

smart
RETAIL

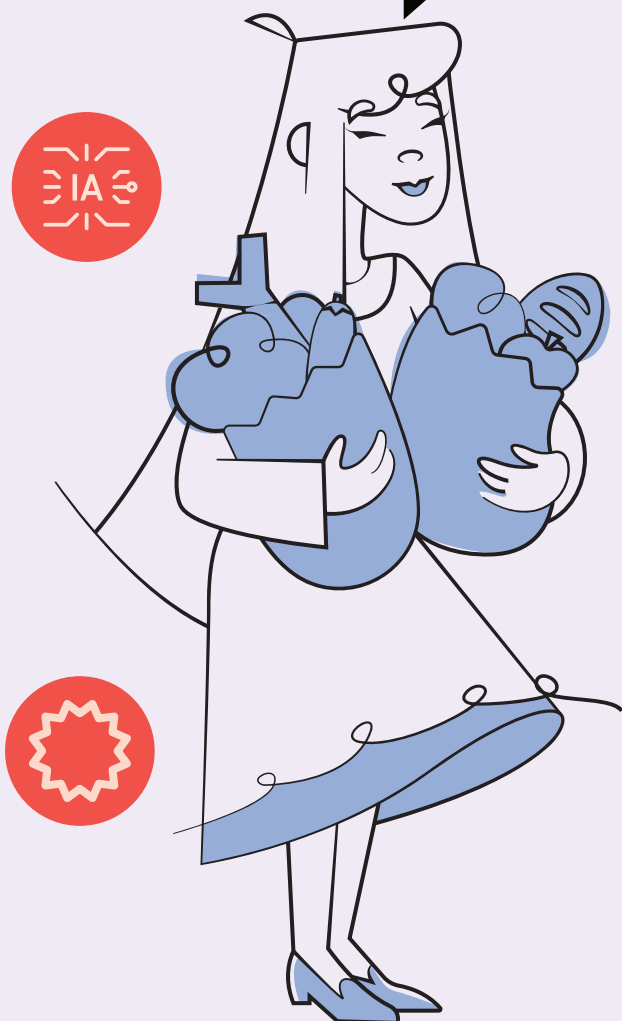
IA ACCELERATOR

IA
secrets

GUÍA

15

Guía para
aprender a
crear videos
con IA para tu
negocio



#EuropaSeSiente

Introducción	2
Un nuevo paradigma en la creación audiovisual.....	2
Ejemplos del impacto del vídeo en retail	4
Sora	5
Introducción a Sora	5
Ventajas clave para tu negocio:.....	5
Cómo usar SORA.....	6
Sora 1 Turbo	6
Sora 2.....	9
VEO 3.....	11
Introducción a VEO 3	11
Modos de creación en Flow	11
Funcionalidades avanzadas	12
Comparativa de herramientas de vídeo IA.....	15
Recursos y enlaces útiles	15
Conclusión	16

Introducción

El vídeo es el formato más poderoso para captar la atención del cliente, transmitir emociones y contar historias que conectan. En una era donde las personas hacen scroll sin parar y los primeros tres segundos deciden si alguien se queda o se va, dominar el lenguaje visual es más que una habilidad: es una necesidad estratégica. El vídeo tiene la capacidad de hacer que un producto cobre vida, que un servicio se perciba como una experiencia y que una marca se quede en la memoria.

Durante años, producir contenido audiovisual implicaba una inversión considerable: equipos de grabación, iluminación, actores, localizaciones, edición, postproducción y sonido. Todo ese proceso era costoso y, para muchas pymes o negocios locales, inalcanzable. Hoy, ese paradigma se ha roto. Con herramientas como **Sora y VEO 3**, la creación audiovisual se ha democratizado. Estas tecnologías permiten generar vídeos publicitarios, institucionales, educativos o artísticos simplemente con palabras, descripciones y creatividad. Ya no es necesario saber editar ni disponer de recursos técnicos avanzados; basta con una buena idea y una guía clara.

En definitiva, este documento será tu mapa para pasar de las ideas al impacto visual. Te enseñará a **transformar una simple descripción en una pieza audiovisual profesional, cargada de emoción y propósito**. Y, sobre todo, te mostrará cómo aplicar estas tecnologías de forma práctica dentro de tu empresa o tienda: desde la creación de anuncios de temporada hasta la elaboración de vídeos de producto, testimonios o historias de marca. Con el enfoque adecuado, podrás aumentar tu visibilidad online, atraer más clientes y fortalecer la identidad de tu negocio de manera sencilla, accesible y creativa.

Un nuevo paradigma en la creación audiovisual

La primera parada de esta guía es comprender el cambio profundo que la inteligencia artificial está provocando en la producción de contenido visual. Tradicionalmente, las empresas dependían de equipos de rodaje, agencias externas o costosas productoras para crear sus materiales promocionales. Cada pieza requería una planificación compleja: rodajes, iluminación, edición, sonido, actores, permisos, y semanas de trabajo antes de obtener un resultado final.

Hoy, ese modelo ha cambiado radicalmente. Gracias a la inteligencia artificial generativa, una sola persona puede crear un vídeo profesional en cuestión de minutos. Ya no es necesario disponer de cámaras, plató o personal especializado. La IA permite que cualquier negocio, sin importar su tamaño o presupuesto, pueda producir contenido visual con calidad cinematográfica y adaptado a su identidad de marca.

Cómo la inteligencia artificial redefine la creación de contenido visual

Las nuevas herramientas de generación de vídeo —como Sora (de OpenAI) o VEO 3 (de Google)— combinan modelos de visión artificial, generación de texto, simulación física y procesamiento del lenguaje natural para interpretar descripciones escritas y transformarlas en secuencias realistas. Cada palabra del prompt se convierte en una instrucción visual: el modelo comprende la luz, la atmósfera, la textura y el movimiento de cámara, construyendo una escena coherente en segundos.

Esto significa que un emprendedor, sin experiencia técnica, puede producir vídeos que antes solo estaban al alcance de grandes marcas.

Por ejemplo, desde su propio portátil puede generar:

- una pieza publicitaria para redes,
- una demostración de producto,
- un vídeo institucional,
- o una campaña completa de temporada.

La IA no sustituye la creatividad humana; la potencia y amplifica. El usuario se convierte en director, guionista y productor al mismo tiempo, guiando a la IA con su visión y estilo.

Imagina que antes necesitabas un fotógrafo, un editor y un especialista en animación para mostrar tu producto en acción.

Hoy basta con escribir un prompt como este:

“Una cámara recorre lentamente una estantería llena de productos locales mientras entra luz natural por la ventana. En el fondo, se oye una melodía suave y la voz en off dice: ‘Hecho a mano, con alma.’”

En menos de un minuto, tendrás un clip completo: con iluminación realista, movimiento de cámara, profundidad de campo y música de fondo. El resultado transmite calidez, profesionalismo y una narrativa visual coherente con tu marca.

El vídeo es el formato con mayor retención y conversión

El vídeo es, hoy, el formato de comunicación más efectivo. Su capacidad para captar la atención, generar emoción y transmitir experiencias supera con creces a la imagen o el texto. En plataformas como Instagram, TikTok o YouTube, los vídeos generan hasta 10 veces más interacciones que las publicaciones estáticas.

En el sector retail, esto se traduce directamente en más tráfico, más tiempo de visualización y más ventas. Los vídeos permiten al cliente “sentir” el producto antes de comprarlo:

- ver cómo se mueve una prenda,
- imaginar el olor de un café recién hecho,

- o experimentar la atmósfera de una tienda física desde su móvil.

El vídeo es también el formato que más fideliza y refuerza la identidad de marca. Un clip de apenas 15 segundos puede transmitir lo que antes requería una sesión fotográfica, un texto y un anuncio pagado.

Ejemplos del impacto del vídeo en retail

1. Escaparates digitales:

Crea vídeos automáticos de tus productos destacados para mostrarlos en pantallas de tienda o en redes.

Por ejemplo, una tienda de calzado puede generar un clip donde unos zapatos “giran” lentamente sobre una base iluminada con luz cálida, acompañados por música suave y un fondo difuminado que resalta el producto.

2. Storytelling de marca:

La IA puede ayudarte a contar tu historia de forma visual. Una panadería puede recrear el amanecer, mostrar cómo se amasa el pan y cerrar con la imagen de un cliente sonriendo al recibir su barra recién horneada. No se necesita grabar nada: basta con describir la escena.

3. Promociones estacionales:

Diseña campañas visuales para Navidad, verano, San Valentín o rebajas con solo ajustar el prompt. Pide a la IA que añada decoración temática, música acorde o mensajes personalizados. Por ejemplo: *“Escaparate con luces cálidas y copos de nieve cayendo, un cartel de ‘Rebajas de Invierno’ y una melodía navideña de fondo.”*

4. Lanzamientos rápidos:

¿Nuevo producto o menú? Genera en minutos un clip de 10 segundos destacando sus atributos: color, textura, brillo o movimiento. Puedes publicarlo el mismo día en redes o proyectarlo en pantalla en el punto de venta.

Vídeos de unboxing o demostraciones.

Muestra cómo se abre, utiliza o combina tu producto. Por ejemplo, una joyería puede generar una animación donde una caja se abre lentamente mostrando un anillo brillando bajo luz natural.

5. Experiencias sensoriales en hostelería

Cafeterías y restaurantes pueden recrear vídeos del café cayendo en cámara lenta, del vapor saliendo de la taza o del chef emplatando un postre, transmitiendo sensaciones que las palabras no logran.

6. Vídeos para campañas emocionales

Una floristería puede crear un clip con flores abriéndose lentamente mientras se escucha un piano suave, transmitiendo calma y delicadeza. O una tienda de juguetes puede mostrar escenas alegres y coloridas con niños sonriendo y luces cálidas.

7. Decoración y hogar: tours virtuales con IA

Crea recorridos que muestren cómo los productos decoran un espacio. Por ejemplo, “una cámara se mueve lentamente por un salón moderno donde se ve un sofá de lino beige iluminado por luz natural”.

8. Moda: pasarelas digitales

Genera pasarelas o combinaciones de outfits con modelos digitales. La IA puede adaptar las prendas a diferentes tipos de cuerpo, edades o estilos, permitiéndote mostrar diversidad sin organizar una sesión de fotos.

Sora

Introducción a Sora

Sora, el modelo de OpenAI que transforma texto en vídeo con realismo cinematográfico, representa un salto tecnológico en la forma en que las empresas pueden crear contenido visual. Se trata de una tecnología avanzada de generación de vídeo mediante inteligencia artificial, diseñada específicamente para transformar descripciones escritas en escenas de alta calidad y estética profesional. Sora combina algoritmos de difusión, comprensión semántica y simulación física para interpretar cada palabra como una instrucción visual. Es capaz de comprender la luz, el color, la textura, la atmósfera y el movimiento de cámara, además de captar la emoción detrás de la escena que se describe.

Para un empresario del sector retail, esto significa que puede crear en minutos vídeos con apariencia profesional para su tienda, catálogo o redes sociales sin depender de productoras ni equipos técnicos. Sora interpreta las descripciones y las convierte en vídeos donde los productos cobran vida: un bolso que gira sobre un fondo elegante, un escaparate iluminado al atardecer o un cliente sonriendo al recibir su compra.

Ventajas clave para tu negocio:

- Permite generar vídeos con realismo cinematográfico sin cámaras, actores ni rodajes.
- Adapta la iluminación y el ambiente al tipo de negocio: desde un estilo moderno y minimalista hasta un entorno cálido y artesanal.
- Crea contenido coherente con la identidad visual de la marca.
- Reduce costes y tiempos de producción drásticamente.

- Facilita la creación de anuncios, vídeos de producto, contenido para escaparates digitales o redes sociales con solo unos prompts bien diseñados.

En la práctica, Sora actúa como una **productora audiovisual virtual** dentro de tu empresa, capaz de transformar ideas o descripciones breves en contenido atractivo que transmite profesionalismo y emoción al cliente.

Cómo usar SORA

Sora 1 Turbo

Qué es y por qué importa

Sora 1 Turbo (habitualmente presentada simplemente como “Sora” en web) es el primer modelo de vídeo de OpenAI. Se lanzó a finales de 2024 y desde febrero de 2025 está disponible comercialmente en la Unión Europea, Reino Unido, Suiza, Noruega, Islandia y Liechtenstein. También está operativa en EE.UU. y Canadá.

Para un equipo europeo, Sora 1 Turbo es importante porque ya sustituye, en gran parte, un rodaje básico de producto. Permite:

- Crear clips verticales para redes sociales (Reels, TikTok, Shorts) en cuestión de minutos.
- Generar loops continuos de producto (por ejemplo, un perfume girando con iluminación de lujo) para pantallas en tienda física.
- Producir pequeñas piezas narrativas que explican el uso o valor del producto paso a paso sin grabar físicamente a una persona.

En otras palabras: convierte el equipo de marketing en un pequeño estudio interno de vídeo.

Acceso y condiciones comerciales

El acceso a Sora 1 Turbo se hace vía web en sora.chatgpt.com (o sora.com). Para entrar, basta iniciar sesión con la misma cuenta de ChatGPT y contar con una suscripción de pago.

Hay dos niveles de suscripción relevantes:

- ChatGPT Plus (~20 USD/mes): pensado para social media y producción rápida. Genera clips cortos (típicamente 5 a 10 segundos), hasta 720p, con marca de agua. Es suficiente para contenido orgánico o pruebas de campaña.
- ChatGPT Pro (~200 USD/mes): pensado para marketing formal. Ofrece prioridad de proceso, 1080p, hasta 20 segundos de duración por clip, exportación sin marca de agua y más volumen de generación mensual. Es el nivel adecuado cuando el material se va a proyectar en tienda física o en soportes pagados.

En la práctica, una marca pequeña que quiera alimentar redes a diario puede operar con Plus. Una marca que necesite activos “finales” para retail físico, pantallas en punto de venta o inversión en paid media suele migrar a Pro.

Cómo se trabaja dentro de Sora 1 Turbo

La interfaz web se organiza en cuatro áreas operativas:

1. Explorar / Tendencias. Un feed con resultados generados por otros usuarios. Aquí se puede inspeccionar el tipo de lenguaje utilizado en los prompts y la estética que está funcionando. Es útil como biblioteca de referencia visual.
2. Biblioteca. Un área privada donde se guardan las generaciones propias: versiones anteriores, variaciones, borradores y clips aprobados. También permite clasificar por cliente, campaña o canal.
3. Crear. El generador. Desde aquí se decide qué tipo de clip se va a crear:
 - Vídeo desde texto.
 - Vídeo a partir de una imagen fija (por ejemplo, una foto de producto que ya tengamos).
 - Transformación de un clip existente (por ejemplo, adaptar un vídeo genérico al entorno navideño de la marca).
4. Storyboard. Un modo específico para construir una pequeña historia de producto con varias escenas enlazadas.

Parámetros creativos básicos

En cada generación Sora 1 Turbo solicita cuatro decisiones que conviene estandarizar internamente para evitar incoherencias entre piezas de campaña:

- Formato: vertical 9:16 para redes sociales, horizontal 16:9 para web/campañas tradicionales, cuadrado 1:1 para feed clásico.
- Duración: clips de 5 a 10 segundos son estándar en Plus; Pro permite hasta 20 segundos cuando se necesita una mini-narrativa (por ejemplo, un “antes y después”).
- Resolución: 720p es suficiente para redes sociales; 1080p se recomienda para pantallas en tienda y soportes premium.
- Variaciones: la plataforma permite generar varias versiones del mismo prompt en paralelo. A nivel de flujo de trabajo interno, esto equivale a pedir a un equipo creativo que entregue 3-4 propuestas visuales de una sola vez. Se revisan, se elige la mejor y se archivan las demás.

Tipos de generación más útiles para retail

a) Texto a vídeo. Escribir una instrucción en lenguaje natural y obtener un clip final. Esta es la vía más rápida para crear un anuncio corto de producto.

Ejemplo de instrucción típica en retail belleza: un plano cercano del frasco sobre mármol negro, luz cálida dorada y sensación de lujo aspiracional.

b) Imagen a vídeo. Sube una fotografía de producto existente (por ejemplo, un render oficial del packaging o una foto de catálogo) y Sora genera una animación realista a partir de esa imagen. Esta modalidad es especialmente valiosa cuando la compañía ya tiene fotografía aprobada por legal y no quiere modificar el diseño del producto.

c) Vídeo a vídeo. Permite transformar un activo que ya existe. Por ejemplo, convertir un spot veraniego en versión de Navidad cambiando la localización, la decoración de fondo y la atmósfera, pero manteniendo exactamente el mismo producto en primer plano.

d) Storyboard. Construcción de una pequeña secuencia narrativa en varias escenas: caja cerrada → caja abriéndose → producto en mano. Este recurso es ideal para demostraciones de uso, unboxing o mensajes “antes / después”. Es donde Sora 1 Turbo se acerca más al lenguaje publicitario tradicional.

Herramientas internas de edición en Sora 1 Turbo

Sora 1 Turbo incorpora varias herramientas pensadas para marketing, no para cineastas. Las más relevantes son:

- **Remix.** Permite rehacer un clip existente modificando solo lo que te interesa (por ejemplo, cambiar el color del producto de azul a rojo sin rehacer el resto de la escena). Es útil para adaptar el mismo material visual a distintas líneas de producto o temporadas.
- **Loop.** Genera un bucle perfecto. Esto es clave para pantallas en tienda física, ferias, escaparates digitales y visuales de escaparate. El producto parece estar en movimiento constante, sin salto visible.
- **Extend.** A partir de un clip corto, genera continuidad adicional. Es útil cuando el clip es bueno pero demasiado corto para el soporte donde va a emitirse.
- **Blend.** Fusiona dos escenas o dos estilos visuales de forma gradual. Se usa sobre todo en piezas más aspiracionales, por ejemplo, al hacer transición de “laboratorio tecnológico” a “estilo lifestyle cálido”.

Por último, Sora 1 Turbo genera sonido ambiente (ruido de fondo, música ligera, efectos sencillos como el roce de una caja al abrirse). Sin embargo, no genera discurso hablado con labios sincronizados. Es decir, no vamos a obtener todavía a una persona “hablando a cámara” de forma creíble. Esa funcionalidad existe en Sora 2.

Limitaciones que hay que conocer

Para planificar bien es importante entender qué no hace (todavía) Sora 1 Turbo:

- La duración por clip está limitada. Incluso en Pro, la pieza máxima es de 20 segundos.
- Aunque la simulación física es buena, en escenas complejas aún puede producir pequeños errores en manos, reflejos o contacto entre objetos. Conviene revisar planos detalle antes de publicar.
- Mantener exactamente el mismo personaje a lo largo de varios clips separados no siempre es consistente. Sora 1 Turbo no incorpora aún una función nativa de “portavoz recurrente”.
- El texto que aparece dentro de la imagen (por ejemplo, carteles, rótulos o packaging con letras minuciosas) no siempre es legible. Es preferible añadir textos corporativos más tarde en edición.

Aun con estas limitaciones, para un equipo europeo es una herramienta madura. Cubrir redes sociales, pantallas de tienda y campañas estacionales ya no requiere un rodaje completo.

Sora 2

Disponible actualmente solo en EE.UU. y Canadá, pero muy pronto estará disponible para todas las cuentas en Europa, de ahí la necesidad de conocer su funcionamiento para estar a la última en herramientas de IA para la creación de videos.

Qué aporta frente a Sora 1 Turbo

Sora 2 es la evolución directa y supone tres avances estratégicos:

1. Portavoces humanos generados por IA. Sora 2 puede generar personajes que hablan mirando a cámara, con voz y labios sincronizados. En términos prácticos, esto significa grabar un “anuncio hablado” sin actriz ni estudio.
2. Cameo. La función Cameo permite capturar el rostro (y opcionalmente la voz) de una persona real, por ejemplo, el CEO, la directora de tienda o el responsable comercial, y reutilizarlo en escenas generadas. Esa persona puede “aparecer” presentando el producto en una tienda flagship, en una fábrica o en un plató inexistente, sin haber estado allí físicamente. A día de hoy Cameo requiere la app iOS oficial de Sora 2, y solo está disponible en EE.UU. y Canadá.
3. Control narrativo más preciso. El Storyboard de Sora 2 es más consistente que el de Sora 1 Turbo: mantiene mejor al mismo personaje escena tras escena, suaviza la transición visual entre planos y permite una narrativa más cercana a un spot televisivo corto.

Además de lo anterior, Sora 2 genera más rápido (aprox. un 30% menos de espera) y simula mejor la física (peso, contacto, equilibrio). Esto se nota mucho en manos sosteniendo producto y en gestos corporales.

Modelo de acceso

Sora 2 se distribuye de dos maneras:

- Aplicación iOS oficial de Sora. Interfaz tipo TikTok: feed vertical, botón “crear”, vista de variaciones del mismo prompt deslizando lateralmente y biblioteca personal. Esta app actualmente exige iPhone con iOS reciente y cuenta registrada en EE.UU. o Canadá.
- Versión web en sora.com. Ofrece herramientas más avanzadas de edición, Storyboard largo (hasta ~20 segundos), exportación a resolución alta y control escena por escena. Para entrar se necesita invitación y una suscripción activa (Plus o Pro).

Uso típico en marketing

Sora 2 cambia la forma de producir “portavoces de marca”. Antes, para lanzar una campaña en la que alguien mira a cámara y dice “Esta es nuestra nueva colección”, había que coordinar rodaje, imagen personal, permisos y varias rondas de aprobación legal.

Con Sora 2 el flujo es distinto:

1. Se graba a la persona de referencia una única vez dentro de la app (proceso Cameo). Son 10-15 segundos de capturas de cara y voz con buena luz.
2. Ese perfil queda guardado. A partir de ahí se pueden generar escenas nuevas donde esa misma persona aparece hablando con distintos mensajes, en distintos espacios, sin volver a grabar.
3. El texto que dice el portavoz puede definirse en el prompt, y Sora 2 generará voz, gesto facial y sincronización labial acorde.

Esto permite personalizar comunicación a segmentos concretos: el director regional saludando a “clientes de Valencia”, el equipo de dirección anunciando una campaña de rebajas, etc., sin convocar grabaciones físicas para cada versión.

En Europa esto no está todavía habilitado de forma oficial. Por eso, cuando una empresa europea pregunta “¿podemos tener a nuestro director hablando en vídeo sin grabarlo?”, la respuesta hoy es: legalmente con Sora 1 Turbo no, con Sora 2 sí, pero solo en EE.UU./Canadá.

VEO 3

Introducción a VEO 3

VEO 3 es el modelo de generación de vídeo de Google orientado a resultados **cinematográficos**, con especial fortaleza en **movimiento de cámara realista**, **consistencia entre planos** y **audio integrado** (música, efectos y, en algunos casos, diálogos). Frente a Sora, VEO 3 destaca cuando necesitas **narrativa audiovisual completa** (imagen + sonido) y un control más fino del lenguaje cinematográfico.

Cómo acceder desde Google AI Studio, Flow o Gemini (paso a paso)

Veo 3.1 está disponible principalmente a través de dos plataformas:

- Gemini (web y app): ideal para generación rápida mediante texto conversacional.
- Flow (flow.google): entorno profesional de filmmaking con control total sobre cámara, personaje, escena y audio.

Modos de creación en Flow

Text to Video (Texto a Video)

Esta funcionalidad permite convertir una simple descripción escrita en un vídeo realista y visualmente impactante. El usuario solo necesita redactar un texto —conocido como *prompt*— explicando la acción que desea mostrar, el entorno donde sucede y el estilo visual (por ejemplo, profesional, cinematográfico, natural o artístico). A continuación, selecciona el modelo (**Fast** o **Quality**), define la duración del clip (entre 4 y 8 segundos) y el tipo de movimiento de cámara, como panorámica, zoom o dron.

El sistema analiza esas instrucciones y crea un vídeo que imita la lógica de una grabación profesional: ajusta el enfoque, la iluminación, el color y la profundidad para que el resultado parezca filmado con una cámara real.

En términos prácticos, este modo permite a cualquier empresa producir clips promocionales, anuncios o presentaciones sin depender de rodajes físicos ni estudios externos. Es especialmente útil para mostrar productos, espacios o conceptos de marca en redes sociales o páginas web con un acabado visual de alta calidad.

Frames to Video (Fotogramas a Video)

Este modo genera movimiento a partir de una o dos imágenes estáticas, creando animaciones fluidas y naturales. Es una herramienta ideal para mantener la coherencia visual de personajes o escenas que aparecerán en diferentes vídeos de una misma campaña.

Cómo funciona: el usuario crea primero una imagen base —por ejemplo, de un producto o portavoz— utilizando **Imagen 4**. Luego, describe en el prompt la acción deseada (como “la persona sonríe y levanta el producto”) y el sistema produce una animación continua.

Además, el último fotograma de un clip puede usarse como punto de partida del siguiente, lo que permite desarrollar secuencias más largas sin pérdida de continuidad.

El resultado es un flujo visual coherente que puede servir para crear embajadores virtuales de marca, influencers digitales o personajes animados que transmitan la identidad visual de una empresa sin necesidad de actores reales.

Assets to Video (Recursos a Video)

Esta función combina varias imágenes en una secuencia narrativa. El sistema analiza las conexiones entre cada imagen y genera transiciones automáticas con una estética profesional.

Por ejemplo, una marca puede mostrar en un mismo vídeo la evolución de un producto: desde su diseño inicial hasta su uso final, pasando por la fabricación y el empaquetado. Flow ajusta el ritmo, la iluminación y el estilo para que todas las escenas se perciban como parte de un mismo relato visual.

Ingredients to Video (Ingredientes Consistentes)

Veo 3.1 permite guardar elementos visuales —como personajes, productos o estilos— en una biblioteca de “ingredientes” que se pueden reutilizar tantas veces como se desee. Esto facilita que las campañas mantengan una imagen uniforme y reconocible.

Además, la versión 3.1 añade audio sincronizado a estos ingredientes, garantizando que las voces, tonos o efectos de sonido se mantengan estables en diferentes vídeos. Es una herramienta muy valiosa para empresas que desean desarrollar un portavoz digital o una identidad audiovisual coherente en todas sus comunicaciones.

Extend (Extensión de Clips)

El modo Extend permite alargar un vídeo ya creado, generando una continuación natural de la acción. La herramienta analiza los últimos segundos del clip y crea una prolongación suave del movimiento y del audio.

Gracias a esta función, es posible transformar clips cortos en piezas narrativas más extensas, adecuadas para vídeos institucionales, demostraciones o campañas que necesitan mayor duración.

Funcionalidades avanzadas

Add Object (Agregar Objeto)

Esta herramienta inserta nuevos elementos dentro de un vídeo ya generado. Los objetos añadidos se integran de forma coherente con la escena, manteniendo iluminación, color y perspectiva realista. Por ejemplo, una marca puede incluir su logotipo, un nuevo producto o elementos decorativos sin volver a crear el vídeo desde cero.

Remove Object (Eliminar Objeto)

Permite eliminar objetos o personas no deseadas de una escena, reconstruyendo automáticamente el fondo. Es ideal para limpiar errores, eliminar distracciones o ajustar el entorno visual sin afectar el resto del contenido. La función estará disponible en futuras actualizaciones de Flow.

Trim (Recortar)

Facilita la edición de la duración de los clips. El usuario puede ajustar el inicio o el final del vídeo con precisión de fotograma, sin consumir créditos adicionales. Es una herramienta sencilla para adaptar los clips a diferentes formatos o sincronizarlos con música o texto.

Upscale (Mejora de Resolución)

Convierte vídeos creados en resolución **720p** a **1080p** con un solo clic, mejorando la nitidez y la calidad visual sin generar costes adicionales. Aunque el proceso no añade nuevos detalles, ofrece un resultado más profesional y atractivo para redes sociales o presentaciones corporativas.

Camera Controls (Control de Cámara)

Ofrece control cinematográfico sobre el movimiento de cámara. Los usuarios pueden definir panorámicas, travellings, movimientos de grúa o tomas aéreas simuladas, además del tipo de lente y velocidad de desplazamiento.

Ejemplo: “Acercamiento lento al producto sobre mesa de madera con luz cálida, fondo desenfocado”.

Estos controles permiten crear vídeos con el mismo lenguaje visual que una producción grabada con cámaras profesionales.

Scenebuilder (Constructor de Escenas)

Scenebuilder es la herramienta que integra todos los clips generados en una línea de tiempo editable. Permite unir escenas, recortarlas, extenderlas o cambiar su orden con facilidad.

Las funciones principales incluyen:

- **Extend:** prolonga vídeos existentes para dar fluidez narrativa.
- **Jump To:** genera nuevas escenas que conectan con la anterior de forma coherente.
- **Trim:** recorta segmentos no deseados.
- **Save Frame as Asset:** guarda fotogramas clave como referencia o ingrediente reutilizable.
- **Rearrange:** cambia el orden de las escenas arrastrando y soltando.

- **Download Scene:** exporta la secuencia final en MP4 con sonido incluido.

Scenebuilder ofrece una experiencia intuitiva, pensada para usuarios no técnicos, que pueden construir vídeos completos sin recurrir a software de edición adicional.

Audio y Diálogo Sincronizado

Veo 3.1 incorpora un sistema avanzado de audio que genera sonidos y voces de manera natural. No solo añade música o efectos, sino que produce **diálogos perfectamente sincronizados con el movimiento de los labios** de los personajes.

Tipos de sonido disponibles:

- **Voces y diálogos:** ideales para spots publicitarios, testimoniales o presentaciones corporativas.
- **Efectos de sonido:** pasos, viento, puertas, objetos o ambiente de tienda.
- **Música ambiental:** melodías adaptadas al tono del vídeo (corporativo, inspirador, relajado).
- **Sonido continuo entre escenas:** mantiene coherencia auditiva en toda la pieza.

Ejemplo de prompt:

Diálogo: “Descubre nuestra nueva colección” + portavoz sonriente en tienda moderna, tono cálido y voz femenina natural.

Recomendaciones: redactar frases cortas, definir el tono de voz (profesional, amable, emocional), especificar idioma o acento y cuidar la relación entre voz e imagen. Esta capacidad convierte a Veo 3.1 en una herramienta de comunicación integral, capaz de generar vídeos que no solo se ven bien, sino que también **suenan bien** y transmiten emociones de forma realista. Una de las mayores innovaciones de Veo 3.1 es su **capacidad para generar audio nativo integrado**. El modelo no solo produce efectos sonoros y música ambiental, sino también **diálogos completamente sincronizados con el movimiento labial** de los personajes.

Tipos de audio generables:

- **Diálogos:** conversaciones, narraciones o mensajes publicitarios con tono y ritmo naturales.
- **Efectos de sonido:** pasos, viento, puertas, agua, objetos o ambiente urbano.
- **Música ambiental:** bandas sonoras adaptadas al contexto visual (corporativa, relajada, inspiradora, dinámica).
- **Sonido continuo entre escenas:** mantiene coherencia auditiva a lo largo del vídeo completo.

Comparativa de herramientas de vídeo IA

Característica	Sora	VEO 3	Synthesia	Runway Gen-3	Pika Labs
Interfaz amigable	Requiere prompts detallados	Técnica	Muy accesible	Simple	Muy accesible
Control cinematográfico real (lentes, cámaras, movimientos)	Muy alto (cine)	Muy alto (cine)	No	Limitado	Estándar
Consistencia de estilo	Depende del prompt	Depende del prompt	Muy limitada	Limitada	Limitada
Enfoque principal	Cine / publicidad premium	Cine / publicidad premium	Vídeo corporativo / formación	Vídeo viral creativo	Vídeos cortos sociales
Soporte para voz y subtítulos	No integrado	No integrado	Sí, nativo	Requiere extra	Requiere extra
Curva de aprendizaje	Alta	Alta	Baja	Baja	Media
Uso corporativo serio	Sí	Sí	Sí	Depende caso	Social

Recursos y enlaces útiles

- [Guía oficial de Sora para generación de video](#)
- [Manual técnico de Sora – Documentación completa](#)

- [Centro de ayuda de OpenAI – Sora y video generativo](#)
- [Blog OpenAI – Actualizaciones y novedades de Sora](#)
- [Políticas de uso de contenido generado por OpenAI/Sora](#)
- [Página oficial de Sora \(OpenAI\)](#)
- [Documentación técnica/API de Sora 2](#)
- [Canal oficial de OpenAI en YouTube](#)
- [Guía oficial de Veo 3 para video generativo](#)
- [Centro de desarrollo y registro en Google Veo 3 \(Vertex AI\)](#)
- [Blog oficial de Gemini/Veo – Actualizaciones y casos de uso](#)
- [Políticas y seguridad de contenido generado por Google](#)
- [Veo 3 en Google AI Studio](#)
- [Guía de generación de videos con Veo en Vertex AI](#)
- [Documentación de novedades y parámetros Veo 3.1 en Gemini](#)
- [Canal oficial de YouTube Google AI](#)

Conclusión

La inteligencia artificial ha redefinido por completo el concepto de producción audiovisual. Dejamos atrás la época en la que crear un vídeo requería rodajes extensos, equipos técnicos costosos y semanas de trabajo. Hoy, herramientas como ChatGPT, Sora y VEO 3 permiten que cualquier empresa, independientemente de su tamaño o presupuesto, pueda dar vida a sus ideas en cuestión de minutos. Lo que antes era un privilegio de las grandes marcas se ha convertido en una capacidad accesible para todos los negocios que deseen comunicar con impacto.

La clave de este cambio no está solo en la tecnología, sino en el enfoque estratégico con el que se utiliza. Sora, con su potencia para generar clips visuales a partir de texto; VEO 3, con su dominio cinematográfico y sonoro; y ChatGPT, como guionista y aliado creativo, forman un ecosistema que democratiza la producción audiovisual. En conjunto, convierten a cualquier profesional del marketing, comunicación o retail en un auténtico director creativo capaz de producir contenido con valor, coherencia e identidad.

El vídeo generado por IA se ha convertido en el lenguaje visual de una nueva era. No reemplaza la creatividad humana: la multiplica. Permite que las ideas cobren forma rápidamente, que los mensajes se adapten a distintos públicos, y que las marcas mantengan una presencia constante sin depender de procesos lentos o caros. Esta revolución no es meramente técnica; es cultural y estratégica. Las empresas que se

formen y adopten estas tecnologías ahora estarán un paso adelante en un mercado donde la atención es breve y la imagen lo es todo.

Para las pymes, la oportunidad es histórica. Los vídeos generados por IA no solo mejoran la visibilidad, sino que también humanizan la comunicación y elevan la percepción de profesionalismo ante los clientes. Un escaparate digital animado, un vídeo de producto con luz realista o una historia visual contada con emoción pueden marcar la diferencia entre pasar desapercibido o destacar frente a la competencia.

La integración de la inteligencia artificial en la comunicación visual no debe verse como una moda, sino como una inversión estratégica y continua. Formar equipos internos que dominen estas herramientas, crear bancos de prompts reutilizables y establecer un estilo visual propio será tan importante como definir la identidad corporativa o el tono de la marca. En esta nueva etapa, la narrativa visual se convierte en un activo empresarial tan relevante como la estrategia comercial o el servicio al cliente.

En resumen, la unión entre creatividad humana e inteligencia artificial marca el inicio de una nueva era en la comunicación empresarial. Una era donde las ideas tienen movimiento, las marcas se expresan con autenticidad y los recursos se optimizan como nunca antes. Adoptar el vídeo con IA no solo es una forma más eficiente de producir contenido: es una declaración de modernidad, visión y liderazgo.