Generación de Audio y Video con IAG

La Nueva Frontera de la Creación de Contenido

TABLE OF CONTENTS

01
¿Qué es la IAG Multimedia?

02
Oportunidades y Aplicaciones Prácticas

03
Desafíos y Consideraciones Éticas

04
Taller Práctico: las Herramientas

05
Ejercicios Prácticos

06
Conclusiones y Próximos pasos

01

¿Qué es la IAG Multimedia?

Entendiendo los nuevos Copilotos Creativos.

La IA que Crea Sonido y Vídeo

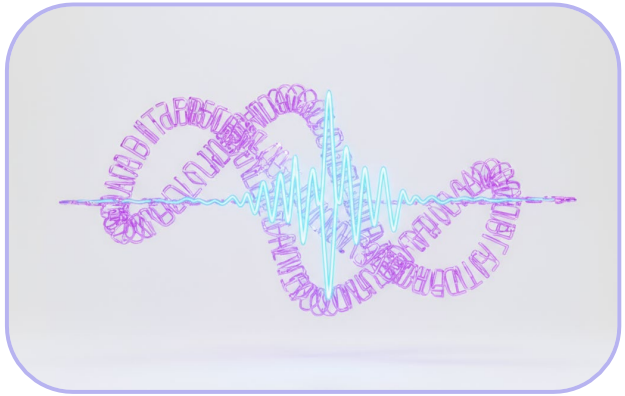
A diferencia de la IA tradicional que analiza o clasifica, la **IA Generativa** puede crear **contenido audiovisual** completamente **nuevo**. Ha sido entrenada con millones de horas de contenido para aprender **patrones**, **ritmos**, **estilos** y **narrativas** del mundo.

Actúa como **copiloto creativo**: le damos una idea con palabras y nos devuelve una pieza única de música, voz o clip de video.



De las Palabras al Audio: Text-to-Audio

Funciona como un **diálogo creativo** con un técnico virtual. Le pedimos algo en lenguaje natural y él lo compone.



1. **La petición (Prompt)** Describimos género musical, ritmo, instrumentos y tono que necesita nuestra comunicación.
2. **La Creación** El modelo interpreta la petición y compone una pieza musical original, generando pistas de cada instrumento.
3. **El Resultado** Obtenemos un archivo de audio único listo para usar en un proyecto.

De la Idea a la Escena: Text - to - Video

Esta tecnología nos convierte en **directores de proyectos visuales** Nos permite crear una **maqueta virtual** de campañas o infraestructuras antes de que existan, utilizando simples palabras

1. **La petición (Prompt)** Describimos los personajes, las acciones, el entorno, el ángulo de cámara y el estilo visual.
2. **La Creación:** La IA interpreta este guión y construye fotograma a fotograma un video coherente.
3. **El Resultado:** Un vídeo corto ideal para ilustrar un concepto, usar en redes sociales o presentar una propuesta de forma clara y convincente.



Imágenes Animadas: Image - to - Video

Esta técnica nos permite **dar vida a imágenes estáticas**. Es especialmente útil para añadir **dinamismo** a contenido que ya tenemos como fotografías o ilustraciones

1. **El Punto de Partida:** Seleccionamos una imagen (fotografía tomada o generada).
2. **La Petición (Prompt)** Le damos una instrucción de movimiento. Podemos pedirle que anime la cámara o que anime elementos de la imagen.
3. **El Resultado:** Un vídeo corto que transforma una imagen plana en una pieza visual dinámica, ideal para redes.



02

Oportunidades y Aplicaciones Prácticas

Transformando la comunicación, el marketing y la formación

¿Por qué es una Herramienta Estratégica?



Agilidad en la Comunicación

Permite responder a **necesidades imprevistas** y lanzar **campañas urgentes** (avisos, eventos) en minutos, no en días.



Eficiencia

Optimiza los recursos públicos, **reduciendo** la **dependencia** de proveedores externos para la creación de contenido audiovisual.



Autonomía para Departamentos

Permite que áreas como Turismo, Urbanismo o Cultura generen sus **propios materiales** de comunicación profesional.



Impacto en el Ciudadano

Facilita la creación de contenido más visual y comprensible para explicar proyectos complejos, mejorando la transparencia y **participación ciudadana**

Promoción Turística y Cultural

La Concejalía de Turismo necesita un **vídeo** para promocionar las **fiestas patronales**. Quieren una **música emotiva** y **única**. Podemos generar en minutos una banda sonora que capture la **esencia mediterránea** y **festiva** del evento.

Prompt (Suno):
Música instrumental para un vídeo promocional de las fiestas patronales. Estilo mediterráneo, alegre y emotivo con guitarras españolas y percusión suave.



Servicios al Ciudadano y Accesibilidad

El **mensaje** de la centralita **telefónica** del Ayuntamiento está **obsoleto**. Además, se necesita hacer **accesibles** las noticias de la web municipal para personas con **discapacidad visual**. Deseamos generar **locuciones** de alta calidad en segundos.



Prompt (ElevenLabs):
(Tono calmado y profesional) "Ha contactado con el Ayuntamiento de Madrid. En estos momentos, nuestras oficinas están cerradas. Nuestro horario de atención es de lunes a viernes de 9 a 14 horas.

Formación Interna y Comunicación

Es obligatorio que todo el personal conozca sus **responsabilidades** según la **Ley** de **Protección de Datos**. Sin embargo, los documentos oficiales son **largos**, **complejos** y su lectura es **tediosa**. Utilizaremos un asistente de IA para que analice los documentos y genere un **podcast-resumen** con los puntos más importantes en **formato conversacional**

Prompt (NotebookLM):
Crea una conversación de audio que resuma las obligaciones más importantes que un empleado debe conocer según estos documentos de la Ley de Protección de Datos.



03

Desafíos y Consideraciones Éticas

Una mirada crítica a los riesgos y responsabilidades

Grandes Poderes, Grandes Responsabilidades



Desinformación y *Deepfakes*: El riesgo de suplantación de identidad nos obliga a una **verificación humana extrema** antes de cualquier publicación.



Calidad y Coherencia: La IA puede generar imperfecciones. Es crucial **revisar cada detalle** para mantener la imagen profesional de la institución.



Licencias y Uso del Contenido: No todo lo generado es de uso libre. Es obligatorio **revisar los términos de cada herramienta** para asegurar el cumplimiento legal.

El Reflejo de Nuestros Sesgos

La IA aprende de Internet, por lo que puede **heredar y amplificar estereotipos sociales** (de género, raza, etc.) si no se la guía correctamente.

Como servicio público, tenemos la **responsabilidad** de usar *prompts* inclusivos para generar contenido que represente la **diversidad real de nuestra sociedad**



La Regla de Oro: Supervisión Humana

La IA es **copiloto**, no el piloto. El juicio crítico, el contexto y la responsabilidad final de lo que publicamos en nombre de la institución **siempre** deben recaer en una persona.



04

Taller Práctico: Conociendo las Herramientas

Recorrido por los modelos de Audio y Video

Compositor de Bandas Sonoras: Suno AI

Suno es la herramienta líder para la **creación** de **canciones** completas desde cero. Funciona como un **estudio de grabación virtual completo**. Es capaz de **componer**, **interpretar**, **producir** y **cantar** una canción basándose únicamente en una descripción de texto.

Ideal para:

- ★ Crear **bandas sonoras** originales.
- ★ Generar **jingles pegadizos** para campañas de radio y redes.
- ★ Componer **música ambiental**.

Debes tener en cuenta:

- ➢ Necesitas una **suscripción** para tener los **derechos**
- ➢ No genera locuciones ni discursos, es puramente **musical**

URL acceso https://suno.com/

SUNO

Música para Campaña de Empadronamiento

El censo **municipal** está **desactualizado** Muchos nuevos **residentes** no se empadronan al llegar, lo cual afecta directamente a la financiación que recibe el municipio del Estado y a la planificación de servicios Queremos **modernizar** la **campaña** creando un **jingle** para **radio** local.



Prompt: Jingle pop alegre en español, con un estribillo que diga ‘Tu ciudad, tu casa, empadrónate ya’. Ritmo animado con guitarras acústicas y palmas.

ElevenLabs: La Voz Institucional

ElevenLabses el estándar mundial en generación de **voz ultra-realista**. Es la herramienta perfecta para una **comunicaciónrinstitucional** clara, **profesional** y **accesible**

Idealpara:

- ★ **Realismo** Genera voces naturales que capturan la emoción delmensaje.
- ★ **Soporte multilingüe:** Perfecto para la comunicación en las lenguas cooficiales.
- ★ **Clonación de voz:** Identidad sonora consistente para la institución.

Debestener encuenta:

- ➢ **Uso ético:** La clonación de voz **siempre** requiere consentimiento explícito.
- ➢ El **coste** se basa en los caracteres generados.

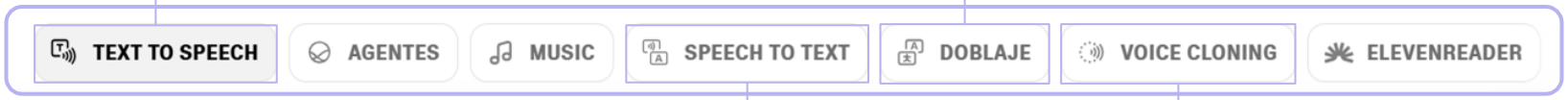
URLaccesohttps://elevenlabs.io/es

IIElevenLabs

Estudio de Sonido Profesional y Sencillo

Text to Speech: Locutor Virtual. Escribe texto y lo convierte en un audio con voz realista.

Doblaje: Traductor Instantáneo. Sube un vídeo y la herramienta lo traduce y dobla.



Speech to Text: Transcriptor Automático. Sube un archivo de audio y la IA lo convierte en texto.

Voice Cloning: Permite crear una réplica de una voz (necesita consentimiento y suscripción) para utilizarla en comunicaciones

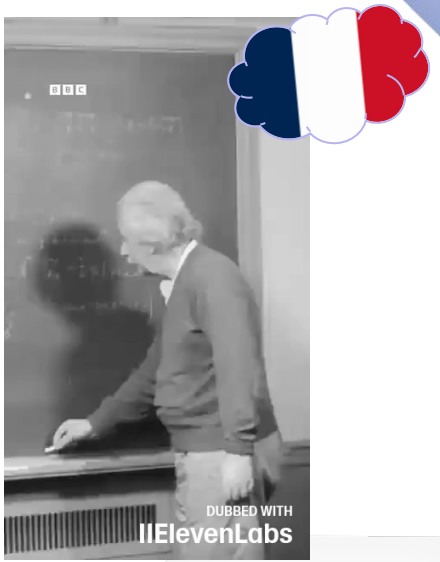
Comunicación Multilingüe

La Concejalía de Cultura está organizando una "Semana de Ciencia" para fomentar el interés por la divulgación científica en la ciudad. El Ayuntamiento tiene un hermanamiento con una ciudad francesa (Burdeos) y quiere compartir con ellos los vídeos explicativos

La **solución** que propone **ElevenLabs**

1. Creamos el **vídeo** en **castellano**
1. Se utiliza la herramienta de "**Doblaje**" de ElevenLabs.
1. Se genera una versión del **video** doblada al **francés** manteniendo un tono de voz similar al original.

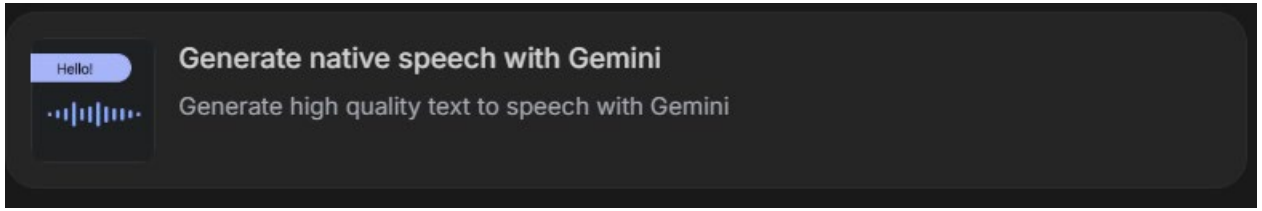




AI Studio de Google

Siquiereshaceruso de la opción text to speech..

https://aistudio.google.com/



Veo3: El Cineasta de Google

Esel **modelo** de vídeo más avanzado de **Google**, diseñado para crear **escenas** de alta fidelidad con una **calidad cinematográfica** y un entendimiento profundo del mundo real. Debemos seleccionar la Herramienta "**Videos con Veo**".

Ideal para:

- ★ **Realismo cinematográfico:** Domina estilos visuales como "tomas de dron", "timelapse" o "plano aéreo".
- ★ **Comprensión de la Naturaleza** Excepcional generando paisajes, agua y vegetación.
- ★ **Coherencia** Mantiene la consistencia de los objetos y entorno del video.

Debes tener en cuenta:

- ➢ **Acceso limitado:** Solo disponible con suscripción Pro (4 usos diarios).
- ➢ **Coste Computacional** Su potencia implica un servicio con coste elevado.
- ➢ **Errores** Aunque es avanzado, puede producir rarezas físicas.

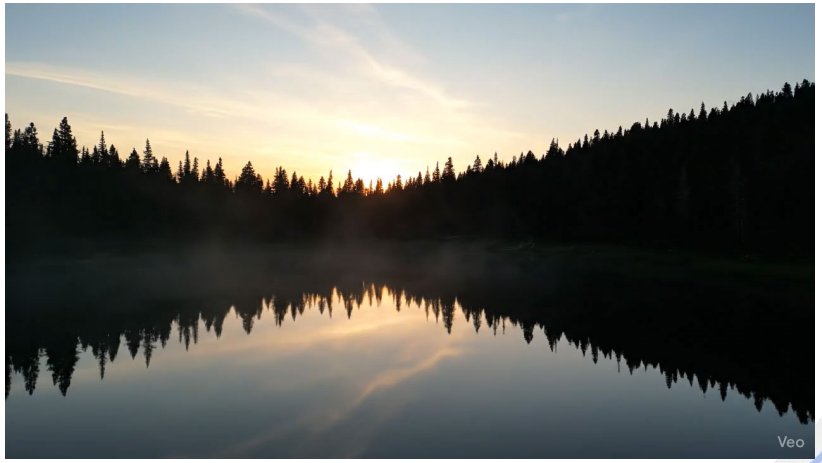
URL acceso https://gemini.google.com



Promoción de Parques Naturales

La Concejalía de Medio Ambiente quiere crear un **vídeo impactante** para promocionar el **parque natural** del municipio. Un rodaje con drones y equipos especializados es logísticamente complejo. Queremos utilizar **Veo3** para generar un montaje de vídeo espectacular que capture la **belleza** del **parque** en diferentes momentos del día.

Prompt: Planeo aéreo del atardecer sobre un lago de montaña, con el sol reflejándose en el agua. Se aprecia un bosque de pinos con niebla flotando entre los árboles.



Visualizando el Futuro de la Ciudad

La Concejalía de Urbanismo va a presentar un **proyecto** para construir una **nueva rotonda ajardinada** Un plano técnico es difícil de entender para la ciudadanía y puede generar rechazo Queremos generar una **maqueta virtual** en video **ultra-realista** de cómo quedará la rotonda **integrada** en un **entorno real**.



Prompt: Video aéreo con dron de la nueva rotonda ajardinada a la entrada del municipio. Día soleado, tráfico ordenado. Estilo hiperrealista, colores vivos.

El Creador de Mundos de OpenAI

Es el **revolucionario modelo** de OpenAI, capaz de generar **escenas complejas** con múltiples personajes, movimientos precisos y una asombrosa capacidad para contar **historias visuales**

Ideal para:

- ★ **Complejidad y Narrativa:** Genera escenas con múltiples personajes e interacciones de forma coherente.
- ★ **Consistencia de Personajes** Mantiene la apariencia de un personaje
- ★ **Interpretación del Lenguaje** Entiende prompts detallados

Debes tener en cuenta:

- ➢ **Acceso Exclusivo** En fase de pruebas con un grupo muy reducido.
- ➢ **Física Compleja** Tiene dificultades con simulaciones físicas complejas
- ➢ **Intensivo en Recursos** Uso muy costoso y reservado

URL acceso https://sora.chatgpt.com
Modelo Sora 2 (no disponible sin invitación)



Kling: La Alternativa Creativa

Desarrollado por Kuaishou (creadores de Kwai, rival de TikTok), Kling es un modelo **potentísimo** que compite directamente con Sora destacando por su **gran calidad** y su **versatilidad** para crear **diferentes estilos** visuales

Ideal para:

- ★ **Flexibilidad de Estilos:** Excelente tanto en estilos de animación como fotorrealismo.
- ★ **Física Realista** Simula movimientos complejos con una precisión asombrosa

Debes tener en cuenta:

- ➢ **Acceso en Lista de Espera** Su acceso es limitado, requiere registro en lista de espera
- ➢ **Texto en Vídeo:** Su punto débil es la generación de texto legible integrado.

URL acceso https://app.klingai.com/

KlingAI

Campañas Creativas: Hábitos Saludables

La Concejalía de Salud Pública quiere lanzar una **campaña** en **redes** para que los **niños** y **niñas** coman más **fruta** y **verdura**. Un vídeo tradicional puede ser aburrido o no captar su atención. Queremos crear un **vídeo** de **fantasía** corto y **divertido** relacionado con verduras y que sea **mágico** y **emocionante**.



Prompt: Estilo animación 3D de Pixar. Un brócoli con una capa de superhéroe vuela por una cocina. A su lado, una zanahoria con gafas de sol le saluda. El ambientes es brillante y alegre.

Seedance: El Movimiento Humano

Seedance es un modelo de IA de ByteDance que genera video a partir de texto o imágenes. Su punto fuerte es la **creación** de **movimientos humanos** increíblemente **realistas** y **fluidos**. Ideal cuando su **protagonista** es una **persona** y el **movimiento** es clave.

Ideal para:

- ★ **Movimiento Humano:** Ideal para animar personajes de forma realista y fluida.
- ★ **Coherencia de Personajes** Mantiene la consistencia de los personajes
- ★ **Entiendo el contexto:** Interpreta prompts complejos

Debes tener en cuenta:

- ➢ **Fuerte Sesgo Cultural:** Si no se especifica etnia y contexto cultural, el modelo tiende a generar figuras de rasgos asiáticos
- ➢ **Dependencia de un buen Prompt:** La calidad del resultado depende directamente de lo específico y detallado que es el texto.

Seedance

Grok Imagine: Integrado en X

Permite generar **imágenes** y **vídeos cortos** de forma casi instantánea a partir de texto o animando imágenes existentes. Es una solución ágil para **contenido** en **redes sociales**

Ideal para:

- ★ **Creación rápida:** Genera vídeos de menos de 15 segundos para reaccionar a tendencias.
- ★ **Prototipado Visual:** Permite crear borradores y *storyboards* de campañas
- ★ **Contenido Dinámico** Anima imágenes estáticas.

Debes tener en cuenta:

- ➢ **Calidad Inconsistente:** Tiene resultados poco realistas
- ➢ **Riesgo de Privacidad** Todo el contenido generado es público por defecto.
- ➢ **Audio Genérico** El sonido se añade automáticamente.

URL acceso https://grok.com/imagine

Grok

Campaña de Reciclaje para Familias

Queremos crear un **vídeo corto** y **amigable** para una **campaña** de concienciación sobre **reciclaje** en **redes sociales** Se busca un tono positivo y educativo. Utilizando un **estilo** de **ilustración** para minimizar los posibles errores visuales.

Prompt Imagen: Una familia sonriente (padre, madre, niño y niña) pasea por una calle peatonal, limpia y soleada. Estilo de animación 3D de Pixar. Iluminación cálida de tarde.

Prompt Animación: Usando la imagen generada, añade un suave *travelling* de zoom hacia la madre y el contenedor azul. Al mismo tiempo, anima el papel para que vuele en un pequeño arco desde su mano hasta el contenedor. Cuando el papel entre, haz que un sutil destello de luz aparezca sobre el contenedor.



Meta AI: La Creatividad Integrada

Meta AI no es una herramienta de vídeo independiente, sino un potente **animador** de **imágenes**. Su función "**Animate**" permite **transformar** cualquier **fotografía** estática en un **vídeo** corto y dinámico, haciéndolo ideal para la comunicación ágil en redes sociales.

Ideal para:

- ★ **Accesibilidad y Rapidez**: Fácilmente integrado en aplicaciones de uso cotidiano.
- ★ **Función Clave** (Image-to-Video): Anima imágenes existentes.

Debes tener en cuenta:

- ➢ **Herramienta muy básica**: Ofrece poco control sobre el tipo de movimiento.
- ➢ **Generador de Video**: Es capaz de generar videos muy sencillos a partir de texto.
- ➢ **Acceso** desde la pestaña "Crear > Video".

URL acceso https://www.meta.ai/media

Meta

Dando Vida a Nuestra Imagen

El Ayuntamiento tiene un gran archivo de fotografías de alta calidad. Son **recursos estáticos** y quiere **maximizar** su **impacto** en redes. Utilizaremos una imagen como punto de partida y añadiremos **movimiento** de cámara en un video corto y dinámico.



Prompt: Añade un efecto lento de *travelling zoom* hacia el edificio más alto (la Giralda de Sevilla).



Runway: Una Caja de Herramientas

Runway, además de generar vídeo, es una **plataforma** de **edición completa** potenciada por IA. Es la herramienta más accesible y versátil del mercado para la **comunicación audiovisual diaria**.

Puntos fuertes:

- ★ **Herramientas IA:** Incluye Text-to-Video, Image-to-Video, eliminar objetos, expande encuadre y mucho más.
- ★ **Acceso inmediato:** Disponible y público.
- ★ **Flujo de Trabajo Integral:** Permite generar un clip, editarlo, añadir efectos…

Debes tener en cuenta:

- ➢ **Duración Limitada:** Los clips generados son cortos (aptos para redes)
- ➢ **Modelo de Créditos:** Uso limitado por créditos, requiere planificación

URL Acceso https://app.runwayml.com/video-tools/

runway

Más opciones gratuita s

- vheer.com → Una herramienta sencilla que se especializa en convertir imágenes en animaciones. Es gratuita, ilimitada y no requiere registro. **No añade marcas de agua** a los videos generados.

- Magi → También de origen chino, permite generar videos de forma gratuita mediante un sistema de créditos. Su interfaz tipo "canvas" permite organizar proyectos visualmente. Es necesario registrarse con un correo electrónico.

- Qwen → Esta plataforma ofrece generaciones ilimitadas, sin marcas de agua y sin límites de exportación de video. La calidad es buena para ser una opción gratuita, aunque los videos pueden ser un poco irregulares. Necesitas registrarte con tu cuenta de Google.

Ejercicio ÚNICO: COMPARACIÓN DE MODELOS

Crea un video a partir de un prompt pensado para tu trabajo utilizando todas las herramientas vistas en la sesión de hoy.

Envía a j.aparicio@umh.es

Nombre herramientas UTILIZADAS	Prompt utilizado	Opinión sobre el resultado generado

NO MANDES EL VIDEO GENERADO

05

Ejercicios Prácticos

Una mirada crítica a los riesgos y responsabilidades

Ejercicio 1: Dando Voz a la Comunicación Oficial

Tienesel texto de una **nota** de **prensa** importante que debe ser **comunicado**de forma clara y **accesible**



Tarea Utilizar una herramienta Text-to-Voice para convertir este comunicado en un archivo de audio con una voz profesional y un tono institucional.

Objetivo: Crear un recurso de audio para la radio local, el canal de WhatsApp del ayuntamiento para complir con las normativas de accesibilidad de la web municipal.

Ejercicio 2: Campaña Movilidad Sostenible

Esta semana es la “Semana Europea de la **Movilidad**” y la Concejalía de Medio Ambiente demanda un **vídeo corto** y llamativo para Instagram con la finalidad de **fomentar** el uso de la **bicicleta**.

Tarea Crear un corto animado que muestre de forma positiva y energética a ciudadanos de todas las edades usando la bicicleta en la ciudad. El estilo de animación es libre.

Objetivo: Generar contenido atractivo y compatible que conecte con una audiencia joven y concienciada, promoviendo hábitos saludables y sostenibles.



Ejercicio 3: Campaña Tenencia de Mascotas

El Ayuntamiento quiere lanzar una **campaña** para **concienciar** sobre la importancia de recoger los excrementos de las **mascotas** en parques y jardines municipales.



Tarea Genera un corto fotorrealista que muestre a un dueño paseando a su perro de forma responsable y concienciado con el estándar de limpieza urbano.

Objetivo: Crear material visual de alta calidad, positivo y relatable para la campaña.

Ejercicio 4: Felicitación Navideña

Encuentra una **fotografía navideña**, ya sea de la plaza de tu Ayuntamiento o del paseo de Málaga. Quieres crear una **felicitación navideña** para las **redes sociales** que sea más **dinámica** que una simple imagen.



Tarea Anima la fotografía. Genera también una canción navideña o villancico con el estribillo que se te ocurra para la ocasión.

Objetivo: Utilizar el archivo fotográfico para crear contenido emotivo, maximizando el impacto de las comunicaciones

06

Conclusiones y Próximos pasos

Integrando la IA Generativa en nuestro flujo de trabajo

La IA Generativa como Copiloto



Es una Herramienta de Apoyo, no de Reemplazo La IA acelera la creatividad y la producción, pero el juicio crítico, el contexto y la responsabilidad final siempre deben ser humanos.



La Clave es Saber Qué Usar y Cuándo No existe "la mejor IA". El éxito reside en entender la misión específica y elegir la herramienta adecuada para esa tarea.



La Supervisión Humana es Innegociable Somos los responsables finales de lo que publicamos. Debemos revisar cada resultado para detectar errores, sesgos y asegurar que el mensaje se alinea con los valores del servicio público.

Recomendaciones para Empezar

- **Empieza con Herramientas Accesibles** Explora las versiones gratuitas. Familiarízate con la dinámica de los prompts sin coste.
- **Define tu Misión:** Antes de empezar, pregúntate: ¿Qué necesito exactamente? La respuesta te dirá qué herramienta es la más eficiente.
- **Invierte Tiempo en el Prompt:** La calidad del resultado depende de la calidad de tus instrucciones Sé descriptivo, usa adjetivos, define el estilo y refina la petición.
- **Revisa Siempre las Licencias** Antes de usar un contenido, dedica cinco minutos a revisar los términos de uso, especialmente si es para una campaña pública.
- **Úsala para Inspirarte:** ¿Atascado en una idea para una campaña? Pídele a la IA que te genere 3 conceptos diferentes.