GS SHOP

When Video Became the Signal: How GS SHOP Drove a 57.5% Lift in Ordering Customers

Most "video recommendation" systems are product recommendations with a video attached. The video is decoration — the engine underneath is still blind to what a customer actually watched. GS SHOP's Search & Recommendation team decided to fix that: they made the video itself the signal. The result was a 57.5% lift in ordering customers, a 29.4% conversion rate increase, and broadcast search time cut from hours to seconds.

Why This Matters Now

Short-form video is no longer an experiment in commerce. It is the dominant interface. Today, Korean home shopping platforms like GS SHOP rely on 30-second clips to drive discovery and sales. However, there is a fundamental disconnect: the recommendation engines underneath these feeds were built for a text-and-image era. They analyze basic data like price and category, but ignore the actual video content. Every time a customer stops scrolling, the video holds the context as to why, but legacy systems miss the signal entirely. To push past current conversion ceilings, platforms need engines that actually understand the video.

THE NUMBERS

RESULT	WHAT IT MEANS
+57.5% Lift in Total Ordering Customers	AI-driven visual recommendations captured high-intent shoppers who were previously missed by legacy, product-first logic.
+29.4% Conversion Rate Lift	A significantly higher share of video views translated directly into revenue, proving the relevance of the recommendations.
+21.7% More Unique Clicks	Deeper initial engagement as shoppers reacted to highly contextual video suggestions.
6.3s → **8.0s** Average Watch Time	Customers lingered longer, indicating that the recommended content felt genuinely tailored to their interests.
1–2 Hours → **Seconds** per Broadcast	Natural-language video search eliminated the daily manual scrub-through for producers, returning hours of time to creative work.

The Problem GS SHOP Recognized

After every one-hour live broadcast, GS SHOP's producers (PDs) manually scrubbed the recordings. Their job was to isolate highlight moments worth turning into a 30-second "Short Pick" clip, while cutting transient content like weather updates, expiring promotions, and swapped-out SKUs.

Speech-to-text (STT) helped catch spoken words, but flashing on-screen captions and vital visual context routinely slipped through. This manual curation took one to two hours per broadcast. Searching for a specific visual—"Where was the scene where the host put kimchi on the boiled pork?"—routinely consumed a producer's entire afternoon.

When GS SHOP's Search & Recommendation team audited their "Short Pick recommendation" engine, they realized the pipeline was strictly running on structured product data. It applied collaborative filtering to surface items based on purchase history and category tags. Only after a product was selected did the system check if a Short Pick video was attached.

The video was a thumbnail, not a signal. If a customer watched a clip featuring a refrigerator, the system simply surfaced more refrigerators. It couldn't see the real pull of the video: the interior design context, the kitchen styling, or the premium atmosphere.

Because the system was blind to the visual content, it couldn't follow the customer's latent interest toward high-margin home decor or furniture items that could drive a cross-category conversion.

What the team needed was a recommendation engine that could read video the way a customer reads it. It had to register the activity, mood, and context of the moment, and translate that understanding into a rich visual signal the recommender could actually score against.

The Strategic Shift: **The Human Behind the Decision**

Lee Chang-min leads the Search & Recommendation Part at GS SHOP, the team responsible for discovery surfaces across the company's mobile and web commerce. Alongside data scientists Jo Min-gyu and Lee Sae-yoon, he framed the challenge not as a recommendation-tuning exercise, but as a definition problem: their system had no way to represent what a video actually meant to a customer. No amount of refining product-based signals would fix that.

The team had been watching short-form video reshape consumer behavior outside their app, with people sharing restaurant clips instead of describing the food. They began asking how their recommendation system could be brought up to the same standard.

> "When we asked Marengo, in Korean, for a scene of a shopping host placing kimchi on boiled pork and eating it with chopsticks, and it came back in seconds with the right clip, we stopped treating video understanding as something we would have to build ourselves. TwelveLabs already had it ready on Bedrock. That let us focus on the real problem: how to turn that understanding into a recommendation signal."
>
> **Lee Chang-min**
> Part Lead, GS SHOP Search & Recommendation Part

What GS SHOP Built with TwelveLabs

GS SHOP rebuilt the Short Pick recommendation system as a hybrid. The existing product-based collaborative filtering engine remained intact, with a new video-context signal layered on top.

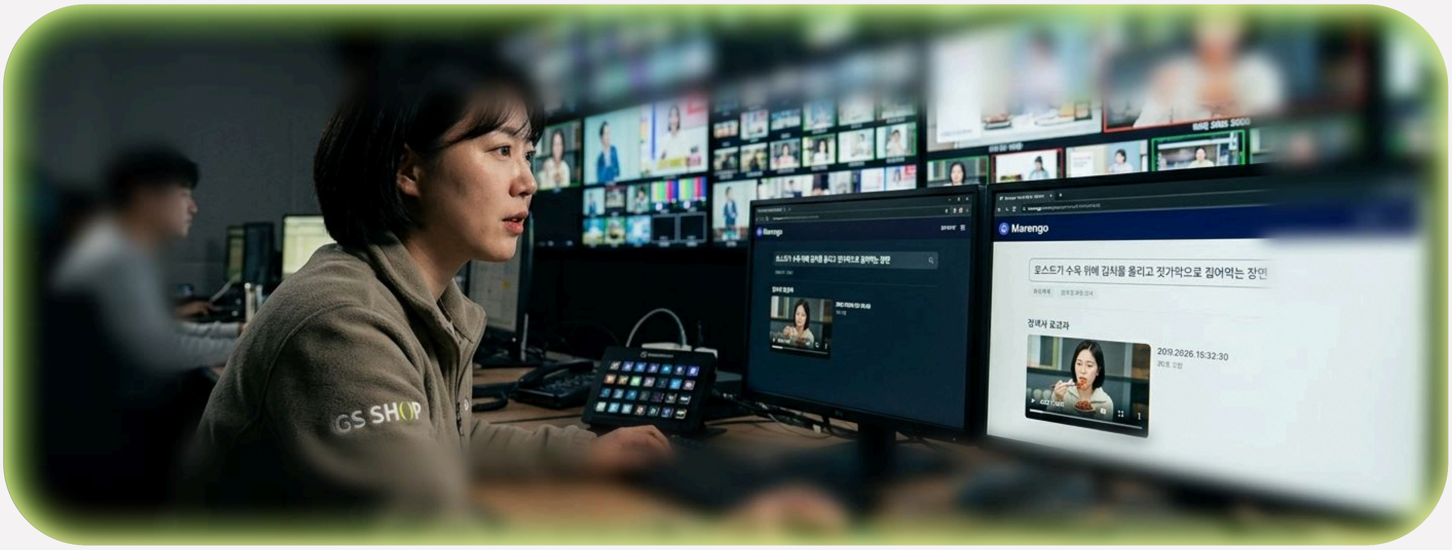
This new signal runs through a model chain on Amazon Bedrock. First, each Short Pick video is analyzed end-to-end by TwelveLabs Pegasus, which generates a contextual description of what the video means, not just the product it features. Anthropic's Claude Sonnet 4 distills that description into a structured set of "Appeal Points" across four axes: functionality, style/sentiment, use situation, and practicality. Finally, Cohere Embed Multilingual turns those Appeal Points into vectors, which are indexed in Amazon OpenSearch.
In parallel, a customer's recent interactions—weighted by intent strength and recency—are held in Amazon ElastiCache. At recommendation time, both signals are scored together alongside the existing product logic.

For example, a customer who clicks a Short Pick featuring running shoes used to see a feed full of more running shoes. After the rebuild, that same click surfaces leggings, windbreakers, and athletic socks—any item whose Appeal Points share an outdoor activity, lightweight design, and commute-ready practicality. The category boundary is no longer the limit. The reason the customer reacted to the video is.

WHAT THIS UNLOCKS ACROSS GS SHOP'S BUSINESS:

01	**Cross-category discovery in the Short Pick carousel**	Customers are introduced to relevant items outside their immediate search history, breaking the repetitive content loop that previously capped session engagement and cross-selling revenue.
02	**Natural-language broadcast search**	Producers can locate specific visual moments in seconds rather than hours, reclaiming thousands of operational hours per year and accelerating the content production pipeline.
03	**Automated expired content removal**	The platform automatically flags and filters out dead promotions or out-of-stock notices, protecting the brand from consumer friction and compliance risks after a live broadcast ends.
04	**Intent-weighted personalization**	The recommendation engine automatically prioritizes active buying signals over passive scrolling, ensuring the digital storefront adapts dynamically to real-time consumer intent.
05	**A data-driven content flywheel**	High-converting visual hooks automatically dictate the creative direction for future video assets, maximizing the financial return on every production hour.

Why TwelveLabs:
What the POC Showed



Before committing to a total recommendation rebuild, the GS SHOP team ran a Phase 1 experiment: an internal video search dashboard powered by TwelveLabs Marengo on AWS Bedrock.

They deliberately stress-tested the system in Korean with a culturally specific, multi-layered action query

> **"쇼핑호스트가 수육에 김치 한 점을 올려 젓가락으로 맛있게 집어먹는 모습"**
> **"a shopping host placing a piece of kimchi on boiled pork and eating it with chopsticks"**

to verify if the model could simultaneously identify a subject, ground itself in complex Korean food vocabulary, and recognize a continuous action across frames. Against a raw, one-hour broadcast tape, Marengo successfully returned the exact, hyper-specific clip in seconds.

That precise validation moved the project from a cautious search experiment to a full-scale recommendation rebuild. In evaluating how to scale this capability, the team explicitly calculated the cost of building or fine-tuning their own proprietary video models and ruled it out. Labeled training data pipelines, dedicated GPU infrastructure, specialized ML headcount, and an endless model upgrade cycle stacked up poorly against utilizing managed foundation models.

Because TwelveLabs models, Marengo for search and Pegasus for reasoning, were already natively callable from the team's existing Bedrock environment, GS SHOP could deploy state-of-the-art video understanding without changing its contract or underlying infrastructure pattern.

Ultimately, the team concluded that they didn't need to take on the operational burden of a new vendor stack to achieve true video intelligence; they simply needed to decide what signals to extract from the video and where to inject them into their existing recommendation logic.

From POC to Production

THE PROJECT PROGRESSED RAPIDLY THROUGH TWO DISTINCT PHASES.

PHASE 1

Korean-language video understanding validation.

Validated native Korean-language video understanding by deploying an internal search dashboard built on TwelveLabs Marengo. The step required only a single Amazon Bedrock API call against raw GS SHOP broadcast footage stored in Amazon S3.

PHASE 2

Completely rebuilt the Short Pick recommendation pipeline for production.

The production stack uses TwelveLabs Pegasus as the upstream video understanding engine, Claude Sonnet 4 to extract specific Appeal Points, Cohere to embed those points, and Amazon OpenSearch paired with Amazon ElastiCache to serve the recommendations in real time.

GS SHOP built and launched both phases entirely self-serve via the Amazon Bedrock API, completely bypassing the need for an external systems integrator, new infrastructure overhauls, or tedious model fine-tuning. The hybrid architecture was designed to be strictly additive; the pre-existing product recommendation logic remained untouched. This isolated the video signal's performance data for a clean A/B test and guaranteed that if the video pipeline ever faces a delay, the system degrades gracefully back to product data rather than breaking the user interface.

With production stabilized, the team has publicly committed to their next phase: turning validated Appeal Points—the exact visual hooks that consistently drive clicks and purchases—into inputs for future Short Pick video production. By leveraging generative video models on Bedrock (with Amazon Nova currently slated as the primary candidate), GS SHOP plans to use recommendation insights to inform how future Short Pick videos are created, creating a tighter feedback loop between content performance and content production.

Results

The team that used to run on product IDs and category trees now reasons about video the way a customer does. A walking-shoe clip is no longer just a walking-shoe clip. It is a clip about being on the move, with overlapping signals around activity, comfort, and outdoor situations.

The Short Pick carousel reads those signals and finds clips with the same emotional and functional resonance, not just the same SKU classification. Producers who once scrubbed broadcasts manually can now query them in natural language. And the team has created the conditions for a content flywheel, where every customer reaction becomes a structured signal about what the next clip should be. A year ago, that loop was not even framable.

Recommendation systems have hit the ceiling of what structured data alone can do. The teams that break past it will be the ones that can represent video, including its rhythm, context, and meaning, as a signal their existing engines can score against. GS SHOP just showed what that looks like in production.

TECHNICAL ARCHITECTURE

Platform	GS SHOP Short Pick, the short-form video recommendation carousel in the GS SHOP mobile app
AI Models	**Marengo 2.7** (Phase 1: natural-language video search validation) **Pegasus 1.2** (Phase 2: production video understanding for Appeal Point extraction)
Other models	Anthropic Claude Sonnet 4 (Appeal Point keyword extraction); Cohere Embed Multilingual (vectorization)
Partner / SI	Amazon Bedrock. TwelveLabs models were accessed completely self-serve through the Bedrock managed API. No separate TwelveLabs implementation engagement or external
Infrastructure	Amazon Bedrock (model serving) Amazon S3 (video/broadcast VOD storage) Amazon Redshift (broadcast data, product catalog, customer behavior logs) Amazon OpenSearch Service (Appeal Point embedding index + similarity search) Amazon ElastiCache / Redis (real-time per-user preference scores)
Integration	Sits as an additive hybrid layer on top of existing product-based collaborative filtering (no engine replacement). **Data Flow:** Shotpick videos and broadcast VOD are stored in S3; product and behavior data are kept in Redshift. Appeal Point vectors are indexed in OpenSearch, while user top-N preference scores sit in ElastiCache. The Recommendation API combines video, product, category, review, and behavior signals using a **five-stage scoring pipeline with MMR (Maximal Marginal Relevance) re-ranking.**
Content scope	Short Pick short-form clips (~30 seconds each), generated daily from one-hour live TV broadcasts. Dozens are exposed daily in the mobile carousel (total publishable volume TBC).
Notable	100% self-serve Bedrock API deployment. Featured in the AWS Tech Blog (May 15, 2026); presented at the AWS Seoul Summit (May 2026) with TwelveLabs as the named technology partner; covered natively in Korean tech press.

About the **Companies**

TwelveLabs

TwelveLabs builds video foundation models that transform how machines understand video. Its multimodal AI platform enables enterprises to search, analyze, and generate insights from video at scale. Based in San Francisco and Seoul, TwelveLabs serves customers across media, security, advertising, and enterprise software.
twelvelabs.io

GS SHOP

GS SHOP began as one of South Korea's representative TV home shopping businesses and has evolved into an online shopping platform spanning TV live channels, data home shopping, and mobile live commerce. Its Short Pick short-form video experience presents product highlights in approximately 30-second clips, bringing content and commerce together in a personalized mobile shopping journey. Part of GS Retail.
gsshop.com