



Inteligencia Artificial en la gestión de las ONGD

Un acercamiento práctico y crítico a la IA

Noviembre 2025

**Cooperar,
transformar
y compartir**

**Somos una cooperativa de
comunicación que trabaja
desde los principios del diseño
social y la economía solidaria.**

**Free
press**
S. Coop. Mad.

Jose Ibañez
jose@freepress.coop

freepress.coop

Sesión 1

Fundamentos, Ética y Aplicaciones Prácticas

- Qué es la IA
- Riesgos, sesgos y responsabilidades
- Aplicaciones prácticas en ONG
- IA para la gestión y mejora de procesos

Sesión 2

Estrategia, Uso Avanzado, Automatización Ética

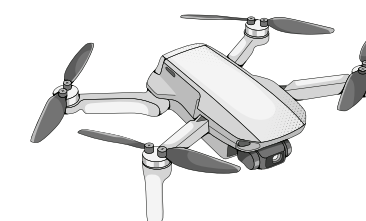
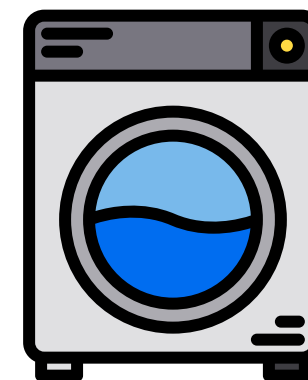
- Configuración avanzada, interfaz y buenas prácticas
- Escritura avanzada con IA
- IA para planificación estratégica y redes sociales
- Evaluación y análisis automatizado



Qué es la IA y cómo funciona

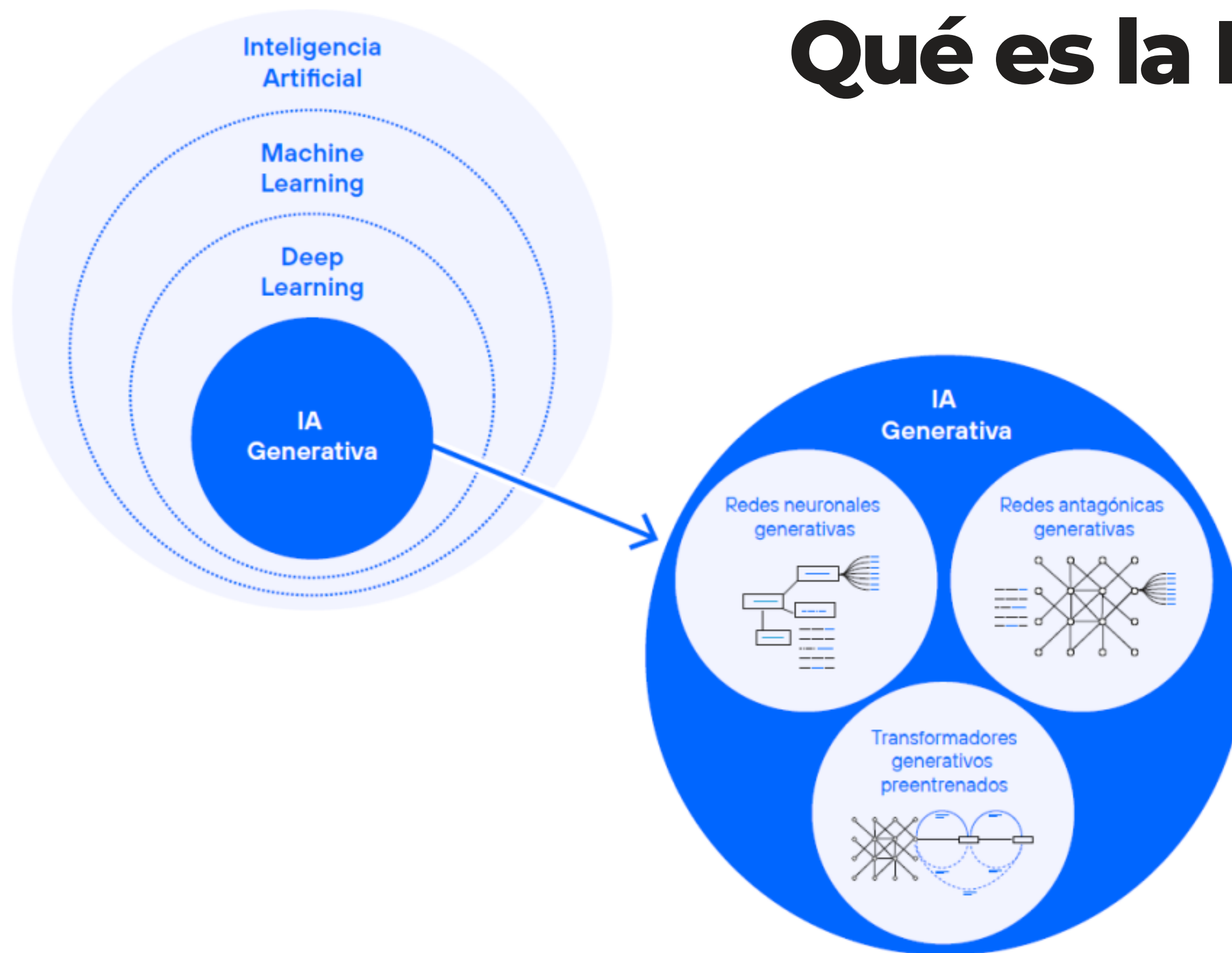
INTELIGENCIA ARTIFICIAL

- └— IA simbólica / basada en reglas
- └— Machine Learning
 - └— Aprendizaje supervisado
 - └— No supervisado
 - └— Reinforcement Learning
- └— Deep Learning
 - └— Redes neuronales
 - └— Transformers (GPT generative pre-trained transformer)
- └— NLP (Procesamiento de lenguaje natural)
 - └— Modelos de lenguaje (LLM)
- └— IA generativa
 - └— Texto
 - └— Imagen
 - └— Audio
 - └— Vídeo
- └— Robótica e IA física
- └— Sistemas multiagente / agentes LLM
- └— Sistemas de recomendación





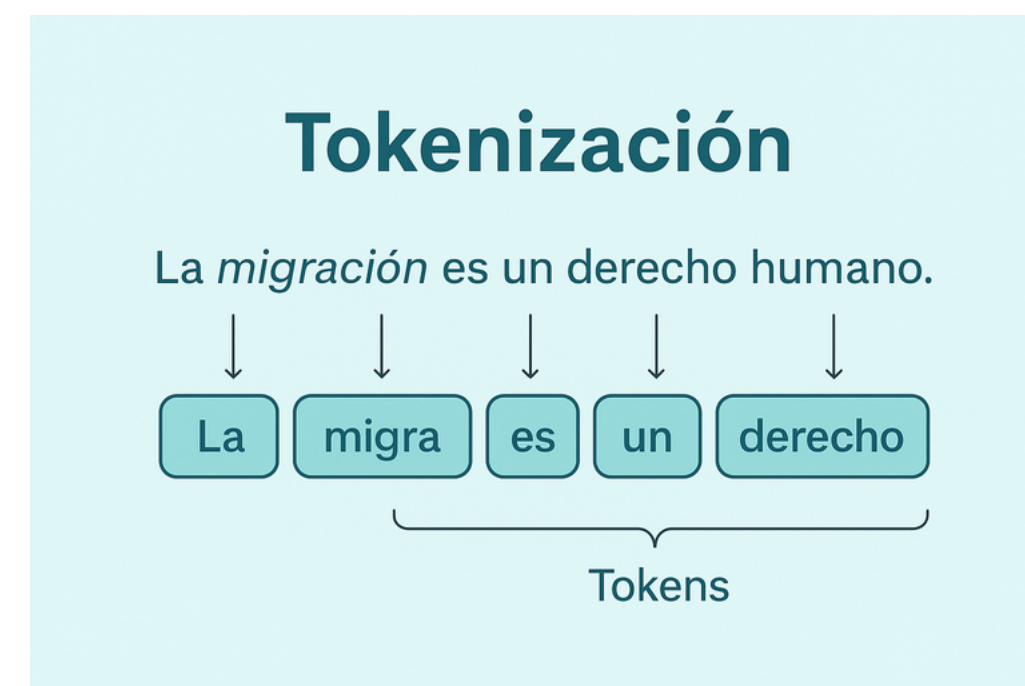
Qué es la IA y cómo funciona





IA Generativa

LLM > **Maching Learning** > Aprendizaje por **predicciones** basadas en **probabilidad** dentro de un conjunto de datos



Contexto o memoria

Es la cantidad de texto que la IA puede tener “en mente” a la vez. Cada modelo tiene un límite de contexto (una especie de memoria temporal).

Si el texto cabe dentro del límite → buena coherencia

Si el texto lo supera → empieza a olvidar partes → errores



¿Cómo aprende una IA generativa?

La IA ajusta los números internos para minimizar errores
Los humanos definen qué es “bueno” y qué es “malo”

Preentrenamiento

El modelo predice
partes de
millones de textos

Fine-tuning

Humanos enseñan
a seguir
instrucciones

RLHF

Humanos clasifican
respuestas como
“mejores” o “peores”

★ ¿Qué no aprende nunca una IA?

Verdad

Intención

Empatía real

Ética

Contexto social profundo

Realidades políticas

Valores humanos propios



Parámetros

Los parámetros son ponderaciones, un número interno del modelo que determina cómo relaciona unas palabras con otras.

- cuánta importancia dar a una palabra
- cómo se combina con otras
- qué estilo de frase es más probable
- qué relaciones tiene (“pobreza” → “ingresos”, “derechos”, “migración”...)
- cómo prioriza tonos o estructuras
- cómo generaliza patrones

GPT-3: 175.000 millones

GPT-4: estimado más de 1 billón

GPT-5: no se ha publicado, se estima entre 1,5 y 5 billones

Claude 3 Opus: comparable a GPT-4

Modelos pequeños (locales): entre 7 y 70 mil millones

Más parámetros = más potencia, pero también más sesgos heredados y más consumo energético.





Los parámetros y el aprendizaje son un factor clave en el consumo energético

Consumo energético

◆ Entrenamiento de modelos grandes (estimación)

Modelo	Estimación de energía de entrenamiento	Equivalencia aproximada
GPT-3	~1,287 MWh (MIT, 2021)	120 hogares/año
PaLM (Google)	~3,300 MWh	300 hogares/año
Modelo multimodal grande (estimación)	5.000–20.000 MWh	una ciudad pequeña durante 1 día

◆ Generación de contenido (uso cotidiano)

Acciones	Consumo (estimación)
20–50 prompts complejos	consumo similar al de cargar un móvil
1000 consultas al día en una ONG	consumo similar al de una nevera doméstica

◆ Consumo de agua para refrigeración (según estudios recientes)

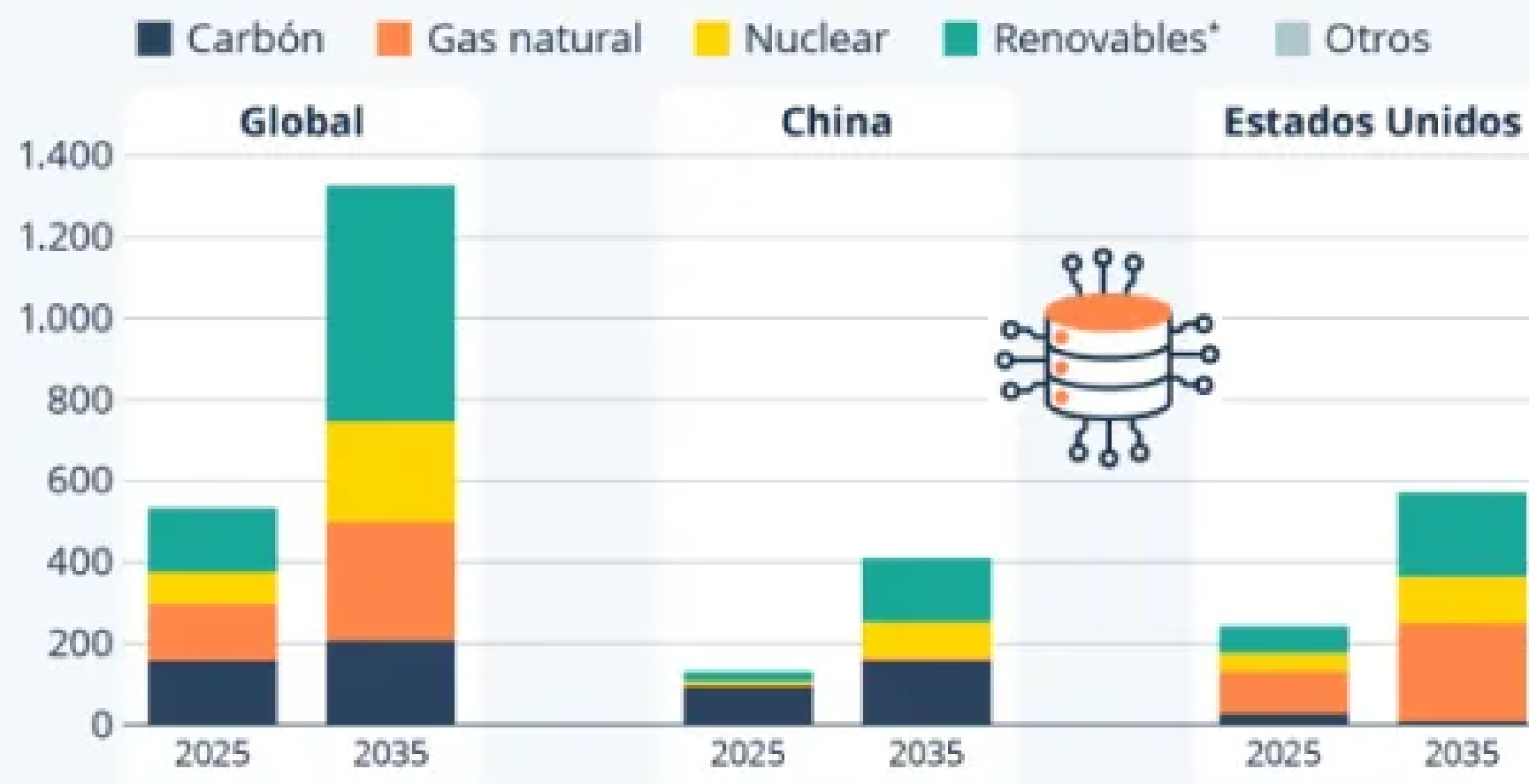
Entrenar un modelo grande puede consumir hasta 700.000 litros de agua (Universidad de Riverside).
Ejecutar ChatGPT en horas punta puede requerir 500 ml por conversación en refrigeración indirecta.





La IA dispara el consumo energético global

Generación de electricidad para abastecer los centros de datos, por fuente de energía (en teravatio-hora)



* Solar fotovoltaica, eólica y otras renovables.

Con base en la combinación de combustibles de la electricidad consumida físicamente por los centros de datos y no en la combinación contractual de los distintos operadores.

Fuente: IEA



statista



Qué es (y qué no es) la IA generativa

✓ Qué es

La IA generativa es un tipo de tecnología capaz de **producir contenido nuevo** a partir de patrones **aprendidos** en millones de datos existentes.

Puede generar:

- textos (como un informe o un post),
- imágenes,
- audio,
- vídeo,
- código,
- resúmenes, listas, tablas, etc.

Su base son modelos matemáticos muy complejos –no ideas ni razonamientos humanos– **entrenados con enormes cantidades de datos**.

✗ Qué no es

• No piensa.

No razona, no comprende, no interpreta emociones ni contextos sociales.

• No sabe.

No tiene conocimientos propios, solo patrones estadísticos.

• No tiene conciencia, intención o valores.

No diferencia entre verdad y mentira.

No distingue si un mensaje es ético, dañino o discriminatorio.

• No puede decidir por sí misma.

Solo responde a lo que se le pide, aunque la respuesta suene muy convincente.



La IA generativa no interpreta la realidad: la imita



Familias de IA generativas según los medios de entrada y de salida

Texto a texto



Traducir, resumir, buscar información, aprendizaje en general, corregir textos.

ChatGPT, Peer, Perplexity, Bard, Copilot, Copymatic, Jasper, TutorAI, Bing.

Texto a imagen



Generar imágenes, inspirarse, crear arte, crear avatares, crear logos.

Midjourney, DALL·E, Bing Image Creator, Stable Diffusion, Adobe Firefly, Fotor, Craiyon.

Texto a audio



Generar ficheros de audio, audiolibros, pódcast, crear locuciones con voces de otras personas o la propia, traducir vídeos, crear temas musicales con parámetros seleccionados, modificar otros temas musicales.

AudioLM, Whisper, Jukebox, Murf, Mubert, AudioStrip (per separar música i veu), Boomy.

Text a vídeo



Generar vídeos con características específicas seleccionadas, editar vídeos, traducir vídeos.

Phenaki, Sundify, Synthesia

Texto a código



Generar o mejorar el propio código, documentarlo.

Alphacode, Codex, GitHub, Copilot, Ghostwriter, Tabnine, SourceAI.

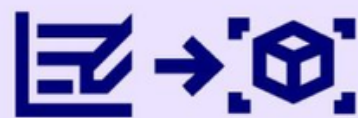
Texto a ciencia



Obtener respuestas con base científica, hacer una lluvia de ideas con la investigación existente, encontrar artículos científicos relacionados con las respuestas de la IA.

Elicit, Consensus, Scite.

Texto a 3D



Generar imágenes para utilizar en videojuegos, obtener una imagen más completa y profunda de un objeto, mejorar las simulaciones o el diseño de elementos de realidad virtual.

Dreamfusion, Magic3d.

Imagen a texto



Proveer descripciones más precisas o resumidas de imágenes, mejorar la accesibilidad en el caso de discapacidades visuales.

Flamingo, VisualGPT.

Otras



Descubrir nuevos algoritmos, realizar multitud de tareas de IA más allá del nivel de especialización que ahora mismo tienen las herramientas de IA. También el caso de las IA multimodales que pueden procesar diferentes tipos de datos

Alphatensor, GATO, CoDI.

* El panorama en cuanto a herramientas soportadas con inteligencia artificial cambia deprisa, con algunas que se posicionan y después quedan obsoletas. Los ejemplos proporcionados pueden variar y casi todas tienen versión gratuita o periodo de prueba



Modelos around the world



perplexity



deepseek



Estados Unidos

OpenAI → GPT, Sora

Google → Gemini

Anthropic → Claude

Meta → LLaMA

Microsoft → Copilot

Europa

Mistral (Francia)

Aleph Alpha (Alemania)

Stability AI (Reino Unido)

HuggingFace (Francia/EE.UU.)

China

Baidu → ERNIE

Alibaba → Qwen

Huawei → Pangu

Deepseek → DeepSeek AI



Alucinaciones, sesgos y errores

ALUCINACIONES

Cuando la IA **inventa la información** que parece real pero no lo es

- Citar leyes que no existen
- Atribuir frases a activistas que nunca dijeron eso
- Inventar estadísticas
- Crear historias coherentes pero falsas
- Fabricar datos sobre proyectos de ONG

¿Por qué ocurre?

Porque el modelo prefiere generar algo plausible antes que decir “no lo sé”.

SESGOS

La IA repite y amplifica:

- Estereotipos de género
- Racismo o colonialidad
- Narrativas asistencialistas
- Miradas occidentales o elitistas
- Imágenes de pobreza estereotipadas
- Enfoques capacitistas

Si los datos tienen sesgo → en la IA también.

ERRORES ESTRUCTURALES

- La IA no sabe si algo es cierto.
- No tiene acceso actualizado a todos los datos.
- Los modelos en la nube pueden confundirse entre versiones.
- No entiende ironías, dobles sentidos ni contextos culturales profundos.





Riesgos principales de la IA

1. Alucinaciones (información inventada)

- La IA puede crear datos falsos pero plausibles.
- Genera leyes, cifras o historias inexistentes con total seguridad.
- **Riesgo:** decisiones basadas en información no verificada.

2. Sesgos (reproducción de discriminaciones)

- Si Internet tiene racismo, machismo o visión colonial → la IA también.
- Imágenes y textos pueden reforzar estereotipos sobre pobreza, migración, género, discapacidad, ruralidad, etc.
- **Riesgo:** comunicación dañina o estigmatizante.

3. Dependencia, pérdida de criterio o de identidad

- Uso excesivo = pérdida del estilo propio y de pensamiento crítico.
- **Riesgo:** uniformidad narrativa, empobrecimiento del lenguaje.

4. Privacidad y datos sensibles

- Subir datos personales o información de beneficiarios puede violar GDPR.
- **Riesgo:** vulneración de derechos.

5. Desinformación y manipulación

- La IA puede producir contenidos persuasivos muy rápidamente.
- **Riesgo:** contribuir sin querer a narrativas falsas o dañinas.





Responsabilidades y buenas prácticas

✓ 1. Control humano significativo

Nunca publicar texto generativo sin revisión humana.

Dos ojos extra mínimo en contenidos sensibles.

✓ 2. Prompting seguro

- Evitar lenguaje estigmatizante.
- Pedir autodiagnóstico de riesgos narrativos.
- Incluir siempre enfoque de derechos.

✓ 3. Política interna de uso de IA

Toda ONG debería tener:

- Normas sobre qué se puede subir a la IA.
- Modelos permitidos y no permitidos.
- Criterios éticos para imágenes.
- Procedimientos de verificación.
- Límites de uso en campañas sensibles.

✓ 4. Protección de datos

No subir:

- nombres
- historias personales
- fotos reconocibles
- números de expediente
- información vulnerable

✓ 5. Transparencia

Indicar cuándo un texto está creado o asistido por IA cuando sea relevante.

✓ 6. Revisión con perspectiva crítica

Antes de publicar, preguntar:

- ¿Refuerza estereotipos?
- ¿Puede dañar a alguien?
- ¿Es fiel a la misión de la ONG?
- ¿Puede inducir desinformación?



Prompting

Principal técnica para comunicarse con una IA para obtener resultados útiles, precisos y alineados con un enfoque concreto.

A través del prompt le decimos a la IA quién debe ser, qué debe hacer, cómo debe hacerlo y bajo qué criterios.

- reduce error y alucinación
- disminuye sesgos
- evita lenguaje paternalista
- mejora coherencia y tono
- permite generar contenido seguro
- reduce consumo energético
- ayuda a mantener identidad institucional
- evita que la IA “improvise” narrativas dañinas





Prompting: estructura

1) Rol

Define quién es la IA.

Ej.: “Actúa como experta en comunicación con enfoque de derechos.”

2) Contexto

De qué organización y sobre qué trata.

Ej.: “Somos una ONG que trabaja con población migrante.”

3) Objetivo

Qué quieres conseguir.

Ej.: “Queremos redactar un mensaje para redes evitando estereotipos.”

4) Tarea concreta

Lo que debe generar.

Ej.: “Escribe 3 versiones breves para Instagram.”

5) Formato

Cómo debe entregarlo.

Ej.: “Formato tabla: versión, tono, riesgos detectados.”

6) Restricciones

Límites éticos, legales o estilísticos.

Ej.: “Evita paternalismo, imágenes de sufrimiento, lenguaje salvador.”





Prompting: tipos

■ 1. Prompt directivo (el básico)

Da una instrucción clara.

Ej.: “Redacta un texto claro y respetuoso explicando nuestra campaña de vivienda digna.”

■ 2. Prompt iterativo (el mejor para ONG)

Se trabaja en varias rondas.

Ej.: “Dame una versión inicial. Luego la revisamos juntos.”

■ 3. Prompt reflexivo

Pide a la IA que detecte riesgos y sesgos.

Ej.: “Genera el texto y luego identifica posibles sesgos o estereotipos.”

■ 4. Prompt evaluador

La IA analiza contenido preexistente.

Ej.: “Revisa este post e indica qué debe corregirse para evitar lenguaje de salvador.”

■ 5. Prompt comparativo

Genera varias versiones según tono/público.

Ej.: “Tres versiones de este mensaje: juvenil, institucional, divulgativa.”



Prompting: buenas prácticas

- ✓ Pedir revisiones y autodiagnóstico de sesgos
 - ✓ Dar contexto social y organizativo
 - ✓ Establecer límites éticos y legales
- ✓ Prohibir inventar datos (ponerlo explícito)
- ✓ Forzar transparencia (“cita riesgos narrativos”)
 - ✓ Definir siempre el público objetivo
 - ✓ Usar lenguaje inclusivo en el propio prompt
- ✓ Trabajar en varias iteraciones, no usar la primera respuesta



Definir la estrategia de uso de IA

Necesitamos una estrategia dentro de cada organización y como red de entidades del tercer sector

La IA no es una moda; es un cambio estructural

1) Objetivos

¿Qué queremos lograr con IA?

2) Límites y exclusiones

¿Qué NO se va a permitir?

3) Criterios éticos y narrativos

¿Cómo queremos que suene la organización?

4) Casos de uso autorizados

Aplicaciones permitidas

5) Herramientas apropiadas

Según uso y criterios propios

6) Roles y responsabilidades

Definir quién hace qué





Definir la estrategia de uso de IA

Dimensiones de evaluación

“En comunicación social, no necesitamos 200 herramientas: necesitamos 6 bien escogidas.
Lo importante no es la tecnología, sino cómo la usamos y con qué criterios éticos.”

1) Tipo de contenido que genera (texto, imagen, vídeo, datos)

reproducción de sesgos y estereotipos
invención de datos
generación de narrativas
riesgo de desinformación
deepfakes

3) ¿Permite verificación y trazabilidad?

citan fuentes, muestran referencias o
permiten ver el razonamiento interno

2) Transparencia del modelo y control del usuario

herramientas abiertas (open-weight) permiten
auditar parámetros, revisar funcionamiento, ajustar
filtros

herramientas cerradas y poco transparentes no
permiten saber cómo se entrenaron o sus fuentes

4) Potencial de causar daño narrativo o de representación

citan fuentes, muestran
referencias o permiten ver el
razonamiento interno





Definir la estrategia de uso de IA

Modelos genéricos: para todo

Herramienta	Mejor para...	Ética / seguridad	Notas clave
Claude 3 Sonnet / Opus	Análisis crítico, textos complejos, revisiones éticas, resúmenes largos, guiones	Muy alta	Recomendación #1 para ONG por prudencia y calidad.
ChatGPT (GPT-4.1 / 5)	Creatividad, campañas, ideas, versiones de estilo, narrativa, redes	Media/alta	Excelente para creatividad; más riesgo de inventar datos.
Gemini (Google)	Integración con Google Docs, tablas, presentaciones, análisis multimodal	Alta	Ideal si vuestra ONG ya usa Workspace.
Perplexity	Investigación con fuentes, verificación de datos, contexto	Muy alta	Imprescindible para evitar alucinaciones. No es un generador creativo.
LLaMA 3 (open source)	Soberanía tecnológica, uso local, datos privados	Alta (si se configura bien)	Requiere conocimientos técnicos; muy útil para ONG grandes.
Mistral / Mixtral	Tareas rápidas, eficiencia, servidores locales	Alta	Alternativa europea open-weight.





Definir la estrategia de uso de IA

Herramientas visuales: ¿usar generadores de imágenes o no?

Herramienta	Funcionalidad	Riesgo ético	Uso recomendado en ONG
Canva (Magic Media)	Prototipos, ilustraciones, presentaciones	Bajo/medio	Mejor opción visual para ONG; no genera hiperrealismo peligroso.
Ideogram	Ilustración tipográfica, cartelería creativa	Medio	Usar para prototipos de carteles conceptuales.
Midjourney	Arte y fotorrealismo de altísima calidad	Alto	Desaconsejado para imágenes de personas. Solo para arte conceptual.
DALL·E 3 (OpenAI)	Integración con ChatGPT, control por texto	Medio/alto	Igual que Midjourney: prototipos, nunca material final.
Stable Diffusion XL (local)	Control total, ejecución local	Variable	Solo para equipos técnicos y con políticas estrictas.





Definir la estrategia de uso de IA

Vídeo generativo: ¿cuándo y cómo usarlo?

Herramienta	Uso	Riesgo	Recomendación ONG
Runway Gen-2	Clips cortos conceptuales	Alto	Solo prototipos internos; nunca vídeos institucionales.
Pika Labs	Vídeo animado	Medio	Bueno para animaciones estilo motion graphics.
Sora (OpenAI)	Vídeo hiperrealista	Muy alto	No recomendada para ONG salvo análisis técnico.





Definir la estrategia de uso de IA

Análisis y evaluación de contenidos

Herramienta	Funcionalidad	Uso en ONG
Claude	Análisis profundo, lectura de PDFs extensos	Ideal para evaluar informes y campañas.
ChatGPT	Análisis narrativo + propuestas	Bueno para revisión de estilo.
HuggingFace Spaces	Clasificación, sentimiento, toxicidad	Muy útil para análisis rápido de contenido.
Surge AI	Etiquetado y clasificación de texto	Para ONGs con mucho contenido (blogs, informes).





Ejercicio

“Diseña la estrategia inicial de IA de tu organización”

Paso 1 – Definir objetivos (5 min)

Paso 2 – Definir límites y exclusiones (5 min)

Paso 3 – Definir criterios éticos y narrativos (3 min)

Paso 4 - Definir cómo se va a usar: roles y políticas (5 min)

Paso 5 – Seleccionar 1-2 herramientas para empezar y crear cuentas (10 min)

Paso 6 – Redactar 2 o 3 “primeros casos de uso” (5 min)



Gracias