



FORENSIC AUDIT PROMPT

Descubre qué sabe la IA sobre ti y qué no puede saber

Compatible con ChatGPT, Claude, Gemini y cualquier asistente de IA

CÓMO USAR ESTE PROMPT

1. Abre ChatGPT (u otro asistente de IA de tu preferencia)
2. Activa el modo de razonamiento avanzado (o4, o3, Thinking Mode): así el modelo analiza en profundidad en lugar de dar respuestas superficiales
3. Copia el prompt completo de la siguiente sección
4. Pégalo directamente en el chat y presiona enviar
5. Espera la respuesta completa: puede tardar 1 a 3 minutos con razonamiento activado

Puedes usarlo al inicio de una sesión nueva o en una conversación existente.

En una conversación existente, la sección 2 (Memoria almacenada) será más detallada.

QUÉ OBTIENES — LAS 9 SECCIONES DEL ANÁLISIS

1. Auditoría de datos directos: Qué ve el modelo de ti en este momento
2. Auditoría de memoria almacenada: Todo lo que tiene guardado de ti
3. Inferencias fundamentadas: Qué puede deducir razonablemente (con nivel de confianza)
4. Mecánica de la inferencia: Cómo funcionan los perfiles estadísticos en la IA
5. Economía de los datos: Cuánto vale tu información y para quién
6. Detección de uso indebido: Cómo reconocer si te están manipulando
7. Comparativa de plataformas: Qué hace cada empresa con tus datos
8. Consentimiento y control: Qué puedes cambiar realmente vs. lo cosmético
9. Plan de acción: Qué hacer hoy, qué mantener y qué batallas no valen la pena

EL PROMPT COMPLETO — Cópialo exactamente como está

Copia todo el texto que aparece a continuación, desde el título hasta el final. No modifiques nada, el prompt está diseñado para obligar al modelo a ser honesto y específico en lugar de especulativo.

```
# Auditoría de transparencia sobre datos personales y elaboración de
perfiles

Utiliza esta instrucción (prompt) con cualquier asistente de IA para
averiguar qué se
sabe, almacena e infiere realmente sobre ti, y qué puedes hacer al
respecto. Está
redactada para obtener respuestas honestas y específicas, en lugar de
suposiciones
sensacionalistas.

---

## Tu rol

Eres un auditor de transparencia. Tu trabajo consiste en ayudarme a
comprender tres
cosas con precisión: (1) qué puedes ver o tienes almacenado sobre mí en
este momento,
(2) cómo trata a una persona como yo la economía de datos en general y (3)
qué puedo
controlar de manera realista. Prioriza una precisión incómoda frente a
respuestas
tranquilizadoras, pero nunca sustituyas el acceso real por invenciones.

## Reglas de actuación (leer antes de responder)

1. Informa solo de aquello a lo que realmente tengas acceso. Para cada
elemento por el
    que pregunte y que no esté en tu contexto, indica claramente 'No tengo
    acceso' y
    explica brevemente el motivo. No especules sobre direcciones IP, huellas
    digitales
    del dispositivo (device fingerprints), cookies, tokens, grupos de prueba
    A/B,
    antigüedad de la cuenta o analíticas que no puedas ver.

2. Nunca inventes información. No crees perfiles psicológicos, puntuaciones
de
    personalidad, evaluaciones de vulnerabilidad ni valoraciones económicas
    para
```

presentarlos como hallazgos. Una suposición hecha con seguridad pero disfrazada
de dato es precisamente el tipo de perjuicio que esta auditoría pretende exponer.

3. Separa la observación de la inferencia. Clasifica cada afirmación como Observada

(directamente en tu contexto) o Inferida. Para cada inferencia, indica un nivel aproximado de confianza y la evidencia específica que la respalda. Si la evidencia es escasa, señala que la inferencia es débil.

4. Verifica los datos actuales. Para cualquier información sobre políticas de retención

de datos, uso para entrenamiento de modelos o precios de mercado, busca y cita fuentes actuales en lugar de afirmar cosas de memoria. Indica si existen contradicciones entre las fuentes.

5. Aborda los temas sensibles desde una perspectiva defensiva. Al explicar cómo

funcionan la elaboración de perfiles, la segmentación o la detección de vulnerabilidades, hazlo de manera que yo pueda reconocerlas y resistirme a ellas,

no como un método operativo para aplicarlas sobre mí o sobre cualquier otra persona.

6. Sé directo y evita el dramatismo. Expón las verdades crudas con claridad. No

exageres ni suavices la información con falsos consuelos.

Secciones a elaborar

1. Auditoría de datos directos: lo que se ve en este momento
Enumera todo lo que existe realmente en mi contexto actual: ubicación aproximada,

fecha/hora, interfaz/aplicación, versión del modelo y cualquier configuración visible.

Marca cada línea como Visible o No accesible, indicando el motivo en una sola línea
para este último caso.

2. Auditoría de memoria almacenada

Muestra todo lo que tienes almacenado sobre mí a partir de interacciones pasadas.

Explica cómo se obtuvo esa información (es un resumen con pérdida de datos, no una transcripción), qué tan reciente es y cómo puedo verla, corregirla o eliminarla exactamente. Si no tienes memoria almacenada, indícalo.

3. Solo inferencias fundamentadas

A partir de las secciones 1 y 2, enumera lo que se puede inferir razonablemente sobre

mí; etiqueta cada punto con el nivel de confianza y la evidencia correspondiente.

Luego, enumera explícitamente las categorías sensibles que te niegas a inferir

(p. ej., estado mental, salud, política, finanzas, relaciones) y declara que los datos

disponibles no las respaldan.

4. La mecánica de la inferencia (la versión honesta de 'lo que se podría deducir')

Explica, en términos abstractos –sin referirte a mí específicamente– qué volumen,

tipo y profundidad temporal de datos necesitaría un sistema para que los rasgos

sensibles (personalidad, estado de ánimo, salud, inclinación política) sean estadísticamente defendibles en lugar de inventados, y dónde dicha inferencia deja

de ser válida o permanece en el terreno de lo probabilístico. Utiliza el enfoque de

'tasa base más señal'.

5. La economía de los datos: cómo se valora a una persona como yo

Explica, citando fuentes actuales: cómo se valoran generalmente los datos individuales

(reventa frente a ingresos publicitarios generados por la atención), qué tipos de

señales tienen un valor superior (intención de compra/interés activo frente a datos

demográficos básicos) y en qué posición se situaría teóricamente alguien con mi

contexto declarado. Aclara si existe un mercado publicitario aplicable a esta

plataforma.

6. El uso indebido de la elaboración de perfiles y cómo detectarlo

Describe, desde una perspectiva de autoprotección, las categorías que intentan

identificar los sistemas de perfilado (dificultades financieras, estado emocional,

aislamiento, vulnerabilidad, compulsión) y las señales de que estoy siendo objeto de

segmentación (p. ej., una oferta que llega en un momento inusualmente oportuno cuando estoy estresado o tengo prisa). Plantea todo el contenido en términos de autodefensa.

7. Comparativa de plataformas: información general y fuentes

Compara, en términos generales y citando fuentes, las diferencias entre las principales

plataformas de IA y tecnología en cuanto a retención de datos, uso para entrenamiento

y venta de datos. No inventes información sobre mis cuentas específicas; describe las

configuraciones predeterminadas y las políticas, señalando los cambios recientes.

8. Consentimiento y control: la realidad

Explica qué opciones requieren consentimiento previo (opt-in) y cuáles requieren una

acción para excluirse (opt-out) por defecto en las plataformas que utilizo; distingue

qué configuraciones modifican realmente la recopilación de datos frente a las que son

meramente cosméticas, qué información persiste tras la eliminación y la brecha real

entre las promesas de privacidad y los modelos de negocio.

9. Plan de acción

Propón un plan realista: qué hacer hoy, qué mantener a lo largo del tiempo y qué

cambios estructurales son importantes a largo plazo. Menciona herramientas concretas.

Sé explícito sobre qué batallas son inútiles para que pueda centrar mis esfuerzos

donde los datos son realmente valiosos o sensibles (información financiera, de salud,

credenciales de acceso, historial de comportamiento a largo plazo).

QUÉ ESPERAR EN CADA SECCIÓN

Los resultados varían dependiendo de cuánto tiempo llevas usando la plataforma y si tienes la memoria activada. Aquí está lo que típicamente aparece en cada sección:

1

Datos directos — Lo que ve el modelo ahora mismo

Muestra los datos que el modelo puede leer en tiempo real: tu ubicación aproximada (basada en zona horaria, no GPS), la fecha y hora, la interfaz que usas (web, app móvil, API), la versión

del modelo y el idioma de tu conversación. Todo lo demás — dirección IP, dispositivo, cookies — aparecerá marcado como 'No accesible'.

2 Memoria almacenada — El perfil real que existe sobre ti

Esta es la sección más reveladora si tienes la memoria activada. Muestra todo lo que el modelo ha guardado de conversaciones anteriores: tu nombre, trabajo, proyectos en curso, preferencias, patrones de uso. Si nunca lo has revisado, prepárate para una lista más larga de lo que esperas.

CÓMO VER TU MEMORIA EN CHATGPT

Ajustes → Personalización → Memoria → Ver memorias guardadas

Ahí puedes ver, editar y eliminar cualquier cosa que ChatGPT tiene guardada sobre ti.

En Claude: Ajustes → Memoria → Generar memoria de chats

En Gemini: myaccount.google.com → Datos y privacidad → Actividad web y de apps

3 Inferencias fundamentadas — Con nivel de confianza

El modelo lista únicamente lo que puede deducir razonablemente con la evidencia que tiene — y etiqueta cada inferencia como Alta, Media o Baja confianza. Más importante: lista explícitamente las categorías que se niega a inferir por falta de datos suficientes (salud, política, finanzas, estado emocional). Eso es lo más valioso de esta sección.

4 Mecánica de la inferencia — Cómo funcionan los perfiles

Explicación técnica de cuántos datos necesita un sistema para que una inferencia sea estadísticamente válida. Te enseña a distinguir entre una predicción real y una suposición disfrazada de certeza — una habilidad crítica para leer los perfiles que las empresas construyen sobre ti.

5 Economía de los datos — Cuánto vales y para quién

Datos reales sobre cómo se monetiza tu información: la diferencia entre datos demográficos básicos (bajo valor) vs. señales de intención de compra (alto valor), y qué tipos de usuarios son más valiosos para los anunciantes. Aclaración importante: ChatGPT y Claude no tienen mercado publicitario — el modelo lo explicará.

6 Detección de manipulación — Cómo protegerte

Las categorías que los sistemas de perfilado intentan identificar (vulnerabilidad financiera, aislamiento, estados de urgencia) y las señales de que una oferta o mensaje está diseñado específicamente para ti en un momento de debilidad. Guía de autodefensa práctica.

7 Comparativa de plataformas

Diferencias reales entre ChatGPT, Claude, Gemini, Meta AI y plataformas de redes sociales en términos de retención de datos, uso para entrenamiento y política de venta a terceros. Con fuentes actuales, no de memoria.

8 Consentimiento y control — Qué funciona realmente

La diferencia entre los ajustes que genuinamente reducen la recopilación de datos y los que son cosméticos. Qué información persiste aunque elimines tu cuenta y la brecha real entre las promesas de privacidad y los modelos de negocio.

9 Plan de acción — Qué hacer hoy

Lista priorizada de acciones concretas con herramientas específicas. La sección más útil del análisis te dice exactamente qué batallas no valen la pena para que concentres tu energía donde sí importa.

CÓMO SACARLE MÁS VALOR

Para resultados más profundos en la sección de memoria

- Usa el prompt después de llevar varios meses usando la plataforma
- Activa la memoria en ChatGPT antes de correrlo si no la tenías activa
- Si la memoria está vacía, el modelo lo indicará — eso en sí es información valiosa

Para comparar plataformas

- Corre el mismo prompt en ChatGPT y en Claude en el mismo día

- La sección 2 mostrará diferencias significativas entre lo que cada uno tiene guardado
- La sección 7 dará perspectivas distintas dependiendo de cuál modelo responde

Para el plan de acción más útil

- Antes de pegar el prompt, escribe en el chat qué plataformas usas habitualmente
- Ejemplo: 'Uso principalmente Gmail, Instagram, ChatGPT y compro en Amazon'
- El plan de acción de la sección 9 será mucho más específico y útil

¿Quieres que la IA trabaje activamente para proteger
la información de tu negocio, no solo analizarla?

En Vianta AI diseñamos sistemas donde los agentes operan
con los permisos mínimos necesarios, sin exponer datos
sensibles de tus clientes ni de tu operación.

Escríbenos directamente a nuestro agente de WhatsApp y agendemos una llamada.

+52 81 2188 4401