Rights-Compliant Model Training Through Depth-Selective Routing

by Nick Clark | Published March 27, 2026 | PDF

Every major AI company faces lawsuits over training data rights. The core technical problem is that standard training provides no mechanism to control how deeply content integrates into model parameters or to trace which training data influenced which model behaviors. Depth-selective gradient routing addresses this structurally: content owners specify integration depth, the training loop enforces it through gradient routing, and provenance is maintained through the training process, enabling rights compliance that is verifiable rather than merely promised.

The rights crisis in AI training

Standard AI training treats all data equally: every training example contributes gradient updates across all model layers with equal depth. A copyrighted novel and a public domain text contribute identically to parameter updates. There is no mechanism to limit how deeply copyrighted content integrates, no

way to verify after training which content influenced which parameters, and no structural means to enforce licensing terms during training.

The legal consequences are now concrete. Major publishers, news organizations, and creative professionals have filed suits arguing that training on their content without licensing constitutes infringement. AI companies respond that training is fair use. Courts have not settled the question. But regardless of legal outcome, the industry needs a technical mechanism for rights-compliant training, both to satisfy potential legal requirements and to enable voluntary licensing arrangements.

Why opt-out and filtering are insufficient

Current approaches to training data rights are binary: include content or exclude it. Robots.txt, opt-out registries, and data filtering remove content from training sets entirely. But rights holders may want their content used for training under specific terms: limited integration depth, attribution requirements, or usage-based licensing. Binary inclusion/exclusion cannot express these nuanced rights.

Post-training machine unlearning attempts to remove the influence of specific training data after the fact. Research has shown that reliable unlearning is difficult to verify and may degrade model quality. It is far more tractable to govern integration depth during training than to remove influence after training.

How depth-selective gradient routing addresses this

Depth-selective gradient routing controls how deeply each training example's gradients propagate through the model's layers. Content licensed for shallow integration contributes gradients only to the model's upper layers, influencing style and surface patterns without deeply embedding in the model's core representations. Content licensed for deep integration contributes gradients across all layers.

Each training example carries a governance profile specifying its permitted integration depth, derived from its rights status. Public domain content permits full-depth integration. Licensed content integrates to the depth specified in the license. Unlicensed content is either excluded or restricted to the shallowest integration level.

Provenance tracing tracks which training examples contributed to which parameter updates at which depths. After training, the model's provenance record can answer specific questions: which copyrighted works influenced the model's behavior in a specific domain? How deeply did a particular work integrate? This provenance enables both rights compliance verification and usage-based royalty computation.

Memorization detection identifies when a training example has integrated deeply enough to be reproducible from the model's outputs. Governance constraints can prevent integration beyond the memorization threshold, ensuring that licensed content influences model behavior without being extractable from model outputs.

What implementation looks like

An AI training organization deploying depth-selective routing modifies its training pipeline to evaluate each training example's governance profile before computing gradients. The gradient routing layer controls which model layers receive updates from each example, enforcing depth limits specified by the content's rights status.

For publishers negotiating training licenses, depth-selective routing provides the technical mechanism that makes licensing practical. A publisher can license content for shallow integration at one rate and deep integration at another, with structural enforcement rather than contractual trust.

For AI companies, rights-compliant training reduces legal exposure while enabling broader access to high-quality training data through licensing arrangements that content owners can trust because compliance is structurally verifiable.

Training Governance All 21 steps →

Govern what the model learns, at what depth, with what provenance.

Primary Technical Disclosure
○ Depth-Selective Training Governance for Machine Learning Systems
Secondary Technical
○ Training Examples as Proposed Semantic Mutations○ Entropy-Band-Indexed Training Depth Profiles○ Depth-Selective Gradient Routing for Governed Training○ Training-Level Memorization Detection○ Differential Privacy Through Depth-Selective Routing○ Governed Fine-Tuning With Verifiable Provenance○ The Training Loop as a Governed Execution Environment○ Policy-Governed Knowledge Retention and Suppression○ Provenance-Traceable Training Dynamics○ Curriculum-Integrated Depth Scheduling○ Affect-Modulated Training Depth○ Training-Inference Governance Integration○ Training Governance for Human-Relatable Agents
Applications (General)
● Rights-Compliant Model Training Through Depth-Selective Routing○ Regulated Industry Model Governance With Provenance○ Training Governance for Medical AI○ Training Governance for Legal AI○ Training Governance for Financial Model Training○ Training Governance for Defense AI○ Training Governance for Educational AI Models○ Training Governance for Creative AI
Applications (Specific)
○ OpenAI's Training Pipeline Has No Depth-Selective Governance○ Constitutional AI Training Lacks Depth-Selective Control○ Stable Diffusion's Training Has No Provenance Layer○ Midjourney Trains Aesthetics Without Governed Depth○ Scale AI Labels Data Without Governing What Models Learn○ Labelbox Manages Annotation Workflows, Not Learning Dynamics○ Snorkel AI Programs Labels but Does Not Govern Gradient Depth○ Weights & Biases Tracks Experiments, Not Learning Governance○ Determined AI Orchestrates Compute, Not Learning Depth○ MosaicML Optimizes Training Efficiency, Not Learning Governance
Training Governance overview →
AQ
deterministic
autonomy

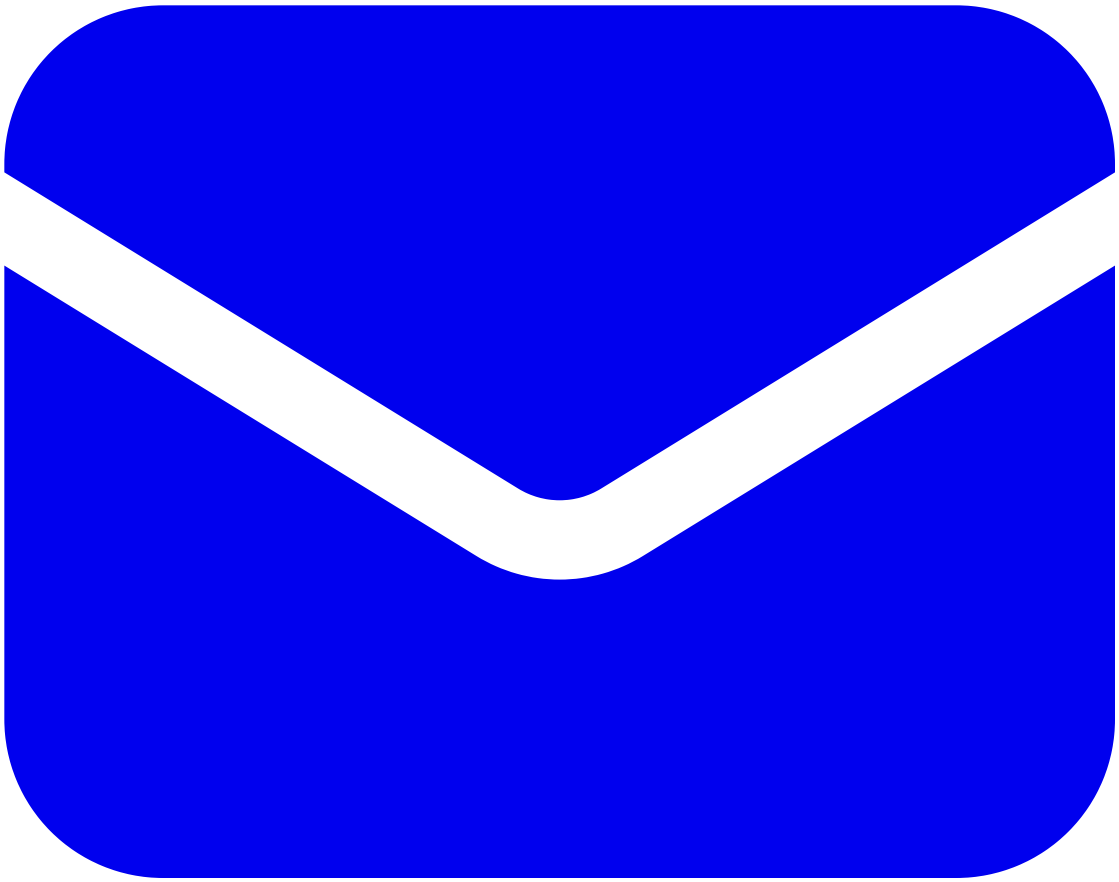
Legal

Subject to one or more pending U.S. and international patent applications, see Patents for the current list and status. No license, express or implied, is granted. Any use requires a separate written agreement—see Licensing. Patent applications referenced on this site are pending. Claim scope, if any, is subject to examination and may issue in altered form or not at all. See Legal for terms and conditions.

Adaptive Query™ is a trademark of Nicholas Clark. U.S. federal registration is pending. federal registration. AQ™ , AQ Inside™ , Adaptive Index™ , Adaptive Network™ , Semantic Agent™ , @AQ™ , AQID™ , and Adaptive Coin™ are used as trademarks in connection with the Adaptive Query platform and brand. Other names may be trademarks of their respective owners.

Platform operated by Adaptive Query LLC, which provides patent and trademark licensing services. Copyright © 2025-2026 Nicholas Clark. All rights reserved.

Last updated: 2026-03-03



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

Invented by Nick Clark | Founding Investors: Devin Wilkie