Concept Paper

The Synthetic Media Exchange: When Lineage Becomes Currency

Chaoyue He * , Xin Zhou , Di Wang , Hong Xu , Wei Liu , Chunyan Miao

Posted Date: 27 March 2026

doi: 10.20944/preprints202603.2210.v1

Keywords: synthetic media; provenance; agentic multimedia; multimodal foundation models; multimodal agents; content credentials; rights metadata; multimedia markets

 doi:10.20944/preprints202603.2210.v1

Concept Paper

The Synthetic Media Exchange: When Lineage Becomes Currency †

Chaoyue He [1,*], Xin Zhou [1], Di Wang [1], Hong Xu [1], Wei Liu [2] and Chunyan Miao [1]

[1] Alibaba–NTU Global e-Sustainability CorpLab (ANGEL), Singapore, Singapore
[2] Alibaba Group, Hangzhou, China
* Correspondence: cyhe@ntu.edu.sg
† Proceedings of the 34th ACM International Conference on Multimedia (MM'26), November 10–14, 2026, Rio de Janeiro, Brazil.

Abstract

Multimodal foundation models and multimodal agents are turning media creation into programmable supply chains. The bottleneck is shifting from synthesis quality to transaction quality: which synthetic asset can be trusted, licensed, recomposed, and audited at low cost by downstream humans and agents? We argue that a generated image, video, audio clip, or mixed-media package should be treated not as a standalone file but as an asset bundle comprising payload, provenance graph, rights closure, task-conditioned trust profile, and derivative-settlement policy. We call the corresponding market layer the *Synthetic Media Exchange* (SMX). The key scientific distinction is between *authenticity* and *market legibility*: an asset may carry an authenticity signal yet remain costly to reuse because permissions, evidentiary strength, or unresolved uncertainty are not computationally usable. To sharpen this claim, we introduce *provenance capital*, *provenance debt*, and *remixability as option value*. We synthesize literature on multimodal foundation models, multimodal agents, provenance standards, media forensics, watermarking, rights expression, transparency design, and platform economics to argue that current synthetic-media ecosystems exhibit an Akerlof-style lemons problem. We then formalize SMX and develop a research agenda for lineage-priced media spanning multi-resolution provenance graphs, rights reasoning, agent-facing retrieval and recommendation, derivative attribution and settlement, benchmark design, and governance.

Keywords: synthetic media; provenance; agentic multimedia; multimodal foundation models; multimodal agents; content credentials; rights metadata; multimedia markets

CCS Concepts: • **Information systems** → **Multimedia information systems**; **Security and privacy** → *Digital rights management*; *Digital forensics*; • **Computing methodologies** → *Artificial intelligence*; • **Human-centered computing** → *Collaborative and social computing*.

1. Introduction

Multimodal foundation models are increasingly evaluated not only on static image understanding, but also on cross-modal reasoning, expert-level problem solving, interleaved document understanding, and video analysis at scale [1**?** –7]. In parallel, large multimodal agents are moving beyond one-shot prompting toward planning, tool use, verification, revision, and human-in-the-loop execution [8,9]. This is one of the clearest current AI shifts: media creation is becoming *agentic*, *compositional*, and *iterative*. A campaign video, educational clip, synthetic presenter, or interactive explainer can now be assembled from prompts, retrieved assets, model calls, edits, approvals, and subsequent regeneration steps. The final output is only the visible surface of a much larger computational history.

However, most media platforms still transact generated outputs as if they were standalone files. Discovery surfaces emphasize semantics, aesthetics, or engagement. Stock-style catalogs expose thumbnails and coarse licenses. Provenance systems focus on authenticity disclosure. Watermarking

 doi:10.20944/preprints202603.2210.v1

systems focus on marking generation. Forensic systems focus on detecting manipulation. These layers are valuable, but they remain weakly connected to the economic operations that determine whether a synthetic asset is discoverable, reusable, legally safe, appropriately compensated, or worth preferring over a visually similar alternative. As generation quality improves, the next bottleneck is not producing plausible media; it is deciding which media can be safely reused, recomposed, and audited.

This paper argues that generative multimedia now needs a new unit of exchange. We call this unit the *synthetic media asset bundle*: the media payload together with its multimodal provenance graph, rights closure, trust profile, and derivative-settlement policy. Building on that unit, we propose the *Synthetic Media Exchange* (SMX), a provenance-native marketplace in which ranking, pricing, licensing, auditing, and downstream compensation depend on verifiable lineage, computationally legible rights, and safe remixability – not only on attention. In one phrase, the thesis is simple: *price the lineage*.

The deeper claim is not that provenance alone solves synthetic-media governance. Content Credentials, watermarking, and deepfake detection are necessary but incomplete [10–15]. Their shared limitation is that they are often treated as post-hoc trust aids rather than as variables in discovery, licensing, or compensation. In practice, markets still struggle to answer economically decisive questions: Can this asset be reused commercially in another pipeline? What derivative obligations follow from its ancestors? Which branch of its history was human-reviewed? Should a multimodal agent prefer this asset over an aesthetically stronger but poorly documented alternative?

We therefore distinguish *authenticity* from *market legibility*. An asset is market-legible when a human or agent can infer, at acceptable cost, four kinds of operationally relevant information: where it came from, what downstream transformations are allowed, which evidence supports those claims, and what uncertainty remains unresolved. An asset can be authentic but not market-legible, for example when origin is signed but consent scope or derivative obligations remain unclear. Conversely, an asset can be descriptively labeled but not verifiable. This distinction matters because synthetic-media markets increasingly trade *future options for reuse*, not just present perception.

Our argument is grounded in three linked observations. First, multimodal agents create a new class of machine consumers of media; provenance is becoming an input to planning, not merely an explanation shown after publication [8,9]. Second, information economics predicts that when high-lineage and low-lineage assets are hard to distinguish, careful creators are underpaid and under-supplied [16–18]. Third, recent disclosure and provenance-UX studies show that disclosure design matters: users respond differently to process-based versus harm-based labels, to "AI-generated" versus "AI-modified" wording, and to different levels of detail [19–26]. The field therefore needs something richer than a badge and more operational than a detector.

The paper makes four contributions. It reframes synthetic-media provenance as a market-design problem rather than only a verification problem. It develops a sharper vocabulary around *market legibility*, *provenance capital*, *provenance debt*, and *remixability as option value*. It formalizes the SMX abstraction and argues that provenance should be treated as a market primitive. It also lays out a research agenda spanning multi-resolution lineage modeling, rights reasoning, agent-facing retrieval and recommendation, derivative attribution and settlement, benchmark design, and governance.

2. From Authenticity to Market Legibility

2.1. Agentic Media Supply Chains

The dominant mental model of generative media remains "prompt in, file out." That picture is becoming obsolete. Multimodal agents can decompose a creative objective into a sequence of actions: retrieve reference material, compare candidate assets, generate alternatives, inspect provenance, edit or regenerate failing segments, request human approval, and package the result for publication [8,9]. Even when the visible product is a single image or short video, the economically meaningful object is often a cross-modal derivation graph with multiple ancestors, execution traces, and policy checks.

 doi:10.20944/preprints202603.2210.v1

This shift matters because the value of a synthetic asset increasingly depends on facts that are *not* visible in the final pixels or waveform. Which model version produced it? Did an agent use licensed source material or scrape an untracked reference? Were likeness or voice-consent constraints satisfied? Did a human reviewer sign off? Was provenance preserved across editing and transcoding steps? Can a downstream agent crop, dub, or remix the asset into another composition without reopening rights questions? These questions do not sit naturally inside file-centric marketplaces, yet they directly affect whether an asset can be reused in advertising, journalism, education, entertainment, or enterprise workflows.

A useful way to state the shift is that synthetic media is moving from *consumable output* to *programmable intermediate good*. A generated clip is often not the endpoint of a pipeline; it is an ingredient in later retrieval, editing, or composition. Once this happens, provenance and rights stop being peripheral annotations. They become part of the computational state that determines what the next model, tool, or agent is permitted to do.

2.2. Why Labels & Detectors Are Not Enough

At first glance, better authenticity indicators might seem sufficient. In fact, recent work shows both the promise and the limits of that view. Provenance-enabled media and cryptographic authenticity cues can positively affect trust and accuracy judgments when presented well [27,28]. Related work also suggests that provenance data carries distinct value for creatives and audiences [29]. Yet adjacent work on disclosure design shows that interface semantics matter substantially: process-based and harm-based labels can have different effects, users interpret "AI-generated" and "AI-modified" differently, and increased detail can change perceived transparency without uniformly improving decision quality [19,21–24]. Generic disclosure can also induce blanket skepticism or collapse meaningfully different workflows into the same social signal [20,25,26].

This literature points to a deeper limitation of single-label solutions. A label can tell a viewer that something is AI-generated, AI-modified, or provenance-enabled, but it rarely exposes the *structure* needed for downstream decision making. It typically does not specify which ancestors materially matter, what obligations attach to reuse, whether the claim is backed by verifiable evidence, or what uncertainty remains unresolved. Detectors and watermarks have a complementary limitation: they can contribute evidence about likely generation or manipulation, but they do not by themselves resolve rights inheritance, derivative compensation, or task-specific trust valuation [12–15].

The design implication is subtle but consequential. The problem is not only to tell users that media is synthetic. The problem is to make synthetic media *computationally legible*. A buyer or agent needs structured answers to questions like "what is the chain of custody?", "what transformations remain allowed?", "what evidence supports those claims?", and "what uncertainty is unresolved?" A provenance-native exchange is one attempt to make those answers available in a form that directly influences retrieval, ranking, pricing, and settlement.

2.3. The Synthetic-Media Lemons Problem

Information economics offers a powerful lens on this situation. In Akerlof's classic "market for lemons," hidden quality differences create adverse selection: buyers discount prices because they cannot distinguish good products from bad ones, and high-quality sellers may exit the market [16]. Synthetic media markets now face a structurally similar problem. Creators or platforms may know whether an asset has complete provenance, clean rights, and reproducible lineage. Downstream buyers often do not. Because current interfaces foreground surface quality and engagement far more than origin, policy cleanliness, or reuse feasibility, visually convincing but poorly documented assets can compete directly with provenance-rich assets.

This distortion has consequences that go well beyond misinformation. A brand licensing an AI-generated spokesperson clip, a newsroom selecting an image, a game studio assembling voice assets, or an agent sourcing ingredients for a derived video all face hidden downstream risks. Those risks include legal uncertainty, consent ambiguity, compliance cost, reputational exposure, and avoidable

 doi:10.20944/preprints202603.2210.v1

regeneration work. When such risks are weakly expressed in pricing and ranking, safe-to-reuse assets are systematically undervalued.

Spence's theory of signaling sharpens the point [17]. Provenance can act as a separating signal when it is sufficiently costly or difficult to fake relative to the benefit of doing so. Reputation models add a longer-horizon view: creators and platforms that repeatedly provide well-documented, reusable assets can command premiums over time [18]. The missing step is to make such signals economically actionable. Without an exchange layer, provenance remains a weak side channel instead of a core determinant of allocation.

3. Why the Problem Is Irreducibly Multimodal

The proposed exchange is specifically about multimedia, not a generic metadata marketplace. The core difficulty is that provenance, rights, and trust no longer stay inside one medium. A text prompt can produce images that become keyframes in a video; a voice model can dub the video; retrieved music alters licensing obligations; subtitles and translated captions change disclosure requirements; and a human editor may splice segments from several sources into a final package. Once this happens, the economically meaningful object is neither purely textual nor purely visual. It is a coupled cross-modal lineage whose value depends on how evidence and permissions propagate across modalities.

This coupling creates at least three scientific challenges. The first is *cross-modal alignment*. A rights obligation may attach to a two-second audio segment, a cropped image region, a subtitle string, or a style-transfer operation inherited from earlier prompts. Provenance systems therefore need region-, segment-, and time-aware lineage rather than only whole-file ancestry. The second challenge is *semantic heterogeneity*. Different modalities carry different kinds of risk: images and video raise copyright and manipulation concerns, voice and music raise consent and performance-rights issues, and text layers influence disclosure, translation, and contextualization obligations. The third challenge is *evidentiary fusion*. Strong cryptographic provenance for a video may coexist with weak evidence about an added soundtrack or imported image. Trust estimation cannot simply average these modalities; it must reason over the weakest relevant link for the intended task.

There is also a second, less discussed form of multimodality: synthetic-media lineage spans *multiple levels of abstraction*. A final asset may have fine-grained local ancestry – prompt, edit, crop, dub, render – and much more distant upstream dependencies on pretrained models, datasets, retrieval indices, safety filters, and documentation artifacts. Full training-data provenance for a frontier model is not realistically exportable at output time. What SMX therefore needs is *multi-resolution lineage*: exact local traces for near-field generation steps, plus signed or documented abstractions for far-field dependencies such as model cards, datasheets, AI usage cards, or model/dataset attestations [30–32]. In this sense, the exchange object should not be a flat manifest, but a layered graph whose granularity depends on what a downstream decision actually requires.

This observation also links multimedia research to software supply-chain security. Systems such as in-toto and Sigstore show how ecosystems can use step-level attestations, signer identities, and public verification infrastructure to make complex artifact pipelines more trustworthy without forcing universal disclosure of every internal detail [33,34]. Synthetic media differs from software in decisive ways – it adds copyright, consent, aesthetics, and multimodal recomposition – but the analogy is scientifically useful. It suggests that provenance-native media markets should borrow the idea of layered attestations while extending them to rights closure, partial disclosure, and derivative settlement.

The recent trajectory of multimodal AI evaluation makes this especially timely. Benchmarks such as MMBench, MMMU, and Video-MME already stress heterogeneous image types, interleaved documents, and long-context video reasoning [4–6], a trend now expanding into complex domain-specific systems and sustainability frameworks [35–38]. What they do not yet test is whether a model or agent can reason about provenance-bearing media as *transactional infrastructure*. A system that

 doi:10.20944/preprints202603.2210.v1

can answer a visual question but cannot determine whether an asset may be safely dubbed, cropped, relicensed, or partially regenerated is still missing a central capability for agentic media economies.

4. Adjacent Literature & the Missing Exchange Layer

Table 1 summarizes the strongest neighboring literatures. The key takeaway is not that the ingredients are absent; it is that they remain disconnected. The field already has graph formalisms for provenance, authenticity standards, strong multimedia forensics, emerging watermarking methods, rich work on multimodal AI, human-centered evidence on disclosure, and even adjacent models of supply-chain attestation. What is still missing is a joint mechanism for *search*, *pricing*, *licensing*, *derivative attribution*, *and settlement* over those signals.

Table 1. Adjacent literatures already provide many ingredients for a provenance-native synthetic-media ecosystem. The missing piece is the *exchange layer*: a mechanism that treats provenance, rights, and trust as first-class variables in retrieval, pricing, licensing, and compensation.

Area	Representative work	What it contributes	What remains missing for lineage-priced media
Multimodal foundation models and evaluation	[1–6]	Broad cross-modal generation, reasoning, and evaluation across images, interleaved documents, and video	Generated assets are rarely treated as reusable economic objects with provenance, rights inheritance, or derivative value
Large multimodal agents	[8,9]	Planning, tool use, revision loops, execution traces, and human-in-the-loop workflows	Agent traces are not yet first-class market objects for later procurement, audit, or settlement
Provenance formalisms and authenticity standards	[10,11,39–41]	Graph models for provenance; signed manifests; transport-aware authenticity workflows; manifest repositories and redaction paths	Strong support for verifiable claims, but little guidance on ranking, pricing, or rights-aware recommendation among multiple valid assets
Image provenance analysis and media forensics	[15,42–46]	Reconstruction of edit histories, manipulation traces, and forensic evidence	Excellent for diagnosis and detection, weak on licensing, derivative accounting, and market allocation
Watermarking and generation-time marking	[12–14]	Detectable signals for generated outputs across text and images	Marking outputs does not determine whether an asset is legally reusable, commercially safe, or compositionally attributable
Rights metadata and documentation	[30–32,47]	Machine-readable policy expression; model/data documentation; AI/ML guidance for datasets, models, and outputs	Rights and documentation are rarely propagated across multimodal derivation graphs or integrated into search and recommendation
Disclosure design and verified-media UX	[19–25,27–29]	Evidence that labels, provenance cues, and interface choices alter trust and interpretation	Human-facing cues are studied separately from creator markets, licensing, machine-readable reuse, and autonomous agents
Supply-chain attestations and signing	[33,34]	Step-level attestations, signer identity, public verification, and layered trust for complex artifact pipelines	Multimedia lacks parallel mechanisms for rights-aware recomposition, provenance debt, and derivative settlement over creative assets
Economics of information goods and platforms	[16–18,48,49]	Theoretical tools for hidden quality, signaling, reputation, and multi-sided platforms	Multimedia systems rarely operationalize provenance completeness or remixability as market variables
Governance and epistemic risk of synthetic media	[50–54]	Framing of disclosure, legal obligations, public accountability, and liar's-dividend risks	Governance requirements are not yet translated into reusable, queryable market objects for assets and agents

Three observations follow from this synthesis. First, provenance is already naturally represented as a graph problem. PROV-DM models entities, activities, and agents [39]; multimedia forensics reconstructs manipulation histories and provenance graphs [42–44]. This is a direct conceptual match for agentic synthetic media, where the important object is a derivation graph rather than an isolated file.

 doi:10.20944/preprints202603.2210.v1

Second, authenticity infrastructure is now sufficiently mature to support more ambitious market design. AMP proposed certified media pipelines [10]; C2PA standardizes manifests, assertions, ingredients, update records, hard bindings, soft bindings, and manifest repositories that can survive across tools and workflows [11,41]; and work on transport-aware authenticity shows that provenance can remain relevant beyond authoring-time capture [40]. This does not solve the exchange problem, but it removes the excuse that machine-readable provenance is still purely speculative.

Third, neighboring literatures already hint at a broader design space. Documentation practices such as model cards, datasheets, and AI usage cards attempt to summarize upstream assumptions and risks [30–32]. Supply-chain systems such as in-toto and Sigstore demonstrate that step-level attestations can become operational infrastructure [33,34]. The missing step for multimedia is to join these ingredients into an exchange layer that can reason over future transformation rights, derivative obligations, and task-specific trust.

5. The Synthetic Media Exchange

5.1. Asset Bundle, Provenance Capital, & Provenance Debt

We define a synthetic media asset in SMX as the tuple

$$\mathcal{A} = \langle x, G, \Pi, \mathbf{t}, \Sigma \rangle, \tag{1}$$

where x is the media payload; G is a multimodal provenance graph over source assets, prompts, model and tool invocations, human interventions, and transformations; Π is the rights state; $\mathbf{t}$ is a trust profile; and Σ is a settlement policy for derivative attribution and compensation.

Two design decisions are fundamental. First, *trust should be a profile before it becomes a score*. A single scalar trust score is too coarse for real markets. A newsroom, a brand, a platform, and a video-editing agent care about different evidence. We therefore represent trust as

$$\mathbf{t}(\mathcal{A}) = [t_{\text{crypt}}, t_{\text{forensic}}, t_{\text{source}}, t_{\text{policy}}, t_{\text{human}}], \tag{2}$$

capturing cryptographic validity, forensic consistency, source accountability, policy cleanliness, and human oversight. A downstream application can then compute a task-conditioned score $T(\mathcal{A}, q) = w_q^\top \mathbf{t}(\mathcal{A})$ for query or use case q.

Second, *rights are a closure problem, not a tag*. In agentic multimodal pipelines, permissions, prohibitions, and obligations may be inherited across footage, stock ingredients, model licenses, likeness consents, jurisdiction-specific disclosure duties, and platform policies. The practical question is not merely "what is the license label?" but "which downstream transformations remain feasible, given the full ancestry of this asset?"

This motivates a useful positive quantity, *provenance capital*, which captures why visually similar assets can deserve very different prices:

$$K_{\text{prov}}(\mathcal{A}) = \lambda_c C(G) + \lambda_p P(G) + \lambda_r \Gamma(\Pi, G) + \lambda_i I(\mathbf{t}), \tag{3}$$

where $C(G)$ denotes lineage completeness, $P(G)$ portability across tools and platforms, $\Gamma(\Pi, G)$ the degree of rights closure over the derivation graph, and $I(\mathbf{t})$ the inspectability of the trust evidence. This is not intended as one universal scalar imposed by a single authority. It is a scientific decomposition of why provenance-rich assets reduce downstream search, verification, and regeneration costs.

A useful counterpart is *provenance debt*: the future verification, legal, or regeneration burden pushed downstream by missing or ambiguous lineage.

$$D_{\text{prov}}(\mathcal{A}) = \rho_m M(G) + \rho_u U(\Pi, G) + \rho_h H(\mathbf{t}), \tag{4}$$

where $M(G)$ captures missing lineage, $U(\Pi, G)$ unresolved rights state, and $H(\mathbf{t})$ hidden or weakly inspectable uncertainty. Provenance capital and provenance debt are not just descriptive metadata.

 doi:10.20944/preprints202603.2210.v1

They are opposing economic forces: the first expands safe action space, while the second transfers due-diligence and regeneration cost to later actors.

5.2. Remixability as Option Value

The most important economic intuition behind SMX is that synthetic-media buyers are often purchasing *future adaptation capacity*. In compositional workflows, the value of an asset depends not only on what it can do now, but also on what it can still become later. A reusable image with auditable lineage and clean derivative rights may be worth more than a slightly better-looking alternative whose consent scope, editing history, or ancestry is unclear.

We can express this by treating remixability as expected feasible downstream utility:

$$R(\mathcal{A};\mathcal{U}) = \sum_{u \in \mathcal{U}} p(u)\, \mathbf{1}[\Pi(\mathcal{A}) \models u] g(u, \mathcal{A}), \tag{5}$$

where $\mathcal{U}$ is a set of plausible downstream uses, $p(u)$ their likelihood, $\mathbf{1}[\Pi(\mathcal{A}) \models u]$ indicates whether rights closure permits use u, and $g(u, \mathcal{A})$ measures utility contribution to that use. This makes remixability a real-option term: when future composition needs are uncertain, assets with larger feasible action sets deserve a premium.

Proposition 1 (Remixability premium under uncertain future use)**.** *Consider two assets $\mathcal{A}_1$ and $\mathcal{A}_2$ with equal current-task utility and equal task-conditioned trust for a present query q. If the feasible downstream use set of $\mathcal{A}_1$ strictly contains that of $\mathcal{A}_2$, and there exists at least one excluded use with positive probability and positive marginal utility, then $\mathcal{A}_1$ has strictly higher expected exchange value.*

The importance of this proposition is conceptual rather than technical: synthetic-media markets increasingly price *future safe adaptability*, not just current appearance.

A simple decomposition of exchange value is therefore

$$\begin{aligned} V(\mathcal{A}, q) = {} & \alpha U(\mathcal{A}, q) + \beta A(\mathcal{A}, q) + \gamma T(\mathcal{A}, q) \\ & + \delta R(\mathcal{A};\mathcal{U}_q) + \xi K_{\text{prov}}(\mathcal{A}) - \eta D_{\text{prov}}(\mathcal{A}). \end{aligned} \tag{6}$$

where U is task utility, A is demand or attention, T is task-conditioned trust, R is remixability, K_{prov} is provenance capital, and D_{prov} is provenance debt. Contemporary platforms overweight A and partially capture U. The SMX thesis is that T, R, K_{prov}, and D_{prov} must become first-class terms in media discovery and pricing.

Proposition 2 (Provenance as a market-enabling signal)**.** *Consider two classes of synthetic assets: H with verifiable lineage, clean rights closure, and low downstream risk, and L with incomplete lineage or unresolved rights. If buyers cannot distinguish H from L, pooled pricing can fall below the reservation value of H sellers, causing under-supply of high-lineage assets. If verifiable provenance-rights signals separate the two classes, type-contingent pricing becomes possible and trade in H assets can be sustained.*

The significance of this proposition is allocative rather than merely epistemic. Provenance changes which assets remain viable to produce, how they are discovered, and whether careful creators are rewarded for documentation and safe composition.

5.3. Agents as First-Class Market Participants

A distinctive implication of contemporary multimodal systems is that the buyer on SMX may not be a human browsing a dashboard. It may be another agent tasked with building a slide deck, product video, training module, or mixed-media story package [8,9]. This changes the design target of media marketplaces. Interfaces that are barely sufficient for human browsing are not sufficient for agentic

 doi:10.20944/preprints202603.2210.v1

procurement, because the consuming system needs structured guarantees about origin, rights, and permitted transformations.

Given task query q, intended downstream action u, candidate pool $\mathcal{C}$, and budget B, an agent can be modeled as solving

$$S^\star = \arg\max_{S \subseteq \mathcal{C}} \Big[\mathrm{Rel}(S,q) + \mu T(S,q,u) + \nu R(S,\mathcal{U}_q) + \xi K_{\mathrm{prov}}(S) - \zeta D_{\mathrm{prov}}(S) - \lambda C(S) \Big] \tag{7}$$

subject to $\Pi(S) \models u$ and $C(S) \leq B$, where Rel denotes semantic and perceptual relevance and $C(S)$ is cost. The exact objective can vary. The important point is that provenance and rights become *decision variables* in autonomous media planning.

This yields a deeper insight: in agentic environments, provenance is also a form of *external memory*. A strong provenance graph does not only answer "where did this come from?" It helps an agent answer "which editable subpart should I reuse?", "what can I legally keep?", "what must I regenerate to satisfy a new policy?", and "which human-approved branch should I continue from?" The same graph that supports auditing can therefore support iterative creativity, selective regeneration, and cross-session collaboration among humans and agents.

5.4. A Multi-Sided Exchange, Not Just a Content Repository

SMX should also be understood as a multi-sided platform [49]. The relevant sides include creators, licensors, model providers, agents, end buyers, platforms, and auditors. Each side values provenance differently. Creators want credit and derivative compensation. Buyers want reduced hidden risk. Platforms want compliance and reliable recommendation. Agents want structured guarantees that can be used in planning. Auditors want inspectable histories. The design challenge is to create positive cross-side externalities: the side that pays the cost of richer provenance should be rewarded by the sides that benefit from it.

This is where signaling and platform design meet. Provenance creation is not free. It requires capture, preservation, validation, and interface work. If marketplaces do not reward it, high-integrity creators may rationally underinvest. Conversely, if provenance is created but remains hard to query, buyers and agents will not use it. A practical implication is that early deployments of SMX may need explicit incentives such as discoverability boosts, lower compliance friction, better default ranking, or favorable settlement rules for assets with stronger provenance capital and lower provenance debt.

6. Architecture & Operational Path

Figure 1 summarizes the architecture. The key design choice is to insert SMX between generation and downstream use. Instead of treating provenance as a sidecar attached only after publication, SMX treats the asset bundle as the exchange object from the moment a mixed-media artifact is created.

 doi:10.20944/preprints202603.2210.v1

Figure 1. Synthetic Media Exchange (SMX). SMX introduces a provenance-native exchange layer between agentic media generation and downstream use. The top layer shows generation via multimodal agents and foundation models. The middle layer organizes assets through four components—provenance graph, rights and licensing, trust and evidence, and derivative settlement—supporting retrieval and ranking, pricing and licensing, and settlement and audit. The bottom layer shows consumption and governance across creators, platforms, model providers, and regulators. Synthetic media is thus valued not only by perceptual quality, but also by verifiable lineage, usable rights, evidentiary trust, and attribution.

Creation-time capture.

At generation time, the system should capture cross-modal ingredients, model and tool versions, prompts or prompt hashes, human interventions, and relevant rights assertions. This does not imply full raw-trace disclosure. Selective disclosure is often essential for privacy, confidentiality, and competitive reasons. The scientific challenge is therefore to record enough structure for verification and reuse while allowing redaction, abstraction, or privacy-preserving proofs where necessary. C2PA's implementation guidance already points in this direction through update manifests, manifest repositories, redaction, and soft-binding-based recovery of detached credentials [11,41].

Upstream model and data references.

A practical exchange cannot stop at provenance-time metadata alone. Output-level bundles should be able to reference higher-level documentation about the models and datasets they depend on. Model cards, datasheets, AI usage cards, and AI/ML-oriented content credentials are useful far-field evidence even when they do not fully determine output-level legality or truth [30–32]. The design problem is to decide which upstream descriptors should be inherited, summarized, or ignored for a given downstream decision.

Exchange-time retrieval and ranking.

At search time, platforms and agents should be able to query not only semantic relevance but also rights-feasible transformations, provenance completeness, trust thresholds, and unresolved uncertainty. This changes retrieval from "find the most relevant asset" to "find the most useful asset that is also verifiable and safely reusable for the intended operation." The difference is especially important when a downstream agent must choose among visually similar assets with different rights or lineage profiles.

 doi:10.20944/preprints202603.2210.v1

Pricing, licensing, and settlement.

Pricing mechanisms should incorporate provenance capital, remixability, and risk. A high-demand asset with poor rights closure may rationally be worth less than a slightly less popular asset that can be safely reused across many contexts. Licensing likewise shifts from flat labels to structured computation over the derivation graph. Once synthetic media is understood as composition over graphs, attribution and compensation must also become compositional. The research challenge is not just bookkeeping; it is designing rules that are incentive-compatible, computationally tractable, and difficult to game through selective omission of lineage.

Persistence through delivery and transformation.

A market layer is only useful if provenance survives ordinary media handling. Transcoding, resizing, platform-specific packaging, and streaming should not automatically destroy legibility. Emerging work integrating content authenticity with DASH video streaming suggests that provenance can be made relevant at the transport and delivery layer as well [40]. This matters because media exchange does not end at upload time; it continues through distribution, transformation, and later reuse.

Table 2. From file-centric platforms to lineage-priced synthetic media.

Dimension	File-centric logic	Lineage-priced logic
Unit of value	File or clip	Asset bundle with lineage, rights, trust, settlement
Trust	Badge or detector output	Task-conditioned trust profile with inspectable evidence
Rights	Flat license label	Rights closure over the derivation graph
Buyer	Human browser	Human or autonomous agent
Ranking	Relevance + attention	Relevance + trust + remixability − risk
Compensation	One-off sale or platform share	Derivative-aware attribution and settlement
Reuse	Manual due diligence	Machine-queryable feasible transformations

Table 2 distills the conceptual shift. The exchange does not abolish attention, aesthetics, or popularity. Instead, it places them alongside variables that become decisive in generative and agentic environments: provenance completeness, task-specific trust, future reusability, and the avoidance of provenance debt.

A common objection is that such an exchange would require universal coordination before any value appears. We argue the opposite. A layered deployment path is plausible. The first layer is creation-time capture and manifest preservation. The second is provenance-aware filtering and ranking for humans. The third is agent-facing APIs for trust thresholds, rights-constrained search, and permitted transformations. Only later does one need full derivative settlement and market simulation. Each layer produces standalone utility while moving the ecosystem toward lineage-priced media.

7. Canonical Transaction Types & Mechanism Design Questions

The exchange becomes more concrete when viewed through recurring transaction types. These scenarios are not exhaustive, but they clarify why a single "AI-generated" label is too weak and why the same asset bundle can support very different decisions.

Single-asset procurement.

A brand, newsroom, educator, or creator often needs to choose one asset from many candidates that appear similarly attractive at first glance. In a file-centric marketplace, that choice is dominated by visual quality, engagement history, and coarse licensing labels. In a lineage-priced marketplace, the same choice also depends on provenance completeness, consent scope, human review state, and the likely cost of later reuse. The asset with the best thumbnail need not be the asset with the highest downstream utility. Procurement therefore becomes a ranking problem over *future safe usability*, not only immediate perceptual appeal.

 doi:10.20944/preprints202603.2210.v1

Graph composition.

A second transaction type is compositional procurement. Here the buyer – increasingly an autonomous agent – does not need one terminal asset, but a set of ingredients that can be assembled into a new mixed-media product. An agent building a narrated explainer video may need a background image, a short stock clip, a voice track, music, subtitles, and a disclosure card. Each ingredient carries its own lineage, rights constraints, and trust evidence. This means the exchange problem is no longer independent scoring of isolated items; it is constrained optimization over a bundle whose legal and technical feasibility emerges only at composition time. Segment-level and region-level provenance become important because the agent may want to preserve some components and regenerate others.

Post hoc audit and contestation.

A third transaction type arises after publication. A platform moderator, auditor, journalist, researcher, or legal team may need to inspect a contested asset: Was it generated? Which components were synthetic? Which claims are cryptographically supported? What edits occurred after initial publication? What obligations followed from the asset's ancestors? In this scenario, the exchange is not primarily about procurement, but about contestability and evidentiary reconstruction. Yet the same bundle still matters economically, because disputed or poorly legible media can trigger costly takedowns, retractions, manual review, and reputational damage.

These transaction types differ in objective but share a common substrate. Single-asset procurement emphasizes ranking under hidden risk. Graph composition emphasizes compositional feasibility and derivative settlement. Post hoc audit emphasizes inspectability, uncertainty, and accountability. Designing one infrastructure that can support all three is precisely what makes SMX a multimedia systems problem rather than a narrow disclosure feature.

Treating provenance as a market primitive also raises mechanism-design questions. How should a platform reward provenance-rich assets without punishing creators who cannot reveal every internal detail? When should ranking use hard constraints versus soft premiums for market legibility? How should provisional trading work under partial lineage, and what evidence should be sufficient for escrow, lower confidence tiers, or agent abstention? How can settlement rules remain stable when derivation graphs branch aggressively, or when different ancestors contribute different modalities? These questions are not implementation details; they determine whether a provenance-native market is incentive-compatible or easy to game.

8. Evaluation Agenda

A useful vision paper should still imply concrete evaluation tasks. SMX suggests at least six benchmark families.

1. **Provenance graph completion and consistency.** Given partial cross-modal lineage, recover missing nodes and edges, verify temporal and causal consistency, and estimate uncertainty under stripped or corrupted metadata.
2. **Rights-constrained retrieval.** Retrieve assets that are not only semantically relevant but also feasible for a specified downstream operation, such as commercial remix, educational reuse, dubbing, or jurisdiction-specific publication.
3. **Trust-profile calibration.** Fuse cryptographic evidence, forensic signals, provenance completeness, and human-review state into calibrated task-conditioned trust estimates under adversarial conditions.
4. **Derivative attribution and settlement.** Infer which ancestors materially contributed to a derivative asset and compute attribution or compensation under explicit settlement rules.
5. **Multi-resolution provenance abstraction.** Compress full workflow traces into exchange-facing summaries without losing the information needed for retrieval, licensing, or audit.

 doi:10.20944/preprints202603.2210.v1

6. **Market simulation and welfare.** Simulate synthetic-media platforms with and without provenance-aware pricing to estimate hidden-risk reduction, trade efficiency, creator surplus, agent failure rates, and incentives to document lineage.

Several methodological consequences follow. First, existing multimodal benchmarks demonstrate the community's ability to evaluate broad capability across heterogeneous images, interleaved media, and video [4–6]; the next step is to evaluate *transactional utility*. Instead of asking only whether a model understands a chart or answers a visual question, we can ask whether a media pipeline retrieved the most useful rights-feasible asset, avoided hidden-risk options, preserved enough evidence for later audit, or correctly decided to abstain and regenerate when provenance debt was too high.

Second, benchmark construction does not require a globally deployed exchange. Controlled generation workflows, synthetic provenance logs, agent traces, and curated mixed-media compositions can already instantiate the core problems. A benchmark item can pair an asset with a partial derivation graph, policy labels, intended downstream use, and adversarial perturbations such as missing manifests, conflicting watermarks, stripped metadata, or incompatible licenses. Evaluation should then measure not only detection accuracy or retrieval relevance, but *decision quality*: did the system choose the best reusable asset, know when uncertainty was too high, and expose enough explanation for a human or agent to recover?

Third, the community may need new metrics beyond pure task accuracy. Examples include provenance completeness under budget constraints, rights-feasibility precision, derivative-attribution fairness, legibility under transformation, settlement stability, and downstream rework avoided. These are not replacements for standard multimedia metrics; they are complements suited to a world in which generated media is increasingly recomposed, contested, and audited rather than simply consumed once.

9. Risks, Governance, & Failure Modes

The promise of SMX is broad. Creators could monetize not only polished outputs but also reusable provenance-rich assets. Platforms could promote safer media without banning synthetic content wholesale. Model providers could expose accountable lineage rather than relying only on post-hoc claims. Journalists, courts, and researchers could inspect chain-of-custody signals for disputed media. Regulators and compliance teams could work with machine-readable records instead of screenshots and informal attestations.

The idea also carries substantial risks. *False certainty:* provenance is evidence about recorded lineage, not truth about the world. A signed history can still describe deceptive or selectively disclosed content. *Centralization:* if validation infrastructure or trust scoring is concentrated in a few firms, provenance could become an infrastructural bottleneck that disadvantages open ecosystems. *Privacy and labor visibility:* detailed traces may leak prompts, sensitive source material, or invisible labor that contributors did not consent to expose. *Provenance inequality:* large organizations may be better able to generate polished manifests than small creators, potentially turning documentation quality into a new source of market power. *Normative ossification:* trust and rights scoring may encode assumptions that privilege certain genres, jurisdictions, languages, or creator communities.

There is also a specifically epistemic and political danger. As provenance becomes more salient, the absence or stripping of provenance can itself be weaponized. Deepfake research has long emphasized broad epistemic risks [50,51]; the liar's-dividend literature shows how actors may exploit uncertainty by dismissing genuine media as fake or manipulated [52]. A provenance-native exchange should therefore avoid binary thinking in which undocumented media is treated as automatically worthless or fraudulent. Contestability, partial evidence, uncertainty-aware interfaces, and selective disclosure are essential.

These risks do not weaken the case for SMX; they define its research frontier. The right goal is not a universal oracle of truth, but a market infrastructure that makes origin, policy state, and uncertainty more legible to both humans and agents while preserving pluralism, privacy, and due process.

 doi:10.20944/preprints202603.2210.v1

10. Conclusion

This paper argued that multimodal foundation models and multimodal agents are changing not only how media is generated, but what media *is* as an economic object. In agentic and compositional environments, a synthetic asset is best understood as a bundle of payload, lineage, rights, trust, and derivative obligations. The central claim is therefore simple: provenance should be treated not only as an authenticity aid, but as a market primitive. The Synthetic Media Exchange provides one way to formalize that shift. Its novelty lies in turning provenance from a post-hoc badge into infrastructure for retrieval, pricing, licensing, settlement, and audit. In agentic multimedia, the scarce resource is no longer generation itself, but reliable and reusable lineage.

Acknowledgments: This research is supported by the RIE2025 Industry Alignment Fund (Award I2301E0026) and the Alibaba–NTU Global e-Sustainability CorpLab.

References

1. Li, C.; Gan, Z.; Yang, Z.; Yang, J.; Li, L.; Wang, L.; Gao, J. Multimodal foundation models: From specialists to general-purpose assistants. *Foundations and Trends in Computer Graphics and Vision* **2024**, *16*, 1–214.
2. Yin, S.; Fu, C.; Zhao, S.; Li, K.; Sun, X.; Xu, T.; Chen, E. A survey on multimodal large language models. *National Science Review* **2024**, *11*, nwae403.
3. Liang, P.P.; Goindani, A.; Chafekar, T.; Mathur, L.; Yu, H.; Salakhutdinov, R.; Morency, L.P. Hemm: Holistic evaluation of multimodal foundation models. *Advances in Neural Information Processing Systems* **2024**, *37*, 42899–42940.
4. Liu, Y.; Duan, H.; Zhang, Y.; Li, B.; Zhang, S.; Zhao, W.; Yuan, Y.; Wang, J.; He, C.; Liu, Z.; et al. Mmbench: Is your multi-modal model an all-around player? In Proceedings of the European conference on computer vision. Springer, 2024, pp. 216–233.
5. Yue, X.; Ni, Y.; Zhang, K.; Zheng, T.; Liu, R.; Zhang, G.; Stevens, S.; Jiang, D.; Ren, W.; Sun, Y.; et al. Mmmu: A massive multi-discipline multimodal understanding and reasoning benchmark for expert agi. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2024, pp. 9556–9567.
6. Fu, C.; Dai, Y.; Luo, Y.; Li, L.; Ren, S.; Zhang, R.; Wang, Z.; Zhou, C.; Shen, Y.; Zhang, M.; et al. Video-mme: The first-ever comprehensive evaluation benchmark of multi-modal llms in video analysis. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2025, pp. 24108–24118.
7. Ji, Y.; Zhang, J.; Xia, H.; Chen, J.; Shou, L.; Chen, G.; Li, H. Specvlm: Enhancing speculative decoding of video llms via verifier-guided token pruning. In Proceedings of the Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, 2025, pp. 7216–7230.
8. Xie, J.; Chen, Z.; Zhang, R.; Wan, X.; Li, G. Large multimodal agents: A survey. *arXiv preprint arXiv:2402.15116* **2024**.
9. Jiang, B.; Xie, Y.; Wang, X.; Yuan, Y.; Hao, Z.; Bai, X.; Su, W.J.; Taylor, C.J.; Mallick, T. Towards rationality in language and multimodal agents: a survey. In Proceedings of the Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), 2025, pp. 3656–3675.
10. England, P.; Malvar, H.S.; Horvitz, E.; Stokes, J.W.; Fournet, C.; Burke-Aguero, R.; Chamayou, A.; Clebsch, S.; Costa, M.; Deutscher, J.; et al. AMP: Authentication of media via provenance. In Proceedings of the Proceedings of the 12th ACM Multimedia Systems Conference, 2021, pp. 108–121.
11. Coalition for Content Provenance and Authenticity. Content Credentials : C2PA Technical Specification. Technical specification, 2026.
12. Kirchenbauer, J.; Geiping, J.; Wen, Y.; Katz, J.; Miers, I.; Goldstein, T. A watermark for large language models. In Proceedings of the International conference on machine learning. PMLR, 2023, pp. 17061–17084.
13. Dathathri, S.; See, A.; Ghaisas, S.; Huang, P.S.; McAdam, R.; Welbl, J.; Bachani, V.; Kaskasoli, A.; Stanforth, R.; Matejovicova, T.; et al. Scalable watermarking for identifying large language model outputs. *Nature* **2024**, *634*, 818–823.
14. Huang, H.; Wu, Y.; Wang, Q. Robin: Robust and invisible watermarks for diffusion models with adversarial optimization. *Advances in Neural Information Processing Systems* **2024**, *37*, 3937–3963.

 doi:10.20944/preprints202603.2210.v1

15. Wang, T.; Liao, X.; Chow, K.P.; Lin, X.; Wang, Y. Deepfake detection: A comprehensive survey from the reliability perspective. *ACM Computing Surveys* **2024**, *57*, 1–35.
16. Akerlof, G.A. The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics*; Elsevier, 1978; pp. 235–251.
17. Spence, M. Job market signaling. In *Uncertainty in economics*; Elsevier, 1978; pp. 281–306.
18. Shapiro, C. Premiums for high quality products as returns to reputations. *The quarterly journal of economics* **1983**, *98*, 659–679.
19. Burrus, O.; Curtis, A.; Herman, L. Unmasking AI: Informing authenticity decisions by labeling AI-generated content. *Interactions* **2024**, *31*, 38–42.
20. Altay, S.; Gilardi, F. People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. *PNAS nexus* **2024**, *3*, pgae403.
21. Wittenberg, C.; Epstein, Z.; Péloquin-Skulski, G.; Berinsky, A.J.; Rand, D.G. Labeling AI-generated media online. *PNAS nexus* **2025**, *4*, pgaf170.
22. Jung, Y.; Hua, P.; Bao, J.; Sundar, S.S. AI-Generated or AI-Modified? User Reactions to Labeling AI Use in Social Media Posts. In Proceedings of the Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 2025, pp. 1–7.
23. Gamage, D.; Sewwandi, D.; Zhang, M.; Bandara, A.K. Labeling synthetic content: User perceptions of label designs for AI-generated content on social media. In Proceedings of the Proceedings of the 2025 CHI conference on human factors in computing systems, 2025, pp. 1–29.
24. Chen, J.; Wang, T.Y.; Williams, M.; Jordan, N.A.; Shao, M.; Zhang, L.; Fussell, S.R. Examining the Impact of Label Detail and Content Stakes on User Perceptions of AI-Generated Images on Social Media. In Proceedings of the Companion Publication of the 2025 Conference on Computer-Supported Cooperative Work and Social Computing, 2025, pp. 270–275.
25. Venkatraj, K.P.; Morosoli, S.; Cools, H.; Naudts, L.; de Vreese, C.; Helberger, N.; Cesar, P.; El Ali, A. Understanding AI Disclosure Needs for News Production and Journalism. In Proceedings of the Proceedings of the 24th International Conference on Mobile and Ubiquitous Multimedia, 2025, pp. 202–208.
26. Morosoli, S.; Naudts, L.; Cools, H.; Venkatraj, K.; Helberger, N.; de Vreese, C. Public accountability and regulatory expectations for AI in journalism: qualitative evidence from focus groups with Dutch citizens. *AI & SOCIETY* **2025**, pp. 1–13.
27. Feng, K.K.; Ritchie, N.; Blumenthal, P.; Parsons, A.; Zhang, A.X. Examining the impact of provenance-enabled media on trust and accuracy perceptions. *Proceedings of the ACM on human-computer interaction* **2023**, *7*, 1–42.
28. Francis II, E.; Barwulor, C.; Monroe, A.R.; Morales, K.O.; Potlapalli, S.; Brown, K.; Jose, J.; Sidnam-Mauch, E.; Mcgregor, S.E.; Caine, K. Usable News Authentication: How the Presentation and Location of Cryptographic Information Impacts the Usability of Provenance Information and Perceptions of News Articles. In Proceedings of the Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 2024, pp. 1–20.
29. Moruzzi, C.; Tallyn, E.; Liddell, F.; Dixon, B.; Collomosse, J.; Elsden, C. Content authenticities: A discussion on the values of provenance data for creatives and their audiences. In Proceedings of the Proceedings of the 2025 Conference on Creativity and Cognition, 2025, pp. 128–140.
30. Mitchell, M.; Wu, S.; Zaldivar, A.; Barnes, P.; Vasserman, L.; Hutchinson, B.; Spitzer, E.; Raji, I.D.; Gebru, T. Model cards for model reporting. In Proceedings of the Proceedings of the conference on fairness, accountability, and transparency, 2019, pp. 220–229.
31. Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J.W.; Wallach, H.; Iii, H.D.; Crawford, K. Datasheets for datasets. *Communications of the ACM* **2021**, *64*, 86–92.
32. Wahle, J.P.; Ruas, T.; Mohammad, S.M.; Meuschke, N.; Gipp, B. Ai usage cards: Responsibly reporting ai-generated content. In Proceedings of the 2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL). IEEE, 2023, pp. 282–284.
33. Torres-Arias, S.; Afzali, H.; Kuppusamy, T.K.; Curtmola, R.; Cappos, J. in-toto: Providing farm-to-table guarantees for bits and bytes. In Proceedings of the 28th USENIX Security Symposium (USENIX Security 19), 2019, pp. 1393–1410.
34. Newman, Z.; Meyers, J.S.; Torres-Arias, S. Sigstore: Software signing for everybody. In Proceedings of the Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, 2022, pp. 2353–2367.

 doi:10.20944/preprints202603.2210.v1

35. He, C.; Zhou, X.; Wu, Y.; Yu, X.; Zhang, Y.; Zhang, L.; Wang, D.; Lyu, S.; Xu, H.; Wang, X.; et al. EsGenius: Benchmarking LLMs on Environmental, Social, and Governance (ESG) and Sustainability Knowledge. In Proceedings of the Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, 2025, pp. 14623–14664.
36. Zhang, L.; Zhou, X.; He, C.; Wang, D.; Wu, Y.; Xu, H.; Liu, W.; Miao, C. MMESGBench: Pioneering Multimodal Understanding and Complex Reasoning Benchmark for ESG Tasks. In Proceedings of the Proceedings of the 33rd ACM International Conference on Multimedia, 2025, pp. 12829–12836.
37. He, C.; Zhou, X.; Yu, X.; Zhang, L.; Zhang, Y.; Wu, Y.; Xiao, L.; Li, L.; Wang, D.; Xu, H.; et al. SSKG Hub: An Expert-Guided Platform for LLM-Empowered Sustainability Standards Knowledge Graphs. *arXiv preprint arXiv:2603.00669* **2026**.
38. He, C.; Zhou, X.; Wang, D.; Xu, H.; Liu, W.; Miao, C. PCA-OS: A Planetary Climate Adaptation Operating System. *Preprint* **2026**.
39. Belhajjame, K.; B'Far, R.; Cheney, J.; Coppens, S.; Cresswell, S.; Gil, Y.; Groth, P.; Klyne, G.; Lebo, T.; McCusker, J.; et al. Prov-dm: The prov data model. *W3C Recommendation* **2013**, *14*, 15–16.
40. Petrangeli, S.; Wang, H.; Fisher, M.; Kozma, D.; Mahamli, M.; Blumenthal, P.; Parsons, A. Integrating content authenticity with DASH video streaming. In Proceedings of the Proceedings of the 15th ACM multimedia systems conference, 2024, pp. 492–498.
41. Coalition for Content Provenance and Authenticity. C2PA Implementation Guidance. Technical guidance, 2026.
42. Moreira, D.; Bharati, A.; Brogan, J.; Pinto, A.; Parowski, M.; Bowyer, K.W.; Flynn, P.J.; Rocha, A.; Scheirer, W.J. Image provenance analysis at scale. *IEEE Transactions on Image Processing* **2018**, *27*, 6109–6123.
43. Jin, X.; Lee, Y.; Fiscus, J.; Guan, H.; Yates, A.N.; Delgado, A.; Zhou, D.F. MFC-Prov: Media forensics challenge image provenance evaluation and data analysis on large-scale datasets. *Neurocomputing* **2022**, *470*, 76–88.
44. Zhang, K.; Kong, C.; Wang, S.; Rocha, A.; Li, H. Image Provenance Analysis via Graph Encoding With Vision Transformer. *IEEE Transactions on Information Forensics and Security* **2025**.
45. Verdoliva, L. Media forensics and deepfakes: an overview. *IEEE journal of selected topics in signal processing* **2020**, *14*, 910–932.
46. Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Morales, A.; Ortega-Garcia, J. Deepfakes and beyond: A survey of face manipulation and fake detection. *Information fusion* **2020**, *64*, 131–148.
47. Iannella, R.; Villata, S. ODRL Information Model 2.2. W3C Recommendation, 2018.
48. Varian, H.R. *Markets for information goods*; Vol. 99, Institute for Monetary and Economic Studies, Bank of Japan Tóquio, 1999.
49. Rochet, J.C.; Tirole, J. Platform competition in two-sided markets. *Journal of the european economic association* **2003**, *1*, 990–1029.
50. Chesney, B.; Citron, D. Deep fakes: A looming challenge for privacy, democracy, and national security. *Calif. L. Rev.* **2019**, *107*, 1753.
51. Fallis, D. The epistemic threat of deepfakes. *Philosophy & technology* **2020**, *34*, 623.
52. Schiff, K.J.; Schiff, D.S.; Bueno, N.S. The liar's dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review* **2025**, *119*, 71–90.
53. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). Official Journal of the European Union, 2024.
54. European Commission. Code of Practice on marking and labelling of AI-generated content. Shaping Europe's Digital Future, 2025.