Getty Images Built the World's Largest Licensed Image Library. Image Identity Still Depends on Metadata.

by Nick Clark | Published March 28, 2026 | PDF

Getty Images built the world's largest commercially licensed image library with comprehensive metadata, rights management, and content moderation. The licensing infrastructure is mature. But image identity in Getty's system depends on attached metadata: file names, IPTC data, custom identifiers, and database records. If metadata is stripped during download, sharing, or re-encoding, the image loses its connection to its license. The structural gap is between metadata-based image identity and content identity derived from the image's own structural properties.

Getty Images' licensing infrastructure and editorial standards serve critical functions in media and publishing. The gap described here is about image identity architecture, not about licensing quality.

Metadata-dependent identity is strippable identity

Getty images carry embedded metadata identifying the image, its creator, and its license terms. But metadata is routinely stripped by social media platforms, content management systems, and image editing tools. An image that loses its metadata loses its provable connection to its license. The identity was in the metadata, not in the image.

Reverse image search is probabilistic, not structural

Getty uses reverse image search to detect unauthorized use. This is valuable enforcement. But reverse image search is probabilistic: it finds similar images, not structurally identical content. It can miss modified images and generate false positives. The identification is approximate because the image has no intrinsic identity to verify against.

What content anchoring provides

Content anchoring derives image identity from the image's own structural entropy: spatial frequency patterns, entropy distribution, and structural signatures. This identity persists through metadata stripping, re-encoding, and standard transformations. A Getty image would carry intrinsic identity that survives every transformation that preserves visual content. License verification would check the content's structural identity, not its metadata.

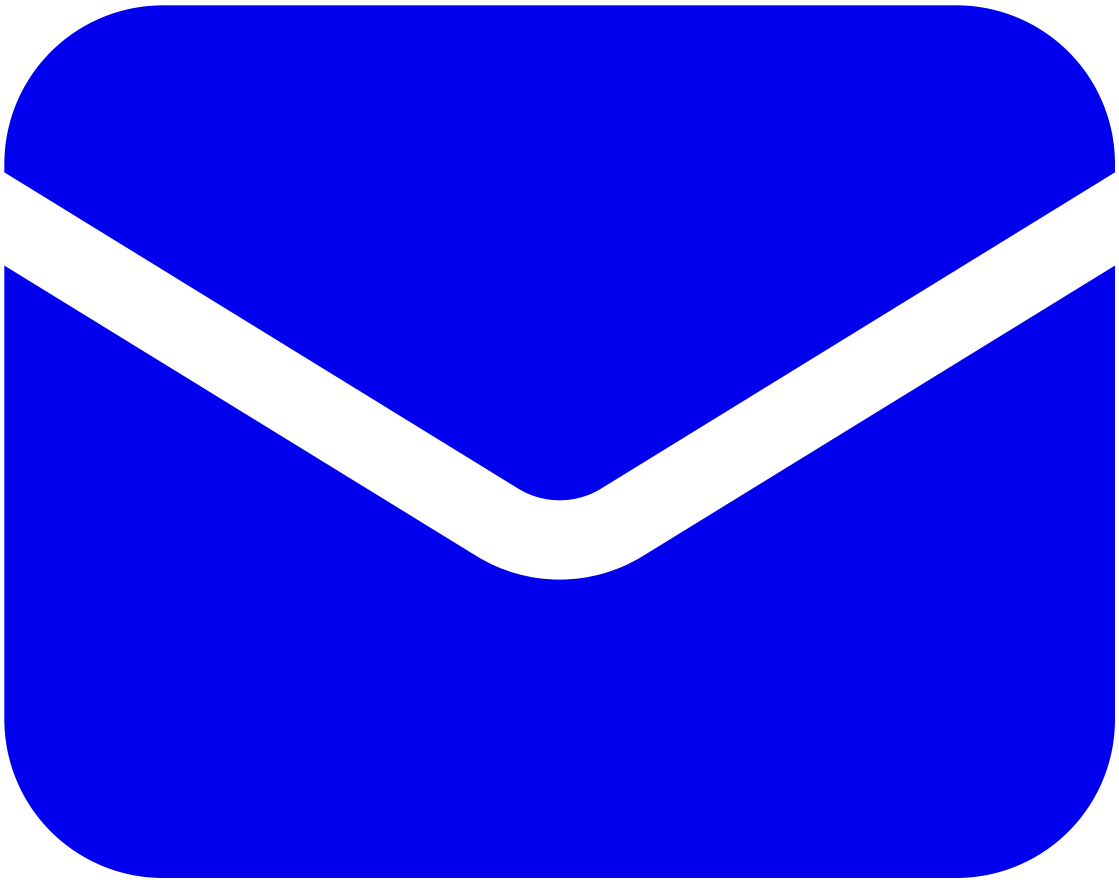
Content Anchoring All 21 steps →

Computable identity for media. Provenance from structural entropy.

Patent
US 63/808,372 · provisional
Primary Technical Disclosure
○ Content Anchoring: Computable Identity for Media That Changes
Secondary Technical
○ Multi-Axis Entropy Vector Extraction: Nine Dimensions of Structural Content Identity○ Quadrant Decomposition: Spatial Sub-Region Fingerprinting for Partial Similarity Detection○ 320-Bit UID Construction: Multi-Segment Hashing for Negligible Collision Probability○ Structure Signature: Background-Invariant Matching Through Gradient-Only Descriptors○ Constellation Signature: Geometry-Invariant Matching Across Crop, Scale, and Occlusion○ Five-Band Entropy Classification: Content Routing by Structural Complexity○ Entropy Saturation-Governed Cache Eviction: UID Density Replacing Static TTL○ Multi-Root Composite Lineage Graphs: Provenance Through Entropy Vector Similarity○ Multi-Modal Content Identity: Unified Pipeline Across Image, Audio, Text, and Video○ Rights-Grade Pre-Release Admissibility: Policy Evaluation Before Content Commitment○ Training Corpus Governance: Verifiable Lineage From Training Data to Model○ Consultation Event Logging: Deterministic Records of Every Generation Reference○ Model Output Provenance Fingerprint: Structural Proximity Without Model Access○ Creator Attribution and Compensation Routing: Payment From Consultation Lineage○ Adversarial Robustness and Deepfake Detection: Content Identity as Detection Substrate○ Client-Side Execution Architecture: Privacy-Preserving Entropy Computation on Device○ UID Resolution Query Protocol: Distributed Lookup Across Anchor Node Networks○ Orientation Canonicalization: Rotation-Invariant Processing Through Gradient Normalization○ Cross-Band Resolution Pathfinding: Traversal Between Entropy Bands Under Mutation
Applications (General)
○ Rights-Grade Generative AI: How to Pay Creators, Exclude Forbidden Content, and Prevent Infringement Before Release○ Deepfake Detection Through Structural Provenance○ Creator Economy Attribution Without Platform Intermediaries○ Content Anchoring for Journalism Verification○ Content Anchoring for Academic Research Integrity○ Content Anchoring for Legal Evidence Chains○ Content Anchoring for Insurance Claims Evidence○ Content Anchoring for Real Estate Documentation○ Content Anchoring for Art Authentication
Applications (Specific)
○ C2PA Attaches Provenance to Content. The Content Itself Has No Identity.○ Google SynthID Watermarks AI Output. Watermarks Are Not Identity.○ Shutterstock Tracks Licensed Media. The Media Itself Cannot Prove Its Own Identity.○ Spotify Tracks Every Stream. The Music Itself Has No Computable Identity.● Getty Images Built the World's Largest Licensed Image Library. Image Identity Still Depends on Metadata.○ Adobe Stock Integrates Licensed Content Into Creative Workflows. Content Identity Is Still External.○ YouTube Content ID Matches Audio and Video. The Content Has No Intrinsic Identity.○ Audible Magic Identifies Audio Content. The Audio Has No Self-Identifying Properties.○ Digimarc Embeds Invisible Watermarks. The Watermark Is Added, Not Intrinsic.○ Irdeto Protects Digital Content Through DRM. The Protection Is Applied, Not Intrinsic.
Content Anchoring overview →
AQ
deterministic
autonomy

Legal

Last updated: 2026-03-03



-
- nick@qu3ry.net
- 72 28 14 36 01

Invented by Nick Clark | Founding Investors: Devin Wilkie