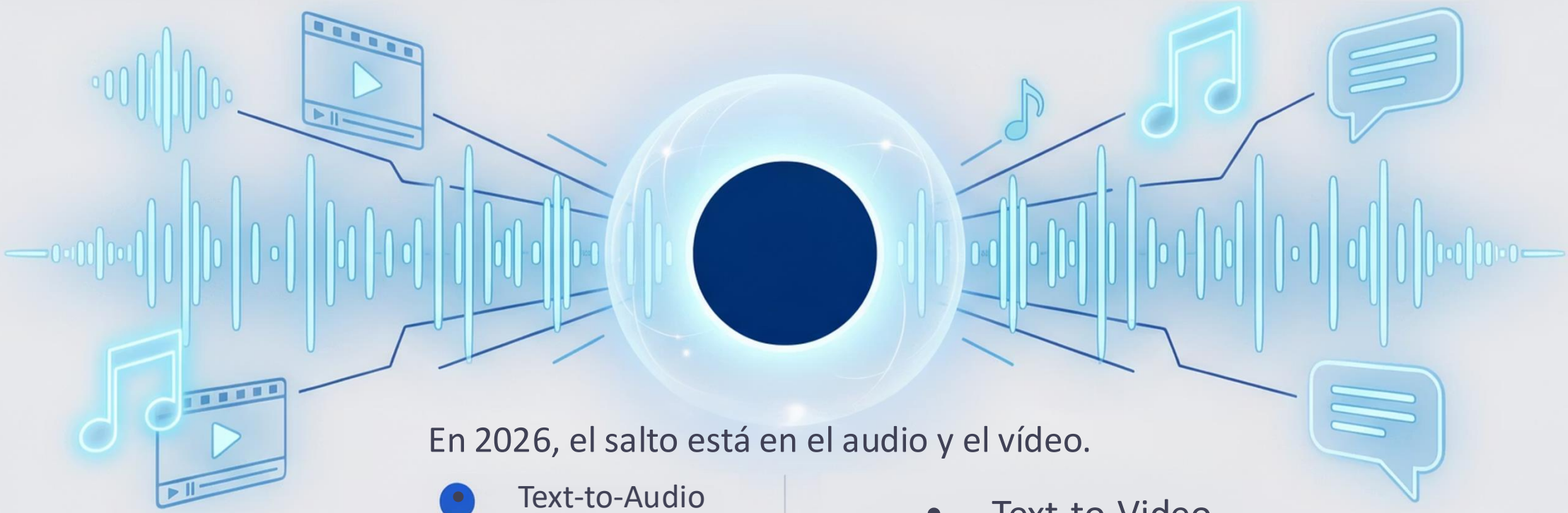


Generación de Audio y Vídeo con IA Generativa en las Administraciones Públicas

De la idea al contenido audiovisual: comunicación, accesibilidad y servicio público

¿Qué es la IA generativa multimedia?



En 2026, el salto está en el audio y el vídeo.

- Text-to-Audio
- Text-to-Speech
- Text-to-Music

- Text-to-Video
- Image-to-Video
- Video-to-Video

- Doblaje automático
- Edición conversacional

De crear contenido a dirigir dirigir contenido

La persona empleada pública actúa como:

- Guionista
- Directora de escena
- Revisora ética
- Supervisora legal
- Responsable de publicación



La IA como copiloto, no como piloto



La IA acelera, pero no decide. El criterio institucional es siempre humano.



El criterio institucional es siempre ..

Marco normativo

RGPD · AEPD · AI Act · Transparencia

- No subir datos personales innecesarios
- No usar rostros o voces sin base legal
- Señalar contenido sintético cuando proceda
- Revisar condiciones de uso de cada herramienta
- Mantener trazabilidad del proceso



Riesgos principales

- Deepfakes y suplantación de identidad
- Sesgos de representación
- Errores visuales o sonoros
- Licencias y derechos de autor
- Falta de accesibilidad
- Uso indebido de datos personales
- Pérdida de confianza institucional

Casos reales en administraciones

- Campañas turísticas · Avisos ciudadanos
- Vídeos de obras y urbanismo
- Accesibilidad web · Formación interna
- Comunicación de emergencias
- Contenido multilingüe · Participación ciudadana





















Bloque 2

El arte de preguntar

Prompt Engineering Multimedia



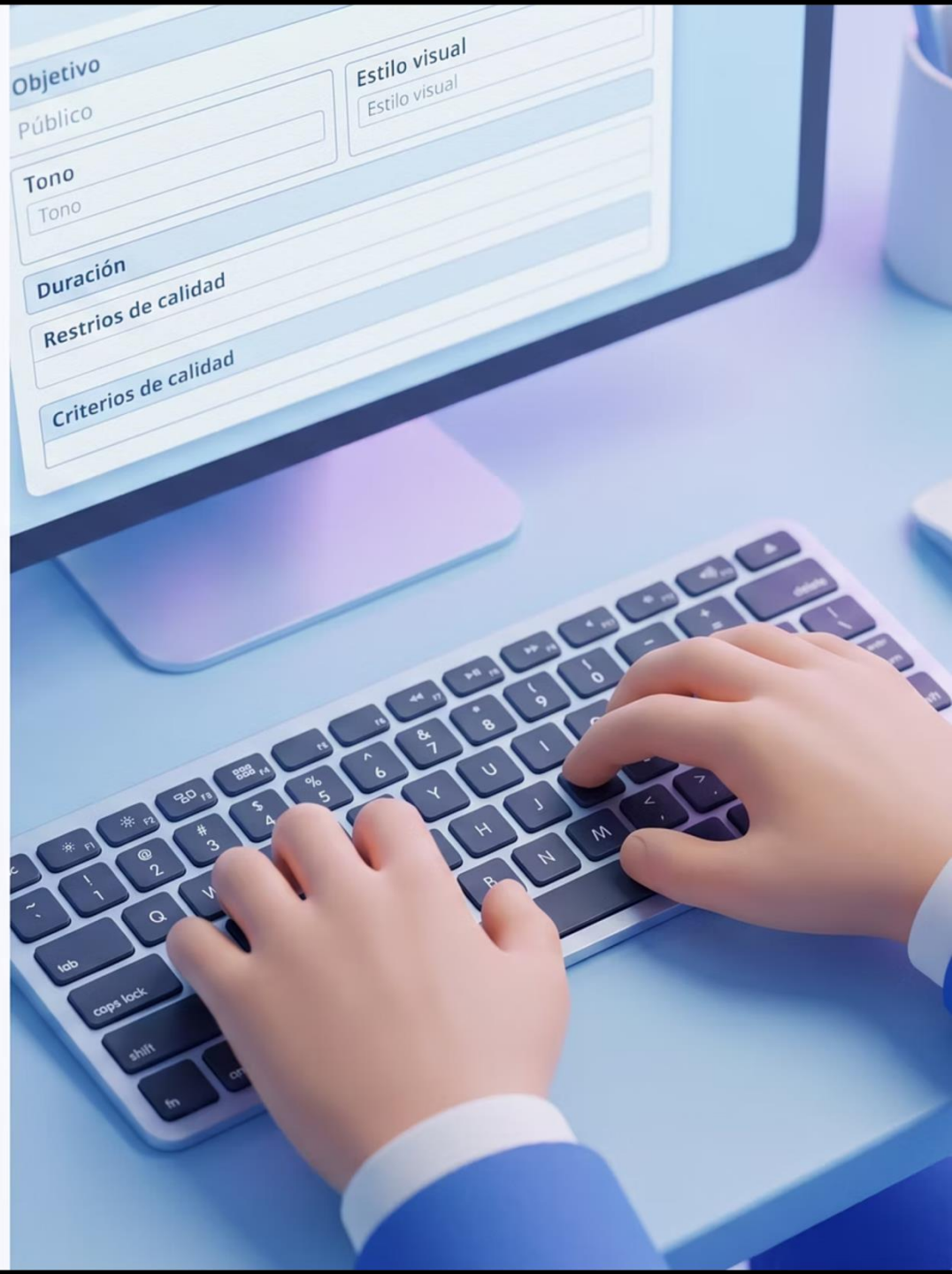


¿Qué es un prompt multimedia?

Un prompt multimedia define:

Un prompt multimedia define:

- Estilo visual
- Cámara
- Sonido-Idioma
- Restricciones
- Criterios de calidad



Fórmula base del prompt audiovisual



Ejemplo

Responsable de comunicación municipal. Vídeo vertical 10 s para Instagram sobre Instagram sobre reciclaje orgánico, familias, tono positivo, ilustración 3D, personas personas diversas, sin texto ilegible.

Prompt para audio

- Tipo: locución, jingle, podcast, aviso
- Idioma · Voz · Emoción · Ritmo · Pausas

- Duración · Contexto institucional
- Contexto institucional



Ejemplo

Voz institucional en valenciano, tono calmadod media, pausas breves, para informar de corte de agua sin generar alarma.

Prompt para vídeo

- Quién aparece · Qué ocurre · Dónde
- Movimiento de cámara · Iluminación · Estilo
- Duración · Relación de aspecto · Audio
- Restricciones éticas



Prompt para Gemini Omni

8 s

No solo genera vídeo desde texto: es creación y edición conversacional multimodal.











Ejemplo

Usa esta foto del paseo marítimo. Crea vídeo vertical 8 s para movilidad sostenible. Cámara hacia adelante, personas diversas en bicicleta, ambiente mediterráneo, sin rostros reconocibles.



Taller práctico 1

Diseñar el prompt base

Elige un servicio municipal y redacta:

- **Prompt para locución**
- **Prompt para vídeo**

- **Prompt de revisión ética**

3

Bloque 3

Audio generativo, voz sintética y accesibilidad

Ecosistema de audio generativo



- Locución institucional · Mensajes telefónicos telefónicos
- Podcasts internos · Resúmenes sonoros
- Doblaje · Música para campañas
- Jingles · Accesibilidad

Voz sintética institucional

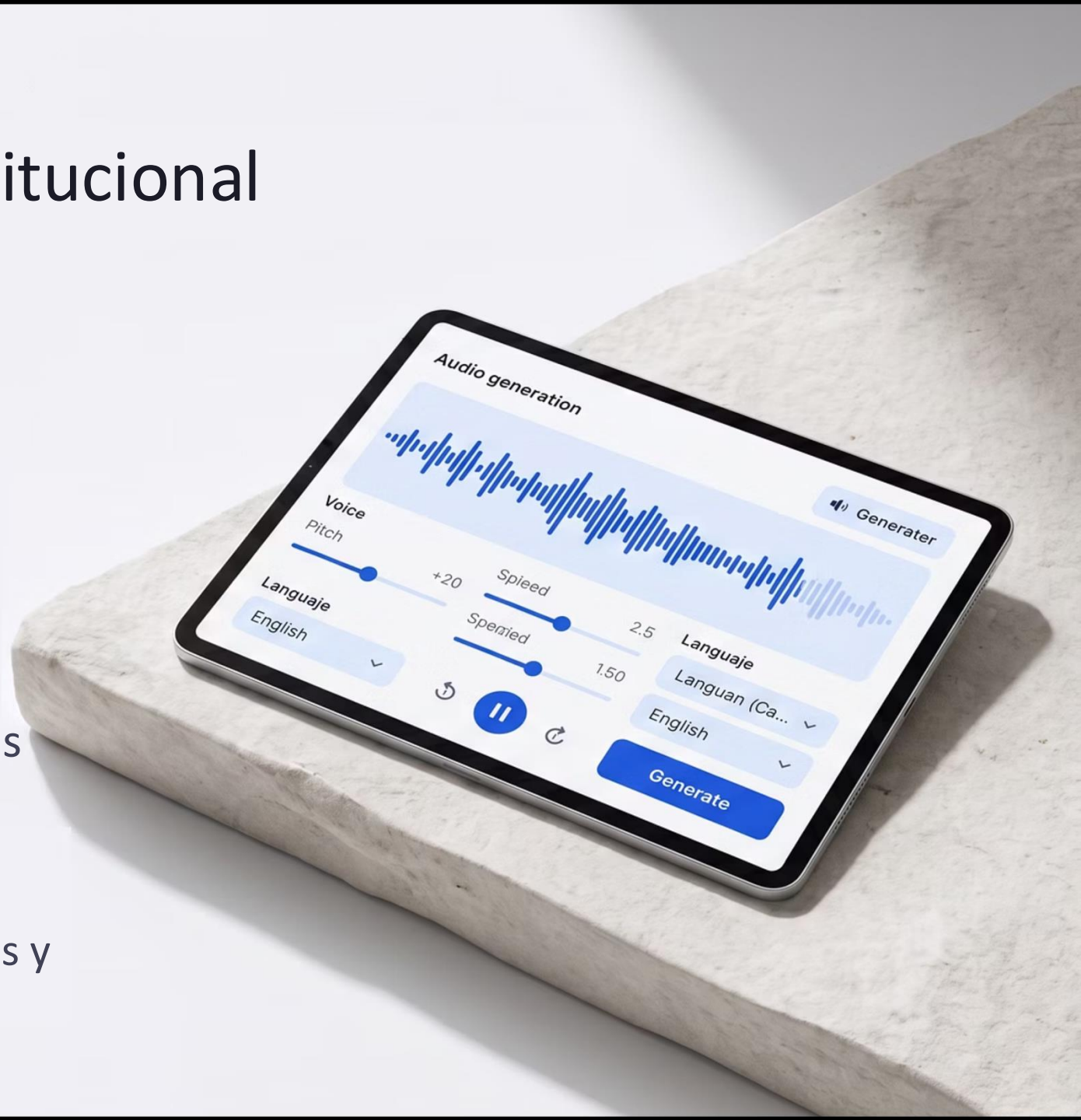
Herramientas

- ElevenLabs · Google AI Studio
- Gemini / NotebookLM

Herramientas

- Herramientas corporativas integradas

Mejora accesibilidad, rapidez y multilingüismo en notas de prensa, bandos y bandos y campañas.



ElevenLabs: locución, doblaje y clonación

- Text-to-Speech · Doblaje multilingüe
- Control emocional · Voces institucionales
- Clonación solo con consentimiento



⚠ No clonar voces de cargos públicos ni ciudadanía sin consentimiento claro y finalidad legítima.

Caso práctico: aviso ciudadano

Corte temporal de agua

Crear tres versiones:

- Formal institucional
- Cercana para redes sociales
- Calmado para centralita telefónica



Música generativa: Suno y alternativas

- Jingles de campañas municipales
- Música para vídeos turísticos
- Sintonías de podcasts · Música para eventos



⚠ Revisar siempre licencias y planes de uso comercial.

Caso práctico: jingle de de empadronamiento

Ejemplo de prompt

Jingle pop alegre en valenciano y castellano. 30 castellano. 30 segundos, ritmo optimista, estribillo fácil, tono inclusivo, dirigido a nuevos nuevos residentes.



NotebookLM: documentos en audio

Usos en el sector público

- Formación interna · Resumen de normativas
- Manuales de procedimiento · Pliegos
- Guías de atención ciudadana

Primera aproximación o material formativo, sin sustituir la lectura jurídica.



Doblaje automático y municipios turísticos

Idiomas de aplicación

[Francés]

[Árabe y otros]

Multiplica el alcance de una campaña sin
sin rehacer el contenido.







Accesibilidad: audio como servicio público público

- Personas con discapacidad visual
- Personas mayores y baja alfabetización digital
- Ciudadanía extranjera
- Materiales escuchables en móvil

Taller práctico 2: locución y música

- Tarea A: locución institucional de 30 segundos
- Tarea B: música instrumental para vídeo municipal

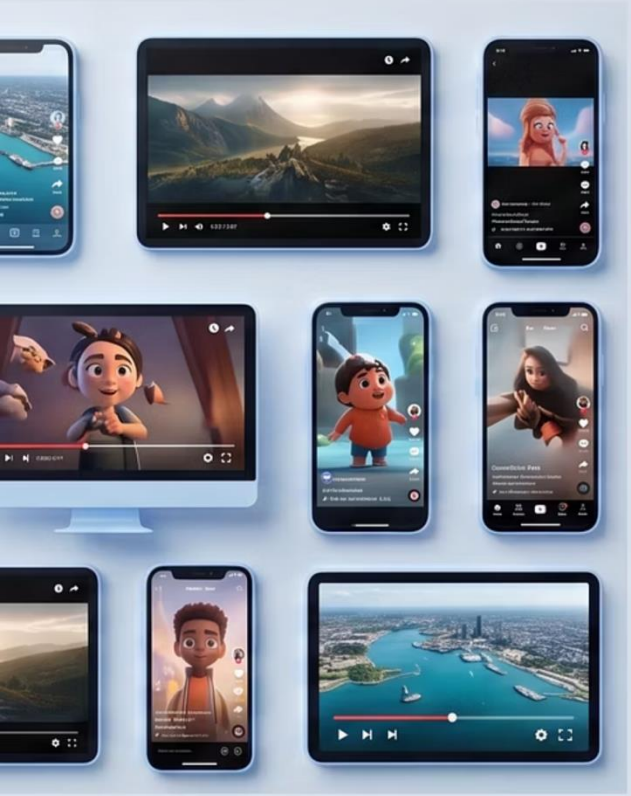


Entrega: prompt + valoración del resultado

Bloque 4

Vídeo generativo y Gemini Omni

Panorama actual del vídeo generativo

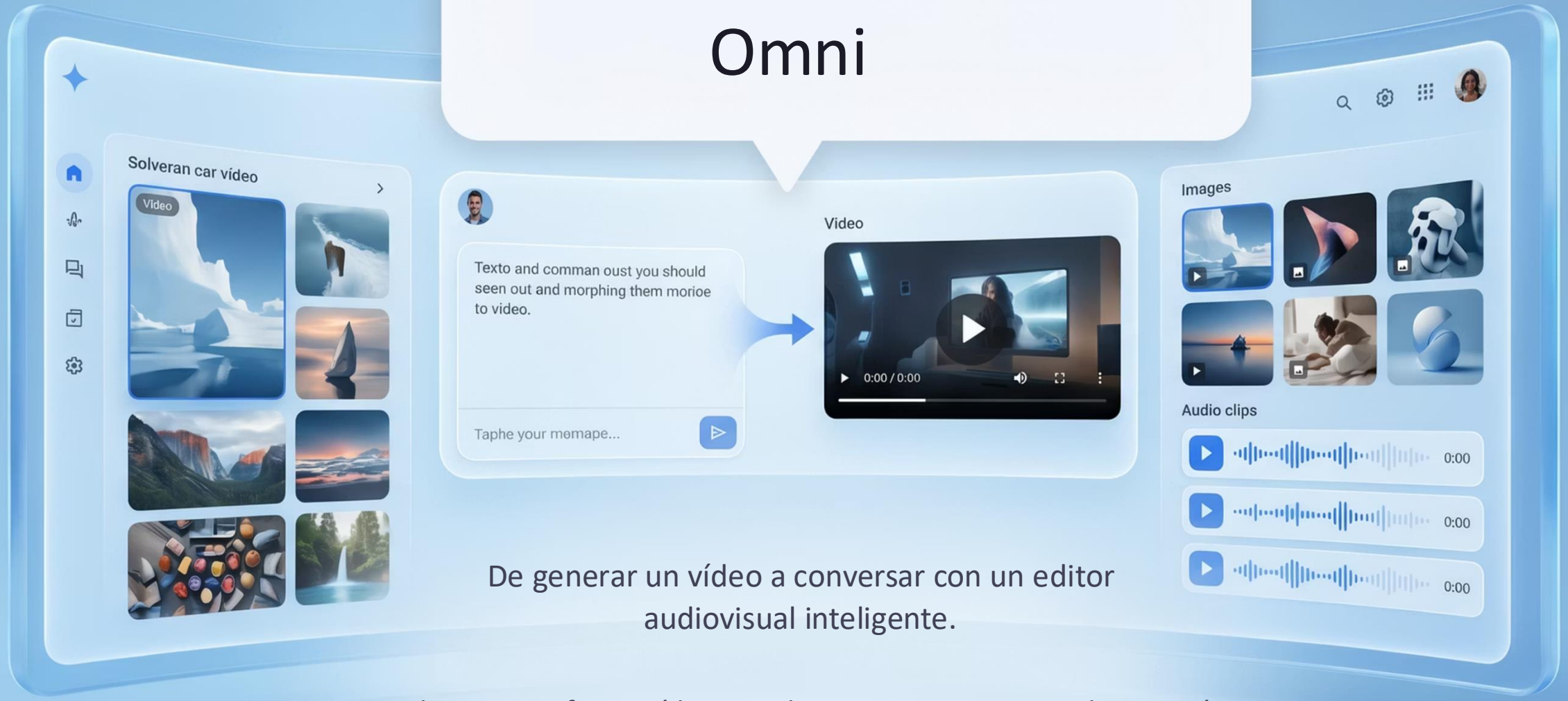


- Kling · Hailuo / MiniMax · Luma
- Pixverse · Grok · YouTube Shorts / Flow



De la curiosidad a la herramienta de preproducción y comunicación institucional.

La gran novedad: Gemini Omni



De generar un vídeo a conversar con un editor audiovisual inteligente.

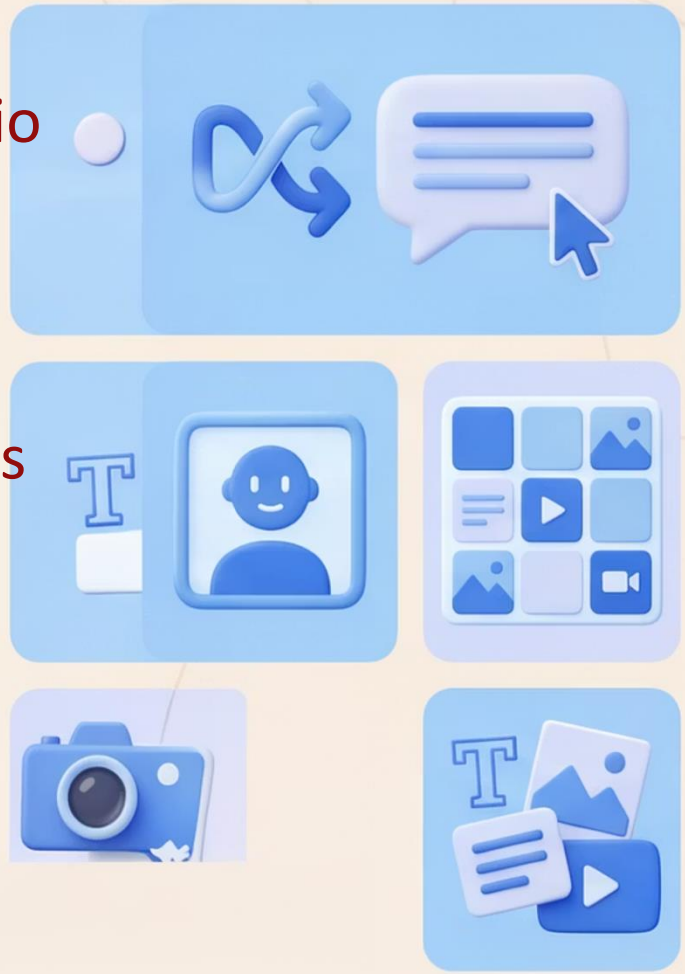
Combina texto, fotos, vídeo y audio en una experiencia de creación multimodal conversacional.

¿Qué puede hacer Gemini Omni?



- Crear vídeo desde texto
- Desde fotos de referencia

- Remix de material propio propio · Plantillas
- Avatares IA personalizados



- **Combinar texto, fotos y vídeo en una pieza**

Veo 3.1 vs. Gemini Omni

Veo 3.1	Gemini Omni
Generación de vídeo de alta fidelidad con audio nativo.	Creación y edición conversacional multimodal.
Entrada: prompt, imagen o API.	Entrada: texto, fotos, vídeo, audio.
Entornos: Gemini API, Flow, Vertex AI, Google Vids.	Entornos: App Gemini, Flow, YouTube Shorts.

Caso Gemini Omni: movilidad movilidad sostenible



Prompt

Vídeo vertical 10 s, Instagram, municipio mediterráneo. Personas diversas en bicicleta, luz de luz de mañana, tono positivo e inclusivo, cámara cámara suave, sin matrículas ni rostros reconocibles reconocibles.

“by Google Flow”

Caso Gemini Omni: animar archivo municipal

Prompt

Anima esta imagen con travelling suave. Mantén arquitectura original, respeta el carácter histórico, sin elementos modernos. Versión de 8 segundos para redes.



Reutiliza patrimonio visual y lo convierte en contenido dinámico.

Caso Gemini Omni: vídeo a vídeo

Mejora este clip para campaña institucional. Estabiliza, Estabiliza, mejora la luz, elimina distractores del fondo. fondo. No cambies el mensaje ni altera la identidad de identidad de las personas.





⚠ Con personas reales: consentimiento, finalidad y no manipulación engañosa.

Veo 3.1: vídeo de alta fidelidad

- Realismo cinematográfico · Audio nativo generado
- 720p, 1080p o 4K · Estilos cinematográficos
- API, Flow, Google Vids, Vertex AI
- Piezas controladas o producción técnica



Caso Veo: patrimonio natural y turismo



Prompt

Vídeo cinematográfico del Penyal d'Ifac al amanecer. Luz dorada, agua tranquila, gaviotas, cámara aérea suave, tono turístico elegante, sin texto, 8 s, horizontal.

Kling, Hailuo y Runway: cuándo usarlos



- Kling — Realismo físico y animación compleja (revisar servidores externos)
- Hailuo / MiniMax — Clips rápidos y estilos para redes (menor control institucional)
- Runway — Edición avanzada y producción profesional (coste y licencias)
- Gemini Omni — Creación conversacional multimodal (revisar condiciones)

Aplicaciones en urbanismo y servicios

- Previsualización de obras · Peatonalización
- Campañas de limpieza · Señalización temporal
- Explicación de nuevos servicios
- Simulación de espacios públicos
- Comunicación previa a la ciudadanía





Vídeo generativo para accesibilidad

- Subtítulos automáticos · Versiones con locución
- Lenguaje claro · Versiones multilingües
- Formatos verticales para móvil
- Piezas breves para WhatsApp, web y redes



Taller práctico 3: creación de vídeo

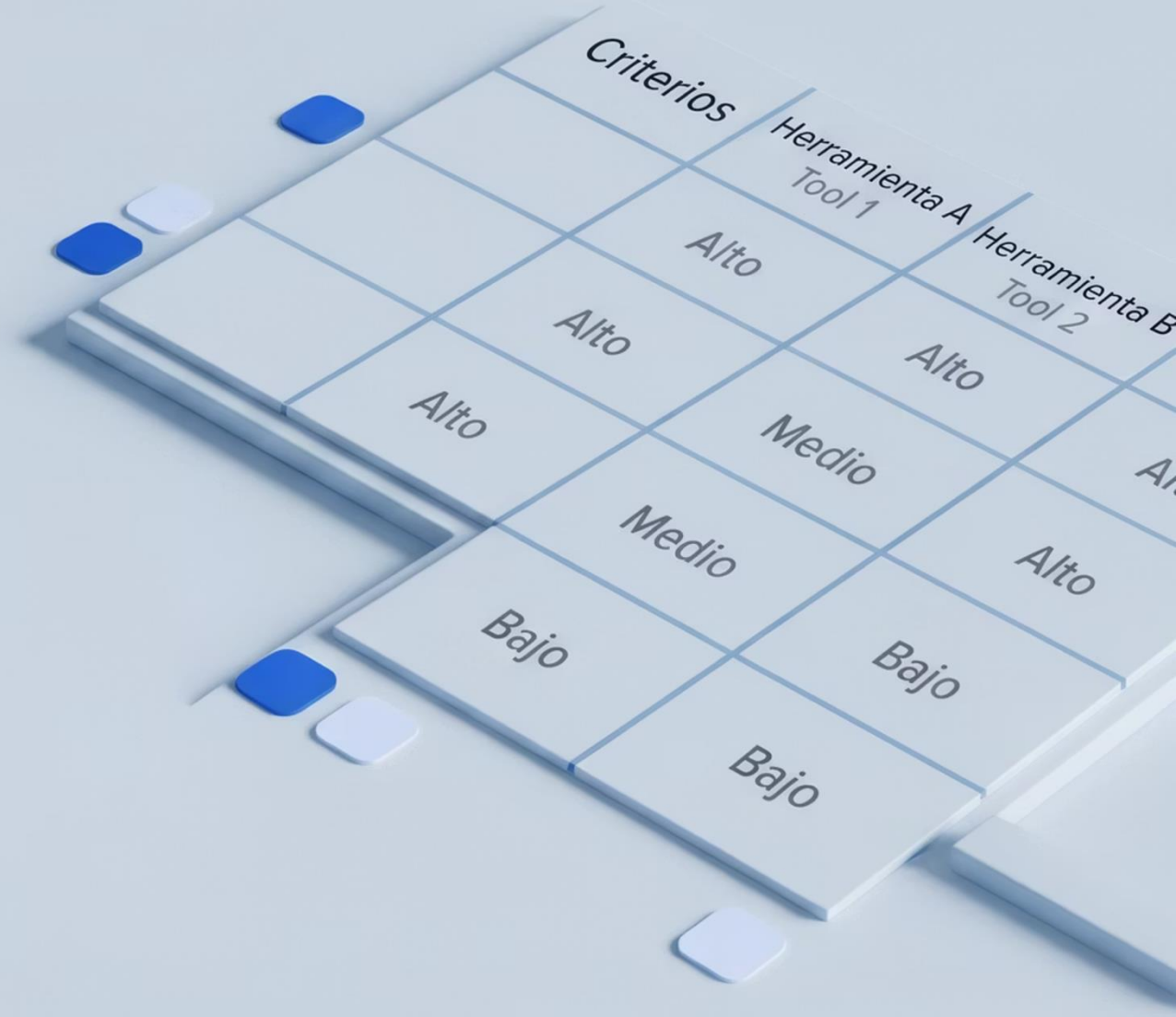
- Tarea A: vídeo 8–10 s sobre reciclaje, movilidad o turismo
- Tarea B: versión animada de una imagen municipal
- Tarea C: prompt de mejora — más institucional, más inclusivo, menos publicitario



Comparativa de resultados

Evalúa cada herramienta según:

- Fidelidad al prompt · Realismo · Inclusión
- Errores visuales · Calidad del audio
- Facilidad de uso · Riesgo legal
- Uso recomendado en contexto público





Bloque 5

Proyecto integrado, ética aplicada

Flujo de trabajo completo



- Definir objetivo · Redactar guion
- Crear locución · Generar música o ambiente
- Crear o animar vídeo · Montar pieza final
- Revisar accesibilidad · Revisar licencias
- Validar jurídicamente · Publicar Publicar con transparencia

Proyecto final integrado

Diseña una campaña municipal breve:

- Tema: movilidad, reciclaje, empadronamiento, cultura...
- Público objetivo · Prompt de vídeo · Prompt de audio
- Prompt de vídeo
- Prompt de audioos · Riesgos detectados
- Herramientas elegidas
- Riesgos detectados
- Versión final propuesta



Checklist antes de publicar

- ¿Hay datos personales?
- ¿Aparecen menores?
- ¿Hay rostros reales?
- ¿Hay voz clonada?
- ¿Tenemos derechos de uso?
- ¿Puede inducir a error?
- ¿Representa diversidad?
- ¿Es accesible?
- ¿Revisado humanamente?
- ¿Debe etiquetarse como generado con IA?





Transparencia y etiquetado

Si el contenido puede parecer real y condicionar la percepción ciudadana, actúa con transparencia.

La ciudadanía tiene derecho a no ser confundida por contenido sintético de una administración pública.

Recomendaciones para empezar el lunes

1. Empezar con casos de bajo riesgo
2. No subir datos personales a herramientas externas
3. Crear prompts inclusivos y revisados
4. Revisar licencias antes de publicar
5. Usar IA para prototipos, no para decisiones finales
6. Documentar qué herramienta se ha usado
7. Validar siempre antes de publicar



El futuro: estudios audiovisuales públicos ligeros

De herramientas aisladas a entornos integrados donde texto, imagen, vídeo, audio y edición se mezclan en una sola conversación.



Gemini Omni representa precisamente esa dirección.



Cierre

La IA generativa audiovisual no sustituye la comunicación pública.
La amplía.

Permite explicar mejor, llegar a más personas y producir materiales más accesibles. El criterio, la ética y la responsabilidad siguen siendo humanos.