FTI Delta's Gen AI Content Licensing Deal Tracker Reveals Hurdles for Premium Video Publishers

FTIDELTA

FTI Delta's Gen AI Content Licensing Deal Tracker Reveals Hurdles for Premium Video Publishers

Authors:
Phil Schuman Senior Managing Director
Sumeet Gupta Senior Managing Director
Daniel Punt Senior Managing Director
Russell Simon Senior Consultant

Of the over 30 content licensing deals done to-date between premium content providers and Gen AI platforms, only eight have included video assets; none feature premium filmed-entertainment IP. Instead, video deals have primarily been for UGC (user-generated content, e.g., Shutterstock licensed content) and libraries of more niche video content. The recent deal between Lionsgate and Runway, while ground-breaking in signaling the intent of a Hollywood studio to explore Gen AI productivity gains, does not grant Runway access to train its primary Gen AI engine on Lionsgate video content.

Date	AI Developer & Publisher		Data Type (Video, text, image, audio, code)	Additional Info
08-29-2024	AVAIL	BRM	VIDEO	Access to ~1 million hours of video content from top YouTubers for AI training and real-time chatbot retrieval
08-20-2024	OpenAI	CONDÉ NAST	TEXT	Access to content from brands including Vogue, The New Yorker, and Vanity Fair into ChatGPT and SearchGPT
06-27-2024	OpenAi	TIME	TEXT	Access to TIME's current and historical content from the last 101 years to enhance OpenAI's products
06-17-2024	Adobe	Tiktok	AUDIO	Integration of TikTok's commercial music library of 1 million pre-cleared, top songs into Adobe Express
06-13-2024	Picsart	Gettyimages	IMAGE	Picsart to launch an AI image generator trained exclusively on Getty Images licensed creative content
06-04-2024	Reka	Shutterstcok	VIDEO, IMAGE	Access to Shutterstock's licensed photo and video assets to inform the creation of multimodal models
05-29-2024	OpenAI	The Atlantic	TEXT	Access to The Atlantic's vast archives to use as training data
05-29-2024	OpenAI	VOXMEDIA	TEXT	Access to Vox Media's portfolio of properties, including The Verge and Eater, to use as training data
05-22-2024	OpenAI	NewsCorp	TEXT	Access to content from News Corp publications, including the WSJ and MarketWatch, as training data
05-16-2024	OpenAI	Reddit	TEXT	Access to Reddit's Data API to train OpenAI products with real-time, structured, and unique content
05-08-2024	Microsoft	Informa	TEXT	Access to Informa's specialized data to advance Microsoft's AI capabilities
05-07-2024	OpenAI	Dotdash meredith	TEXT	Access to content from 40+ media brands including PEOPLE and Investopedia as training data for models
05-06-2024	OpenAI	Stack overflow	CODE	Integrate OverflowAPI into OpenAI models to deliver trusted technical content to developers
04-29-2024	Microsoft	Axel Springer	TEXT	Access to content from Axel Springer brands, including Politico and Business Insider, as training data
04-29-2024	OpenAI	Financial Times	TEXT	Access to Financial Times content as training data
04-17-2024	Meta	Individual Creators (e.g., Snoop Dogg)	VIDEO, TEXT, AUDIO	License likenesses from creators to build AI chatbots for direct messaging between influencers and their followers

Date	AI Developer & Publisher		Data Type (Video, text, image, audio, code)	Additional Info
04-11-2024	**Apple**	**Shutterstock**	**VIDEO**, TEXT, AUDIO	Access to Shutterstock's media collection of images, videos, music, and metadata to train AI models
04-11-2024	**Adobe**	**Individual Creators**	**VIDEO**, IMAGE	Photographers and videographers are training Adobe's new AI text-to-video generator on a pay-per-use rate
04-05-2024	OpenAI, Google, Meta, Microsoft, Apple, Amazon	**Defined.AI**	**VIDEO**, TEXT, IMAGE, AUDIO	Access to Defined.ai marketplace of licensed content to train models
Unspecified, as reported 04-05 by Reuters	**Unspecified** (at least 2 firms)	**EyeEm**	IMAGE	EyeEm (purchased by Freepik in 2023) is licensing its users' photos to train AI models
03-13-2024	**OpenAI**	**Le Monde**	TEXT	Access to French content from Le Monde as training data
03-13-2024	**OpenAI**	**Prisa**	TEXT	Access to Spanish content from Prisa Media as training data
02-29-2024	**Google**	**Stack Overflow**	CODE	Integrate Gemini for Google Cloud with Stack Overflow to provide coding assistance capabilities to developers
02-29-2024	**Google**	**Reddit**	TEXT	Access to Reddit's data API to train Google AI models and improve Google Search
02-05-2024	**Microsoft**	**SEMAFOR**	TEXT	News content from Semafor used to create AI-powered news search tool "MISO," aggregation tool for journalists
12-13-2023	**OpenAI**	**Axel Springer**	TEXT	Access to content from Axel Springer brands, including Politico and Business Insider, as training data
12-04-2023	**Runway**	**Getty images**	TEXT	Getty-licensed IP informs a new video model for enterprise customers
09-08-2023	**BRIA**	**Getty images**	IMAGE	AI image generator using licensed Getty Images
07-13-2023	**OpenAI**	**AP**	TEXT	Access to the Associated Press archive dating back to 1985 as training data for models
07-11-2023	**OpenAI**	**Shutterstock**	**VIDEO**, IMAGE, AUDIO	Access to Shutterstock's media collection of images, videos, music, and metadata to train AI models
01-26-2023	**OpenAI**	**BuzzFeed**	TEXT	Access to Buzzfeed's quiz content to generate new quizzes and content for users
01-12-2023	**Meta**	**Shutterstock**	**VIDEO**, IMAGE, AUDIO	Access to Shutterstock's collection of images, videos, and music to train Meta's ML capabilities

Source: Variety Intelligence Platform Content Licensing Deals for AI Training, FTI Delta analysis

On the licensing front, we believe that premium content owners may have points of leverage in Gen AI negotiations given their unique and high-quality IP. To date, Gen AI platforms have handled content more as a commodity. Gen AI platforms are attempting to claim "fair use" protection on their exploitation of copyrighted material that is scraped from the internet and tend to approach any licensing with "equal playing field" unit economics (e.g., a small, flat rate per minute of video or per word of text).

Over time, Gen AI platforms will be more dependent on high-quality content, as content informed by high-quality training data will be a major competitive differentiator. However, in these early stages, several hurdles have emerged[1] that are preventing content licensing deals on the premium video front.

1. **Lack of Deal Precedents**

   There are few existing deal precedents to serve as norms or comparisons, and premium video publishers are seeking to understand how data is used for training. As it stands, once video is used to train models, that IP will continue to be present in future models and cannot be "recovered" if a content deal ends, which presents further complications.

2. **How to Value Content Licensed to Gen AI Platforms**

   The most common model today is an input-based deal, where platforms value and pay based on the volume and quality of data received from an IP owner. While this is the most prevalent model today, it has been applied mostly to existing text-based and news licensing agreements but likely isn't the ideal long-term structure to maximize value for IP holders. We have developed our internal view of the deal archetypes that can maximize value for IP holders and are advising our clients accordingly.

3. **Concerns Around IP Rights and Protections**

   Studios own the rights to license IP as "completed works," but contractual rights likely do not extend to AI training, which may be considered a "derivative work." Additionally, components such as soundtracks and actor likenesses will complicate Gen AI licensing and will require clearances and separate compensation mechanisms. Guild requirements and individual talent contracts play a significant role here. Based on the new SAG-AFTRA agreement, talent has broad approval rights through each step of the Gen AI process when AI is used by producers to create TV episodes and movies, and talent approval would play a role in any attempts to license content to AI platforms.

4. **Fears of Enabling Competition**

   IP owners that participate in content licensing to Gen AI platforms can harvest initial benefits, including partnership revenue, the launch of new consumer-facing products, and a voice in setting initial precedents. However, there is some concern among IP providers that allowing Gen AI platforms to train models to produce cinematic-quality outputs will enable eventual technological competitors. Our view is that multimodal Gen AI outputs, which feature video with contextual sound, are coming with or without content licensing deals, and premium IP holders have leverage now to reap financial upside in the growth of Gen AI platforms via effective licensing deals.

The views expressed herein are those of the author(s) and not necessarily the views of FTI Consulting, Inc., its management, its subsidiaries, its affiliates or its other professionals.

FTI Consulting, Inc., including its subsidiaries and affiliates, is a consulting firm and is not a certified public accounting firm or a law firm.

This article has been updated to include attribution to Variety Intelligence Platform as the source for the content deals tracker. This information was inadvertently omitted in the original version.

---

[1] Variety Intelligence Platform, "Why studios still haven't licensed movies and TV shows to train AI", https://variety.com/vip/why-studios-still-havent-licensed-movies-and-tv-shows-train-ai-1236109292/

FTI DELTA