights approach with comprehensive ex ante regulation [4].

3.1. Text and Data Mining Approaches

Under U.S. copyright law, the fair use doctrine (17 U.S.C. §107) provides potential flexibility for machine learning applications through a four-factor test. However, the legal landscape has become increasingly fractured and fact-specific. Thomson Reuters v. Ross Intelligence (February

2025) initially suggested narrow protection, with the court rejecting fair use defense for an AI legal research tool that directly competed with original products [9]. The court found Ross's purpose was to create a market substitute for Westlaw's legal headnotes, harming potential licensing markets.

Subsequent rulings in June 2025 dramatically complicated this analysis. In Bartz v. Anthropic and Kadrey v. Meta, Northern District of California courts found that training large language models on copyrighted works can constitute ”spectacularly” and ”highly transformative” fair use [25]. These courts distinguished their findings from Ross, reasoning that while Ross created direct market substitutes, the LLMs used copyrighted books for the fundamentally different purpose of learning to understand and generate human language [22].

Critically, the Bartz ruling introduced data provenance as a key factor. The court ruled that Anthropic's use of legally purchased books for training was fair use, but creation of a digital library from pirated sources was not [17]. This establishes that lawfully acquired training data receives stronger fair use protection than pirated materials.

The emerging U.S. framework now balances three critical factors: (1) the degree of transformativeness (learning and generating novel outputs versus simple repackaging), (2) the provenance of training data (legally acquired versus pirated), and (3) demonstration of concrete, non-hypothetical market harm [15]. AI systems generating fashion works from André Kim's designs would face this complex, highly fact-dependent analysis rather than categorical prohibition.

The EU codifies TDM exceptions in CDSM Articles 3 and 4. Article 3 provides mandatory exception for research organizations conducting scientific research, while Article 4 allows broader TDM but permits rights holder opt-out through machine-readable means. This creates uncertainty for commercial AI development as rights holders can retroactively restrict training use.

3.2. Transformative Use vs. Enumerated Exceptions

U.S. courts recognize transformative use adding new expression or meaning as favoring fair use (Campbell v. Acuff-Rose Music) [30]. This potentially protects AI systems generating creative outputs serving different purposes than training data. European copyright law lacks equivalent doctrine, providing only closed list of exceptions (quotation, parody, teaching). AI-generated content outside these narrow exceptions may constitute infringement regardless of transformative nature.

3.3. Platform Liability

Section 512 DMCA provides U.S. platforms safe harbor through notice-and-takedown procedures. Article 17 CDSM requires EU platforms to obtain authorization or ensure unavailability of specific works, effectively mandating content filtering. The Digital Services Act adds systemic risk assessment requirements for larger platforms.

3.4. South Korea's Hybrid Framework

South Korea's comprehensive approach, crystallized in June 2025 guidance from the Ministry of Culture, Sports and Tourism and Korea Copyright Commission, offers a bridge between U.S. flexibility and EU prescriptiveness [14]. The ”Generative AI Works Copyright Registration Guide” and ”Copyright Dispute Prevention Guide for Generative AI Results” provide concrete standards for distinguishing human versus AI contributions [18].

The framework requires ”evident human creative input” for copyright protection. Pure AI outputs lack copyright, but works demonstrating human modification, selection, arrangement, or highly creative iterative prompting processes may qualify [21]. The guidance emphasizes documenting the creative process to establish authorship and recommends AI businesses implement technical safeguards including filtering and deduplication, secure licenses for training data, and clarify liability in terms of service [18].

This directly applies to the André Kim project: simple prompt-based outputs remain unprotected, but designer-modified works incorporating creative choices receive protection. The Korean framework's emphasis on process documentation aligns with emerging international consensus on establishing human creative control.

The divergence necessitates careful project design. Research initiatives benefit from clearer exceptions, while commercial platforms require differentiated architectures for compliance. South Korea's graduated approach suggests emerging consensus around human creativity primacy, transparency requirements, and proportionate liability, informing the harmonization framework presented next.

4. Framework for International Harmonization

The regulatory divergence creates significant challenges for AI systems in cultural heritage contexts. This section proposes harmonized principles enabling international interoperability while respecting jurisdictional differences, adopting a ”minimum common denominator” approach that facilitates cross-border cooperation without requiring fundamental legal restructuring.

Table 1. Comparative Analysis and Implications for WCCA André Kim Project

Legal Dimension	United States (Fair Use)	European Union (Authors' Rights)	South Korea (Hybrid)	Implications for WCCA (André Kim Legacy)
Training Data Use	Highly fact-dependent and contested. Non-generative, market-substitutive use disfavored (Ross) [9]. Transformative use for LLM training may be fair use, especially with legally acquired data (Bartz, Kadrey) [25]. Pirated data source is high-risk [17]	Prescriptive with opt-out. Research exempt (Art. 3 CDSM). Commercial use subject to machine-readable opt-out (Art. 4) [10]. GPAI models must provide sufficiently detailed summary of training data (AI Act, Art. 52) [2]	Promotional but requires license. AI Act promotes data use, but Copyright Act applies. June 2025 guidance urges licensing for commercial use and public disclosure of training data sources [18]	High risk for commercial use without license. Research use of dataset likely permissible. Commercial model to sell new fashion designs faces complex, fact-specific fair use analysis in US, requires navigating Art. 4 opt-out in EU, and likely requires license in SK. Data provenance is critical
Authorship of Output	Requires substantive human creative control. Prompts alone are insufficient. Human modifications, selection, or arrangement of AI output are key (Jan 2025 USCO Report) [31]	Human originality required. Work must reflect the author's own intellectual creation and personality. Pure machine outputs are unprotectable	Requires evident human creative input. Pure AI outputs uncopyrightable. Creative prompting, selection, and modification can confer authorship (June 2025 KCC Guide) [18]	Documentation is paramount. To claim copyright in Kim-inspired designs, the WCCA project must rigorously document the human designer's iterative prompting, selection rationale, and post-processing modifications. The SK/US standards are nearly identical and provide clear roadmap
Platform Liability	DMCA safe harbor (reactive). Notice-and-takedown procedures for infringing content	Proactive obligations. Art. 17 CDSM requires best efforts to prevent infringement [10]. AI Act adds systemic risk assessments for large platforms	Proactive obligations. AI Act requires user notification for generative AI and proportionate risk management [14]	A platform hosting a Create your own André Kim design tool faces proactive filtering/risk management obligations in EU/SK, but reactive takedown duties in US. Legal burden varies dramatically by region

4.1. Principle 1: Transparency Through C2PA Standard

The Coalition for Content Provenance and Authenticity (C2PA) standard provides foundation for AI content transparency through tamper-evident provenance chains [5]. Implementation requires addressing technical and infrastructural challenges beyond standard adoption.

C2PA employs Content Credentials tracking model identification, generation metadata, and training data provenance. However, AWS's Sinclair Broadcasting implementation revealed architectural trade-offs between containerized services and serverless functions affecting cost, latency, and scalability [1].

Significant technical and practical challenges limit C2PA's effectiveness. A 2025 RAND Corporation analysis argues that C2PA's threat model is outdated, lacks rigorous security analysis against GAI-specific threats, and falls short of addressing many pressing risks currently threatening media including GAI content [23]. The critical vulnerability that metadata can be easily stripped by non-compliant platforms or users creates fundamental weakness in the provenance chain [7]. This necessitates external manifest stores with robust security, potentially using distributed ledger technology.

For cultural heritage contexts like the André Kim dataset, additional challenges emerge. High-resolution scans of design documents and archives cannot embed C2PA manifests directly. Instead, they require separate sidecar files, complicating provenance tracking and increasing risk of manifest-asset dissociation [6]. This technical limitation particularly affects non-media file types central to cultural preservation projects.

Effectiveness depends on achieving end-to-end ecosystem compliance from camera manufacturers to social media platforms, representing a massive socio-technical hurdle [23]. User interfaces must make provenance meaningful to non-technical audiences, as trust in AI content depends on transparency [16]. The Atlas framework proposes tracking entire training pipelines including dataset composition and model architectures [29], enabling nuanced copyright analysis but raising computational and privacy concerns. C2PA should therefore be positioned as a vital component of a broader trust and safety strategy rather than a turnkey solution.

4.2. Principle 2: Protected Research Activities

Harmonized protections for non-commercial research should apply to universities, public institutes, and cultural heritage institutions. Exception covers reproductions required for machine learning when primary purpose remains scientific investigation. Clear demarcation between research and commercial activities with transitional provisions is essential. International coordination through WIPO and bilateral agreements can develop model provisions adaptable to different legal systems.

4.3. Principle 3: Tiered Compensation Architecture

Two-stage framework addresses training-phase use and output-phase generation. At training level, collective licensing mechanisms for commercial use operate similarly to existing management organizations, with special provisions for cultural heritage datasets. At output level, similarity detection triggers automated attribution when substantial resemblance detected, with compensation pathways for commercialization.

4.4. Principle 4: Evidentiary Framework for Creative Control

Moving beyond abstract standards, this principle operationalizes emerging official government policies. The U.S. Copyright Office's January 2025 report explicitly requires "substantive human creative control" and rejects mere prompts as sufficient for authorship [31]. The Office states that copyright protects works only where there is sufficient human authorship and creative control, with simple provision of prompts to generative AI systems generally insufficient to warrant protection [8].

South Korea's June 2025 guidance provides parallel standards for "evident human creative input," specifying that modification, selection, arrangement, or highly creative iterative prompting processes can confer authorship [18]. The Korea Copyright Commission emphasizes documenting this process to establish authorship claims [21]. Research reveals successful practitioners engage in complex iterative processes that align with these regulatory standards [20]. The legally defensible "work" becomes the documented creative process, not merely final output.

Creators should implement comprehensive documentation practices including versioned prompt histories with timestamps, parameters, and creative intent annotations. This extends to curatorial documentation explaining selection rationale from generation batches, alongside post-processing records through layered files showing human modifications. Additionally, artistic intent statements connecting choices to expressive goals and process authentication through blockchain timestamping or cryptographic hashing establish verifiable creative control.

For the André Kim project, documentation might track prompt evolution from "Korean fashion" to "structured shoulders with flowing silk, inspired by traditional hanbok proportions, emphasizing André Kim's signature dramatic silhouettes," explaining specific selections and modifications to achieve cultural authenticity.

4.5. Principle 5: Risk-Based Compliance

Projects should align with ISO/IEC 23053 (AI trustworthiness), ISO/IEC 23894 (AI risk management), and NIST AI Risk Management Framework. Assessment proportionate to deployment scale: basic documentation for research, comprehensive systems for commercial platforms. Implementation through technical standards, industry self-regulation, and policy frameworks with safe harbors for compliant systems incentivizes adoption.

The European Commission's July 2025 GPAI Code of Practice, while voluntary, provides influential guidance for complying with AI Act obligations that became enforceable August 2, 2025 [32]. General-purpose AI models must now provide "sufficiently detailed summary" of training data under Article 52 [2]. This comprehensive regulatory activity demonstrates that the proposed principles are not merely theoretical but reflect active international policy consensus.

The framework provides pragmatic approach to managing regulatory divergence while fostering innovation in AI-driven cultural heritage preservation. The increased fragmentation and uncertainty in the global legal landscape, particularly following divergent U.S. court rulings, makes this harmonization framework essential for researchers, developers, and cultural institutions navigating this rapidly evolving domain.

5. Conclusion

The rapid evolution of copyright law in 2025, from the restrictive Ross Intelligence ruling to the transformative use findings in Bartz v. Anthropic, from the U.S. Copyright Office's human authorship requirements to the EU's GPAI Code of Practice, demonstrates that the legal landscape for AI-generated content remains in active formation. This paper has analyzed these developments to propose a framework for navigating the intersection of artificial intelligence and cultural heritage preservation.

5.1. Key Findings and Contributions

Our analysis reveals that the U.S. legal framework has become increasingly fragmented rather than converging toward clarity. The June 2025 rulings establish that AI training's legality depends on multiple intersecting factors: the transformativeness of use, the provenance of training data, and demonstrable market harm. This fact-specific approach creates uncertainty but also flexibility for cultural heritage applications that can demonstrate transformative purpose and utilize legally acquired datasets.

The international regulatory landscape shows surprising convergence around core principles despite surface-level divergence. The U.S. Copyright Office, European Commission, and Korean Copyright Commission all require demonstrable human creative control for copyright protection. All three jurisdictions emphasize process documentation and transparency. This emerging consensus suggests that our proposed harmonization framework aligns with regulatory trajectories across major markets.

Technical architecture choices carry profound legal implications. The decision between GANs and diffusion models, between curated datasets and web-scale corpora, between open-source and proprietary systems fundamentally shapes copyright risk profiles. Cultural institutions must integrate legal considerations into technical design from project inception rather than treating compliance as an afterthought.

5.2. Theoretical and Practical Implications

This research contributes to understanding copyright adaptation to technological change. Generative AI challenges anthropocentric copyright assumptions, requiring authorship reconceptualization when creative choices are algorithmically mediated. The idea-expression distinction becomes problematic when AI reproduces stylistic patterns at previously unprotectable abstraction levels.

For practitioners, comparative framework analysis enables informed decisions regarding architecture, governance, and procedures. The harmonization principles offer cross-jurisdictional operation roadmaps minimizing legal risk. Cultural institutions can structure projects maximizing legal protection while advancing preservation goals through transparency, documentation, and stakeholder engagement.

5.3. Limitations and Future Research

Several limitations suggest future research directions. Analysis excludes jurisdictions like China, Japan, and India developing distinctive AI governance approaches. The harmonization framework requires empirical validation through implementation testing. Emerging capabilities in multimodal generation and synthetic data raise unaddressed legal questions requiring continuous technological monitoring.

5.4. Concluding Remarks

The André Kim fashion legacy project stands at the intersection of technological possibility and legal uncertainty. Can AI systems trained on five decades of Korean haute couture generate designs that honor cultural heritage while respecting intellectual property? The answer depends not on technology alone but on our collective ability to evolve legal frameworks that recognize both human creativity's primacy and AI's transformative potential.

The 2025 legal developments analyzed in this paper reveal a system in flux, where courts struggle to apply analog-era concepts to digital realities. Yet within this uncertainty lies opportunity. The convergence around human creative control, transparency requirements, and risk-based compliance suggests that international harmonization remains achievable. The framework proposed here offers a path forward that neither stifles innovation nor abandons creator rights.

As cultural institutions worldwide digitize their collections and explore AI applications, the stakes could not be higher. The decisions made today about training data use, output ownership, and platform liability will determine whether AI democratizes cultural heritage or concentrates creative power. The window for thoughtful regulation remains open, but rapid technological advancement demands urgent action. The principles articulated in this paper, grounded in recent legal developments and practical implementation challenges, provide a foundation for ensuring that AI serves cultural preservation's highest aspirations while respecting the creative legacy of artists like André Kim whose works inspire future generations.

Acknowledgments

This research was supported by IITP and MSIT of Korea through the Graduate School of Metaverse Convergence Program (RS-2022-00156318), and by MCST of Korea through the KOCCA grant (RS-2023-00219237) as part of the Culture, Sports, and Tourism R&D Program.

References

[1] Amazon Web Services. Ensuring media authenticity by running c2pa on aws, 2025. Implementation case study with Sinclair Broadcasting. 5

[2] Bird and Bird LLP. Ai act from timelines to tensions: A mid-2025 round-up, 2025. EU AI Act implementation and Article 52 requirements. 5, 6

[3] N. Carlini. Privacy and copyright in generative models, 2025. Research on model memorization and extraction attacks. 3

[4] M. W. Carroll. The new copyright directive: Digital services act, platforms and upload filters. *Berkeley Technology Law Journal*, 37:1–45, 2022. 3

[5] Coalition for Content Provenance and Authenticity. C2pa: Verifying media content sources, 2025. Technical specification for content credentials. 5

[6] Coalition for Content Provenance and Authenticity. Guidance for artificial intelligence and machine learning, 2025. C2PA Specifications 2.2 on sidecar files for non-media assets. 5

[7] Coalition for Content Provenance and Authenticity. C2pa implementation guidance, 2025. Public draft implementation guidance. 5

[8] Congressional Research Service. Generative artificial intelligence and copyright law. Technical report, Library of Congress, 2025. CRS Report LSB10922. 6

[9] Davis and Gilbert LLP. Thomson Reuters v. Ross Intelligence: Court Rules AI Training Not Fair Use, 2025. Federal court ruling on AI training and copyright. 4, 5

[10] European Parliament and Council. Directive (EU) 2019/790 on Copyright in the Digital Single Market, 2019. OJ L 130, 17.5.2019, p. 92-125. 5

[11] G. Franceschelli and M. Musolesi. Copyright in the age of generative artificial intelligence. *arXiv preprint arXiv:2302.01390*, 2023. 3

[12] J. S. Gans. Copyright policy options for generative artificial intelligence. Working Paper 32106, National Bureau of Economic Research, 2025. 3

[13] J. Grimmelmann. Three theories of copyright in ratings. *Ohio State Law Journal*, 85:1–50, 2024. 2

[14] Hall Booth Smith PC. South korea passes ai law, 2025. Analysis of South Korea's comprehensive AI Act. 4, 5

[15] Husch Blackwell LLP. Recent decisions clarify fair use doctrine in ai context, 2025. Emerging factors in AI fair use analysis. 4

[16] M. Johnson and K. Williams. Public opinions about copyright for ai-generated art. *arXiv preprint arXiv:2407.10546*, 2025. 5

[17] Jones Day. Two u.s. courts address fair use in generative ai training cases, 2025. Analysis of data provenance in fair use determination. 4, 5

[18] Kim and Chang. Ai issues and implications in h1, 2025, 2025. South Korean AI copyright guidance and registration requirements. 4, 5, 6

[19] K. Lee, A. F. Cooper, and J. Grimmelmann. Talkin' 'bout ai generation: Copyright and the generative ai supply chain. *Journal of the Copyright Society USA*, 71, 2023. 3

[20] A. Miller and R. Davis. Generative artificial intelligence, human creativity, and art. *PNAS Nexus*, 3(3): pgae052, 2025. 6

[21] N. Palmieri. South korean copyright office issues ai guidance, 2025. Korea Copyright Commission guidance on AI-generated works. 4, 6

[22] Publishers Weekly. Federal judge rules ai training is fair use in anthropic copyright case, 2025. Court finding of transformative use in LLM training. 4

[23] RAND Corporation. Overpromising on digital provenance and security, 2025. Critical analysis of C2PA threat model and GAI-specific risks. 5

[24] Reddit Machine Learning Community. Advantages of gans over diffusion models in image generation, 2025. Discussion of GAN benefits for specialized applications. 3

[25] Ropes and Gray LLP. A tale of three cases: How fair use is playing out in ai copyright lawsuits, 2025. Analysis of Bartz v. Anthropic and Kadrey v. Meta rulings. 4, 5

[26] P. Samuelson. Generative ai meets copyright. *Science*, 381(6654):158–161, 2023. 2

[27] Sapien AI. Gans vs. diffusion models: In-depth comparison and analysis, 2025. Technical comparison of generative AI architectures. 3

[28] Skadden, Arps, Slate, Meagher and Flom LLP. Copyright office weighs in on ai training and fair use, 2025. Analysis of U.S. Copyright Office guidance on AI. 3

[29] J. Smith and L. Chen. Atlas: A framework for ml lifecycle provenance and transparency. *arXiv preprint arXiv:2502.19567*, 2025. 5

[30] Supreme Court of the United States. Campbell v. Acuff-Rose Music, Inc., 1994. 510 U.S. 569 (1994). 4

[31] U.S. Copyright Office. Copyright and artificial intelligence, part 2: Copyrightability report. Technical report, Library of Congress, 2025. 5, 6

[32] Vinson and Elkins LLP. Build once, comply twice: The eu ai act's next phase, 2025. Analysis of EU AI Act enforcement and GPAI Code of Practice. 6