Elevam

GEO

El SEO para la era de la IA



Cómo aparecer en ChatGPT cuando tus clientes buscan lo que vendes

GEO

El SEO para la era de la IA

Cómo aparecer en ChatGPT cuando tus clientes buscan lo que vendes

Asier López Ruiz

Elevam

Edición digital: octubre de 2026

© 2026 Asier López Ruiz

Todos los derechos reservados.

Queda rigurosamente prohibida, sin la autorización escrita del titular del copyright, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, así como la distribución de ejemplares mediante alquiler o préstamo públicos.

Edita: Elevam (Fullseo S.L.U.)

Contacto con el autor: elevam.es

Índice

Agradecimientos 6

Introducción: por qué este libro y por qué ahora 8

PARTE I EL CAMBIO DE PARADIGMA

1 Desprogramando al SEO 15

2 Cómo «piensa» una IA cuando responde 26

3 Lo que GEO NO es 40

PARTE II LA INGENIERÍA DETRÁS DE LA VISIBILIDAD

4 RAG 101 54

5 Del keyword al vector 66

6 Context windows y tipos de consulta 77

PARTE III ARQUITECTURA WEB PARA LA ERA IA

7 RAG avanzado 91

8 Tu «Source of Truth» 105

PARTE IV MEDICIÓN: LO QUE NO MIDES, NO EXISTE

9 Las cinco métricas que importan 118

10 Test Suite 131

11 Auditoría GEO completa 139

PARTE V PRIORIZACIÓN Y EJECUCIÓN

12 Impacto vs. Ruido 153

13 Escribir para la IA sin dejar de escribir para personas 165

14 Cada token cuenta 175
15 Schema avanzado para GEO 184

PARTE VI ENTIDAD, REPUTACIÓN Y AUTORIDAD
16 Consolida tu entidad 195
17 Off-Site 206
18 Reputación y sentimiento 215

PARTE VII ESTRATEGIAS POR PLATAFORMA Y VERTICAL
19 Domina los motores: Google, Perplexity, ChatGPT y Gemini 223
20 GEO por verticales 232

PARTE VIII EJECUCIÓN, AUTOMATIZACIÓN Y NEGOCIO
21 Roadmap operativo 242
22 Automatización y Tooling 251
23 Packaging y venta de servicios GEO+SEO 262

PARTE IX SEGURIDAD, ÉTICA Y FRONTERAS
24 Seguridad, ética y lo prohibido 273
25 Labs GEO 281
Apéndice diferencial: Anatomía de una auditoría GEO real 291
Epílogo: Tu siguiente movimiento 305
Glosario GEO de Elevam 308
Bibliografía y referencias 326
Índice analítico 330
Sobre el autor 333
Tu siguiente movimiento 335

Agradecimientos

Escribir sobre algo que todavía se está construyendo tiene una dificultad particular: no puedes apoyarte en lo que ya existe, porque casi no existe nada. Este libro nació en ese espacio incómodo, y no habría tomado la forma que tiene sin las personas que aparecen aquí.

La primera es mi padre. Él ha sido la influencia más directa en cómo Elevam Labs entiende la investigación: la importancia de tener una metodología sólida antes de sacar conclusiones, de construir un baseline antes de medir, de no confundir ruido con señal. En un campo donde todo el mundo opina y casi nadie tiene datos, ese criterio es la diferencia entre hacer algo que dura y hacer algo que suena bien dos semanas. Gran parte de lo que encontrarás en este libro tiene su huella, aunque él nunca lo firmaría.

A Elena, Adriana y Juan David les debo la prueba de que esto no es solo una idea mía. Han estado dentro del proceso, cuestionando cosas que yo daba por sentadas y poniendo trabajo real en páginas que el lector verá terminadas sin saber el estado en que estaban antes.

Lidia Moreno llegó cuando Elevam era otra cosa y se quedó mientras se convertía en lo que es ahora. Ha sido

la persona que ha traducido criterio en proceso, intuición en metodología, caos en sistemas que funcionan. Mucho de lo que hoy llamamos «forma de trabajar de Elevam» tiene su origen en conversaciones con ella.

Y luego están Anna, mi madre y mi hermana. Construir algo desde cero tiene un coste que no siempre se ve desde fuera. Ellas tres han sido el lugar al que volver cuando todo pesaba demasiado, cada una a su manera, sin pedirles que estuvieran. La calma que generan no aparece en ningún capítulo, pero este libro no existiría sin el espacio que crean.

A todos, gracias.

Introducción: por qué este libro y por qué ahora

El dato que lo cambia todo

El 88% de las fuentes citadas por asistentes de IA no aparece en el Top 10 de Google para ese mismo prompt.

Lee eso otra vez. Ochenta y ocho por ciento.

Llevas años peleando por posiciones en el buscador, invirtiendo en SEO, en contenido y en enlaces. Ahora resulta que la IA responde a tus potenciales clientes y ni siquiera mira esa carrera del mismo modo. Elige sus fuentes con criterios diferentes: relevancia semántica, densidad informativa y autoridad percibida en contexto. No le importa si estabas primero en Google. Le importa si tu contenido responde mejor a la pregunta.

Mientras tanto, el 60% de las búsquedas terminan sin que nadie haga clic en nada. El usuario obtiene la respuesta directamente de Google, de ChatGPT o de Perplexity. Tu web puede estar generando valor que Analytics nunca captura, o peor, puede estar siendo completamente ignorada mientras la IA recomienda a otro.

Si estos dos datos no te hacen replantear todo lo que creías saber sobre visibilidad digital, no sé qué lo hará.

La cuestión es el negocio

Tienes un producto brutal. Un servicio que resuelve problemas reales. Clientes satisfechos que repiten. Pero en internet eres invisible, o peor, visible en Google pero inexistente para la IA. Tu competidor se lleva a tus clientes porque aparece en las respuestas que la gente consulta antes de tomar decisiones.

Quizá llevas meses quemando presupuesto en anuncios que dejan de funcionar en cuanto apagas el grifo, o invirtiendo en SEO sin ver resultados proporcionales. Notas que algo ha cambiado pero no sabes exactamente qué. Los clientes ya no llegan por los mismos canales de antes. Las métricas que siempre miraste ya no cuentan la historia completa.

Spoiler: los clientes ya no buscan como antes y nadie te ha explicado qué hacer al respecto.

Hasta ahora.

Qué es GEO y por qué es la respuesta

El juego ha cambiado porque la gente ya no busca igual. Ahora le preguntan a ChatGPT, a Perplexity, a Google con IA. Y esos sistemas no muestran diez enlaces azules en la SERP. Generan una respuesta y mencionan dos o tres fuentes. Si no eres una de esas fuentes, no existes.

GEO (Generative Engine Optimization) es la disciplina que te convierte en esa fuente. Es la forma de construir tu presencia digital para que los modelos de lenguaje te

elijan, te citen y te recomienden cuando alguien hace la pregunta que importa para tu negocio.

Es un activo digital que trabaja 24/7 para ti. Es la diferencia entre perseguir clientes a puerta fría o que lleguen a tu puerta porque la IA te recomienda como la mejor opción.

Y no, GEO no mata al SEO. Lo amplía. Todo lo que has construido en SEO sigue sirviendo pero ya no es suficiente. Ahora necesitas una capa adicional: entender cómo piensa la IA, qué fuentes elige, por qué te ignora y cómo cambiar eso.

¿Quién soy y por qué deberías escucharme?

Me llamo Asier. Dirijo Elevam, una firma de crecimiento especializada en optimización y posicionamiento para motores de búsqueda (SEO) y para IAs generativas (GEO), junto a otras soluciones orientadas a hacer crecer negocios.

No voy a contarte que llevo X años en marketing digital. Eso lo dice cualquiera. Prefiero contarte algo más honesto: vengo de una etapa muy dura. Acepté proyectos que no debería haber aceptado, pasé por un duro bloqueo cognitivo, sufrí deslealtades empresariales y llegué a un punto en el que casi tiro la toalla. Depresión real, no de LinkedIn. De esas que no se publican.

Salí de ahí reconstruyendo desde cero, con una idea clara: no más parches rápidos, no más clientes que no encajan, no más servicios que no puedo respaldar con re-

sultados. Elevam nació de esa decisión. Una firma enfocada en empresas que quieren construir algo sólido a largo plazo.

Lo que vas a leer no es teoría. Es la misma metodología que aplico con mis clientes cada semana. Clientes que facturan miles de millones al año. He visto empresas pasar de ser completamente invisibles para la IA a aparecer como primera recomendación en su categoría, en plazos de meses, no años. Y he visto lo contrario: empresas que dominaban Google y que la IA ignora porque nadie les explicó que el juego cambió.

Soy joven. Pero me he ganado las cicatrices. Y este libro es lo que me hubiera gustado tener cuando empecé a entender que el SEO que conocíamos estaba a punto de volverse insuficiente.

Lo que sabrás hacer cuando termines este libro

No voy a prometerte que «aprenderás GEO». Eso es vago y no compromete a nadie. Voy a decirte exactamente qué serás capaz de hacer:

Sabrás auditar si la IA te está ignorando o te está citando, con datos concretos y metodología reproducible. Sabrás reestructurar tu web para que los modelos de lenguaje te elijan como fuente en vez de a tu competencia. Sabrás medir mes a mes si estás ganando o perdiendo terreno en las respuestas de ChatGPT, Perplexity y Google AI Overviews. Esa medición es la que automatizamos en Antropus (antropus.io), la plataforma de medi-

ción SEO + GEO que hemos creado en Elevam; aun así, aquí tendrás el método completo para aplicarlo con las herramientas que prefieras.

Sabrás construir el sistema completo: medición, diagnóstico, ejecución, monitorización. No un check puntual, sino un proceso continuo que se integra en la operación de tu negocio o de tu agencia.

Si diriges un equipo o una consultoría, sabrás empaquetar y vender este servicio a clientes que todavía no saben que lo necesitan, con propuesta, contrato, pricing y reporting profesional.

Todo lo que necesitas está en los capítulos que siguen. Sin atajos, sin secretos guardados.

Para quién es (y para quién NO es) este libro

Esto no es un curso de marketing digital para principiantes. No se explica qué es un CMS, ni cómo funciona Google Ads, ni qué significa SEO. Si necesitas esa base, este no es tu libro todavía.

Es para dueños de negocio con presencia digital que sienten que algo se les está escapando. Para SEOs y marketers que quieren liderar la transición en vez de sufrirla. Para consultores y agencias que quieren ofrecer un servicio que sus competidores aún no entienden.

No es para quien busca «trucos de SEO rápidos». No es para quien quiere manipular rankings con atajos. No es para quien espera que un plugin le resuelva la vida.

Es para quien entiende que el cambio de paradigma ya llegó, que la ventana de oportunidad no va a durar para siempre, y que prefiere estar del lado de los que entienden el juego antes de que el juego les pase por encima.

Si te reconoces en eso, sigue leyendo. Este libro está escrito para ti.

La encrucijada

En los próximos capítulos voy a destripar la misma metodología técnica que utilizo con mis clientes. Vas a tener los planos completos: desde la psicología de un modelo de lenguaje hasta la creación de un pipeline automatizado de medición, arquitectura web, schema avanzado, gestión de entidades y reputación.

Llegados a este punto, tienes dos formas de aprovechar este libro. La primera es usarlo como tu hoja de ruta interna: si tienes equipo y capacidad técnica, puedes construir tu sistema GEO desde cero siguiendo estos pasos. La segunda es usarlo para entender exactamente qué necesitas y, si prefieres no invertir tu tiempo en la implementación técnica, delegarlo en nosotros: en Elevam. Decidas lo que decidas, sigue leyendo. Te prometo que lo que viene a continuación cambiará tu forma de entender este ecosistema.

Ahora sí. Capítulo 1.

PARTE I

El cambio de paradigma

Del SEO que conocías al mundo donde la IA decide quién existe.

CAPÍTULO 1

Desprogramando al SEO

Bienvenido al juego que ya no controlas

Del índice al vector: por qué «posición» ya no significa «recomendación»

Durante casi tres décadas, visibilidad en internet significó una sola cosa: estar lo más arriba posible en la lista de resultados de Google. El buscador indexaba páginas por palabras clave y enlaces, calculaba relevancia y autoridad con su algoritmo y te devolvía una lista ordenada de diez enlaces (resultados). Ser el primero era el objetivo supremo. Toda una industria se construyó alrededor de esa premisa.

Ese juego ha terminado. No se ha matizado ni ha evolucionado ligeramente. Ha cambiado de raíz. Sé que con esto me ganaré enemigos, pero no escribo para contentarte. Escribo para transmitirte todo lo que sé, sin filtros.

Un modelo de lenguaje no te enseña diez enlaces. Te entrega una respuesta sintetizada y, con suerte, cita dos o tres fuentes como apoyo. La visibilidad ya no consiste en ocupar el primer puesto en Google. El nuevo reto es aparecer dentro de la respuesta de la IA. Hemos pasado

de un entorno gobernado por índices de búsqueda a otro regido por vectores semánticos y recomendaciones probabilísticas. Tu contenido puede ser extraordinariamente relevante aunque nunca haya entrado en el Top 10 de Google. Lo decisivo es que el modelo te seleccione para construir su respuesta.

La distancia entre ambos criterios no es teórica. Es medible. Ahrefs analizó las URLs que citaban varios asistentes de IA en sus respuestas y encontró que solo el 12% de ellas aparecían también en el Top 10 de Google para el mismo prompt. Dicho de otro modo: el 88% de las fuentes citadas por estos sistemas quedaban fuera del Top 10 para esa misma consulta. Ranking y recomendación han dejado de ir de la mano. Puedes dominar Google y que ningún modelo de lenguaje te mencione. O al revés: un sitio modesto y útil puede colarse de forma consistente en las respuestas de la IA.

Fuente: https://ahrefs.com/blog/ai-search-overlap/



Solo 12 de cada 100 fuentes citadas por la IA están en el Top 10 de Google. Fuente: Ahrefs, «AI search overlap» (ahrefs.com/blog/ai-search-overlap).

Si solo miras rankings orgánicos, estás ciego ante un canal que millones de usuarios ya usan a diario para tomar decisiones.

El ejemplo que no mola (o sí)

Imagina que tu página es la número uno en Google para «mejores herramientas de gestión de proyectos». Durante años, cualquiera que buscara eso te veía primero. Ahora un usuario le pregunta lo mismo a ChatGPT: «¿Qué herramientas de gestión de proyectos me recomiendas para este año?»

El asistente no devuelve diez enlaces. Redacta una respuesta mencionando tres o cuatro herramientas puntuales, con una breve descripción de cada una. Debajo, muestra las referencias de dónde sacó la información: «Fuente: TechnologyAdvice +3». Combinó datos de TechnologyAdvice y de otras tres fuentes para construir esa respuesta.

¿Qué quiero decir con esto? Aunque tu página era la número uno en Google, podría no aparecer citada en ningún sitio. Si el modelo decidió que tu contenido no era necesario, tal vez encontró otra lista más completa, más reciente, mejor estructurada y te dejó fuera. En cambio, otra página, quizás de un blog especialista que ni siquiera rankeaba en el Top 10 de Google, aparece citada con honores en la respuesta.

El modelo ha decidido qué contenidos incluir. Y lo ha hecho de forma completamente independiente al ranking tradicional.

Ser el número uno en Google ya no garantiza nada en el terreno de la IA.

Cuatro errores de mentalidad que siguen costando dinero

Hay cuatro creencias que veo repetirse en prácticamente todos los equipos de marketing. Son creencias que tenían sentido hace tres años. Hoy son trampas.

La primera ya la hemos introducido: creer que ser número uno en posicionamiento orgánico garantiza presencia en IA. No la garantiza. Muchos SEOs siguen obsesionados con «ser primeros en Google» mientras el usuario, en muchos casos, ya ni siquiera ve una SERP tradicional. Alcanzar esa posición no asegura que ChatGPT, Gemini, Claude o Perplexity te citen. Son sistemas distintos, con lógicas y criterios distintos.

La segunda es la obsesión con la keyword exacta. Optimizar solo para la coincidencia textual de la palabra clave, ignorando el contexto semántico más amplio, ya no es suficiente. Un LLM entiende intención, temas relacionados y sinónimos. Si tu contenido es demasiado estrecho, podría preferir otra fuente más completa sobre el tema. El modelo no busca coincidencias de texto; busca la mejor respuesta posible.

La tercera es confiar en que la autoridad histórica de tu marca te protege. «Somos una marca conocida, la IA seguro nos usará.» Los modelos eligen contenido que responda mejor. Si otra página explica lo mismo que tú pero de forma más concisa o mejor estructurada, el LLM la puede preferir aunque tenga menor PageRank o noto-

riedad. La IA no siente reverencia por tu marca. Siente predilección por el contenido útil.

Y la cuarta, y la más peligrosa de todas: no vigilar si apareces en IA. Si ni siquiera compruebas si tus páginas salen citadas en respuestas generativas, puede que estés perdiendo terreno frente a competidores que ni siquiera son mejores que tú. Ellos o tus competidores han entendido cómo funciona este nuevo juego. Volar a ciegas es un error estratégico que se paga caro, porque cada día que pasa sin que actúes es un día que tus competidores pueden estar ocupando el espacio que te corresponde.

Zero-click: adiós al CTR y a la atribución que conocías

La era del zero-click, búsquedas que terminan sin que nadie haga clic en ningún resultado, ya estaba aquí antes de los LLMs. Los featured snippets de Google ya anticipaban el fenómeno: el usuario obtenía su respuesta directamente en la página de resultados y no necesitaba visitar ninguna web. Con los modelos de lenguaje, este comportamiento ha explotado.

Los números no dejan mucho margen para la interpretación. Según Bain & Company, a partir de una encuesta Bain–Dynata realizada en diciembre de 2024 con una muestra de 1.117 consumidores, alrededor del 60% de las búsquedas ya terminan sin que el usuario visite otro sitio. Además, el 80% de los consumidores afirma apoyarse en resultados zero-click o resultados escritos por IA en al menos el 40% de sus búsquedas, lo que está redu-

ciendo el tráfico orgánico entre un 15% y un 25%. Son cifras enormes que confirman algo fundamental: menos clics no significa menos gente obteniendo respuestas. Significa que muchas de esas respuestas ya se resuelven directamente en la interfaz del buscador o del chatbot, sin necesidad de visitar tu web.

Fuente: https://www.bain.com/insights/goodbye-clicks-hello-ai-zero-click-search-redefines-marketing/



La búsqueda sin clic, en cifras. Fuente: Bain & Company, encuesta Bain–Dynata (diciembre de 2024, n = 1.117), «Goodbye Clicks, Hello AI».

Funciona así: el usuario hace una pregunta, la IA le responde usando tu contenido, el usuario toma una decisión basándose en esa respuesta y, como resultado, nunca llega a visitar tu web.

A este fenómeno en Elevam lo llamamos generar decisiones sin clic. El usuario obtiene la respuesta, actúa en consecuencia y tu contenido generó valor. En tus informes de Analytics esa conversión aparecerá como «tráfico directo» o como un origen imposible de rastrear. Si

sigues mirando solo Google Analytics, no te enteras de que tu contenido está trabajando.

El diagnóstico equivocado que mata presupuestos

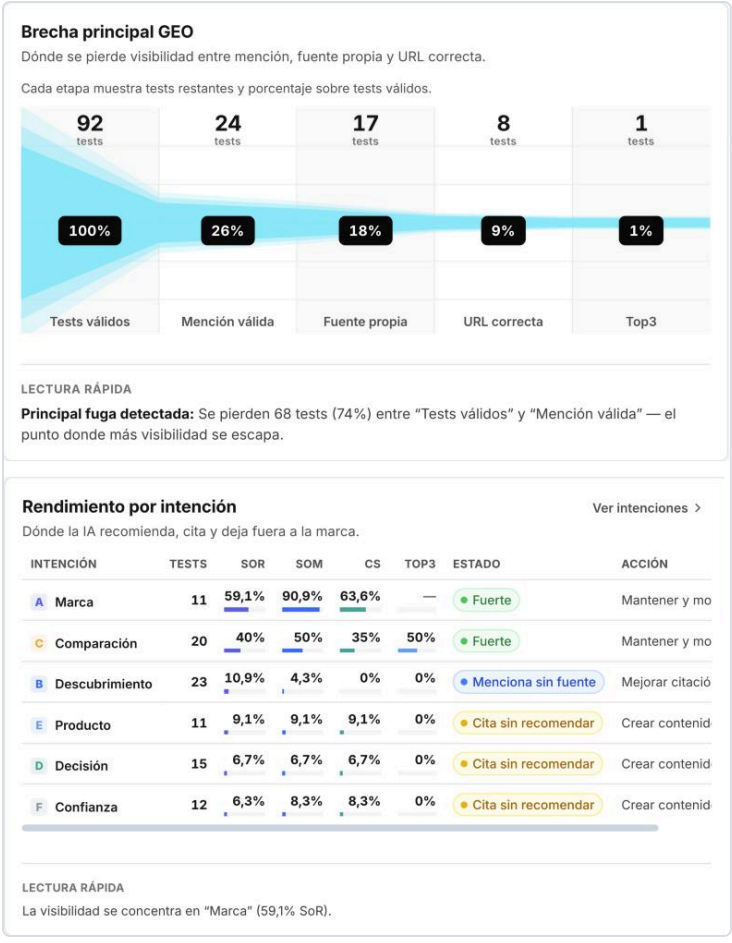
Aquí es donde la cosa se pone realmente peligrosa para los que toman decisiones de negocio.

Es una escena que se repite en docenas de despachos profesionales: el socio director ve que el tráfico orgánico ha bajado dos dígitos en los últimos meses, concluye que «los posts ya no funcionan» y plantea recortar la inversión en contenido.

Sonaba razonable. Y a la vez, era un tiro en el pie.

Lo que esa lectura ignora es que, en paralelo a la bajada de tráfico medible, los análisis jurídicos del despacho podían estar siendo utilizados por la IA para contestar consultas de decenas de empresas.

Es el momento de empezar a hablar de nuevas métricas. El Share of Model mide qué porcentaje de la visibilidad generada por los modelos en tu categoría corresponde a tu marca o a tu contenido. Esa métrica puede ser más útil que el CTR puro para evaluar tu visibilidad real en la era de la IA. En Antropus lo calculamos a partir del Prompt Tracker: el mismo set de prompts lanzado a ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview, midiendo si te mencionan, te recomiendan y te citan frente a tus competidores.



Antropus · Prompt Tracker: dónde se pierde la visibilidad y rendimiento por intención. Proyecto real de Elevam (elevam.es), baseline del 14 de agosto de 2026 en ChatGPT, Claude, Gemini y Perplexity.

La atribución también se complica en formas que no anticipábamos. Hoy en día, un usuario puede descubrirte a través de una respuesta generativa y finalmente sin querer, acaba entrando directamente a tu sitio web, yen-

do a tu tienda física o llamándote por teléfono. Atribuir esa conversión al SEO es prácticamente imposible con las herramientas tradicionales, porque no hubo clic orgánico de por medio. Pero en espíritu, esa conversión existe gracias a tu contenido insertado en la respuesta de la IA.

Si ignoramos esta realidad, caemos en la miopía de marketing más cara de la década: apagar esfuerzos que sí están generando demanda.

En la economía zero-click, la influencia precede a la visita. Primero influyes en la respuesta de la IA, tal vez la visita ocurra después o tal vez ni haga falta la visita porque el usuario tomó acción directa. Tu contenido puede estar trabajando las 24 horas del día aunque tus gráficos de Analytics no lo muestren.

El caso que nadie mide (y debería)

Ponte en situación. Un potencial cliente le pregunta a su asistente de voz: «¿Cuál es la mejor aseguradora de coche actualmente?» El asistente, apoyado por un LLM, genera una respuesta: «Según lo que he encontrado, las aseguradoras mejor valoradas hoy son A, B y C. La más recomendada para calidad-precio es X.» X es tu marca.

El usuario escucha esto, confía en la IA y decide contratar contigo. No hizo ningún clic en tu web. Quizá luego busque tu marca directamente o te llame por teléfono. La conversión ocurre sin pasar por una SERP tradicional. Tu contenido influye en esa respuesta y te gana un

cliente. En tus informes de Analytics esa venta aparecerá como «tráfico directo» o algo peor: sin atribución alguna.

Este es un caso de libro de decisión sin clic. La marca ganó un cliente gracias a su presencia en una respuesta de IA, sin que el usuario visitara la web durante la consulta inicial.

Para captar este tipo de valor, algunas empresas están empezando a mirar métricas alternativas y métodos cualitativos. Por ejemplo, hacer encuestas a nuevos clientes preguntando «¿Cómo nos conociste?» incluyendo opciones como «Asistente de IA» o «Chat online», monitorear menciones de marca en herramientas como Perplexity o los AI Overviews de Google o estimar el Share of Model en un conjunto de preguntas frecuentes de su sector. Esas son formas de sacar a la luz esa influencia que de otro modo permanecería completamente invisible. En Antropus separamos las dos vigilancias: Brand Monitoring sigue las menciones de tu marca en noticias, blogs y foros, y AI Visibility te dice en qué búsquedas te cita el AI Overview de Google, con qué página y frente a qué competidores.

Los errores que más dinero cuestan en la era zero-click

El primero: seguir midiendo éxito exclusivamente con CTR y tráfico orgánico. Si lo haces, vas a infravalorar o directamente ignorar el impacto real de tu SEO. Un contenido puede estar cumpliendo su función perfectamen-

te, siendo visible en las respuestas de IA y posicionando tu marca. Y tú lo ves como un fracaso porque no genera clics. Actualiza tus métricas de rendimiento antes de tomar decisiones basadas en datos incompletos.

El segundo: no educar a quienes toman las decisiones de presupuesto. Si tu jefe o tu cliente sigue pensando exclusivamente en «¿cuántos clics orgánicos trajimos?», necesita entender este cambio. Depende de ti evangelizar. El valor del SEO ahora también es branding y presencia en IA, aunque no siempre genere clic directo a tu web. Si no lo explicas, te seguirán pidiendo lo de antes y juzgando con métricas que ya no cuentan la historia completa.

El tercero: ignorar la atribución asistida. No intentes medir todo con el modelo de last-click de siempre. Empieza a buscar correlaciones y señales indirectas. ¿Aumentó el tráfico directo o las búsquedas de marca tras aparecer en respuestas de IA? ¿Hay preguntas en tus canales de soporte que citan «leí en un chatbot que...»? ¿Aparecen picos de conversión que no puedes atribuir a ningún canal específico? Todo eso son indicios de que tu SEO sí está generando demanda, solo que la conversión ocurre por otro canal.

Y el cuarto, el más crítico de todos: confundir zero-click con zero conversion. El hecho de que un usuario no haga clic no significa que tu contenido no haya influido. Puede haber conversiones sin clic. Puede haber valor sin visita. No ver clics no equivale a que no hubo impacto. Grabarte esto es, probablemente, lo más importante que puedes sacar de este capítulo.

CAPÍTULO 2

Cómo «piensa» una IA cuando responde

Anatomía de la respuesta generativa

Dos cerebros, una respuesta

Para optimizar algo, primero tienes que entender cómo funciona por dentro. Y la respuesta de una IA no es tan simple como parece. Lo que el usuario ve (un párrafo limpio, bien redactado y con un par de fuentes al final) es el resultado de un proceso que tiene dos componentes muy distintos trabajando juntos. Si entiendes estos dos componentes, sabrás exactamente dónde encaja tu contenido en todo esto. Y lo que es más importante: sabrás por dónde puede fallar.

El primer componente es la generación por lenguaje. Un LLM, al recibir una pregunta, empieza a escribir palabra por palabra lo que cree que viene a continuación. Esto se llama predicción de siguiente token. El modelo se apoya únicamente en lo que «aprendió» durante su entrenamiento y a partir de ahí genera lo que estadísticamente parece la mejor opción posible.

Este proceso es probabilístico. El modelo no busca en Google, no consulta ninguna base de datos factual y no verifica nada. Genera lo más plausible según sus parámetros. Y plausible no es lo mismo que cierto.

Por eso, si le preguntas algo que sale del alcance de sus datos internos, puede alucinar: inventarse la respuesta con un aire de convicción total. Le preguntas «¿Quién ganó el partido anoche?» y un ChatGPT estándar, sin acceso a información en tiempo real, puede darte un nombre falso, decirte que no lo sabe o inventarse un marcador que suena perfectamente creíble. No tiene mecanismo para verificar la realidad cuando genera en modo puro.

Aquí es donde entra el segundo componente: la recuperación de hechos.

Cuando la IA sale a buscar: RAG y la capa de retrieval

Las arquitecturas modernas de IA no se fían del todo de la memoria entrenada del modelo. Han añadido un paso previo que cambia las reglas del juego: antes de que el LLM dé la respuesta final, el sistema busca información en fuentes externas relevantes a la pregunta. Puede ser una búsqueda en un índice vectorial, contenido pre-indexado para el asistente o incluso una búsqueda web en vivo.

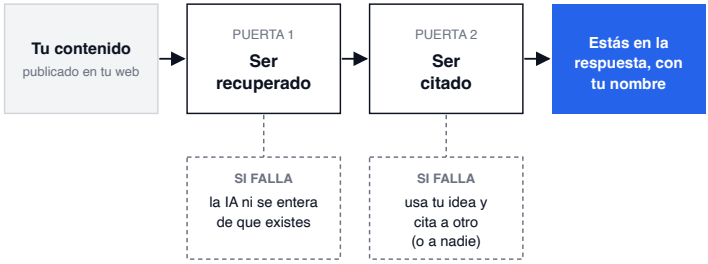
Este enfoque se conoce como RAG: Retrieval-Augmented Generation, o generación aumentada por recupera-

ción. En la práctica, funciona así: el sistema recibe la pregunta, recupera los documentos o fragmentos más relevantes desde un índice, una base de conocimiento o un buscador, y los incorpora al contexto de entrada del LLM para que los utilice al generar la respuesta final.

En herramientas como Microsoft Copilot, Perplexity o Google AI Overviews puedes ver esta anatomía dual con total claridad: te presentan un párrafo de respuesta generado por IA junto con numeritos o etiquetas que citan fuentes. Cada cita corresponde a un fragmento recuperado que el modelo usó para elaborar su respuesta.

Una respuesta de IA «aumentada» combina dos capas distintas: lo que el modelo ya sabe por su entrenamiento y la información adicional que puede recuperar en el momento de responder. Esa información puede venir de la web, de un buscador interno, de una base vectorial, de un índice documental o de documentos añadidos al sistema. Con todo ello, el LLM construye la respuesta final.

¿Por qué te cuento esto con tanto detalle? Porque entender esta anatomía te da algo que la mayoría de profesionales de marketing no tiene: sabes exactamente dónde tienes dos oportunidades distintas de entrar en una respuesta de IA. Puedes ser recuperado y puedes ser citado. Si fallas en cualquiera de las dos, quedas fuera.



Las dos puertas para entrar en una respuesta de IA. Fuente: elaboración propia.

La primera puerta: ser recuperado

Si el modelo no encuentra tu contenido, da igual lo bueno que sea. No lo usará. Así de simple.

Puedes tener la mejor respuesta del mundo en tu web, pero si no logras posicionarla de forma que el sistema de retrieval la considere, el LLM ni se entera de que existes. En la analogía más directa: si no estás en su índice vectorial o no logras salir en su búsqueda web, no existes para la IA.

Esto refuerza algo que a veces se olvida en medio de toda la novedad: el SEO técnico sigue siendo absolutamente fundamental. Indexabilidad, velocidad de carga, contenido optimizado semánticamente, estructura clara. Todo eso no ha dejado de importar. Lo que ha cambiado es que ahora importa para un motivo adicional: no solo para rankear en Google, sino para que el sistema de recuperación de la IA te encuentre.

La segunda puerta: ser citado

Que el modelo te encuentre no significa que te vaya a dar crédito. Esta distinción es crucial y la mayoría de gente no la entiende.

Imagina que la IA recupera un párrafo de tu web con un dato importante, y otro párrafo de la web de tu competidor con otro dato. Si ambos se complementan, la IA podría redactar la respuesta usando tus datos y los del competidor pero citando solo a uno de los dos. O a ninguno.

¿Por qué haría eso? Porque a veces agrupa varias fuentes bajo una sola cita para simplificar. O simplemente, porque tu texto no contenía una frase autocontenida, clara y directamente citable. En ese caso, la IA puede captar la idea general, reformularla con sus propias palabras y no considerar necesario atribuírtela de forma explícita.

Ser citable es un arte. Implica crear contenidos que ofrezcan lo que yo llamo «bits de información»: fragmentos claros, autocontenidos y fácilmente atribuibles. Un dato concreto con su contexto. Una definición precisa. Una conclusión contundente formulada en una sola frase.

Si en tu página esos bits están bien delimitados, es mucho más probable que el modelo diga «Según TuSitio.com, ...» y use tu frase. En cambio, si tus datos están todos mezclados en un texto largo y ambiguo, la IA puede tomar la idea, usarla, y citar a otro que lo expresó de forma más limpia. O no citar a nadie.

Las citas visibles en las respuestas de IA son oro puro. No solo amplifican tu visibilidad sino que te aportan credibilidad. El usuario ve tu nombre asociado a la información que recibió, lo cual genera confianza. En un mundo de respuestas algorítmicas, ser la fuente citada es destacarse como la voz confiable.

Cuando la IA miente sobre ti (o te inventa)

Hay un aspecto de este sistema que produce inseguridad. A veces la IA produce un párrafo que aparentemente es una verdad objetiva pero en realidad contiene errores. Hay casos documentados de LLMs que inventan referencias con nombre de empresa que ni existen.

Si un día ves «Según TuMarca.com...» en una respuesta de ChatGPT, no corras a celebrarlo. Comprueba primero que la referencia es real. Haz clic en el enlace si lo hay, o búscalo. Si descubres que es inventada, toma nota: eso indica que el modelo tenía un hueco de información y se lo imaginó. Tal vez deberías publicar oficialmente esos datos que la IA andaba inventando para que, la próxima vez, tenga una fuente real a la que recurrir.

Monitorizar cómo la IA habla de ti no es opcional. Es parte de la disciplina GEO. En Antropus lo hacemos con el Prompt Tracker: lanza las preguntas sobre tu marca y tu sector a ChatGPT, Claude, Gemini y Perplexity y registra si te mencionan y si te citan, para que revises qué dicen de ti antes de que lo lea tu cliente.

Cómo se ve esto en la práctica

Supongamos que preguntas en Microsoft Copilot: «¿Cuál es el índice de inflación de 2026 y cómo afecta a las hipotecas?» Copilot hace algo inteligente: primero lanza una búsqueda en la web, encuentra la cifra oficial en la web del Banco Central, encuentra alguna explicación en un portal financiero y luego genera la respuesta integrando esos datos. «El índice de inflación de 2026 está en X% según [BancoCentral], lo cual implica que las cuotas hipotecarias subirían aproximadamente Y% [Fuente: DiarioFinanciero].»



Imagen ficticia generada con IA.

Cada dato clave tiene su fuente. Si tu contenido fue el que la IA encontró, tú ganas la cita. Si tu competidor lo publicó mejor, la cita se la lleva él. La pelea ya no es por

quién rankea primero. Es por quién ofrece el dato de forma más clara, más actualizada y más fácil de extraer.

Segundo ejemplo: le pides a ChatGPT que te resuma los anuncios más importantes de la última conferencia de desarrolladores de Apple. ChatGPT no sabe nada de eso de base, pero con su búsqueda web va a internet, busca artículos recientes, saca los puntos clave y redacta un resumen citando las fuentes. Si tu sitio publicó un post «Resumen WWDC 2026: las 5 novedades clave» con bullet points bien estructurados, es muy probable que el LLM tome algunos de esos puntos y te cite como fuente. En cambio, si ChatGPT no tuviera esa capacidad de búsqueda, probablemente soltaría un resumen inventado o hablando de conferencias pasadas.

La moraleja es directa: publicar contenido actualizado, bien estructurado y rico en datos clave te convierte a ti en relevante para la capa de retrieval. Y eso aumenta tus probabilidades de ser citado en la síntesis final.

Los tres tipos de motor que existen hoy

Para entender en qué terreno estás jugando, necesitas saber qué tipo de arquitectura tiene el motor con el que compites. Porque no todos funcionan igual y la forma de optimizar cambia según el tipo.

El primer tipo es el motor determinista. Es el Google de siempre. Opera con un algoritmo predecible: ante una consulta, analiza palabras clave, busca coincidencias en su índice, aplica factores de ranking (autoridad, enlaces,

velocidad) y devuelve una lista ordenada de resultados. Si haces la misma búsqueda dos veces, obtienes prácticamente la misma lista. Buscar «capital de Francia» devuelve «París – Wikipedia» como primer resultado, hoy y mañana. Este determinismo nos dio a los SEOs un terreno firme durante años: podíamos optimizar factores concretos y ver subir nuestra posición de forma más o menos previsible.

El segundo tipo es el motor probabilístico. ChatGPT en modo estándar, sin acceso a buscador, es un ejemplo puro. Recibe la consulta y genera una respuesta con su modelo de lenguaje, tomando decisiones probabilísticas a cada paso. Incluso los LLMs que sí hacen retrieval integran probabilidad en varios momentos del proceso. Un modelo avanzado puede reinterpretar la pregunta en sub-preguntas, lanzar varias búsquedas internas, elegir qué fragmentos de texto usar, y luego redactar la respuesta. Cada sub-paso tiene cierto grado de variabilidad. Por eso, si le haces la misma pregunta dos veces, la respuesta puede no ser idéntica. Ya no existe un «ranking número uno» fijo al que apuntar. El modelo selecciona información dinámicamente y genera texto sobre la marcha.

El tercer tipo, el más relevante hasta el día de hoy, es el motor híbrido. Microsoft Copilot es un modelo de lenguaje que trabaja encima del buscador Bing. Google AI Overviews es un generador soportado por la búsqueda clásica de Google. Ante tu consulta, primero hacen lo de siempre: una búsqueda tradicional determinista para

obtener un conjunto base de páginas relevantes. No abandonan su índice web ni su algoritmo. Toman esos resultados (los mejores 10, 20 o 50) y se los pasan a un modelo generativo, que produce la respuesta final a partir de ellos.

La ventaja del híbrido es que aprovecha lo mejor de cada método. La búsqueda clásica aporta cobertura amplia y capacidad de recuperación. La generativa añade síntesis y lenguaje natural. Pero orquestar este matrimonio no es trivial. El buscador clásico se apoya en términos, índices y señales de recuperación. El modelo generativo trabaja sobre significado, contexto y relaciones semánticas. Son mundos distintos que todavía están aprendiendo a fusionarse.

La tendencia actual es correr dos búsquedas en paralelo: una «sparse» (el Google de siempre basado en coincidencias textuales) y otra «dense» (búsqueda semántica por vectores, que encuentra páginas con contenido afín aunque no contengan exactamente las mismas palabras). Después se fusionan los resultados y de ahí se eligen los fragmentos más relevantes para alimentar al LLM.



Los tres tipos de motor. Fuente: elaboración propia.

Qué significa todo esto para tu estrategia

Si estás ante un sistema híbrido, el SEO técnico y tradicional sigue siendo necesario. La parte determinista del motor sigue ahí. Necesitas que tu página sea rastreable, indexable, que tenga buen contenido optimizado on-page y datos estructurados. Si fallas en lo básico, ni siquiera entras en el pool inicial de candidatos que el sistema va a considerar.

Al mismo tiempo, tienes que pensar en la parte vectorial y semántica. Enriquecer tu contenido tiene que generar embeddings relevantes: contenido profundo y contextual, que toque varios subtemas relacionados. No basta con tener la keyword en el título. Necesitas que tu contenido responda a la intención completa, con matices, datos y perspectivas que un modelo de lenguaje pueda reconocer como valiosas.

Cambia también cómo pruebas y mides resultados. En SEO determinista hacíamos rank tracking: medir tu posición en Google día tras día.

En entornos probabilísticos eso sirve de poco, porque la respuesta puede variar de un día a otro. Aquí necesitas hacer múltiples consultas de prueba y ver en cuántas apareces, promediar menciones y monitorear la volatilidad. Es más complejo que un rank tracker tradicional, pero es el tipo de monitorización que hay que desarrollar. Es exactamente lo que automatizamos en Antropus con el Prompt Tracker: el mismo set de prompts, lanzado a varios motores, con el histórico guardado para comparar mes a mes.

Y hay otro cambio profundo que muchos todavía no han asumido: ya no compites solo página contra página. Compites fragmento contra fragmento. Antes, una URL fuera del Top 10 apenas existía. Ahora, un solo párrafo dentro de una página mediocre puede acabar citado por un modelo si responde con precisión quirúrgica a una pregunta concreta. La optimización deja de ser solo una cuestión de página y pasa a ser también una cuestión de bloques, secciones y fragmentos. Cada parte del contenido debe poder sostenerse por sí sola: tener un encabezado claro, una idea nítida y un valor extraíble sin ambigüedad.

La variabilidad no es un fallo, es una característica

Si haces un experimento y hoy vas a Google AI Overviews y preguntas «¿Cuál es la mejor plataforma de e-commerce para una pyme?», quizá te responda recomendando Shopify, WooCommerce y PrestaShop. Una semana después, haces la misma consulta y esta vez la

respuesta menciona Shopify, Wix y Magento, con fuentes completamente distintas.

No te has vuelto loco ni el sistema está roto. Es variabilidad intrínseca del motor. Quizá en la segunda ocasión el sistema exploró una variación de la consulta, buscando también «plataformas e-commerce comparativa técnica», encontró un foro de programadores con fans de Wix y decidió incorporar esa perspectiva. Ninguna de las dos respuestas está mal. La IA exploró caminos alternativos sobre la misma pregunta.

Para un SEO clásico esto es desconcertante. ¿Cómo que los resultados cambian sin que haya pasado nada? Pero en el mundo de la IA es normal. Y nos obliga a un enfoque diferente: en lugar de perseguir una posición fija que no existe, necesitamos cubrir las distintas posibilidades para aparecer en la mayoría de ellas.

Si quieres que la IA te elija de forma consistente, tu contenido tiene que ser completo y versátil. Que cubra aspectos como el coste, lo fácil o difícil que es ponerlo en marcha, su capacidad para crecer y las integraciones disponibles. También tiene que estar actualizado y ofrecer formatos que la IA pueda sintetizar con facilidad, como pros y contras, tablas comparativas o conclusiones claras. En definitiva, se trata de convertirse en la fuente más completa para esa pregunta, de modo que, sea cual sea la ruta que tome la IA, tu contenido aporte valor y tenga opciones reales de ser citado.

La optimización en entornos probabilísticos es un juego de cobertura: cubrir todos los flancos relevantes para que, aunque el modelo explore variaciones, tu sitio tenga algo valioso que ofrecer.

En el siguiente capítulo vamos a desmontar todo el humo que rodea al GEO (los prompts mágicos, las garantías imposibles, los atajos que no existen) y a construir en su lugar el modelo mental operativo que usarás como brújula a lo largo de todo el libro.

CAPÍTULO 3

Lo que GEO NO es

Desmontando el humo

El campo de minas

En un campo tan nuevo como GEO, es inevitable que aparezcan dos cosas al mismo tiempo: profesionales serios intentando entender las reglas del juego y vendedores de humo prometiendo atajos que no existen. La proporción entre unos y otros, ahora mismo, no es favorable.

Antes de seguir avanzando con el contenido técnico del libro, necesito que ajustes tus expectativas. No porque lo que viene no funcione, sino porque si entras con las expectativas equivocadas vas a perder tiempo, dinero y paciencia persiguiendo espejismos. Y lo que es peor: vas a dejar de hacer las cosas que sí importan.

Así que vamos a limpiar el terreno.

No existen los prompts mágicos

Olvídate del mito de «encontrar el prompt perfecto» para que ChatGPT siempre recomiende tu marca. Olvídate de meter texto oculto en tu web esperando «hac-

kear» al modelo con inyección de prompts. Eso no es una estrategia. Es un callejón sin salida.

Los modelos actuales ignoran en gran medida esos intentos. Y cualquier truco que pudieras explotar temporalmente sería corregido rápidamente por las plataformas de IA. Google tardó años en neutralizar las trampas de SEO clásico. Los modelos de IA aprenden mucho más rápido qué funciona y lo neutralizan en semanas, a veces en días.

GEO no es buscar atajos sucios. Es jugar limpio pero con inteligencia: optimizar datos y contenido a largo plazo para que la IA te elija de forma legítima. Si alguien te ofrece un «hack de GEO», corre en la dirección contraria.

Nadie puede garantizarte el número uno en ChatGPT

En Google clásico ya era cuestionable garantizar una primera posición, salvo contadas excepciones. En entornos generativos es directamente una mentira o una ignorancia de manual.

No existe un Top 1 fijo en las respuestas de IA. Cada respuesta es probabilística, variable, dependiente de contexto. Ni siquiera el mejor optimizador del mundo puede forzar a una IA a mencionarte el 100% de las veces. Lo que sí se puede hacer es aumentar tus probabilidades de ser incluido. Inclinar la balanza a tu favor en cada generación.

Pasamos de un lenguaje de certezas a un lenguaje de probabilidades. Donde antes decías «optimizaremos para que tu página esté siempre en el Top 3 para tal palabra», ahora el planteamiento honesto es: «Optimizaremos para que, de cada 10 veces que alguien le pregunte a la IA sobre este tema, en 8 tu marca esté presente.» Es otro lenguaje. Acepta la variabilidad, pero trabaja para que las probabilidades estén a tu favor.

Regla general para protegerte: cualquier agencia o «experto» que venda «#1 en respuestas de IA garantizado» o no entiende cómo funciona esto, o te quiere estafar. No hay término medio.

GEO no reemplaza los fundamentos de SEO

No tires por la borda todo lo que sabes. GEO es una evolución, no una ruptura total.

La investigación de palabras clave sigue siendo necesaria. Las consultas pueden haberse alargado y acercado al lenguaje natural, pero entender qué quiere saber la gente sigue siendo la base. La arquitectura web y la calidad del contenido siguen siendo los cimientos para que una IA pueda entender, recuperar y reutilizar tu información. Conviene desmontar una fantasía frecuente: pensar que una IA te recomendará aunque no hayas publicado nada relevante. No lo hará. Si no tienes contenido útil, visible y recuperable, lo normal no es que la IA te elogie. Es que ni siquiera te mencione.

GEO construye sobre los fundamentos de SEO. Los amplía, los adapta y les da un nuevo propósito. Pero no los elimina.

Cuidado con las modas que duran tres meses

Hay mucho ruido ahí fuera. Posts y gurús diciendo cosas como «entrena tu propio mini-LLM para SEO y domina» o «inyecta tu marca en el modelo X y posiciona». Suena muy sofisticado. La realidad es que la mayoría de empresas primero deberían ocuparse de lo básico antes de experimentar con cosas exóticas.

¿Ya tienes contenido optimizado orientado a las preguntas que tu audiencia le hace a la IA?

¿Ya usas schema markup para ayudar a los modelos a entender tu información?

¿Monitorizas si apareces en respuestas de IA? Si la respuesta a alguna de estas preguntas es no, ni te plantees entrenar un modelo propio ni comprar la herramienta de moda. Primero lo primero.

En la historia del SEO hemos visto montones de modas y «trucos milagrosos» que luego quedaron en nada. Las meta keywords en los 2000. El PageRank sculpting con nofollow. Más recientemente, el hype de «escribe todo tu contenido con GPT y ya está». Quien lleva años en esto sabe que mantener la cabeza fría es la habilidad más rentable.

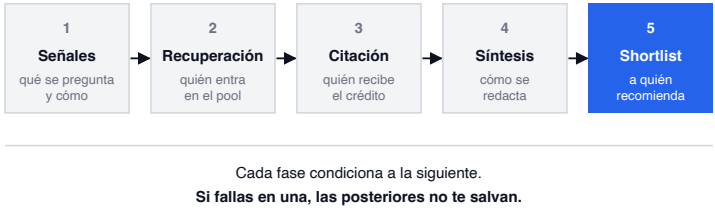
En GEO, aún más: la novedad y la opacidad de los modelos lo convierten en un terreno especialmente fértil para las promesas vacías y la información confusa. Si partes sabiendo lo que no debes esperar, podrás concentrarte en lo que sí funciona. Quienes se pongan las pilas con tácticas reales ahora, mientras otros pierden el tiempo persiguiendo espejismos, van a sacar una ventaja que será muy difícil de cerrar.

El modelo mental que lo ordena todo: señales, recuperación, citación, síntesis, shortlist

Hemos cubierto mucho terreno en los tres primeros capítulos. Nuevas métricas, cómo funciona la IA por dentro, qué tácticas valen y cuáles no. Ahora necesitamos integrar todo en un modelo mental que te sirva de mapa para el resto del libro y, más importante, para tu trabajo diario.

Piensa en una cadena de cinco eslabones. Cada eslabón es una fase por la que pasa una consulta desde que el usuario la formula hasta que tú apareces en la respuesta final. Cada fase condiciona a la siguiente. Si fallas en una, las posteriores no te salvan.

La cadena es esta: Señales → Recuperación → Citación → Síntesis → Shortlist.



La cadena de la respuesta generativa. Fuente: elaboración propia.

Fase 1: Señales

Todo comienza con las señales iniciales que guían al motor. La intención del usuario, el contexto personalizado del usuario y las señales que emergen del contenido disponible.

Los motores generativos modernos detectan señales ricas. Pueden inferir, por cómo está formulada la pregunta, qué cosas valora el usuario. Consideran señales universales como frescura, autoridad o especificidad temática.

¿Qué controlas aquí? Sinceramente, muy poco de lo que hace el usuario. Pero sí puedes alinear tu contenido para aprovechar esas señales. Habla el mismo idioma que tu audiencia. Usa la terminología que tus usuarios usan. Si vendes seguros y la gente pregunta «¿cuál es la mejor aseguradora de coche?», asegúrate de que en tu contenido aparece la palabra «aseguradora» y no solo «compañía de seguros automotrices». El LLM recibe señales de lenguaje, y si tu texto usa términos poco habituales para tu audiencia, puede que no lo asocie a la intención del usuario con la misma fuerza.

Estructura tu contenido con claridad: títulos descriptivos, secciones bien definidas. Eso señaliza al modelo que tu información está organizada y es fácil de extraer. En pocas palabras: habla claro y en sintonía con la intención de tus usuarios.

Fase 2: Recuperación

Con las señales claras, el sistema se pregunta: «¿Qué información candidata hay ahí fuera para resolver esto?» Esta es la fase en la que el motor busca en sus índices (vectoriales, clásicos o ambos) documentos, páginas o los fragmentos que podrían responder la pregunta.

Aquí es donde entran en juego tus páginas. Si has hecho los deberes, alguna de tus páginas será considerada lo suficientemente relevante como para aparecer en el conjunto candidato. Si no, da igual todo lo demás.

¿Qué influye para que te recuperen? Lo clásico y algo más: la calidad del contenido, la coincidencia semántica con la consulta, la autoridad de tu dominio, la actualidad de la información y la estructura. Muy parecido a los factores SEO de siempre, pero con más peso en la relevancia semántica que en la mera keyword exacta.

De miles o millones de páginas posibles, quizás solo 5 a 10 pasan el corte para alimentar al modelo generativo. Tu misión es estar en ese grupo. Y aquí sí tienes bastante control: SEO técnico impecable, contenido indexado, información útil y pertinente y cobertura profunda del tema. No controlas cuántas páginas exactas cogerá el

modelo, pero sí puedes optimizar para ser uno de los mejores candidatos.

Fase 3: Citación

Una vez el sistema tiene en su «bolsa» fragmentos de varias fuentes, debe decidir cuáles menciona explícitamente en la respuesta final. Y esto es importante: que te hayan recuperado no garantiza que te citen.

La IA podría usar tu información pero no mencionarte o podría agrupar tu dato bajo la cita de otra fuente. La decisión depende en parte de las políticas de la plataforma. Microsoft Copilot siempre intenta dar referencias visibles y ChatGPT, cuando responde sin buscar en la web, no cita nada.

Si tu contenido es muy claro y autocontenido en sus afirmaciones, es más fácil que la IA te asigne esa pieza de información y te cite. Si tu contenido está entremezclado con ideas de otras fuentes o no tiene frases que funcionen como unidades independientes, la IA puede usarte sin darte crédito.

Ser citable es prácticamente un nuevo tipo de SEO: el SEO de la atribución. Proveer datos únicos aumenta las probabilidades de que el modelo te cite. Si tu contenido es genérico o redundante con otras fuentes, la IA puede dar esa misma información y citar a cualquiera. O a nadie.

Lograr la cita es el nuevo Top 1. Es tu marca apareciendo en pantalla, asociada a información útil, generando con-

fianza. Eso vale más que un ranking orgánico.

Fase 4: Síntesis

Esta es la etapa en que el LLM redacta la respuesta final. Toma fragmentos de las fuentes recuperadas y los combina con sus propias palabras para construir algo coherente. Añade estructura discursiva, conectores y resúmenes.

¿Por qué nos importa esta fase? Porque incluso si lograste ser recuperado, tu contenido tiene que encajar bien en la síntesis para que no sea descartado durante la redacción.

Si tu contenido viene con frases claras y concisas, se integra mejor en la respuesta final. Si es enrevesado o muy largo, es más probable que la IA prefiera resumirlo o usar directamente la explicación de otra fuente más sencilla. Si varias fuentes dicen lo mismo con distintas palabras, la IA puede amalgamar y citar solo una.

Hay un aspecto importante: el tono. Si la mayoría de fuentes que el modelo ha recuperado tienen un tono neutro e informativo y la tuya es puro marketing «somos los mejores del mundo», puede que la IA decida no usar tu texto.

No controlas cómo sintetiza la IA, pero influyes escribiendo de forma que le resulte fácil integrarte. Contenido claro, conciso, estructurado: más probabilidades de que te use. Contenido confuso, exagerado o disonante: más probabilidades de que te deje de lado.

Fase 5: Shortlist

La shortlist es el conjunto final de resultados o recomendaciones que ve el usuario. En preguntas tipo «¿cuál me recomiendas?» o «¿cuáles son los mejores?», la IA suele dar una lista corta: «Te recomiendo A, B y C.» Esa es la shortlist. El nuevo Top 3.

Estar en la shortlist es el objetivo último de GEO. Y tiene una característica que lo cambia todo respecto al SEO clásico: es binario. O estás o no estás. No es como antes, que estar en la posición 5 te daba algo de tráfico aunque menos que la posición 1. Aquí, si la shortlist son tres opciones y no estás entre ellas, no existes para ese usuario.

Por eso en GEO hablamos de Share of Model en vez de ranking. Qué porcentaje de las veces te incluyen. Es una cuota de presencia, no una posición fija.

No controlas directamente ser shortlist. Es el resultado de haber pasado todas las fases anteriores. Pero sí puedes medirlo, fijar objetivos y trabajar cada eslabón de la cadena para mejorar tu tasa de aparición. En Antropus, el Prompt Tracker te da esa tasa de aparición motor a motor, y el GEO Strategy Builder convierte el diagnóstico en un roadmap priorizado a 30, 60 y 90 días.

Dónde encaja cada táctica

Lo útil de este modelo mental es que te permite entender en qué fase actúa cada mejora. Por ejemplo, mejorar el tiempo de carga de tu web influye sobre todo en la fase de recuperación, porque puede ayudarte a ser en-

contrado a través del SEO clásico. En cambio, reescribir tus párrafos clave con un tono más neutral y factual influye más en las fases de citación y síntesis. Es decir, puede hacer que la IA use mejor tu contenido cuando ya lo ha encontrado, pero no garantiza que vaya a encontrarte más veces.

Cada uno de los capítulos que vienen profundizará en optimizaciones para una o varias de estas fases. Tener clara la cadena completa desde ahora te permite entender el porqué de cada táctica. No estás haciendo cosas al azar, estás trabajando sobre un sistema.

Qué controlas, qué no, y dónde están tus palancas

Para cerrar con claridad el marco que vas a usar a lo largo de todo el libro:

La consulta del usuario, su intención y su contexto personal no los controlas. Ni falta que hace intentarlo. Lo que sí puedes hacer es investigar esas intenciones para preparar contenido alineado. Para eso usamos en Antropus Prompt Intelligence, que organiza las preguntas de tu sector en sets por intención (marca, descubrimiento, comparación, decisión, producto y confianza), y el GEO Prompt Deriver, que convierte tus keywords SEO en prompts realistas.

Las señales que emite tu contenido las controlas indirectamente. Tú decides cómo escribir, qué lenguaje usar, qué estructura dar. Eso envía señales que el mode-

lo puede captar. No controlas qué señales exactas usa la IA, pero sí qué emite tu contenido.

La recuperación es la fase en la que más puedes influir, aunque nunca la controles al cien por cien. Puedes hacer muchas cosas para convertirte en una fuente candidata fuerte, como trabajar bien el SEO, crear contenido relevante y cubrir el tema con profundidad. Aun así, no puedes garantizar que la IA te recupere siempre, porque también intervienen factores externos. Pero, de todas las fases, esta es en la que tus acciones suelen tener más impacto.

La política de citación de la plataforma no la controlas. Si ChatGPT no cita por diseño cuando responde sin buscar en la web, no hay nada que hacer. Si Copilot siempre cita al menos dos fuentes, al menos sabes que compites por estar entre esas dos.

Ser citado lo influencias parcialmente. No decides que te citen sí o sí, pero puedes dar motivos para la cita: datos únicos, frases autocontenidas, análisis originales.

La síntesis no la controlas. La redacción final la decide el modelo. Lo que influyes es tu estilo: si tu contenido se integra bien o no.

La shortlist final es resultado de todo lo anterior. Todo tu trabajo anterior te lleva hasta ahí.



Qué controlas, qué influyes y qué no controlas. Fuente: elaboración propia.

La consecuencia práctica de este marco: concéntrate en lo controlable, sé inteligente en lo influenciable y no malgastes energía en lo que no puedes cambiar. No trates de engañar al modelo. Mejor ocúpate de publicar contenido relevante y bien optimizado. No te frustres si un día no saliste. Saca promedio de muchos intentos. Este foco te ahorrará tiempo y te hará más efectivo.

A partir de aquí empieza el trabajo técnico de verdad. En la Parte II vamos a abrir completamente el capó de los LLMs para entender RAG, embeddings, ventanas de contexto y tipos de consulta. Son los fundamentos de ingeniería que necesitas para dejar de optimizar a ciegas y empezar a tomar decisiones basadas en cómo funciona realmente la máquina.

PARTE II

La ingeniería detrás de la visibilidad

Lo que necesitas entender de los LLMs para dejar de optimizar a ciegas.

CAPÍTULO 4

RAG 101

La recuperación es tu único pase

Tu mejor contenido no existe si la IA no puede encontrarlo

Hay una verdad incómoda que la mayoría de profesionales de marketing todavía no ha asimilado: tu contenido, por brillante que sea, no existe para un modelo de lenguaje si ese modelo no puede recuperarlo. Da igual que hayas escrito el artículo definitivo sobre tu sector. Da igual que tengas la respuesta perfecta a la pregunta que acaba de hacer el usuario. Si esa información no llega al modelo, no aparecerás.

Los LLM modernos no lo saben todo ni responden por arte de magia. Cuando necesitan información actualizada o externa, muchos utilizan un proceso llamado RAG, siglas de *Retrieval-Augmented Generation*. Dicho de forma sencilla: antes de redactar una respuesta, buscan y recuperan datos relevantes de otras fuentes. Después, con esa información como apoyo, generan la respuesta final.

Un LLM no investiga por su cuenta ni verifica como un humano. Si no le aportas buenas fuentes, puede completar los huecos con una respuesta inventada.

Cómo funciona RAG por dentro

El proceso tiene una secuencia clara. Cuando un usuario hace una pregunta, el sistema convierte esa consulta en un embedding vectorial (representación numérica que captura el significado semántico de la pregunta). Después, busca en su base de conocimientos los fragmentos de contenido que están más cerca de ese vector en el espacio semántico. Recupera esos fragmentos, los inserta en la ventana de contexto del modelo generativo, y solo entonces el modelo genera la respuesta apoyándose en esos datos.



Cómo funciona RAG, paso a paso. Fuente: elaboración propia a partir del texto de este capítulo.

La clave está en la recuperación, no en la generación. Si recuperas basura, la respuesta será basura. Si no recuperas nada, el modelo rellenará con sus suposiciones probabilísticas. Eso es exactamente lo que coloquialmente llamamos alucinaciones: el modelo inventando datos con total convicción porque no tenía nada real a lo que agarrarse.

Por eso la mayoría de agencias todavía no dan pie con bola en IA: siguen obsesionadas con keywords y posiciones, cuando la guerra se gana o se pierde en la fase de recuperación.

Las consecuencias de no estar en los datos

Piensa en lo que está ocurriendo ya en muchos procesos de compra B2B: un comprador corporativo pide a un asistente de IA que le recomiende soluciones para su problema, y el modelo, trabajando con datos desactualizados, propone la alternativa open-source del competidor en lugar de tu producto. En su base de entrenamiento no figura la información reciente sobre tu SaaS, así que el sistema tira de lo que tiene: el nombre del competidor. Veneno puro para la venta, y nadie en tu equipo se entera de que perdiste esa oportunidad porque nunca hubo lead que rastrear.

Otro caso, este completamente público y documentado: en febrero de 2024, el Civil Resolution Tribunal de British Columbia condenó a Air Canada a indemnizar al pasajero Jake Moffatt porque su chatbot le había dado información incorrecta sobre las tarifas de duelo. El sistema, al no encontrar la política real actualizada, se la inventó. La aerolínea intentó alegar que el chatbot era «una entidad legal separada», argumento que el tribunal rechazó. Sentencia firme. Indemnización pagada.

Es el primer precedente legal claro de que una empresa es responsable de lo que su IA dice sobre ella. No será el

último.

La idea es simple, pero importante: cuando la IA no dispone de datos fiables para apoyarse, puede rellenar los huecos con conjeturas en lugar de con hechos. Para una marca, eso supone un riesgo real, porque puede dar respuestas inexactas sobre tus productos, servicios o posicionamiento. Y eso puede traducirse en confusión, pérdida de confianza, oportunidades perdidas e incluso problemas legales en algunos casos.

La misión ha cambiado

La solución no tiene glamour, pero es real: obsesionarte con la recuperabilidad de tu contenido en entornos de IA.

Tu misión como profesional del marketing ya no es solo lograr que te encuentren en Google. Ahora también consiste en hacer que un LLM pueda recuperar tu información cuando sea relevante y la considere útil para responder. Eso exige ir más allá de la redacción SEO clásica y empezar a pensar en cómo facilitar la recuperación, la comprensión y el uso de tu contenido.

En la práctica, esto se traduce en varias líneas de trabajo simultáneas.

La primera es tener una base de conocimiento bien estructurada. Tus contenidos importantes deben estar organizados y troceados en fragmentos semánticos accesibles: documentos clave, FAQs y fichas técnicas, todo lo que un algoritmo de búsqueda pueda indexar y recupe-

rar por separado. No vale tener un PDF de 80 páginas sin estructura. Hay que preparar los datos para ser encontrados.

La segunda es asegurar tu presencia en fuentes externas que los modelos consultan. Si usas modelos externos como ChatGPT o Perplexity, trabaja para que tu información esté en las fuentes que esos modelos consideran fiables, por ejemplo: tener una página en Wikipedia actualizada, contribuir a Wikidata, publicar estudios o datos en repositorios que los modelos ingieren. Muchos LLMs dan un peso especial a Wikipedia y a fuentes con datos abiertos y verificables.

La tercera es monitorizar cómo «ven» los modelos tu contenido. Existen APIs para obtener el vector de un texto. Genera embeddings de tus páginas y de las preguntas comunes de tus clientes, calcula la similitud coseno entre ambos. Si están cerca, vas bien. Si están lejos, significa que tu contenido habla de algo diferente a lo que te preguntan. Hay brecha semántica, y esa brecha hay que cerrarla.

Recuperabilidad: el nuevo ranking

En el SEO tradicional competías por aparecer lo más arriba posible en Google. En el nuevo escenario, también compites por seguir siendo visible dentro de las respuestas generadas por IA. Puedes estar muy bien posicionado en las SERP y, aun así, quedar fuera si los LLM no recuperan ni tienen en cuenta tu contenido.

Recuperabilidad (*retrievability*) es qué tan probable es que el modelo te saque como fuente cuando genera una respuesta. En la práctica, viene a reemplazar el concepto de *rankability* de toda la vida. Ya no importa tanto «¿en qué posición salgo en Google?» sino «¿el modelo me recuperará o me ignorará?».

Los LLM no recorren toda la web cada vez que responden. Trabajan con un conocimiento previo y, en algunos casos, con búsquedas limitadas sobre un conjunto acotado de fuentes. No consultan internet de forma abierta, completa y en tiempo real como lo haría una persona. Por eso, si tu contenido no entra en las fuentes que el sistema puede recuperar o tener en cuenta, tus opciones de aparecer se reducen al mínimo.

Si no estás en los datos, no estás en la respuesta. Este principio es la piedra angular de la visibilidad en IA.

Las dos etapas: descubrimiento y elección

Lograr recuperabilidad pasa por dos fases. Primero, ser descubierto por la IA (estar presente en sus índices o en sus fuentes de datos) y segundo, ser elegido por la IA como fuente relevante entre todas las que encontró.



Las dos etapas de la recuperabilidad. Fuente: elaboración propia a partir del texto de este capítulo.

El descubrimiento empieza por algo muy básico: estar presente donde la IA puede encontrarte. Muchas marcas no lo están. No tienen página en Wikipedia, no contribuyen con datos estructurados a knowledge graphs públicos y no aparecen en fuentes que estos sistemas suelen consultar. Si fallas en varios de esos frentes, partes con desventaja. Porque la IA no puede recuperar lo que apenas está presente en su radar.

La elección ya no depende solo de estar presente, sino de resultar realmente útil y relevante. Supongamos que el modelo ya sabe que existes. Aun así, queda una pregunta decisiva: ¿por qué iba a usar tu contenido y no el de otro? Aquí es donde muchas prácticas SEO antiguas pierden fuerza. Hay páginas muy pensadas para posicionar, pero poco útiles para un LLM: repiten palabras clave, alargan el texto sin aportar demasiado y priorizan el SEO antes que la claridad. Un modelo de lenguaje no busca solo coincidencias, sino información útil, concreta y fiable. Si tu página es puro marketing y otra ofrece

una comparativa clara con datos específicos, es mucho más probable que el sistema elija la otra.

En esta etapa también entra la alineación de entidades. Si tu marca o tus términos no están claramente contextualizados, la IA puede confundirlos o ignorarlos. Si tu marca se llama Orange y vendes software, pero tu contenido no deja clarísimo ese contexto, un modelo puede descartar tus menciones para no arriesgarse a confundir con la fruta o con la operadora de telefonía. Tu marca debe estar desambiguada en tus contenidos y, si es posible, en fuentes externas como Wikipedia o Wikidata. En Antropus lo resolvemos con la Ficha de Entidad: nombre canónico, descripción, texto de desambiguación, alias y perfiles oficiales (*sameAs*) en un solo sitio, para que la IA sepa de qué marca hablas y no te confunda con un homónimo.

Tres frentes para mejorar tu recuperabilidad

El primer pilar es la estructura técnica, porque ayuda a que los sistemas entiendan y aprovechen mejor tu contenido. Hablamos de un HTML semántico y limpio, con títulos bien jerarquizados, listas, tablas y encabezados que dejen claro de qué trata cada sección.

También de datos estructurados con schema.org, que aportan contexto adicional sobre la página. La idea es sencilla: no escondas lo importante bajo capas de contenido poco útil. Y hay otro punto relevante, aunque a algunos SEO les resulte menos intuitivo: enlazar a fuentes

fiables y relacionadas puede reforzar la credibilidad de tu contenido.

Cuando una página referencia fuentes reconocidas, transmite mejor que está informando y contextualizando, no solo promocionándose.

El segundo frente es la presencia en fuentes externas. Conviene estar en aquellos lugares que los sistemas suelen tomar como referencia para contrastar o completar información. Si encaja con tu caso y cumples las políticas, una página en Wikipedia puede ayudar. Si aún no tienes suficiente notoriedad, Wikidata suele ser una vía más accesible y también puede aportar señales útiles. Si cuentas con estudios propios, datos originales o análisis relevantes, intenta publicarlos en fuentes académicas, repositorios abiertos o espacios especializados. También es importante aparecer en directorios, comparadores, rankings e informes de referencia dentro de tu sector. Si tu competencia aparece de forma recurrente en ese tipo de fuentes y tú no, es más probable que la IA la perciba como una opción más visible, más reconocida y más fácil de citar.

El tercer frente es la afinidad semántica. Tu contenido debe resultar relevante para la consulta del usuario, no por repetir la misma frase una y otra vez, sino porque trata el tema con el lenguaje, el contexto y los matices que realmente importan. Eso implica cubrir las entidades, conceptos y subtemas relacionados con la intención de búsqueda, incorporar sinónimos y abordar también los puntos que suelen aparecer alrededor de la pre-

gunta principal. Los modelos valoran especialmente el contenido que no se queda en lo abstracto y profundiza con ejemplos, escenarios reales y casos de uso concretos. Además, conviene reforzarlo con evidencia verificable, como datos, comparativas, referencias o estudios, porque eso no solo aumenta la confianza, sino que también enriquece tu embedding (la huella semántica del contenido) y lo conecta con hechos más sólidos.

El test que deberías hacer esta semana

Entra en el chat de tu preferencia, como por ejemplo ChatGPT, Perplexity o Copilot, y haz una búsqueda sobre tu categoría sin mencionar tu marca. Por ejemplo: «¿Cuáles son las mejores opciones de [tu producto genérico]?» o «Recomiéndame un [tipo de servicio que ofreces]». Después, analiza la respuesta con calma. ¿Apareces? Y si no apareces, ¿qué marcas sí lo hacen y por qué? ¿Qué señales, datos o enfoques están mostrando ellas que tú no estás mostrando de forma visible? A veces descubrirás que un competidor destaca por una característica que tú también tienes, pero que apenas has comunicado en tu contenido público. Ahí suele haber una oportunidad clara de mejora.

Este tipo de pruebas te da una pista muy directa de cómo te está viendo la IA y de qué te falta para entrar en la conversación.

Hacer esta prueba una vez abre los ojos; repetirla cada mes es lo que te dice si avanzas. Es lo que automatizamos en Antropus con el Prompt Tracker: el mismo set de

prompts, lanzado a ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview, midiendo si te mencionan, te recomiendan y te citan.

Los límites de todo esto

Implementar RAG no es una bala de plata. No basta con añadir un paso de búsqueda si los datos que recuperas son malos. La recuperación solo funciona con material de calidad. Si tus páginas están desordenadas, la IA podría ignorarlas, o peor, usarlas mal y dar una respuesta incorrecta citándote a ti.

Además, hay factores que escapan a tu control. En algunos temas, el sistema puede dar más peso a ciertas fuentes, como papers, documentación técnica o Wikipedia. También puede ocurrir que un competidor acumule tanta presencia pública y comunitaria que su vector de relevancia sea mucho más fuerte que el tuyo. Y en algunos sectores, la IA puede inclinarse por fuentes neutrales y documentales antes que por la web corporativa de una marca porque las percibe como más objetivas.

Ser recuperable no garantiza ni clics ni conversiones. Muchas veces la IA dará la respuesta usando tus datos y el usuario no necesitará ir a tu web. Es una especie de branding invisible. Tu marca gana presencia mental aunque no tengas un clic medible. Una mención en la respuesta del asistente que resuelve la duda del usuario puede valer más que cien impresiones en Google donde nadie leyó tu snippet.

Aun así, es valioso. Estamos en el juego de la influencia sin clic. Y la alternativa —no estar en los datos, no ser recuperable, ser invisible para la IA— es mucho peor.

En el siguiente capítulo vamos a desmitificar los embeddings y el espacio semántico: cómo las máquinas «entienden» el significado de tu contenido, por qué la densidad de keywords ha muerto, y cómo medir si estás cerca o lejos de donde necesitas estar.

CAPÍTULO 5

Del keyword al vector

Embeddings, distancias y la nueva semántica

El universo invisible donde compites

Hemos hablado de embeddings, espacios vectoriales, similitud coseno. Es el momento de desmitificar todo esto, porque si no entiendes cómo funcionan los vectores, estás optimizando a ciegas. Y optimizar a ciegas en un entorno probabilístico es tirar dinero.

La densidad de palabras clave ha muerto. Los LLMs operan en un universo de vectores semánticos. Para ser relevante en IA, tienes que pensar en coordenadas, no en keywords.

Un embedding es la representación numérica de un texto en un espacio multidimensional. En vez de palabras sueltas, piensa en conceptos situados en un espacio. La posición en ese espacio indica significado. Cuando introduces un término o una frase en un modelo, este lo convierte en un vector. Dos textos con contenido semántico similar producirán vectores cercanos. Dos textos sobre cosas totalmente distintas estarán muy alejados.

La medida más común de cercanía es la similitud coseno. Mide el ángulo entre dos vectores, ignorando su longitud. Si el ángulo es pequeño (coseno cercano a 1) los vectores apuntan en direcciones parecidas: los conceptos están semánticamente alineados. Si el ángulo es grande (coseno cercano a 0) no tienen nada que ver.

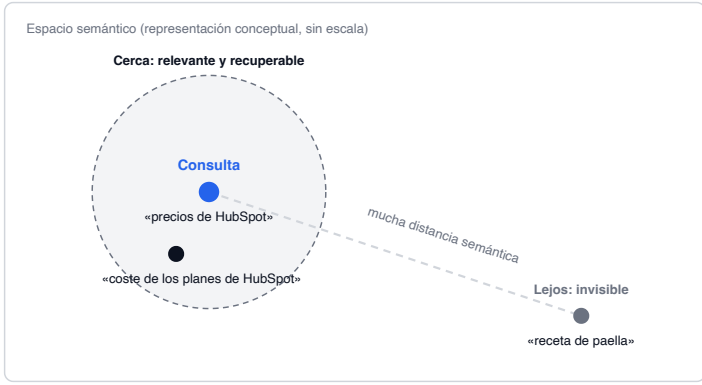


La similitud coseno mide el ángulo, no la longitud. Fuente: elaboración propia a partir del texto de este capítulo.

¿Por qué el coseno y no contar palabras? Porque captura la dirección del significado. «Precios de HubSpot» y «coste de los planes de HubSpot» van en la misma dirección semántica aunque no compartan ninguna palabra exacta. Sus embeddings estarán cercanos. En cambio, «HubSpot» y «receta de paella» estarán en polos opuestos del espacio.

Para ti, el mensaje es simple: lo que importa es la similitud entre el vector de tu contenido y el vector de la consulta del usuario. Si están cerca, eres relevante y puedes ser recuperado.

Si están lejos, eres invisible.



Cerca o lejos de la intención: el espacio semántico. Fuente: elaboración propia con los ejemplos del texto; representación conceptual, sin escala.

El error que cometen los veteranos

Paradójicamente, los profesionales del SEO con más experiencia pueden tener más riesgo de cometer este error. Durante años han trabajado pensando en señales visibles, como la keyword en el H1, en el título o en el texto alternativo de una imagen. Pero en los entornos de IA eso ya no basta. El foco no está solo en repetir una palabra clave, sino en el significado global que transmite el contenido y en cómo lo interpreta el modelo.

He visto páginas que repiten la keyword principal quince veces, pero no incluyen ni un solo sinónimo, concepto relacionado o matiz útil. Hace años, eso podía colar mejor en SEO. Pero para un sistema basado en embeddings, ese contenido transmite una señal semántica débil y queda incompleto.

Un ejemplo concreto: una página titulada «CRM innovador para empresas» que repite constantemente la palabra «innovador», habla de beneficios genéricos, pero no menciona conceptos clave como pipeline de ventas, automatización de marketing o integraciones. Su embedding queda flotando en un espacio semántico demasiado genérico, desconectado de los términos y contextos que realmente definen un CRM. Aunque esa página llegara a posicionar en Google para «CRM innovador», un LLM, al vectorizarla, la interpretará como una fuente incompleta: le faltan demasiadas señales semánticas para representar bien la categoría.

Mientras tanto, otra página menos adornada, pero que sí menciona el pipeline de ventas, las integraciones, la automatización y casos de uso reales, generará un embedding mucho más rico y mejor conectado con lo que realmente define un CRM. Por eso, es más probable que el LLM la considere una fuente más útil y la prefiera.

Optimizar solo el texto visible (keywords exactas, frecuencia) es optimizar para humanos y para el Google antiguo. Optimizar para IA es optimizar tu posición en el espacio semántico. Y eso requiere cubrir temática, subtemas y contextos conceptuales, no repetir la misma palabra sin agregar profundidad.

Cómo medir lo que no puedes ver

La buena noticia es que ya existen herramientas accesibles para visualizar tus vectores. OpenAI ofrece embed-

dings de textos con una simple llamada a su API. Con eso puedes hacer experimentos que valen oro.

Genera el embedding de una pregunta frecuente de tu sector. Genera el embedding de tu página que crees que la responde. Calcula la similitud coseno entre ambos.

Si obtienes un valor bajo —digamos 0.15 en una escala donde 1.0 es idéntico— es alerta roja. Significa que para la IA, tu página habla de otra cosa respecto a esa pregunta. Si ves un valor alto, alrededor de 0.8, vas bien: tu página está semánticamente muy alineada con la consulta.

Otra práctica reveladora: toma la página top de un competidor o de Wikipedia sobre tu tema, saca su embedding y compáralo con el de tu página para la misma consulta. Si tu similitud es muy inferior a la de Wikipedia, ya sabes por qué la IA cita a Wikipedia y no a ti.

Y lo más útil: identificar las diferencias. Muchas veces, esta comparación vectorial revela qué tiene el otro contenido que el tuyo no. Quizás la página competidora menciona un caso de estudio específico o una tecnología que en la consulta estaba implícita y tú ignoraste. Esas ausencias restan puntos a tu embedding. Una vez detectadas, ajustar el contenido es cuestión de incluir ese concepto faltante de forma natural. No es meter keywords por meter. Es sumar profundidad semántica donde falta.

Esa comparación es la que hace el Content Gap Optimizer de Antropus partiendo de un prompt real: le das la

pregunta y tu URL, mira qué páginas están citando los motores y te devuelve qué tienen ellas que la tuya no, punto por punto.

Un truco avanzado: crea un vector «ideal» de intención combinando varias formulaciones de la misma consulta. Promedia los embeddings de «¿qué CRM elegir?» + «mejor CRM 2026» + «CRM recomendaciones». Obtienes un vector que encapsula la intención general de «elegir CRM». Calcula el coseno de tu contenido contra ese vector promedio. Te servirá como un indicador aproximado de cuánto cubres la intención completa.

Rico pero enfocado: la paradoja del buen embedding

Para gustarle a la IA, tu contenido debe ser semánticamente rico pero temáticamente enfocado. Suena contradictorio, pero no lo es.

Rico significa que abordas los subtemas relevantes. Si escribes sobre CRM para ecommerce, asegúrate de hablar de integración con plataformas de tienda online, gestión de carrito abandonado y analítica de clientes. Si solo hablas de CRM genérico, tu embedding se queda a medio camino. No se trata de alargar el texto sin necesidad; a veces basta un buen párrafo para cubrir un concepto que faltaba. Calidad sobre cantidad, también en vectores.

Enfocado significa que tu contenido debe mantenerse centrado en la intención principal de la página. Si estás

escribiendo sobre «CRM para fintech», no conviene meter bloques largos sobre temas que se alejan de esa intención, como una historia extensa de los CRM desde los años 80. ¿Por qué? Porque los modelos suelen trabajar por fragmentos. Si una parte del contenido habla de historia y otra de fintech, cada bloque puede generar señales semánticas distintas. El riesgo es que el sistema entienda mejor tu página como una fuente sobre la historia del CRM que como una fuente útil para elegir un CRM en el sector fintech. En ese caso, podrías aparecer para preguntas que no te interesan y perder fuerza justo en la consulta para la que sí querías ser relevante.

Si una sección no aporta valor directo a la intención principal de la página, conviene revisarla. En algunos casos, será mejor moverla a un contenido independiente. Cuanto más centrada esté cada página en su objetivo, más clara será su señal para buscadores y modelos de IA.

Tu marca en el espacio semántico: la distancia que define tu destino

Ahora viene la parte que duele. Hablemos de tu marca y de una de las intenciones más jugosas que existen: «¿Cuál es el mejor [producto/servicio]?».

Que tu marca sea la líder del mercado no garantiza que un LLM la recomiende como la mejor. Si hay mucha distancia semántica entre cómo la IA percibe tu marca y la idea de «mejor X», te van a saltar. Y el negocio se lo llevará otro.

En el espacio semántico del modelo, una consulta como «¿Cuál es el mejor software de gestión de proyectos?» genera un vector de intención que combina señales como funcionalidades, reputación, opiniones positivas o relación calidad-precio. Tu marca representa otro punto, o incluso otro vector, dentro de ese mismo espacio. La pregunta clave es hasta qué punto ambos quedan cerca: es decir, si el modelo asocia tu marca con esa intención de búsqueda.

Si tu marca queda cerca de esa intención en el espacio semántico del modelo, tendrá más opciones de aparecer en la respuesta. Si, en cambio, el sistema la asocia con otra categoría (por ejemplo, una herramienta básica, una opción gratuita o una solución pensada para microempresa), puede que no entre las principales recomendaciones cuando alguien pregunte por «la mejor». En el mejor de los casos, aparecerá en un contexto más concreto, pero no como una de las opciones destacadas.

En este tipo de sistemas, la distancia semántica puede decidir si entras en la recomendación o te quedas fuera.

De qué depende esa distancia

De los datos que entrenaron al modelo y de las señales contextuales que esos datos contienen. En qué contexto aparece tu marca habitualmente: ¿junto a palabras como «top, líder, ganador de premio»? ¿O más en frases tipo «alternativa barata a...»? Qué describen de ti en reviews y artículos: ¿«muy completo y robusto» o «herramienta sencilla para empezar»? Quién habla de ti y dón-

de: no es lo mismo que te mencione un blog de autoridad comparándote con líderes, que solo aparezcas en foros de usuarios pidiendo ayuda técnica.

La historia importa. Si durante años tu marca no figuró en ningún ranking de «lo mejor de...», el modelo arrastra esa visión rezagada aunque hoy seas excelente. Los LLMs no tienen datos en tiempo real y suelen cargar con la inercia de su entrenamiento.

Cómo acercar tu marca al cluster de «los mejores»

El primer paso es auditar cómo te ve la IA. Haz búsquedas sin nombrarte: «mejores [tu categoría]». Y con tu nombre: «¿[TuMarca] es bueno para...?». Observa qué dice. Lee tu propia entrada de Wikipedia como si fueras la IA: ¿suenas a líder o a uno más del montón? Busca los rankings y listas de tu sector de los últimos años: ¿estás en ellos? Si no, identifícalos como objetivos de PR.

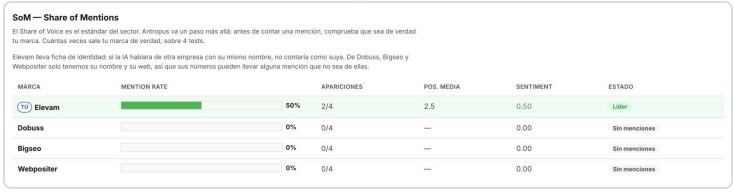
El segundo paso es crear contenido y PR que te reposicionen. Atrévete a publicar contenido comparativo en tu web (tu marca junto a competidores, con datos objetivos). Aunque suene a hacerles publicidad, en realidad te posicionas en el terreno de la comparativa. Cuando el modelo vea tu nombre junto a los otros en un mismo texto, te asociará como jugador de esa liga.

Colabora con analistas para salir en informes «Top 10». Publica casos de éxito donde clientes digan explícitamente por qué te eligieron sobre otras opciones. Destaca

tus logros en fuentes que los modelos ingieren: Wikipedia, comunicados de prensa indexados, LinkedIn.

El tercer paso es monitorizar la evolución de tu visibilidad en IA. Es decir, medir con el tiempo si aumentan tus menciones, citaciones, atribuciones correctas y apariciones en shortlist para las consultas estratégicas de tu categoría. Si tienes medios técnicos, puedes apoyarte en embeddings y similitud semántica. Si no, puedes hacerlo con pruebas manuales y un seguimiento periódico de prompts clave.

Para la parte competitiva usamos en Antropus Competitor Intelligence: cuántas veces sale de verdad tu marca frente a tus competidores cuando la IA habla de tu categoría, en qué posición y con qué tono, y qué rivales reconoce el modelo que quizá tú no tenías en el radar.



Antropus · Competitor Intel: menciones, posición media y sentimiento de Elevam frente a Dobuss, Bigseo y Webpositer. Proyecto real de Elevam (elevam.es), 31 de agosto de 2026. Es una muestra pequeña (4 tests en un motor): sirve para ver la pantalla, no para sacar conclusiones.

Y el cuarto punto, quizá el más incómodo, es aceptar que no todas las batallas se pueden ganar. Si tu marca no destaca objetivamente en ciertos aspectos, es probable que la IA también lo refleje. Estos sistemas suelen matizar sus recomendaciones según el contexto: una opción

puede ser mejor para grandes empresas, otra para pymes y otra para quien busca una solución gratuita. Por eso, si tu fortaleza real está en un nicho o en un caso de uso concreto, lo más normal es que la IA te mencione, sobre todo, en ese terreno. A partir de ahí, tienes dos caminos: asumir ese posicionamiento especializado o trabajar para ampliar tu propuesta de valor y competir en más escenarios.

Evita los atajos oscuros. No intentes inundar la web de artículos falsos que fingen ser neutrales y te promocionan. Eso ya no funcionaba bien en Google y los LLMs tampoco son tontos. Pueden detectar patrones de autopromoción y esos contenidos no tendrán peso o incluso podrían hacerte ver menos confiable. Mejor cinco menciones de calidad en lugares distintos que cincuenta de spam.

En el siguiente capítulo vamos a hablar de algo que parece muy técnico pero que tiene implicaciones directas en cómo estructuras cada página de tu web: las ventanas de contexto de la IA, por qué lo que pones al principio y al final importa más que lo del medio, y los tres tipos de consulta que necesitas dominar para alinear tu contenido con cada intención.

CAPÍTULO 6

Context windows y tipos de consulta

Las reglas del tablero

La memoria limitada de la IA (y por qué te afecta más de lo que crees)

Hay un factor técnico que determina cuánto de tu contenido lee realmente la IA al responder, y qué partes prioriza. Es la ventana de contexto, y si no lo tienes en cuenta, puedes estar haciendo un trabajo impecable de contenido que el modelo nunca llega a procesar.

La ventana de contexto es la cantidad máxima de texto que un LLM puede manejar a la vez, medida en tokens (fragmentos de palabras que el modelo procesa como unidades). Los primeros modelos generativos solo podían procesar unos pocos miles de tokens a la vez; hoy muchos de los modelos más recientes trabajan con ventanas de un millón de tokens.[1] Suena a mucho, pero siempre hay un límite. Ningún modelo procesa texto infinito.

¿Qué pasa cuando le das al modelo más texto del que puede procesar de una vez? Que tiene que quedarse solo

con una parte. Puede recortar, resumir o seleccionar los fragmentos que considera más útiles. Si publicas una guía larguísima en tu web, un LLM no la va a meter entera en la respuesta. Lo normal es que trabaje con los trozos que vea más relevantes. Por eso, no puedes asumir que va a leer todo tu contenido ni que tratará cada parte con la misma importancia.

Y hay algo peor: no toda la parte leída recibe la misma atención.

Primacía, recencia y el agujero negro del medio

Los modelos tienen un sesgo documentado hacia el principio y el final del texto que procesan. Lo que aparece primero influye más en la respuesta (es el efecto de primacía) y lo que aparece al final también tiene importancia (efecto de recencia).

Lo del medio se puede perder. Literalmente. Se conoce como el efecto *lost in the middle* y no es teoría. Un equipo de investigadores de Stanford, UC Berkeley y Samaya AI lo documentó con modelos como los de OpenAI: su rendimiento era mejor cuando la información relevante estaba al principio o al final de la entrada, y empeoraba cuando estaba en la mitad.[2]



Primacía, recencia y el agujero negro del medio (curva conceptual, sin cifras).
Fuente: elaboración propia a partir de Liu et al., «Lost in the Middle» (TACL, 2024).

Las consecuencias para tu contenido son directas. Puede que la IA solo ingiera fragmentos de tu página, o no toda, por el límite de ventana. Y de lo que ingiera, le hará más caso a lo primero y a lo último. Si tu dato clave está escondido en la línea 300 de un texto largo, hay un riesgo real de que el modelo ni lo note o lo considere menos importante que algo que leyó al principio.

Además, muchas implementaciones de retrieval fragmentan los documentos en bloques pequeños y solo recuperan unos pocos fragmentos por consulta. Eso significa que, aunque tu página tenga muchos párrafos, es posible que solo una parte muy limitada llegue al contexto final del modelo. La pregunta importante es si esos fragmentos contienen realmente lo esencial.

La limitación de contexto se parece, en cierto modo, al antiguo crawl budget del SEO. Antes te preocupaba cuántas páginas podía rastrear e indexar Google. Ahora

también importa cuánto de tu contenido puede llegar a entrar en la respuesta que construye la IA.

Prácticas antiguas que ahora te perjudican

Las introducciones largas llenas de paja son un problema serio. Antes se estilaba meter contexto, anécdotas o definiciones amplias en la intro para «enganchar» o para colocar keywords tempranamente. En IA, cada frase irrelevante al inicio es espacio de primacía desperdiciado. Si tus primeros párrafos no dicen nada útil, el modelo les dará peso y no encontrará nada que valga la pena. Peor aún, pueden empujar fuera del contexto datos más útiles que venían después.

Los pillar pages enormes de cinco mil palabras son otro problema. Quizá en Google funcionaban por la amplitud semántica. En un LLM, esa extensión se recorta. Tu mega guía será dividida en trocitos y el modelo solo usará uno o dos. El resto de la guía jamás se considera. Hubiera sido más efectivo tener cinco guías de mil palabras específicas, donde cada una pudiera entrar entera en contexto para su consulta particular.

La información crítica relegada al final o escondida en elementos desplegables es otro error habitual. Muchas webs dejan lo más valioso para la conclusión o lo colocan dentro de bloques que requieren interacción. En contextos largos, un LLM puede no llegar a recuperar esa parte o darle menos peso, especialmente si el sistema trabaja con fragmentos limitados. Y, si la informa-

ción importante depende de estructuras HTML complejas o poco accesibles, también existe el riesgo de que no se procese con la misma fiabilidad que un bloque de texto claro y visible.

Un punto que nadie mide: el peso en tokens de tu página. Me encuentro con sitios donde una página que visualmente tiene mil palabras, en HTML real tiene 50.000 caracteres por scripts, menús, comentarios de código y JSON incrustados. Un LLM que scrapee eso gasta tokens en basura. Los SEO solían preocuparse del peso en kilobytes para velocidad. Ahora preocúpate del peso en tokens para contexto.

Cómo optimizar para la ventana de contexto

Pon lo más importante al inicio. Parece obvio, pero hay que romper viejos hábitos. Si la pregunta del usuario es «¿Cuál es el mejor CRM para fintech?» y tu artículo es «Los 5 CRM top para fintech», abre con la respuesta resumida. «Para fintech, los CRMs líderes son AcmeCRM y FinTechForce por X razón, aunque DephtiCRM destaca en startups pequeñas.»

Si el LLM lee solo tu intro, ya pescó lo esencial y puede citarlo. Escribe pensando que quizás solo lean tu primer párrafo. Si no dice nada sustancioso, reescríbelo.

Estructura tu contenido con secciones bien delimitadas. Usa subtítulos con sentido claro, listas para enumeraciones, conceptos clave destacados. Cuando las herramientas de chunking procesan una página, detectan seccio-

nes y puntos destacados. Un contenido bien segmentado se trocea en fragmentos lógicos. Páginas con buen formato son más amigables para el scraping y la extracción.

Mantén párrafos y oraciones de longitud razonable. Cuando un bloque es demasiado largo, el modelo tiende a resumirlo y puede dejar fuera detalles importantes. En cambio, si repartes el contenido en varios párrafos breves, cada uno con una idea clara, aumentas las opciones de que la IA conserve al menos uno de ellos con su significado esencial.

Considera dividir los contenidos demasiado extensos en piezas más enfocadas. Muchas veces funciona mejor tener varias páginas especializadas que una sola página enciclopédica. Si la IA necesita información sobre CRM en fintech, le resultará más fácil trabajar con una página centrada en ese tema que con una guía enorme que mezcla demasiadas cosas. En este contexto, dividir bien el contenido no debilita: ayuda a que cada pieza gane claridad y fuerza semántica.

Añade resúmenes intermedios tras secciones largas. Si tienes una explicación técnica profunda de ocho párrafos, incluye al final un «En resumen…» con dos o tres puntos clave. Si el modelo salta los detalles, quizás capture el mini-resumen. Esos resúmenes actúan como anclajes dentro de tu contenido.

Cierra con una conclusión potente. Ya abriste con lo importante, perfecto. Pero también cierra con un remate.

Si tu última frase es «En conclusión, X es la mejor opción para [caso de uso] por [razón]», es altamente probable que si el modelo lee los últimos tokens, agarre exactamente eso. Muchos asistentes priorizan reproducir conclusiones cuando están claramente marcadas.

Conviene que cada sección pueda entenderse incluso si se lee de forma aislada. Por eso, es mejor no abusar de referencias como «como vimos arriba» o «más adelante hablaremos de ello», ya que la IA puede recuperar solo una parte del texto y dejar fuera el resto. Cuanto más autosuficiente sea cada bloque, más fácil será que mantenga su sentido cuando se extraiga o se cite.

Los tres tipos de consulta que debes dominar

No todas las búsquedas son iguales, y esto es más cierto con IA que nunca. Un error frecuente es pretender que la misma estrategia de contenido sirve para todo. La IA entiende la intención del usuario, y tú también necesitas entenderla para alinear tu contenido de forma precisa.

Hay tres tipos macro de consultas, y cada uno requiere un enfoque diferente. Para no medirlo todo con el mismo rasero, en Antropus el GEO Prompt Deriver convierte tus keywords en prompts realistas separados por intención (informativa, comparativa, de decisión, transaccional...), y así ves en qué tipo de consulta apareces y en cuál no.

Consultas de conocimiento

«¿Qué es el algoritmo PageRank?» El usuario busca información, definición y explicación. El LLM actúa como enciclopedia. Recuperará definiciones y datos objetivos de fuentes fiables y generará una respuesta concisa, neutra y factual.

Para destacar en este tipo de consulta, tu contenido debe aspirar a ser una referencia clara, útil y basada en hechos. Conviene ofrecer la definición directa en las primeras líneas y añadir después, de forma concisa, los detalles esenciales: de qué se trata, para qué sirve, cómo funciona o cuál es su origen si resulta relevante. El formato de pregunta y respuesta puede ayudar cuando encaja con la intención de búsqueda. Hay una regla importante: en las consultas puramente informativas, conviene reducir al mínimo el tono comercial. Si un contenido sobre «¿Qué es PageRank?» mezcla la explicación con mensajes promocionales, es más probable que el sistema perciba otra fuente como más neutra y más útil para responder. Primero debe ganar la claridad. Ya habrá otros contenidos pensados para vender.

Si tu producto se apoya en un concepto complejo, te interesa explicar ese concepto mejor que nadie en tu propia web. Cuando alguien pregunte a la IA qué es ese concepto y qué empresas ofrecen soluciones relacionadas, el sistema tenderá a fijarse en las fuentes que mejor lo hayan explicado. Quien demuestra comprensión, claridad y utilidad gana autoridad en esa conversación. Y si

esa referencia no eres tú, probablemente será uno de tus competidores.

Consultas de decisión

«¿Qué CRM tiene mejor integración con Slack?» Aquí el usuario está en modo investigación. No quiere la definición de CRM, quiere saber cuál elegir según un criterio concreto. El LLM asume rol de analista. Buscará en reseñas, artículos comparativos, foros con experiencias y fichas de producto. Te entregará algo sopesado, con pros y contras.

Para destacar aquí, tu contenido debe ser el mejor consejero del tema. Si el usuario busca «el mejor CRM con Slack», tener un artículo titulado «Los 5 CRM con mejor integración con Slack (Comparativa)» es casi obligatorio. Si no lo tienes, estás regalando esa respuesta a quien sí lo tenga.

Ese contenido debe ser honesto y con datos. Evalúa varias opciones (incluidos tus competidores) con criterios claros. Tablas comparativas son especialmente valiosas porque son concisas y la IA puede extraer información de ellas fácilmente. Incluye casos de uso concretos: «En nuestras pruebas, integrar Slack con AcmeCRM tomó 5 minutos y permitió enviar notificaciones de leads a canales; con MegaCRM tardó 15 minutos y requería un plugin de terceros.» Esos detalles van más allá de la página oficial de cada producto y demuestran análisis real. Un LLM preferirá citar a quien tiene insights verificables.

Adopta un tono casi de reviewer independiente: «Nuestra experiencia mostró...» en vez de «Nuestro producto es el mejor». Es un desafío para marketers, pero es lo que funciona. Mantén ese contenido actualizado: las consultas de decisión evolucionan rápido y la IA prefiere información fresca. Y vigílalo: con el Prompt Tracker de Antropus puedes seguir mes a mes si los motores siguen citando tu comparativa o si ya se han ido a la de otro.

Consultas de acción

Son aquellas en las que el usuario no busca solo información, sino completar una tarea. Por ejemplo: «Comprar una licencia de AcmeCRM» o «Reservar un hotel en el centro de Madrid». En estos casos, la IA actúa más como asistente que como buscador. Si puede interactuar con sistemas externos para completar la tarea, intentará hacerlo. Y si no puede, lo habitual es que ofrezca instrucciones claras, opciones concretas o un enlace para continuar.

Para aprovechar este tipo de consultas, conviene crear páginas claramente orientadas a la acción, con llamadas claras, pasos evidentes y URLs fáciles de entender. Si la IA no está conectada a tu servicio para completar esa gestión por sí misma, lo más probable es que ofrezca un enlace al usuario. Por eso, es mejor que encuentre directamente la página donde la acción puede realizarse, y no una portada general desde la que el usuario aún tenga que buscar.

Usa lenguaje accionable en tu sitio: «Regístrate gratis», «Comprar ahora», «Iniciar prueba». He visto a ChatGPT decir «Puedes comprar [producto] en su web oficial» replicando literalmente el texto del botón de la web. Si tu botón dice «Más información», la IA dirá «visita la web para más información» en vez de «compra ahora». Es sutil, pero importa.

Publica guías de «Cómo hacer [acción] en [tu servicio]» con los pasos claros. Un LLM que no pueda ejecutar la acción directamente quizás lea tus instrucciones y las reproduzca como respuesta. Mejor que la IA sepa tu flujo de compra a que no lo sepa y mande al usuario a otro sitio.

La intención híbrida

A veces una misma consulta mezcla decidir y actuar. En una búsqueda como «mejor plataforma CRM para comprar», la IA no solo necesita saber qué opción recomendar, sino también cómo ayudar al usuario a dar el siguiente paso. Por eso debes cubrir ambos frentes: contenido comparativo para justificar por qué eres una buena opción y una página de acción clara para comprar, registrarse o pedir una demo. Si solo ofreces uno de los dos, puedes perder visibilidad o la conversión.

La regla que lo resume todo

La tecnología cambia. La intención del usuario, no tanto. Quien mejor la entienda y la resuelva será quien se

lleve la atención, la confianza y la conversión. En el contexto de la IA, eso exige alinear el contenido y los canales con cada tipo de intención.

Para consultas de conocimiento, conviértete en una fuente fiable dentro de tu nicho. Para consultas de decisión, actúa como el mejor analista: alguien que ayuda a comparar, entender y elegir, y donde tu producto destaca porque realmente ofrece una buena solución. Para consultas de acción, conviértete en la opción más fácil de contratar, reservar o comprar, con llamadas a la acción claras e integraciones que hagan el siguiente paso obvio.

Resumiendo: **para conocimiento, sé la fuente que mejor explica. Para decisión, sé la opción que mejor se justifica. Para acción, sé la opción más fácil de ejecutar.**

	Conocimiento	Decisión	Acción
Ejemplo	«¿Qué es el algoritmo PageRank?»	«¿Qué CRM tiene mejor integración con Slack?»	«Comprar una licencia de AcmeCRM»
Rol de la IA	Enciclopedia	Analista	Asistente
Tu contenido	Definición directa en las primeras líneas, tono neutro, mínimo comercial	Comparativas honestas, con datos, tablas y casos de uso concretos	Páginas de acción, CTA claros, pasos evidentes y URLs fáciles
Sé…	**la fuente que mejor explica**	**la opción que mejor se justifica**	**la opción más fácil de ejecutar**

Intención híbrida: si una consulta mezcla decidir y actuar, cubre ambos frentes.

Conocimiento, decisión y acción: qué hace la IA y qué necesita tu contenido.
Fuente: elaboración propia a partir del texto de este capítulo.

La idea es simple: que el usuario pueda llegar a ti tanto si consulta una IA como si sigue un recorrido más tradicional. Y que, cuando llegue, encuentre justo lo que necesitaba. Ese es el objetivo de la ingeniería de contenidos para LLM.

Con los fundamentos técnicos cubiertos (RAG, retrievability, embeddings, distancias semánticas, ventanas de contexto, tipos de consulta) estamos listos para pasar a la acción. En la Parte III vamos a construir: la arquitectura web que la IA quiere usar como fuente, y la página canónica que se convierte en tu verdad oficial ante los modelos de lenguaje.

NOTAS

1. Google indica que las primeras versiones de los modelos generativos «solo podían procesar 8.000 tokens a la vez» y que muchos modelos Gemini tienen ventanas de 1 millón de tokens o más (Google AI for Developers, «Long context», ai.google.dev/gemini-api/docs/long-context). Anthropic documenta ventanas de 1 millón de tokens en sus modelos Claude más recientes (platform.claude.com/docs/en/build-with-claude/context-windows). Datos consultados en octubre de 2026. ↩
2. Nelson F. Liu et al., «Lost in the Middle: How Language Models Use Long Contexts», *Transactions of the Association for Computational Linguistics*, vol. 12 (2024), pp. 157–173. Preprint: arxiv.org/abs/2307.03172 (julio de 2023). ↩

PARTE III

Arquitectura web para la era IA

Cómo construir un sitio que la IA quiera usar como fuente.

CAPÍTULO 7

RAG avanzado

Recuperabilidad, ranking interno y citación

Cómo decide la IA a quién escuchar

En los capítulos anteriores entendimos qué es RAG y por qué la recuperabilidad es la nueva métrica que importa. Ahora vamos a ir un nivel más profundo. Porque no basta con saber que la IA busca fuentes antes de responder. Necesitas entender exactamente cómo filtra, qué prioriza y por qué tu contenido puede ser ignorado aunque sea técnicamente bueno.

Un motor LLM con RAG funciona más como un filtro semántico que como Google. Convierte todo a vectores, compara significados en vez de keywords y busca documentos cercanos en el espacio semántico. En paralelo, puede hacer una búsqueda clásica de palabras para no perder datos (nombres propios, códigos, cifras exactas) y fusionar ambos resultados. Luego viene un reordenamiento interno donde el sistema evalúa cada candidato con modelos más complejos, comparando contenido con pregunta, para elegir solo unos pocos fragmentos que usará para generar la respuesta.

Lo que suele priorizar este filtrado interno es lo que los investigadores llaman information gain: documentos que aportan información nueva, específica o complementaria a la respuesta que se está construyendo. Si tu página añade un dato, un matiz o una perspectiva que no aparece en otras fuentes, tienes más opciones de ser tenida en cuenta. Si repite lo mismo que otras diez, pierde valor.

En un sistema RAG, la fuente que se elige no es necesariamente la más famosa, sino la que mejor responde a la pregunta y aporta un valor diferencial. La autoridad de dominio clásica puede influir, pero de forma indirecta. La frescura de la información es clave cuando la consulta lo requiere: si el usuario pregunta por algo reciente y el sistema lo detecta, buscará fuentes actualizadas para construir la respuesta.

Si quieres entrar en la selección, tu contenido no solo debe estar bien escrito. Tiene que ser semánticamente relevante, aportar información propia, tener una estructura que la máquina pueda interpretar bien y estar al día cuando la consulta lo requiera. Aquí la visibilidad no depende tanto del clic humano como de que el sistema te recupere como una fuente útil. En muchos casos, eso pesa más que las señales clásicas del SEO, como el CTR o la autoridad de dominio.

Los cinco errores que hacen que la IA te ignore

El primer gran error a tener en cuenta es que el contenido genérico penaliza más en entornos de IA. Si tu página repite lo mismo que muchas otras y no añade valor real, el sistema la percibe como redundante. Puedes seguir posicionando en Google, pero ser irrelevante para un LLM.

La estructura confusa es otro error habitual: introducciones demasiado largas, párrafos interminables y un uso forzado de palabras clave que no aportan nada. Todo eso dificulta que el sistema entienda con claridad de qué trata la página. Si la respuesta que busca el usuario aparece escondida al final, mezclada con mucho texto irrelevante, tendrás menos opciones de ser recuperado.

El tercer error es repartir la información importante entre demasiadas páginas. Un concepto aparece en un post viejo, otro en un PDF y otro en una página secundaria. El problema no es solo que exista, sino que esté fragmentado. Si el sistema no conecta bien esos elementos, ninguna de tus URLs destacará lo bastante. Por eso suele tener ventaja quien ofrece una respuesta más unificada y fácil de recuperar.

El cuarto error es no actualizar el contenido. Una página puede haber sido excelente hace unos años y, aun así, dejar de servir si la consulta exige información reciente. Cuando la IA detecta que el contenido está desfasado, suele darle menos valor o simplemente lo ignora.

El quinto problema es ponerle obstáculos técnicos al crawler. Si tu web depende de estructuras complejas, esconde contenido tras elementos interactivos o bloquea documentos importantes, la IA puede no llegar a ver lo que has publicado. Y si no lo ve, no puede recuperarlo ni usarlo.

Páginas fuente vs. posts vs. landings: la arquitectura que frena tu visibilidad

Otro mito que conviene desmontar: la arquitectura web que durante años funcionó en SEO puede estar frenando tu visibilidad en IA. En el SEO clásico era habitual crear muchos posts para atacar keywords concretas y landings comerciales para convertir. En un entorno dominado por LLMs, esa misma estructura puede convertirse en un laberinto: mucha información, demasiadas rutas y ninguna página claramente reconocible como referencia principal.

Cuando el contenido esencial de una marca está disperso, escondido o fragmentado entre muchas URLs, la IA lo tiene más difícil para recuperarlo y utilizarlo con confianza. Puedes tener decenas o cientos de posts posicionados en Google, pero si ninguno actúa de forma inequívoca como la página de referencia sobre ese tema, el modelo puede terminar apoyándose en otra web mejor organizada y más fácil de interpretar.

En términos de recuperabilidad, muchas veces menos es más. Para un LLM suele aportar más valor una sola página fuente bien construida que cien piezas compi-

tiendo entre sí por explicar lo mismo. No se trata de publicar menos por sistema, sino de concentrar el conocimiento clave allí donde el modelo pueda identificarlo, entenderlo y priorizarlo sin fricción.

Una página fuente es la página central de conocimiento sobre un tema, producto, servicio o entidad dentro de tu sitio web. Es el lugar donde se reúne, de forma ordenada y accesible, la definición, el contexto, los datos clave, las preguntas frecuentes, los detalles técnicos y, en general, todo lo necesario para comprender ese asunto en profundidad. Es, en esencia, la URL que tú mismo enviarías a alguien si te pidiera «explícame esto bien y en un solo sitio».

Un post de blog suele tener otra función. Normalmente es más narrativo, más contextual y más ligado a un momento, una opinión o un enfoque concreto. Eso lo hace muy útil para atraer tráfico humano, trabajar descubrimiento o desarrollar ideas específicas. Pero cuando un LLM necesita responder de forma directa, suele preferir la fuente más clara, completa y autoritativa, no necesariamente la más interesante de leer.

Una landing comercial responde a una lógica distinta: está diseñada para persuadir y convertir. Por eso suele ser más visual, más breve y más centrada en mensajes de marketing, llamadas a la acción y formularios. El problema es que, para una IA que busca resolver una consulta informativa, ese tipo de página a menudo ofrece poca sustancia: promete mucho, pero explica poco. Y

si explica poco, cuesta más que el modelo la considere una fuente útil.



Página fuente, post y landing: tres funciones distintas. Fuente: elaboración propia a partir de este capítulo.

Este patrón se repite más de lo que parece. He visto líderes de mercado prácticamente invisibles en respuestas de IA mientras proyectos open source con documentación pública acumulaban la mayoría de las citas. La diferencia no era necesariamente la calidad del producto, sino la accesibilidad y la estructura del conocimiento. Unos tenían la información repartida, parcialmente cerrada o escondida tras login. Los otros ofrecían páginas fuente claras, públicas, bien organizadas y fáciles de recuperar.

La IA, cuando necesita construir una respuesta fiable, suele inclinarse hacia quien le reduce el trabajo. No «premia» a la marca más grande ni a la web más bonita, sino a la fuente que mejor concentra, organiza y expone

la información necesaria para responder. En este contexto, la arquitectura deja de ser solo una cuestión de SEO o UX. Pasa a ser también una cuestión de legibilidad para máquinas.

Cómo reestructurar tu sitio pensando en recuperabilidad

La primera acción consiste en crear páginas fuente canónicas para tus temas críticos. Cada tema, producto, servicio o entidad clave en tu estrategia necesita una página central que actúe como la referencia oficial dentro de tu web. Es ahí donde debería concentrarse la definición, el contexto, los datos esenciales y las respuestas principales sobre ese asunto. Tu marca, cada producto principal y los conceptos fundamentales de tu sector, cuando tengan peso estratégico, deberían contar con su propia página fuente.

La segunda acción es consolidar contenido fragmentado. Si tienes muchos posts pequeños sobre lo mismo, considera fusionarlos en una sola guía. La canibalización no solo confunde a Google, confunde a la IA también. Es preferible un pilar extenso y completo que diez artículos flojos dispersos.

La tercera acción es interconectar tu contenido de forma inteligente. Los artículos del blog no deberían quedar aislados: deben reforzar a las páginas fuente mediante enlaces internos claros. A su vez, las páginas fuente deben apuntar a los mejores contenidos de apoyo. Así construyes un clúster temático sólido, en el que

la jerarquía resulta evidente también para un crawler semántico: una URL principal que organiza el conocimiento y varias piezas secundarias que lo expanden.

La cuarta acción es enlazar las landings comerciales con contenido informativo. Con frecuencia, estas páginas quedan desconectadas del resto del conocimiento del sitio y se limitan a vender. Cuando incorporas bloques explicativos, mini FAQs con preguntas clave y enlaces hacia guías o páginas fuente, la landing gana profundidad. Deja de ser solo un soporte comercial y pasa a convertirse en una pieza útil también para la recuperación semántica y la comprensión del tema por parte de la IA.

Y la quinta es usar datos estructurados e identificadores: Schema markup, identificadores únicos, sameAs para conectar con tus perfiles oficiales. Es necesario todo lo que ayude a la IA a entender sin ambigüedad quién eres y qué ofreces.

Por qué la IA te usa pero no te cita (y cómo cambiar eso)

Posicionar bien en Google no garantiza que una IA te cite. Puedes ser primero en SEO y, sin embargo, quedar fuera mientras ChatGPT, Perplexity u otro motor generativo recomiendan a un competidor. Visibilidad en buscadores y visibilidad en respuestas de IA no son exactamente lo mismo. Para ver ese desfase en Google, en Antropus usamos AI Visibility: te dice en qué búsquedas el AI Overview enlaza a tu web, con qué página y frente a qué competidores.

La razón es simple y bastante fría: si la IA no te menciona, es porque no te necesita para responder. Puede ocurrir por dos motivos principales: o tu contenido no le aporta nada diferencial frente a otras fuentes, o tu sitio no transmite suficiente claridad y fiabilidad como para ser utilizado como referencia.

Cuando un LLM genera una respuesta con citas, no enlaza fuentes al azar. Después de la fase de retrieval dispone de un conjunto de documentos potencialmente relevantes, pero en la redacción final solo cita aquellos que realmente aportan una pieza útil para construir la respuesta. El objetivo no es mencionar el mayor número de fuentes posible, sino combinar las que mejor se complementan entre sí. Por eso, si tu contenido repite lo mismo que otros sin añadir un matiz, un dato o una formulación más clara, puede quedar fuera por aburrimiento. El sistema tenderá a escoger una sola fuente para esa idea: la que considere mejor expresada, más fiable o más útil para sostener la respuesta.

Estar recuperado no es igual a ser citado. Podrías haber pasado el primer filtro semántico, pero en la fase final el modelo encontró que no tenías una frase directamente citable, o que el dato clave lo tenía mejor expresado otra fuente.

Además, los patrones de citación varían por temática. En salud y medicina, la IA cita fuentes de alta confianza: instituciones, publicaciones científicas, webs oficiales; en tecnología y programación, mezcla documentación oficial con foros como Stack Overflow; en e-commerce,

suele citar blogs de reviews independientes y comparadores, ignorando a los propios e-commerce que solo tienen fichas de producto; y en finanzas y en el ámbito legal, predominan las fuentes oficiales e institucionales.

El patrón común es claro: la IA reparte las citas con una lógica fría. No cita por simpatía ni por notoriedad superficial, sino por confianza, especialización y capacidad de cubrir una parte concreta de la respuesta. En temas sensibles, tiende a apoyarse en las fuentes más fiables. Cuando varias dicen lo mismo, suele quedarse con la más consolidada. Y cuando necesita una pieza específica, prioriza a quien se la ofrece de forma más clara y útil.

Seis tácticas para que la IA te cite

La primera táctica es aportar datos originales. Si en tu sector nadie más ha publicado un dato concreto y tú sí, aumentan mucho las probabilidades de que te conviertas en la fuente que el modelo necesita citar. Encuestas anuales, estudios de mercado propios, benchmarks internos o análisis con datos exclusivos. Antes este tipo de activos servía sobre todo para generar PR y enlaces. Ahora puede convertirte en una fuente especialmente valiosa para sistemas de IA.

La segunda táctica es convertirte en el experto profundo de tu nicho. No basta con publicar contenido generalista o superficial. Las oportunidades de cita aumentan cuando desarrollas guías con rigor técnico y respondes tam-

bién a las preguntas menos obvias, más específicas o más avanzadas. Cuando aparece una consulta compleja y pocas fuentes la resuelven bien, la probabilidad de que la IA recurra a tu página crece de forma natural.

La tercera táctica es redactar contenido fácil de citar. No basta con tener la respuesta. Hay que formularla de manera que el modelo pueda identificarla, aislarla y reutilizarla con facilidad. Definiciones breves, frases autocontenidas para las ideas clave y estructuras claras, como listas numeradas. Si la respuesta existe, pero aparece diluida en un bloque de texto denso, la IA tenderá a preferir otra fuente que la exprese de forma más limpia.

La cuarta es ajustar el tono al nivel de confianza esperado. En salud y finanzas: tono objetivo, documentado y con referencias bibliográficas. En tecnología: tono práctico con ejemplos y código.

Observa qué estilo de contenido está citando la IA en tu campo, y alinea la presentación de tu contenido a esos patrones.

La quinta es observar y aprender de las citas actuales. Monitoriza qué está citando la IA en tu sector. Usa Copilot, Perplexity, pregunta cosas relevantes y fíjate a quién citan y cómo. Identifica huecos: si nadie da un ejemplo práctico para cierta pregunta, ahí tienes tu oportunidad. Si quieres hacerlo de forma sistemática, el Prompt Tracker de Antropus lanza tus preguntas a ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview y re-

gistra, motor a motor, si te mencionan, te recomiendan y te citan.

Y la sexta es marcar y metadatear tus contenidos importantes: Títulos de página que reflejen preguntas, meta descriptions que resuman los puntos clave, schema FAQ y HowTo, y un TL;DR al inicio con los puntos principales. Es decir, todo lo que ayude al modelo a entender en dos segundos que tu página tiene lo que busca.

Diseño RAG-friendly: la estructura que la IA necesita

Tu web puede ser visualmente preciosa, pero si no es RAG-friendly, para la IA es ruido. Una página RAG-friendly está diseñada para que los pasos de retrieval y generación del LLM funcionen sin fricciones.

También conviene que el contenido pueda fragmentarse en unidades autónomas. Una idea por párrafo facilita que cada bloque tenga sentido por sí mismo y pueda competir mejor en procesos de recuperación semántica (HTML limpio y semántico, con etiquetas correctas, jerarquía clara de encabezados, tablas cuando los datos son tabulares y sin ocultar contenido esencial dentro de scripts). La escaneabilidad, además, debería reforzar la parseabilidad: títulos descriptivos, idea principal al inicio de cada párrafo y uso moderado de negritas en los conceptos clave.

Un TL;DR al inicio con bullet points es especialmente valioso. Para humanos resume el contenido; para la IA funciona como un mini mapa que le permite detectar rápidamente si tu página tiene la respuesta que busca.

Formato pregunta-respuesta para FAQs presentado explícitamente: la pregunta como encabezado o en negrita y la respuesta justo después. Si la consulta del usuario coincide con esa pregunta textual, el match es casi perfecto.

Respuesta directa antes que contexto: Si la sección se titula «¿Qué es SEO On-Page?», la primera frase debe ser la definición. No empieces con rodeos históricos. Primero responde, luego expande. Así, si la IA busca la definición, tu primera frase se lleva la cita.

Los datos y las cifras deben aparecer acompañados de su referencia dentro del propio texto. Cada número, porcentaje o fecha puede convertirse en una pieza especialmente valiosa si coincide con lo que el usuario está buscando. Además, indicar la fuente junto al dato refuerza la percepción de rigor y ayuda a que tu contenido resulte más fiable tanto para lectores humanos como para sistemas de IA.

El lenguaje debe ser natural y sencillo. Evita las frases demasiado largas y las estructuras rebuscadas que dificultan la comprensión. Piensa en cómo formularía esa pregunta una persona real y acércate a ese mismo registro al responder. Una buena prueba es leer el texto en voz alta: si suena como una explicación clara dirigida a alguien inteligente, funciona. Si parece una tesis o un texto publicitario inflado, pierde fuerza.

En el siguiente capítulo vamos a construir la pieza más estratégica de toda tu arquitectura web: la Source of

Truth. La página canónica que reúne toda la verdad sobre tu marca o tu tema, diseñada para que la IA prácticamente no tenga más remedio que acudir a ti.

CAPÍTULO 8

Tu «Source of Truth»

La página canónica que la IA necesita

O escribes tu verdad, o la IA la inventará por ti

Si tu sitio no cuenta con una página canónica que concentre la información esencial sobre cada tema o entidad clave de tu negocio, estás dejando espacio para que otros definan el relato. O peor aún: que la IA construya respuestas sobre tu marca a partir de fragmentos dispersos, incompletos o directamente erróneos.

La consigna es clara: concentra la información clave en un solo lugar y conviértela en tu referencia más sólida. Piensa en esa página como tu referencia central, casi como una Wikipedia propia. Si no redactas tú la versión más completa y fiable sobre tu marca o tu especialidad, otros lo harán por ti. Y la IA construirá sus respuestas con esas piezas ajenas, no con el relato que mejor te representa.

Qué es una Source of Truth interna

Es una página o un pequeño conjunto de páginas en tu dominio que funciona como referencia canónica y auto-

rizada sobre un tema determinado. Puede ser tu empresa, tu producto estrella, o un concepto de tu industria en el que quieres establecerte como autoridad.

Imagina una página que anticipa cualquier pregunta que un usuario o la IA pudiera tener y la responde ahí mismo, con datos verificables, actualizados y bien organizados. Esa es tu Source of Truth.

Los LLM tienden a confiar más en el conocimiento coherente que en la información dispersa. Cuando encuentran una página completa que reúne lo esencial sobre un tema, les resulta más fácil interpretarla como una fuente sólida y suficiente. Pero si esa misma información aparece fragmentada en varias URLs, puede que no la ensamblen bien y prefieran recurrir a otra web que ya la ofrece unificada, ordenada y lista para usar.

Tener una Source of Truth interna es como decirle al algoritmo: aquí está la verdad definitiva sobre este tema. No necesitas buscar en otro lado.

Qué debe incluir

Debe incluir todo lo relevante para entender el tema de forma completa. La definición formal. El contexto y la evolución. En el caso de una empresa, su año de fundación y sus hitos principales; en el de un concepto, su origen y desarrollo. También los datos y cifras oficiales que aporten sustancia, como número de clientes, rendimiento, cuota de mercado o reconocimientos. A eso conviene sumar descripciones detalladas de productos o componentes, preguntas frecuentes, comparativas úti-

les con alternativas y enlaces a fuentes externas fiables que validen la información. Y, cuando el tema lo pida, otras capas necesarias para cubrir todas sus aristas: casos de uso, ventajas, limitaciones conocidas y visión a futuro.

Debe ser la página que un tercero citaría si tuviera que explicar ese tema. Cuanto más neutro, informativo y apoyado en hechos concretos sea su tono, mejor. No es un panfleto de marketing. Piensa más bien en la estructura y el tono de un buen artículo de Wikipedia: claro, ordenado, útil y fácil de citar, pero escrito por ti sobre tu propia entidad.

Es verdad, exige un esfuerzo considerable. Ese esfuerzo tiene una lógica clara: en lugar de dejar que la IA construya respuestas sobre tu marca con datos ajenos, le ofreces una fuente completa basada en tu propia información. Además, una página de este tipo puede generar un efecto secundario muy valioso: convertirse también en referencia para otros sitios. Si un medio especializado enlaza tu página «Todo sobre [Producto X]» en lugar de redactar su propia explicación, aumenta la probabilidad de que esa URL gane autoridad y sea percibida como una fuente fiable también por la IA.

Cómo construirla paso a paso

Paso 1: Inventario de hechos y preguntas

Reúne toda la información verificable sobre ese tema en un solo lugar. En la práctica, se trata de construir una

base de conocimiento propia y fiable. Si el tema es tu empresa, esa página debería incluir fechas clave, cifras actualizadas, definiciones oficiales, funcionalidades detalladas de tus productos o servicios, limitaciones conocidas, preguntas frecuentes de clientes con sus respuestas oficiales y comparativas con competidores siempre que estén apoyadas en datos medibles.

No escatimes. Cada hecho en una fila. Año de fundación, número de clientes a fecha, tiempo medio de implementación, principales sectores que atiendes, premios, certificaciones, preguntas repetidas con sus respuestas.

Si hay algo incómodo pero público (brecha de seguridad en el pasado, una limitación conocida de tu producto...) es mejor que lo menciones tú con contexto y solución, a que la IA lo cuente sacado de un blog sensacionalista. Mejor que lo sepan de ti que de terceros.

Paso 2: Crear la página canónica

Con esa materia prima, llega el momento de construir la página. Empieza con un título claro, que identifique sin ambigüedad el tema: «Acerca de [Empresa X]» o «Todo sobre [Producto Y]». Añade al inicio un resumen ejecutivo, en uno o dos párrafos o en formato de viñetas, que explique de forma rápida quién eres y cuáles son los datos más destacados. A partir de ahí, organiza el contenido en secciones lógicas con encabezados descriptivos, como Historia, Productos y servicios, Tecnología, Datos clave, Reconocimientos, Comparativa con alternativas, Preguntas frecuentes o Referencias.

En cada sección, prioriza la precisión y la verificabilidad. Cada afirmación que pueda sonar promocional debería apoyarse en un dato, una fuente o una evidencia concreta. Es preferible escribir «reconocida como líder según [Informe independiente 2026]» que limitarse a afirmar «somos líderes». Y si una idea no puede demostrarse, conviene reformularla en un tono más neutral y descriptivo.

También conviene incluir aquello que, a primera vista, podría parecer negativo, pero explicado con contexto. Una página que solo enumera virtudes suena menos creíble que una que reconoce límites, trade-offs o decisiones de enfoque. Por ejemplo: «Limitaciones: por diseño, nuestro producto no realiza [función], porque prioriza la especialización en [otra capacidad]». O en una FAQ: «¿Por qué es más caro que otros? Porque incluye soporte 24/7, certificación X y garantía Y». Al hacerlo, te adelantas a objeciones legítimas y evitas que la IA tenga que completar esa parte de la respuesta con información tomada de otros sitios.

No temas la extensión, siempre que esté justificada y bien estructurada. Si antes describías tu producto en 500 palabras de marketing, aquí podrías necesitar varios miles para cubrirlo con rigor. Eso no es un problema en sí mismo. Lo importante es que la información esté organizada de forma impecable: índice al inicio, encabezados claros, párrafos breves, listas y tablas cuando ayuden a segmentar. La IA puede aprovechar conteni-

dos extensos, pero solo si encuentra rápido las piezas relevantes.

Paso 3: Diseño orientado a verificación

Cierra la página con una sección de fuentes externas que valide lo que has explicado. Ahí puedes enlazar documentos de terceros que aporten respaldo independiente: la nota de prensa que anunció una ronda de financiación, un artículo en un medio reconocido, un informe de analistas que te menciona o cualquier otra referencia fiable. Acompaña esa estructura con un índice de contenidos enlazado mediante anclajes internos. Cuanto más fácil sea saltar directamente a la sección relevante, más sencilla será también la navegación, la verificación y la recuperación de la información.

Indica fecha de última actualización, autor y cargo. Esto añade transparencia y señales de confianza que tanto humanos como modelos de IA valoran.

Paso 4: Mantén la página viva

Crear la página no es suficiente. Su valor depende de que siga siendo actual. En cuanto los datos cambian, la referencia pierde solidez. Por eso, cualquier modificación importante debería incorporarse lo antes posible: una nueva versión del producto, un ajuste de precios, una actualización de funcionalidades o un nuevo hito de la empresa. Una página fuente solo cumple su función si sigue siendo, de verdad, la versión más fiable y vigente de ese tema.

Además de reaccionar a los cambios, conviene fijar una revisión periódica. Cada tres o seis meses, aunque no parezca haber novedades, es útil revisar la página para detectar datos desactualizados, secciones incompletas o referencias que ya no encajan. Una página que muestra señales de desactualización pierde fuerza como fuente. Y esa pérdida de credibilidad afecta tanto a la IA como a cualquier lector con criterio.

Paso 5: Promueve e interlinkea

No basta con crearla. También tienes que hacerla visible. Enlázala desde los lugares estratégicos de tu web: la página principal, la sección de recursos, el menú o cualquier bloque donde tenga sentido como referencia. Cada vez que en el blog menciones un dato clave, apunta hacia esa URL canónica. Inclúyela también en notas de prensa y asegúrate de que esté presente en el sitemap. Una página fuente solo cumple bien su función si el resto del sitio la reconoce y la refuerza como referencia central.

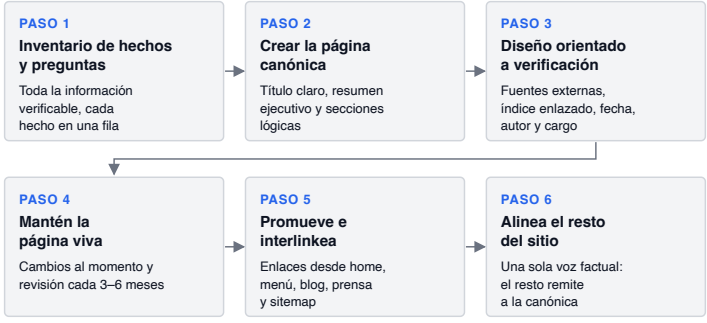
El objetivo de todo esto es claro: que esa página sea fácil de descubrir, quede inequívocamente asociada a tu marca y termine consolidándose como la fuente oficial de referencia.

Paso 6: Alinea el resto del sitio

Esto es crítico: tu página canónica debe convertirse también en la referencia interna. Si antes una ficha de producto incluía cuatro líneas sobre la empresa, ahora debería enlazar a esa URL en lugar de repetirlas. Reprodu-

cir el mismo dato exacto en varias páginas parece inocente, pero introduce un riesgo serio: cualquier discrepancia debilita la confianza. Por eso conviene centralizar la información precisa en la página canónica y dejar que el resto del sitio la resuma, la contextualice o remita a ella cuando sea necesario consultar el dato exacto.

Haz auditoría de contenido existente: reemplaza números obsoletos, quita secciones duplicadas y añade enlaces a la canónica donde haga falta. Toda tu web debe hablar con una sola voz factual. Nada destruye más rápido la confianza de la IA que tu propio sitio contradiciéndose.



Los seis pasos de la Source of Truth. Fuente: elaboración propia a partir de este capítulo.

La auditoría de entidad: sin coherencia, la Source of Truth no sirve

Antes de dar por buena tu Source of Truth, necesitas asegurarte de que tu identidad digital es coherente en

todas partes. No solo en tu web, sino en todo el ecosistema donde la IA busca información sobre ti.

Esto es lo que llamamos auditoría de entidad: verificar que tu marca, tu categoría, tus atributos y tus relaciones están representados de forma consistente en todas las fuentes que la IA consulta.

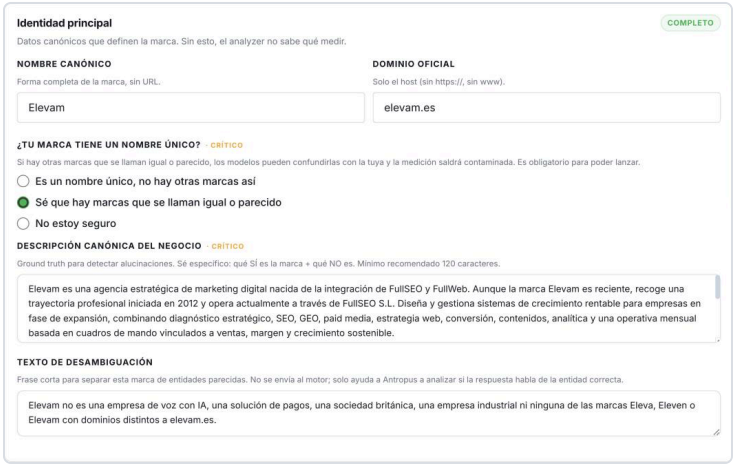
¿Tu nombre aparece igual en tu web, en Wikipedia, en Wikidata, en tu perfil de Google Business, en LinkedIn y en los directorios de tu sector? ¿La descripción de tu empresa dice lo mismo en todos lados, o en uno eres «consultora de marketing» y en otro «agencia de comunicación digital»? ¿Los datos numéricos coinciden?

Cada inconsistencia es una señal de ruido para la IA. Si encuentra información contradictoria sobre ti en distintas fuentes, puede optar por no citarte para no arriesgarse. O peor: puede mezclar datos de distintas fuentes y construir una versión de ti que no es correcta.

La auditoría de entidad cubre cuatro dimensiones. La primera es la consistencia de marca: que tu nombre, descripción y datos básicos sean idénticos en todas partes. La segunda es la categoría: que la IA entienda sin ambigüedad a qué te dedicas y en qué sector operas.

La tercera son los atributos: que las características que te definen (tamaño, ubicación, especialización, logros) estén actualizados y sean coherentes. Y la cuarta son las relaciones: que los vínculos entre tu marca y otras entidades (socios, clientes, competidores, premios) estén correctamente representados.

En Antropus, la Ficha de Entidad recoge esa versión canónica (nombre, descripción, texto de desambiguación, alias y perfiles oficiales enlazados como sameAs) y genera a partir de ella el JSON-LD de schema.org. El resto de herramientas la reutiliza para no confundirte con un homónimo cuando mide.



Antropus · Identidad de marca: la «source of truth» con la que se mide. Proyecto real de Elevam (elevam.es).

Sin esta coherencia de base, tu Source of Truth estará desconectada de lo que el resto de internet dice sobre ti. Y la IA siempre busca consenso.

Las limitaciones que debes aceptar

No puedes controlar por completo la versión que la IA utilizará. Si hay otras fuentes reputadas que ofrecen un dato distinto al tuyo, el modelo puede combinarlas o inclinarse por la versión que aparece con más frecuencia

y consistencia. Imagina que tu página indica «500 clientes en 2026», pero Wikipedia sigue mostrando «300 clientes en 2023» y ese dato ha sido repetido por decenas de blogs y directorios. En ese escenario, la IA puede seguir respondiendo «300», no porque tu cifra sea falsa, sino porque la señal dominante todavía favorece ese dato anterior. Hasta que la información nueva se difunda y gane consenso, el modelo puede seguir anclado en la versión más repetida. La única forma de saber cuándo cambia es preguntar con método: con el Prompt Tracker de Antropus lanzas las mismas preguntas sobre tu marca cada cierto tiempo y comparas lo que responde cada motor con la medición anterior, hasta ver que el dato nuevo se impone.

Wikipedia y otras fuentes externas fuertes no siempre se pueden desplazar con facilidad. En muchos casos, tu Source of Truth no sustituirá de inmediato a esas referencias, sino que convivirá con ellas y las complementará. Si tu dominio es nuevo o todavía tiene poca credibilidad acumulada, necesitará tiempo para ganar tracción. Además, existe otra limitación importante: la IA puede incorporar tu información a sus respuestas sin citarte de forma explícita. Es decir, puede apoyarse en tus datos y, aun así, no considerarte necesario como fuente visible de atribución.

Incluso con estas limitaciones, disponer de una Source of Truth te devuelve una posición mucho más fuerte. Como mínimo, te aseguras de que, cuando la IA necesite comprobar algo sobre ti, exista un lugar claro y fiable

donde encontrar la respuesta correcta. Y eso ya marca una diferencia enorme frente al escenario contrario: que el modelo tenga que reconstruir tu identidad, tu oferta o tus datos a partir de fragmentos dispersos, desactualizados o directamente incorrectos.

La buena noticia es que la mayoría de tus competidores aún no se ha adaptado. No crearán páginas canónicas, no trabajarán su visibilidad en IA y seguirán aplicando lógicas de SEO heredadas como si nada hubiera cambiado. Solo reaccionarán cuando empiecen a perder presencia, tráfico o capacidad de recomendación frente a otros actores mejor preparados. Si tú haces este trabajo, no solo corriges una debilidad: construyes una ventaja. En la nueva era de la información generativa llegar antes importa mucho.

Con la arquitectura web y la Source of Truth construidas, estamos listos para el siguiente paso: medir. Porque lo que no mides, no existe. En la Parte IV vamos a montar el sistema completo para saber si la IA te ve, te cita o te ignora con métricas reales, no con suposiciones.

PARTE IV

Medición: lo que no mides, no existe

El sistema completo para saber si la IA te ve, te cita o te ignora.

CAPÍTULO 9

Las cinco métricas que importan

Medir visibilidad en LLMs sin supersticiones

El tablero de mandos que no tienes (y necesitas)

Puede que la IA esté hablando de tu marca ahora mismo y no lo sepas. Puede que la esté destrozando en silencio. Mientras tú miras Google Analytics buscando clics y conversiones, un modelo de lenguaje le está diciendo a cientos de potenciales clientes qué comprar, a quién contratar y a quién evitar. Y tú ni siquiera lo estás viendo.

Se acabaron las excusas. En este capítulo vamos a montar el sistema de medición que te permite auditar tu visibilidad en IA con la misma frialdad analítica con la que auditas tus rankings y tus logs de servidor. Los números ya no son solo visitas y CTR. Ahora son menciones, posiciones medias, porcentajes de citación, asociaciones contextuales y puntuaciones de sentimiento.

Son cinco dimensiones. Cada una mide algo diferente. Y las cinco juntas te dan la fotografía completa de cómo te ve la IA.

Dimensión 1: Share of Model — ¿Qué porcentaje de la voz es tuyo?

Share of Model es la métrica reina. Indica qué porcentaje de las respuestas de un modelo sobre tu sector incluye tu marca. Si le preguntas a ChatGPT cien veces por «mejores plataformas de ecommerce» con variaciones, y tu marca aparece en treinta de esas respuestas, tu SoM es 30%.

Ya no compites en una página de resultados con diez enlaces azules. La IA no te rankea con una posición numérica visible. Simplemente, o apareces en la respuesta o no. Y cuando apareces, cada mención cuenta.

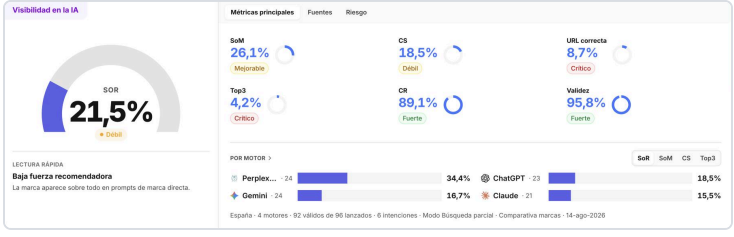
Si tu SoM es bajo, es que la IA apenas te menciona cuando los usuarios preguntan por temas de tu sector. Puedes tener la mejor web del mundo, pero si los modelos generativos no te nombran, para buena parte de tu mercado simplemente no existes. Y si no existes en esa primera respuesta, difícilmente te van a buscar después.

Para medirlo necesitas montar un pequeño sistema. Define un conjunto de prompts relevantes para tu nicho, es decir, consultas representativas de cómo la gente podría encontrarte a través de la IA. Lanza esos prompts al modelo varias veces, variando la formulación. Extrae las entidades mencionadas en cada respuesta y registra los resultados de forma ordenada: el prompt, el modelo, la fecha, la hora y las marcas mencionadas.

Con suficientes datos, calcula porcentajes. Observa en qué porcentaje de las respuestas apareces tú y en qué

porcentaje aparecen tus competidores. Así obtienes tu Share of Model y una referencia clara para compararte con el resto del mercado. Es exactamente lo que automatizamos en Antropus con Competitor Intelligence: el mismo set de prompts lanzado a ChatGPT, Claude, Gemini y Perplexity, con tu Share of Model, tu posición media y el tono frente a cada competidor, comprobando antes de contar una mención que de verdad sea tu marca y no un homónimo.

Un apunte importante sobre interpretación: no entres en pánico si tu SoM baja de un día para otro con pocas muestras. Puede ser ruido estadístico. La IA es probabilística; tus datos también tienen que serlo. Analiza tendencias con suficiente muestra, no resultados de un solo día.



Antropus · Prompt Tracker, visión general: fuerza de recomendación (SoR), menciones (SoM), citación del dominio (CS) y reparto por motor. Proyecto real de Elevam (elevam.es), baseline del 14 de agosto de 2026: 92 respuestas válidas de 96.

Dimensión 2: Shortlist Position — No basta con estar; importa dónde te ponen

Cuando la IA responde a una pregunta tipo «¿Cuáles son los mejores X?», genera internamente una lista de opcio-

nes. A veces la presenta numerada, a veces en prosa. Tu posición en esa lista importa tanto como el hecho de estar en ella.

Si siempre apareces al final, tu visibilidad es casi nula. En una respuesta con cinco opciones, salir quinto es ser invisible para efectos prácticos. La IA tiende a mencionar primero lo que considera más relevante o confiable. El primer nombrado es al que otorga mayor peso. Los siguientes son complementarios.

Para rastrear esto, añade una capa a tu sistema de logging: además de registrar si apareces, registra en qué posición. Si la respuesta enumera cinco opciones y tú no saliste, asígnate un «sexto» para reflejar que quedaste fuera. Con suficientes datos, calcula tu posición media cuando apareces, y tu posición media global incluyendo las veces que no apareciste.

Esa radiografía te dice mucho. Aunque tu SoM sea decente, si casi nunca eres la primera opción que la IA sugiere, estás siendo el farolillo de cola. Y eso tiene causas identificables: contenido insuficiente que no da al modelo razones para ponerte primero, menor confianza percibida frente a competidores más establecidos, o simplemente falta de datos diferenciadores que te distingan.

Dimensión 3: Cobertura de citas — ¿Apareces como fuente o solo nombrado de pasada?

En la era del SEO clásico nos desvivíamos por conseguir backlinks. En la era LLM, la lucha es por conseguir lo que podríamos llamar «frontlinks»: que tu contenido sea citado directamente por el modelo como referencia en sus respuestas.

De nada sirve que la IA mencione tu marca si no muestra ni usa tus páginas como fuente.

Cuando la IA responde al usuario y al final incluye referencias numeradas (como hacen Copilot, Perplexity o los AI Overviews de Google), la pregunta clave es: ¿aparece alguna URL tuya entre esas referencias?

Si la IA habla de ti pero la cita es un artículo de un tercero que te menciona de pasada, el usuario curioso que haga clic terminará en el sitio del tercero, no en el tuyo. Estás regalando autoridad. La IA valida tu existencia con el contenido de otro.

La cobertura de citas mide el porcentaje de respuestas donde tus propias páginas aparecen como fuente referenciada. Para medirla, en tu sistema de registro incluye las URLs citadas en cada respuesta y marca si alguna pertenece a tu dominio. En el Prompt Tracker de Antropus esta comprobación se hace en cada respuesta: extrae cada URL citada, verifica si el dominio es el tuyo y si la página corresponde a la intención del prompt.

Un SoM alto con baja cobertura de citas indica algo específico: la IA menciona tu nombre pero se apoya en

contenidos ajenos para hacerlo. Hablan de ti, pero no recurren a ti para obtener la información. Tu objetivo es que cuando la IA hable de tu tema, tus páginas estén entre las referencias que muestra.

Lo que determina si te citan o no: que tu contenido exista y sea accesible en la web pública, que aparezca en los resultados de la búsqueda interna que hace el LLM, y que ofrezca una respuesta o dato tan pertinente que el modelo lo elija para mostrar. Si tu página tiene el dato exacto que completa la respuesta y nadie más lo tiene, la cita es tuya.

Dimensión 4: Co-Mentions — Dime con quién te nombra la IA y te diré quién eres

Co-mentions mide con qué otras marcas te menciona la IA en la misma respuesta o contexto. Indica que el modelo las asocia en el mismo grupo mental.

Si cada vez que la IA habla de «plataformas de automatización de marketing» menciona a tu competidor archifamoso y no a ti, estás fuera del mapa conceptual. No entraste en la consideración inicial. Juego perdido antes de empezar.

Por el contrario, si la IA siempre te nombra junto al líder del sector, te reconoce en la misma categoría competitiva. Estar mencionado al lado de los grandes significa que al menos te considera un jugador de esa liga.

Para medir co-mentions, construye una matriz de coocurrencia: filas y columnas con nombres de marcas,

incrementando el contador cada vez que dos marcas aparecen juntas en una respuesta. A partir de ahí puedes calcular la probabilidad de que salga una marca dado que salió la otra.

Los patrones que descubras son reveladores. Si jamás apareces cuando mencionan al líder del mercado, la IA no te percibe como competidor directo. Si solo te mencionan con jugadores de nicho y nunca con las marcas top, te tiene encasillado. Si ves que siempre te emparejan con cierto competidor, la IA os ve como intercambiables y eso puede ser bueno o malo dependiendo de cómo te posicione respecto a él. Un detalle que suele sorprender: la IA a menudo te empareja con marcas que ni siquiera tenías en tu lista de competidores. Competitor Intelligence, en Antropus, las detecta y las separa para que las tengas en el radar.

Un matiz importante: si tu SoM es muy bajo, tu matriz de co-mentions estará llena de ceros.

No saques conclusiones precipitadas; primero toca levantar la presencia básica. Co-mentions es más útil cuando ya tienes cierta frecuencia de aparición.

Dimensión 5: Sentiment Score — No es solo salir, es caer bien

«Que hablen de mí, aunque sea mal» es una frase que jamás diría un consultor sensato de GEO. Si la IA habla mal de ti, estás perdiendo en silencio. Tu panel de Analytics no va a reflejar cuántos potenciales clientes te

descartaron porque ChatGPT les dijo «esa empresa es cara y su soporte deja mucho que desear».

Aparecer en la respuesta pero con mala fama es peor que no aparecer. Al menos si no salías, el usuario podía encontrarte después por otra vía. Pero si su primera interacción contigo vía IA planta una percepción negativa, perdiste antes del primer clic.

El Sentiment Score cuantifica el tono con que el modelo describe tu marca. Cada vez que la IA te menciona, clasifica esa mención como positiva, neutral o negativa según las palabras y el contexto. Asigna +1 a lo claramente positivo («excelente opción», «muy recomendado»), 0 a lo neutro («proveedor de software CRM»), y -1 a lo negativo («tiene serias limitaciones», «usuarios se quejan»).

Recopila suficientes respuestas, aplica ese criterio y calcula la distribución. Un 30% positivo, 60% neutral y 10% negativo te da un panorama. Compara con competidores: si la IA siempre alaba al competidor A pero a ti te presenta tibio, estás perdiendo en percepción de valor.

Cómo se ve esto en la práctica

Para que las cinco dimensiones no se queden en abstracto, te enseño un patrón típico que vemos en auditorías iniciales con Antropus.io, la plataforma que usamos en Elevam para automatizar este sistema de medición.

El perfil habitual es una empresa B2B de tamaño medio, con SEO clásico decente (rankea para sus términos principales) que llega convencida de que «la IA tiene que es-

tar mencionándonos». El baseline cuenta otra historia. Sobre 50 prompts representativos del sector, lanzados en los cuatro motores principales (ChatGPT, Claude, Gemini, Perplexity), el Share of Model sale en torno al 4-6%. El competidor líder ronda el 22-28%. Cuando aparecen, lo hacen en posición 4 o 5 de una shortlist de cinco. La Citation Coverage es prácticamente cero: la IA habla de ellos en algunas respuestas, pero nunca enlaza a su web. Y el Sentiment Score es neutro puro: ni positivo ni negativo, simplemente descriptivo.

Lo revelador no es ninguna métrica aislada. Es el cruce de las cinco. Un SoM bajo con cero Citation Coverage te dice que el problema no es solo presencia, es que cuando aparecen, lo hacen citados por otros. Una posición media de 4 sobre 5 te dice que cuando la IA los considera, los considera tarde. Y un sentiment neutro sin adjetivos potentes te dice que no han trabajado la asociación de atributos del Capítulo 17.

Sin estas cinco lentes a la vez, el diagnóstico es opinión. Con ellas, es ingeniería.

Las sub-métricas más útiles son los principales adjetivos asociados a tu marca en las respuestas (un conteo de palabras calificativas te dice cómo te resume la IA en pocas palabras) y el sentimiento condicional: «bueno para PYMEs, no tanto para empresas grandes» no es positivo ni negativo absoluto, pero te da información accionable sobre cómo te encasilla el modelo.

Cómo mejorar cada dimensión

Para mejorar tu Share of Model, necesitas que tu marca esté en más fuentes que la IA consulta. Presencia en Wikipedia, Wikidata, directorios sectoriales, comparativas de terceros, contenido propio optimizado para las consultas de tu nicho.

Para mejorar tu posición en la shortlist, necesitas diferenciación. Datos exclusivos, un ángulo único, un superlativo real que le dé al modelo una razón para ponerte primero.

Para mejorar la cobertura de citas, necesitas contenido técnicamente accesible, indexado y con respuestas tan pertinentes que el modelo las elija para mostrar. Formato RAG-friendly, datos concretos y frases directamente citables.

Para mejorar co-mentions, necesitas contenido comparativo explícito, presencia en rankings y listados del sector, participación en comunidades donde se discute tu tipo de producto y conexiones en knowledge graphs como Wikidata.

Para mejorar el Sentiment Score, necesitas identificar los puntos negativos recurrentes, corregir información falsa o desactualizada, contrarrestar críticas válidas con contexto positivo, y reforzar las fortalezas que la IA ya reconoce. Aquí es donde GEO se cruza con gestión de producto y PR: si el problema es real (soporte lento, interfaz anticuada) la solución no es solo de contenido. Es de negocio.

El Scorecard LLM: tu cuadro de mandos

Las cinco dimensiones juntas forman lo que llamamos el Scorecard LLM. Es la herramienta que deberías rellenar periódicamente (mensual o trimestralmente) con tus datos reales.

El scorecard resume: tu SoM y el de competidores clave, tu posición promedio en las shortlists, tu cobertura de citas con detalle de quién acapara las referencias cuando tú no estás, tu mapa de co-mentions con qué marcas te asocian y cuáles están ausentes, y tu distribución de sentimiento con los principales elogios y críticas que la IA repite.

DIMENSIÓN	QUÉ PREGUNTA	CÓMO SE MIDE
1. Share of Model	¿Qué porcentaje de la voz es tuyo?	% de respuestas que incluyen tu marca
2. Shortlist Position	¿Dónde te ponen en la lista?	Posición media; si no sales, cuentas «sexto»
3. Cobertura de citas	¿Apareces como fuente o solo de pasada?	% de respuestas que citan una URL tuya
4. Co-mentions	¿Con quién te nombra la IA?	Matriz de co-ocurrencia entre marcas
5. Sentiment Score	¿Con qué tono te describe?	+1 positivo · 0 neutro –1 negativo

Juntas forman el Scorecard LLM: se rellena cada mes o cada trimestre.

Las cinco dimensiones del Scorecard LLM. Fuente: elaboración propia a partir de este capítulo.

Con este scorecard, puedes comunicar a tu equipo exactamente dónde están las brechas. Puedes medir progreso: si después de tres meses de trabajo tu SoM pasó de

15% a 30% y tu cobertura de citas de 5% a 15%, sabes que las acciones están funcionando. Si el sentimiento pasó de tener un 20% de menciones negativas a un 8%, sabes que la corrección de narrativa está surtiendo efecto.

Lo que controlas, lo que influyes y lo que aceptas

De todo esto, controlas directamente la calidad y estructura de tu contenido, tu sistema de medición, y la corrección de información en tu web.

Influyes en la narrativa externa sobre tu marca a través de PR y marketing de contenidos, en las co-mentions a través de contenido comparativo y presencia en rankings, y en el sentimiento mejorando tu producto y difundiendo historias positivas.

No controlas el algoritmo interno del LLM, las actualizaciones del modelo, los sesgos de diseño de cada plataforma, las acciones de tus competidores ni las consultas de los usuarios.

Concéntrate en lo accionable. No malgastes energía maldiciendo a la IA por no mencionarte. Dale razones para hacerlo. Mide, actúa y mide de nuevo. El objetivo no es el 100% de citas. Es maximizar tu cuota de presencia en un sistema probabilístico. Habrá casos perdidos que no puedes ganar, y habrá muchos que sí. Si antes te citaban en una de cada cincuenta respuestas de tu sector y tras aplicar todo pasas a una de cada cinco, es un logro enorme.

La visibilidad en LLMs va mucho más allá de «salgo o no salgo». Ahora importa: estás, pero ¿cómo estás? Con estas cinco dimensiones tienes las lentes para verlo con claridad. Se acabó volar a ciegas.

En el siguiente capítulo vamos a construir el sistema de pruebas que alimenta estas métricas: cómo diseñar tu set de prompts, qué metodología usar para que los datos sean reproducibles, y cómo evitar la trampa de confundir correlación con causalidad cuando interpretas los resultados.

CAPÍTULO 10

Test Suite

Prompt Set, metodología experimental y trazabilidad

Medir sin método es coleccionar anécdotas

En el capítulo anterior definimos las cinco métricas que importan. Ahora toca construir el instrumento que las alimenta: tu set de prompts de prueba, el protocolo para ejecutarlos y el sistema de registro que garantiza que cada dato sea reproducible.

Sin un test suite estructurado, tu medición GEO es folklore. Preguntas sueltas a ChatGPT un martes, capturas de pantalla que se pierden en un chat de Slack, conclusiones basadas en dos o tres respuestas que casualmente te gustaron. Eso no es evidencia. Es autoengaño con disfraz de trabajo.

El Prompt Set: tus preguntas base

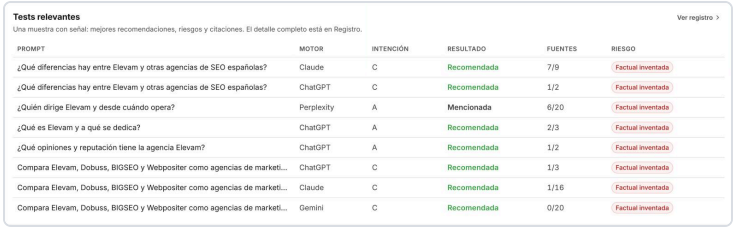
El primer paso es decidir qué consultas vas a monitorizar. No vale improvisar cada vez. Necesitas un dataset fijo, versionado y segmentado que represente las búsquedas reales de tu mercado.

Segmenta por intención. Las consultas informacionales («qué es [concepto de tu sector]», «cómo funciona [proceso]») miden si la IA te considera fuente de conocimiento. Las comparativas («mejor [producto] para [necesidad]», «[tu marca] vs [competidor]», «alternativas a [líder]») miden si entras en la shortlist de decisión. Y las transaccionales («contratar [servicio] en [ciudad]», «precio de [producto]») miden si la IA te recomienda cuando el usuario está listo para actuar.

Segmenta también por plataforma e idioma. El mismo prompt puede dar resultados distintos en ChatGPT, Copilot, Perplexity y Google (AI Overviews o AI Mode). La misma pregunta en español e inglés puede incluir o excluir marcas completamente diferentes. Si operas en mercados multilingües necesitas prompts en cada idioma.

El formato mínimo por entrada: un identificador único, el texto exacto del prompt, la categoría de intención, la plataforma objetivo y el idioma. Un CSV simple basta para empezar. Lo crítico es que sea un documento vivo (lo actualizas conforme detectas nuevas preguntas relevantes) pero versionado, para que siempre sepas qué versión del dataset se usó en cada ronda de medición. En Antropus resolvimos esta parte con Prompt Intelligence: sets de prompts repartidos por categorías de intención, que puedes importar desde CSV o Excel y editar sin sobrescribir, y que reutilizan tanto el Prompt Tracker como Competitor Intelligence.

Un buen punto de partida es de 30 a 50 prompts. Menos de 20 no te da masa estadística suficiente para ver patrones. Más de 100 puede ser inmanejable si lo ejecutas manualmente. Ajusta según tus recursos.



Tests relevantes

Una muestra con señal: mejores recomendaciones, riesgos y citaciones. El detalle completo está en Registro.

Ver registro >

PROMPT	MOTOR	INTENCIÓN	RESULTADO	FUENTES	RIESGO
¿Qué diferencias hay entre Elevam y otras agencias de SEO españolas?	Claude	C	Recomendada	7/9	Factual inventada
¿Qué diferencias hay entre Elevam y otras agencias de SEO españolas?	ChatGPT	C	Recomendada	1/2	Factual inventada
¿Quién dirige Elevam y desde cuándo opera?	Perplexity	A	Mencionada	6/20	Factual inventada
¿Qué es Elevam y a qué se dedica?	ChatGPT	A	Recomendada	2/3	Factual inventada
¿Qué opiniones y reputación tiene la agencia Elevam?	ChatGPT	A	Recomendada	1/2	Factual inventada
Compara Elevam, Dobuss, BIGSEO y Webpositer como agencias de marketi...	ChatGPT	C	Recomendada	1/3	Factual inventada
Compara Elevam, Dobuss, BIGSEO y Webpositer como agencias de marketi...	Claude	C	Recomendada	1/16	Factual inventada
Compara Elevam, Dobuss, BIGSEO y Webpositer como agencias de marketi...	Gemini	C	Recomendada	0/20	Factual inventada

Antropus · Prompt Tracker, tests relevantes del prompt set. Proyecto real de Elevam (elevam.es), baseline del 14 de agosto de 2026.

Las reglas de evaluación: qué cuenta y qué no

Antes de lanzar tus prompts necesitas definir con precisión qué consideras un resultado positivo, negativo o neutro. Sin estas reglas, dos personas del mismo equipo evaluarán la misma respuesta de forma distinta y tus datos serán ruido.

Una mención primaria es cuando la IA te nombra como una de las opciones principales. Estás en el cuerpo de la respuesta, con contexto que indica que te considera relevante. Una mención secundaria es cuando apareces de pasada o como «también existe [tu marca]» sin desarrollo. La ausencia es cuando no apareces en absoluto, aunque la consulta era relevante para ti.

Define también cuándo una mención es positiva, neutra o negativa en tono. «Marca X es reconocida por su soporte excepcional» es positiva. «Marca X ofrece servi-

cios de soporte» es neutra. «Marca X ha tenido problemas reportados con su soporte» es negativa. Estas distinciones alimentan directamente tu Sentiment Score del capítulo anterior.

Y establece si cuentas las citas con URL (la IA enlaza a tu web), las menciones sin enlace (nombra tu marca pero no enlaza), o ambas. Para el cálculo de Share of Model, lo habitual es contar cualquier mención (con o sin enlace) como aparición. Para Citation Coverage, solo las que incluyen URL.

El sistema de registro: trazabilidad total

Cada ejecución de tu test suite debe quedar registrada con suficiente detalle para ser reproducible. Eso significa: modelo usado (ChatGPT, Gemini, Copilot..., con su versión y modo concretos), fecha y hora exactas, prompt exacto enviado, respuesta completa recibida, fuentes citadas con URLs, y tu evaluación (mención primaria/secundaria/ausente, sentimiento).

El formato recomendado es un archivo estructurado (JSON Lines, CSV o base de datos) donde cada fila es una ejecución de un prompt en un motor. Nada de capturas de pantalla sueltas. Nada de «me acuerdo de que salía bien». Si no está registrado con estos campos, no existe como dato.

Esta trazabilidad es lo que convierte tu medición en algo defendible. Cuando un cliente pregunte «¿cómo sabemos que mejoramos?», puedes mostrar: «En octubre, de

40 prompts en ChatGPT, aparecíamos en 4. En enero, en 11. Aquí están los registros de cada uno.» Eso es evidencia. Lo otro es opinión.

La trampa de confundir correlación con causalidad

Este es el error más peligroso del reporting GEO y merece su propia sección. Implementas un cambio en tu web: digamos que añades schema FAQ a cinco páginas. Dos semanas después, tu Share of Model sube del 8% al 14%. Conclusión tentadora: «El schema FAQ nos hizo subir.» Conclusión real: puede ser. O puede ser que ChatGPT actualizó su modelo esa semana y ahora cita más fuentes en general. O que un competidor borró contenido clave y dejó hueco. O simple variación aleatoria.

Los LLMs son probabilísticos. La misma pregunta puede dar respuestas distintas en momentos distintos. Con muestras pequeñas (30-50 prompts), las fluctuaciones son normales. Un cambio de 8% a 14% con 40 prompts significa pasar de 3,2 a 5,6 apariciones. Esa diferencia podría ser ruido estadístico.

Para no caer en esta trampa: nunca atribuyas un cambio a una sola causa sin verificar.

Contrasta entre plataformas. Si solo subiste en ChatGPT pero no en Copilot ni en Perplexity, probablemente no fue tu cambio sino el modelo. Mira tendencias de al menos 2-3 mediciones antes de declarar victoria. Y cuando reportes resultados, sé honesto: «Observamos una mejo-

ra que coincide temporalmente con nuestras acciones, aunque no podemos aislar la causalidad al 100%.»

Esta honestidad no te debilita, te fortalece. Un cliente inteligente respeta más a quien dice «los datos sugieren mejora, pero seguimos verificando» que a quien proclama certezas imposibles en un entorno probabilístico.

Test A/B en contexto GEO

Puedes aplicar una lógica de experimentación controlada a tu medición. La idea es simple: cambias una variable, mantienes el resto constante, y comparas resultados antes y después.

Por ejemplo: tienes una página que la IA no cita nunca. Hipótesis: le falta un resumen ejecutivo al inicio. Acción: añades el resumen. Medición: ejecutas los mismos 10 prompts relevantes antes del cambio (baseline) y 2-3 semanas después (post-cambio), en las mismas plataformas. Si pasas de 0 menciones a 3, y el resto de tus páginas no cambió, tienes un indicio fuerte de que el resumen ayudó.

No es un A/B test puro en el sentido estadístico. No puedes controlar el modelo ni eliminar variables externas. Pero es lo más cercano a experimentación rigurosa que puedes hacer en GEO. Documenta cada experimento como un mini caso: hipótesis, baseline, cambio, resultado y conclusión. Esos casos alimentan tu biblioteca de ejemplos del capítulo 25 y te dan evidencia acumulativa de qué funciona y qué no.



El ciclo baseline → cambio → medición. Fuente: elaboración propia a partir de este capítulo.

Significancia estadística: cuánto es suficiente

Con muestras pequeñas, la tentación de sobreinterpretar es enorme. Aquí van reglas prácticas:

Si tu dataset tiene 40 prompts y pasas de 4 a 6 apariciones, eso es un cambio del 10% al 15%. Suena bien, pero con 40 observaciones esa diferencia puede no ser estadísticamente significativa. Necesitarías ver esa mejora sostenida en 2-3 mediciones consecutivas para tener confianza.

Si pasas de 4 a 12, de 10% a 30%, ahí sí hay señal clara. Cambios grandes son más fáciles de interpretar que cambios pequeños.

La recomendación práctica: no reacciones a fluctuaciones menores entre una medición y la siguiente. Mira la tendencia de 3 meses mínimo. Los promedios móviles son tus amigos. Y cuando un cambio sea ambiguo, amplía el dataset temporalmente. Añade más prompts o repite la medición varias veces en días distintos para ver si el patrón se sostiene.

El baseline: tu punto cero

Tu primera ejecución completa del test suite es tu baseline. Es el punto contra el que medirás todo lo que venga después. Trátalo con la importancia que merece: ejecútalo antes de implementar ningún cambio. Documéntalo formalmente y guárdalo donde no se pierda. El Prompt Tracker de Antropus parte justo de ahí: la primera ejecución es tu baseline y cada medición posterior se compara con la anterior, motor a motor.

El baseline debe incluir: fecha, versión del dataset de prompts, resultados completos por prompt y plataforma y las métricas agregadas (SoM, posición media, sentiment, co-mentions, citation coverage). Este documento es tu «foto del día cero» y lo referenciarás en cada informe posterior.

Sin baseline, no hay progreso demostrable. Solo hay «creemos que antes estábamos peor».

Con el sistema de medición completo —métricas (capítulo 9), test suite (este capítulo) y auditoría (capítulo 11)—, cerramos la Parte IV. Ahora sabes exactamente dónde estás. En la Parte V vamos a priorizar y ejecutar: separar lo que mueve la aguja de lo que solo parece importante.

CAPÍTULO 11

Auditoría GEO completa

Radiografía de tu presencia en IA

Seis capas, un diagnóstico

Antes de tocar una sola línea de código o reescribir una sola página, necesitas saber exactamente dónde estás. No lo que crees. No lo que tu equipo asume. Lo que los datos dicen. Una auditoría GEO profesional tiene seis capas, y cada una ilumina un ángulo distinto de tu visibilidad o tu invisibilidad en los modelos de lenguaje.

Las seis capas son: consultas «money», respuesta generativa, fuentes citadas, entidad, reputación y extractabilidad técnica. Vamos una por una.

1	**Consultas «money»**	¿Sales en «mejores X», «alternativas a Y», «X vs Z»?
2	**Respuesta generativa**	¿Qué cuenta la IA de tu empresa? ¿Es verdad?
3	**Fuentes citadas**	¿En qué fuentes se apoya la IA? ¿Alguna es tuya?
4	**Entidad**	¿Nombre, categoría, atributos y relaciones coherentes?
5	**Reputación**	¿Qué resume la IA de lo que opinan de ti?
6	**Extractabilidad técnica**	¿Puede un bot leer tu contenido sin JavaScript?

Para cada capa: foto actual → brecha con el ideal → plan de acción

Las seis capas de la auditoría GEO. Fuente: elaboración propia a partir de este capítulo.

Capa 1: Auditoría de consultas «money»

Estas son las consultas que generan negocio: «mejores X», «alternativas a Y», «X vs Z», «opiniones de W». Son las preguntas que un usuario hace cuando está a punto de tomar una decisión de compra. Y son exactamente las consultas donde la IA tiene más poder para influir en esa decisión.

Los LLMs responden estas consultas fusionando el consenso web sobre el tema. Internamente, el modelo detecta patrones: qué productos o servicios aparecen una y otra vez en listas top, en reseñas, en comparativas. Si tu marca nunca figura en artículos del tipo «los mejores [tu categoría]» ni en comparativas sectoriales, su vector semántico no estará cerca del clúster de los «top players». Resultado: la IA te omite por falta de asociación, aunque seas relevante en SEO tradicional.

Además, los motores de IA no se limitan a una sola consulta. Suelen desplegar varias búsquedas relacionadas (fan-out): «mejores X», «alternativas a X» o «X vs Y», entre otras, y luego combinan toda esa información en una única respuesta. Por eso, las marcas que salen mejor paradas son las que aparecen de forma consistente en fuentes reputadas a lo largo de todas esas consultas derivadas.

Aquí la auditoría es bastante directa. Pregunta a los LLMs justo lo que quieres saber:

«¿Cuáles son los mejores [tu producto genérico]?», «¿Qué alternativas existen a [TuMarca]?» o «¿[TuMarca] vs CompetidorX, cuál es mejor?». Haz la prueba en varios motores y analiza las respuestas con atención. Fíjate en si te mencionan, en qué posición apareces, con qué contexto se te presenta y qué competidores salen cuando tú no estás. A partir de ahí, podrás detectar con claridad las brechas: en qué conversaciones estás ausente y en cuáles dominan otras marcas. Esa comparativa es la que montamos en Antropus con Competitor Intelligence: qué marcas nombra la IA en tu sector, en qué posición y con qué tono, incluidas las que tú no tenías en el radar.

Los errores más comunes que encuentro: no tener contenido comparativo propio (el vacío lo llenan otros y la IA los cita a ellos), usar lenguaje corporativo que no conecta con cómo busca el usuario real («soluciones integrales ABC» cuando la gente busca «mejor software de

logística»), y sobreestimar el peso del ranking orgánico sin comprobar si la IA te incluye o no.

Capa 2: Auditoría de respuesta generativa

¿Sabes qué cuenta la IA sobre tu empresa? Cuando un LLM habla de ti, está mezclando datos entrenados (lo que aprendió hasta cierta fecha), información recuperada en el momento (si tiene browsing), y —si falta algo— lo rellena con su imaginación probabilística.

Esto no es por malicia. Es simplemente cómo funcionan los modelos: prediciendo la siguiente palabra plausible. Pero una respuesta incorrecta puede ser catastrófica. Es habitual encontrarse con respuestas de este tipo cuando una marca tiene poca presencia en los datos: pides a un LLM información sobre una empresa real y, en lugar de admitir que no tiene datos suficientes, te devuelve una descripción inventada que mezcla sector equivocado, año de fundación incorrecto y servicios que esa empresa nunca ha ofrecido. La IA no miente a propósito. Con poca información, hace un relleno estadístico que suena plausible y falla estrepitosamente.

Haz la prueba ahora mismo con tu propia empresa: pregúntale a ChatGPT, Claude o Gemini «qué es [TuEmpresa]» y compara la respuesta con la realidad. Si tu marca tiene poca presencia digital, los resultados pueden sorprenderte.

El problema se agrava si tu marca es poco conocida o si tu nombre es ambiguo. Si te llamas igual que otra cosa,

el modelo puede confundir entidades y mezclar información. Sin una identidad bien definida en las fuentes, se produce lo que llamamos alucinación de marca: datos falsos presentados con seguridad.

Esta auditoría consiste en preguntar directamente a los LLMs por tu marca: quién es, qué ofrece y cómo la describen. Registra cada respuesta y céntrate especialmente en los fallos: datos incorrectos, atributos inventados, asociaciones erróneas o un tono que no encaja con tu realidad. El siguiente paso es rastrear el origen de cada error. Puede venir de contenido desactualizado, de mensajes ambiguos en tu propia web o de fuentes externas que estén alimentando esa confusión.

Corregir información falsa o desactualizada es una prioridad absoluta. Crea en tu web una página oficial de referencia sobre tu marca que reúna sus datos canónicos, es decir, la información oficial y actualizada que debe tomarse como válida. Corrige también las fuentes externas más importantes, como Wikipedia, Wikidata o los perfiles de negocio. Desmiente de forma proactiva los rumores persistentes con contenido que aborde el error de frente.

Capa 3: Auditoría de fuentes citadas

Cuando la IA responde con referencias, no solo importa el contenido de la respuesta, sino también las fuentes en las que se apoya. Si todas las citas remiten a terceros, tienes un problema de autoridad.

Mapear fuentes citadas implica hacer una lista de preguntas representativas de tu nicho, lanzarlas en herramientas que muestran citas (Copilot, Perplexity), y anotar qué URLs aparecen y qué fragmento de información aportaron. Con 20-30 preguntas distintas tendrás un panorama sólido de tu «share of citations». Para la parte de Google, AI Visibility de Antropus te da esa lista hecha: las búsquedas en las que el AI Overview enlaza a tu web, con qué página, y las páginas que más cita de hasta tres competidores.

Una vez tengas la lista de fuentes dominantes, analiza por qué las está eligiendo. ¿Tienen mejor contenido? ¿Más actualizado? ¿Presentado de forma más clara? ¿Ofrecen un recurso específico (tabla, código, FAQ) que la IA reutiliza? A veces el patrón es evidente: todas las fuentes citadas tienen un FAQ estructurado y tú no, o todas explican con ejemplos concretos mientras tu página es pura teoría.

La acción va en dos frentes. Mejorar tu contenido propio para ser «cite-worthy»: incluir los formatos ganadores, agregar frases contundentes que resuman hallazgos, usar lenguaje preciso en vez de florituras vacías. Y aprovechar fuentes ajenas: si no puedes destronar a quien siempre citan, colabora con ellos. Un guest post, un caso de estudio tuyo publicado en esa fuente de autoridad. Así, la próxima vez que la IA la cite, tú estarás dentro del contenido.

Capa 4: Auditoría de entidad

Si el mundo online no mantiene una imagen coherente de quién eres, la IA acabará construyendo una versión fragmentada y distorsionada de tu identidad. En GEO, la consistencia de entidad no es opcional.

Piensa en tu marca como una entidad singular en el universo digital: nombre, productos, industria, fecha de fundación, sede. Los LLMs dependen de la descripción de esas entidades en sus fuentes de texto. Si hay inconsistencias (tu nombre aparece de tres formas distintas, tu categoría cambia según el sitio, tus datos numéricos no coinciden), el modelo no tiene certezas y puede inventar o barajar versiones.

La auditoría cubre cuatro dimensiones: consistencia de marca (que tu nombre y descripción sean idénticos en todas partes), categoría (que la IA entienda sin ambigüedad a qué te dedicas), atributos (que las características que te definen estén actualizados y sean coherentes), y relaciones (que los vínculos entre tu marca y otras entidades estén correctamente representados).

El proceso: hacer un censo de identidad reuniendo todas las variantes bajo las que existes online, elegir tus verdades canónicas para cada atributo, optimizar tu web como hub de la entidad con JSON-LD Organization completo, y sincronizar todos tus perfiles y menciones externas. Cada inconsistencia corregida reduce la probabilidad de que la IA se confunda o invente.

Capa 5: Auditoría reputacional

Tu reputación online no es solo lo que dicen los usuarios en reviews. Es lo que la IA sintetiza cuando le preguntan por ti. La IA no tiene filtro: mezcla lo bueno y lo malo sin reparos.

La auditoría reputacional empieza recopilando todas las opiniones externas sobre tu marca: búsquedas manuales como «[TuMarca] opiniones» o «[TuMarca] problemas», revisión de plataformas de reseñas, foros, redes sociales y portales como Glassdoor. Los problemas recurrentes que aparezcan en esas fuentes son, con frecuencia, los mismos que la IA resaltará cuando alguien pregunte por tu empresa.

Después, pregúntale directamente a la IA: «¿Qué opinan los clientes de [TuMarca]?» La respuesta te dirá cómo te resume el modelo. Si en ese resumen sale un tema negativo que tú subestimabas, hay que prestarle atención.

Para cada crítica recurrente, la pregunta es operacional: ¿qué estamos haciendo al respecto? Si muchos dicen «el soporte tarda una semana», eso no se arregla con un FAQ en tu web; se arregla mejorando el soporte. Una vez tengas un plan de mejora, hazlo público: publica la mejora, comunícala. La IA eventualmente reflejará ese cambio.

Refuerza lo positivo auténtico: anima a clientes satisfechos a dejar reviews genuinas. Jamás compres opiniones falsas. Los LLMs son sorprendentemente buenos detectando lenguaje repetitivo. Y participa en la conversa-

ción donde hablan de ti: foros, Q&A, comentarios. No para vender, sino para aportar soluciones y mostrar que estás presente.

La única cura real de reputación es mejorar la realidad que refleja. La IA es un espejo. Si la imagen que devuelve no te gusta, el problema no es el espejo.

Capa 6: Auditoría técnica de extractabilidad

Si el modelo no puede extraer tu contenido, da igual lo bueno que sea. Tener la web SEO-friendly para Google no garantiza que sea LLM-friendly.

Los bots de IA son generalmente menos capaces que Googlebot: no ejecutan JavaScript complejo, no esperan mucho a que cargue una página lenta, y puede que no soporten ciertos formatos. Si tu sitio depende fuertemente de JS para mostrar contenido, un bot tipo GPTBot verá la página vacía.

Los puntos críticos a revisar son: el rendimiento (una página que tarda 10 segundos en cargar puede hacer que un bot impaciente se marche); el contenido oculto tras interacciones como logins, muros de pago o acordeones cerrados; la estructura semántica del HTML, que ayuda a los sistemas automáticos a entender la jerarquía de la página; la fragmentabilidad del contenido, es decir, lo fácil que resulta extraer pasajes útiles; el archivo robots.txt y las políticas anti-bot (puede que estés bloqueando a GPTBot sin saberlo), así como el contenido dinámico no textual, como infografías sin texto alterna-

tivo o precios calculados por JavaScript sin una referencia estática.

Revisa tu robots.txt y decide si quieres permitir a los nuevos agentes de IA. Comprueba qué ve un bot sin JS en tu página y asegúrate de que la información clave está en texto HTML accesible, no escondida en scripts o formatos no parseables.

Las seis capas, en acción: un patrón habitual

Vale la pena ver cómo se cruzan estas seis capas en un caso típico. Te describo el patrón que aparece una y otra vez en auditorías iniciales que hacemos en Elevam con Antropus.io, la plataforma donde tenemos automatizado el sistema completo de seis capas.

El perfil habitual es una empresa que llega con la hipótesis equivocada. Cree que el problema es técnico («la IA no nos encuentra porque algo está mal en la web») o de contenido («nos faltan blogs»). El cruce de las seis capas casi siempre cuenta otra historia.

En la Capa 1 (consultas money), descubren que no aparecen en el 80-90% de las preguntas comerciales reales de su sector. En la Capa 2 (respuesta generativa), encuentran que cuando la IA habla de ellos, mezcla datos desactualizados con inventos plausibles: sector mal identificado, año de fundación incorrecto, servicios que no ofrecen. En la Capa 3 (fuentes citadas), la IA se apoya en blogs sectoriales, directorios y comparativas de terceros donde ellos no aparecen. En la Capa 4 (entidad), su

nombre aparece de tres formas distintas entre web, LinkedIn, Wikidata y prensa. En la Capa 5 (reputacional), no hay nada grave pero tampoco hay nada positivo: el silencio reputacional pesa más de lo que parece. Y en la Capa 6 (extractabilidad técnica), descubren que su web es perfectamente accesible. Su problema técnico es cero.

La conclusión es la que el cliente no esperaba: el problema no estaba en el código ni en escribir más blogs. Estaba en el pilar Evidencia (no tenían datos propios publicados que la IA pudiera citar) y en el pilar Consenso (no aparecían en los directorios sectoriales que la IA consultaba como fuente).[3] Sin auditar las seis capas a la vez, ese diagnóstico es imposible.

El marco para no perder la cabeza

Con seis capas de auditoría, es fácil sentirse abrumado. El marco de controlabilidad te ayuda a priorizar.

Lo que controlas directamente: el contenido de tu web, la optimización técnica, tus perfiles oficiales, la corrección de tu Wikipedia, tus comunicaciones públicas. Aquí es donde debes volcar esfuerzos primero, porque son acciones seguras con impacto inmediato.

Lo que puedes influenciar: que terceros hablen de ti, las reviews que dejan los clientes, la posición en que la IA te cita, las discusiones en foros. Trabaja con estrategia y persistencia, pero acepta que los resultados son variables.

Lo que no controlas: los algoritmos internos del LLM, el histórico de entrenamiento, las decisiones de diseño de las plataformas, la existencia de competidores fuertes, las preguntas que hacen los usuarios. No malgastes energía aquí. Asúmelo como el terreno de juego.

Maximiza lo que controlas, influye hasta donde puedas, y acepta con calma lo que escapa de tus manos. La mayoría de competidores ni siquiera harán esto. Tú sí.

De la auditoría al plan de acción

El resultado de las seis capas debe traducirse en un formato operativo: para cada bloque, documenta la foto actual (qué encontraste), la brecha con el ideal (qué falta) y el plan de acción (qué harás para cerrarla).

Ese documento es tu radiografía GEO. Puedes presentarlo a dirección para justificar iniciativas, usarlo internamente como roadmap, y revisarlo cada trimestre para medir progreso. Cada vez que pases de nuevo por las seis capas y veas más puntos en verde, sabrás que la estrategia funciona. En Antropus, el GEO Strategy Builder convierte ese diagnóstico en un documento editable con un roadmap a 30, 60 y 90 días, 6 y 12 meses.

La auditoría no es un evento puntual. Es un ciclo. La IA evoluciona, tus competidores se mueven, las fuentes cambian. Lo que hoy funciona puede no funcionar mañana. Pero con el sistema montado, cada iteración es más rápida y cada mejora se acumula.

Con el diagnóstico completo en la mano, estamos listos para pasar a la ejecución. En la Parte V vamos a separar lo que realmente mueve la aguja de lo que solo parece importante, y a construir el contenido, el schema y la estrategia de optimización que convierten el diagnóstico en resultados.

NOTAS

3. Evidencia y Consenso son dos de los cuatro pilares de visibilidad GEO (Entidad, Extractabilidad, Evidencia y Consenso) que se desarrollan en el capítulo 25. ↩

PARTE V

Priorización y ejecución

Separar lo que mueve la aguja de lo que solo parece importante.

CAPÍTULO 12

Impacto vs. Ruido

Heurísticas para no perder el tiempo

El 80% de lo que haces es ruido

Muchos SEO senior siguen confundiendo actividad con impacto. Se aferran a checklists interminables, convencidos de que cualquier cambio etiquetado como «SEO» suma valor. Pero la realidad es más dura: buena parte de esas tareas no moverá la aguja en GEO. Puedes celebrar un +10 % de tráfico orgánico y, aun así, no aparecer entre los primeros resultados de las respuestas de IA. En ese caso, no estás ganando visibilidad real, sino optimizando la métrica equivocada. De poco sirve ocupar la primera posición en Google si los modelos generativos te ignoran por completo.

Para separar impacto real de ruido necesitas pensamiento de ingeniería, no intuición ni costumbre. La pregunta que filtra todo es esta: ¿reduce el retrievability cost de mi contenido? Si una acción facilita que el modelo te encuentre y te cite, probablemente tiene impacto. Si no cambia nada en tu vector semántico ni en cómo te percibe la IA, es ruido. Añadir definiciones claras y datos verificables en tus páginas fuente baja la fricción de

recuperación. El modelo accede a información estructurada con facilidad. Eso es impacto. Rehacer un meta-tag por capricho estético no cambia nada para la IA. Eso es ruido.

Un baremo útil: si una tarea no aumenta la probabilidad de aparecer en el top 3 de fuentes citadas por la IA, es ruido. Consolidar una entidad dispersa tiene impacto real. Refuerza tu embedding y reduce el vector drift. Cambiar un sinónimo menor en tu copy es ruido.

Corregir un dato crítico que la IA está usando erróneamente es impacto. Micro-optimizar la velocidad de página de 1,3 a 1,1 segundos cuando tu contenido ni siquiera aparece entre las fuentes que el modelo consulta es ruido.



El test de impacto: la pregunta que filtra todo. Fuente: elaboración propia a partir del texto de este capítulo (Asier López Ruiz).

Por qué los equipos tropiezan

Por comodidad y viejos hábitos. Celebran métricas vanidosas (rankings, tráfico) sin comprobar si el modelo generativo los cita. Se autoengañan interpretando correlaciones fantasmas: «tocamos X y subió el SoM, debe ser por X», cuando podría ser ruido estadístico o un update ajeno del modelo. Para no caer en esa trampa necesitas una línea base y un histórico. Es exactamente lo que automatizamos en Antropus con el Prompt Tracker: el mismo set de prompts, lanzado a ChatGPT, Claude, Gemini, Perplexity y Google AI, guardado para comparar el antes y el después de cada cambio.

La falta de datos robustos empuja a muchas marcas a dedicar tiempo a lo irrelevante. Se invierten semanas en ajustar la densidad de palabras clave, como si aún estuviéramos en una fase primitiva del SEO, mientras la IA sigue sin entender de qué va el producto. Se sobrediseñan tareas marginales, se ignora el sesgo de citación de los LLMs, que tienden a apoyarse en fuentes de alta autoridad como Wikipedia o los grandes medios, y se acometen cambios cómodos pero estériles. Mientras tanto, los problemas que sí importan siguen intactos: escasez de citas de calidad, contenido poco extraíble y una entidad difusa o incoherente.

Incluso tareas bien intencionadas pueden generar vector drift si no están alineadas con tu estrategia. Pequeñas ediciones fuera de lugar desvían tus embeddings del tema central. Crees mejorar, pero te alejas del clúster semántico relevante.

Después de la auditoría GEO, con la lista en mano, aplica un test simple a cada iniciativa: ¿afecta directamente a que el modelo te encuentre o te entienda mejor? Si no, al fondo del backlog. ¿Aporta señal fuerte o es solo maquillaje? Si no refuerza tu entidad, tus citaciones o tu contenido base, es ruido.

Priorizar con método, no con instinto

Tu plan posterior a la auditoría no puede ser hacerlo todo a la vez y confiar en que algo funcione. Sin priorización, terminarás con una lista interminable de tareas sin criterio ni rumbo. Creer que puedes abordar cincuenta hallazgos al mismo tiempo no es estrategia, es dispersión. Necesitas cortar el ruido y empezar por lo que de verdad mueve el negocio.

ICE (Impact, Confidence, Effort) y su variante RICE (que añade Reach) te dan un scoring objetivo para cada iniciativa. La lógica es directa: asignas a cada tarea una puntuación de Impacto (cuánto podría mejorar tu visibilidad), Confianza (qué tan seguro estás de ese impacto), Esfuerzo (horas de trabajo y dificultad) y, si usas RICE, Alcance (cuántas páginas o consultas se benefician). Con esos valores calculas un score que te ordena el backlog de mayor a menor valor.

Un ejemplo: implementar schema completo en páginas clave podría tener Impacto 5, Confianza 5, Esfuerzo 2. Score elevado. Reescribir los títulos de tres posts viejos: Impacto 1, Confianza 2, Esfuerzo 2. Score bajísimo. El

framework impone disciplina: obliga a explicitar supuestos y a balancear el ROI técnico.



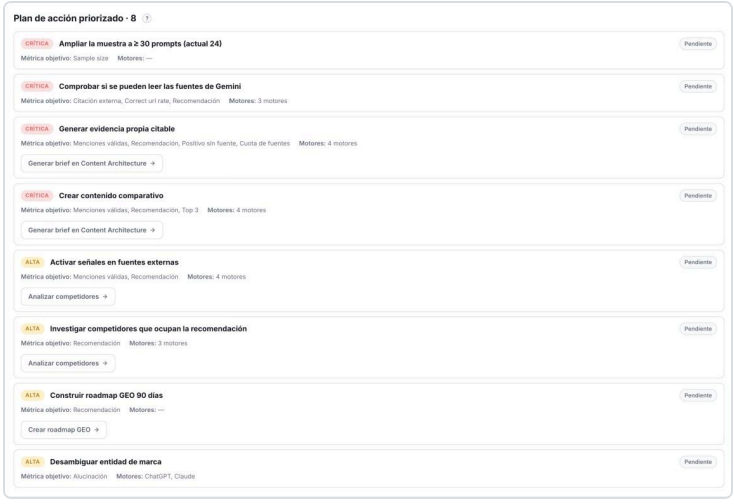
ICE/RICE con los dos ejemplos del capítulo. Fuente: elaboración propia a partir del texto de este capítulo (Asier López Ruiz).

Incorporar Reach es clave en GEO. No es lo mismo arreglar algo que afecta solo a tu página «Sobre Nosotros» que algo que impacta a cien páginas de producto o a veinte intenciones de búsqueda generativa. Tener en cuenta ese alcance evita dar demasiada prioridad a mejoras rápidas pero muy limitadas, con un efecto real mínimo.

Muchos equipos fallan porque convierten este ejercicio en un teatro de puntuaciones vacías. Todo termina con un Impacto 5, ya que cada responsable exagera la relevancia de su iniciativa. Sin honestidad, el framework deja de servir. La solución es sencilla: implicar a perfiles distintos en la evaluación (SEO, ingeniería y producto) y fijar criterios claros antes de puntuar. Un Impacto 5 de-

bería reservarse para cambios que realmente muevan la aguja en SoM o aumenten de forma notable las citaciones. Un Impacto 1, en cambio, debería indicar una mejora menor, marginal o meramente estética. Sin ese marco previo, los números no valen nada.

Ordena las iniciativas por prioridad y marca un punto de corte. Los cinco o diez frentes mejor valorados deben concentrar tu atención inmediata. Detecta las mejoras de alto impacto y bajo esfuerzo, y ponlas en marcha en la primera semana para conseguir impulso rápido. Después, deja para la siguiente fase las acciones con mucho potencial, pero también con una ejecución más costosa o compleja. Es la misma lógica con la que construimos el GEO Strategy Builder de Antropus: diagnóstico por pilares y un roadmap priorizado en horizontes de 30, 60 y 90 días, 6 y 12 meses, que puedes editar a medida que mides.



Antropus · Plan de acción GEO priorizado que genera el informe del Prompt Tracker. Proyecto real de Elevam (elevam.es), baseline del 14 de agosto de 2026.

Quick wins típicos de alto impacto en GEO: corregir referencias inconsistentes de tu entidad, añadir el párrafo de definición que falta en tu página principal, arreglar un dato erróneo que el modelo podría estar usando en tu contra. Cada acción bien alineada con cómo piensa la IA vale por cien movimientos aleatorios sin rumbo.

La secuencia que importa: de los cimientos al tejado

Una de las peores formas de abordar el SEO generativo es hacerlo todo de manera desordenada. Sin una lógica de prioridades, los esfuerzos se dispersan y la estrategia pierde eficacia desde el principio.

GEO no se construye de forma aislada, sino sobre una secuencia lógica de dependencias. Primero hay que consolidar la entidad, para que la IA comprenda con precisión quién eres. Después se construyen las páginas fuente, que actúan como base oficial de contenido para esa entidad. Sobre ellas, el schema ayuda a reforzar su interpretación estructurada. A partir de ahí, las citas externas aportan validación y respaldo. Y, finalmente, la reputación suma la capa de confianza que termina de moldear la percepción de la marca.

Este orden refleja cómo un LLM desambigua y valida información. Primero identifica la entidad canónica. Luego busca sus fuentes oficiales. Lee las señales estructuradas para confirmar detalles. Mira quién más la menciona. Y evalúa sentimiento y confianza.

Cada capa depende de la anterior. Si tu marca no está bien unificada como entidad porque aparece con nombres distintos, no tiene señales claras de correspondencia o carece de una referencia consistente, da igual cuántas citas consigas: el modelo no entenderá que «MiStartup Inc.», «Mi Startup» y tu web oficial son la misma cosa. Y si tampoco tienes páginas fuente sólidas sobre tus productos, de poco sirve añadir schema, porque el bot no encontrará contenido de base lo bastante claro y sustancial como para interpretarlo y apoyarse en él.



De los cimientos al tejado: cada capa depende de la anterior. Fuente: elaboración propia a partir del texto de este capítulo (Asier López Ruiz).

Implementar en orden reduce la entropía que enfrenta el LLM al procesarte. En términos vectoriales, guías al modelo por un camino claro a través del espacio semántico de tu tema, en vez de dejarlo vagando.

Un solo backlog para dos mundos

Separar el SEO clásico del SEO generativo es dividir artificialmente un mismo problema. He visto empresas trabajar con dos listas de prioridades: una para el SEO tradicional y otra para supuestos «proyectos de IA», duplicando esfuerzos, persiguiendo métricas que compiten entre sí y haciendo que los equipos interfieran unos con otros.

Google y los LLMs parten, en buena medida, de la misma materia prima: el contenido y la estructura de tu si-

tio. Por eso, muchas optimizaciones generan un doble beneficio. Mejorar la arquitectura web no solo favorece el rastreo de Googlebot, sino también la capacidad de un LLM para recuperar e interpretar mejor tu contenido. No tiene sentido tratarlo como dos proyectos distintos cuando, en realidad, es una sola mejora con dos efectos complementarios.

Un backlog unificado significa que todas las iniciativas de mejora, tanto las orientadas a Google como las pensadas para ganar visibilidad en IA, se reúnen en una única lista priorizada y se valoran con los mismos criterios.

He presenciado casos de manual: el equipo SEO eliminó cierto contenido «porque no rankeaba y afectaba la UX», sin saber que ese contenido era justo lo que alimentaba a Gemini para mencionar la marca. Desaparece la citación en las respuestas de IA porque alguien «limpió» el sitio pensando solo en Google.

Integra ambos planes en uno. Toma tu backlog SEO actual y los resultados de la auditoría GEO, y combínalos en una única hoja de ruta. Después, aplica una priorización unificada.

Cada tarea debe dejar claro qué se espera conseguir en SEO y qué se espera conseguir en GEO. Por ejemplo: «Reescribir la introducción de la página X». Objetivo SEO: mejorar la relevancia para la palabra clave Y. Objetivo GEO: incorporar una frase con datos y la entidad principal para facilitar la citación. Una sola tarea, dos objetivos complementarios.

Lo que controlas, lo que influyes y lo que dejas pasar

El marco de controlabilidad aplicado a la priorización te ahorra frustraciones.

Hay un ámbito que sí controlas de forma directa: tu sitio, tu marca y todo lo que proyectan. Ahí entran el contenido que publicas, la estructura web, el marcado schema y la coherencia de tu entidad. Son las palancas internas sobre las que puedes actuar sin depender de terceros, y por eso deberían absorber la mayor parte de tu esfuerzo.

Influyes en factores externos sobre los que puedes actuar, pero no garantizar resultados: las citas que otros te conceden, las reseñas de usuarios o las menciones en prensa. Puedes impulsar acciones de outreach, es decir, de contacto y relación con medios, creadores o sitios relevantes, y también mejorar tu producto para favorecer buenas reseñas, pero la decisión final siempre depende de terceros.

No controlas los algoritmos y actualizaciones de los modelos, los sesgos intrínsecos del LLM, las acciones de tus competidores ni los eventos externos fortuitos. No gastes energía obsesionándote con ellos. Mantente informado y prepárate para adaptarte cuando sucedan cambios.

Enfoca tu plan en lo que sí controlas, trabaja con criterio para influir en lo que puedas y acepta con serenidad lo que escapa de tus manos. El plan está vivo, igual que el

mercado y la propia IA. Revisa cada mes el backlog priorizado con datos frescos. Si una acción no genera el impacto esperado, reduce su prioridad en la siguiente ronda. Priorizar no es un acto fijo, sino un proceso dinámico: a medida que implementas, mides y aprendes, recalculas.

Y recuerda: un score no es un contrato con el éxito. Cada implementación es un experimento. La lista está para guiar tus decisiones, no para obligarte a seguir un plan rígido.

En el siguiente capítulo vamos a pasar de la priorización a la ejecución: cómo escribir contenido que la IA quiera usar sin dejar de escribir para personas. Formato, estructura, densidad informativa y el arte de ser citable.

CAPÍTULO 13

Escribir para la IA sin dejar de escribir para personas

El formato decide si la IA te usa o te ignora

Puedes tener el mejor contenido del mundo, pero si lo presentas mal, la IA lo descarta. No es cuestión de estilo ni de gusto. Es mecánica pura: un LLM procesa tu página troceándola en fragmentos, y cada fragmento compite por su cuenta para ser el elegido. Si tus fragmentos son confusos, densos o están mal delimitados, pierdes. Si son claros, autónomos y directos, ganas.

En este capítulo vamos a ver las cinco palancas de formato que determinan si tu contenido es citable o invisible: estructura Q&A, tablas comparativas, densidad informativa, capas de profundidad y optimización para extracción.

Estructura Q&A: cuándo conviene y cuándo es forzar el formato

No todo tu contenido debe convertirse en un interrogatorio interminable. Ignorar por completo el formato pregunta-respuesta sería un error en la era de las IAs.

Un LLM procesa mejor el contenido en formato Q&A porque se asemeja a la forma en que devuelve respuestas. Si tu subtítulo es literalmente la pregunta que haría el usuario («¿Cuánto cuesta el software X y qué planes ofrece?»), el modelo detecta una coincidencia semántica inmediata con la consulta. Internamente, un encabezado interrogativo bien formulado genera un vector muy alineado con la pregunta del usuario, haciéndote más visible para la IA.

Si colocas justo debajo una respuesta breve y directa (uno o dos enunciados que resuman la solución), le estás dando al modelo material listo para citar sin que tenga que rebuscar. Si escondes la respuesta dentro de párrafos densos o das rodeos antes de responder, el asistente saltará a otra fuente más clara. Si quieres saber qué fuente te está ganando en una consulta concreta, el Content Gap Optimizer de Antropus ejecuta ese prompt en ChatGPT, Claude, Gemini y Perplexity, compara tu página con las que sí aparecen y te devuelve un checklist de lo que te falta.

El error más común es añadir al final un batallón de FAQs irrelevantes sacadas de keywords sueltas, sin integrarlas de forma natural. Ese ruido estructural un LLM lo ignora por completo. El otro extremo es redactar títulos en forma de pregunta, pero luego no responder inmediatamente, empezando con vaguedades tipo «Bueno, depende...».

La regla que debemos seguir: convierte tus subtítulos en preguntas reales que haría el usuario y responde justo

debajo de forma concisa. Plantea cada sección como un mini Q&A autónomo: si el LLM recortara solo ese fragmento, debería entenderse por sí mismo. Pero no conviertas todo en un Q&A por sistema. Cuando el contenido adopta una forma narrativa, como en un caso de estudio relatado paso a paso, introducir preguntas constantes puede cortar la fluidez y hacer que el texto suene forzado.

Tablas comparativas: el arma que subestimas

En el SEO tradicional muchos evitaban las tablas. En el SEO para LLMs las abrazamos por su utilidad. Una tabla bien hecha vale oro para comparativas y listados tipo shortlist.

Los LLM leen tablas sin problema porque al final son texto con estructura. Las cabeceras dan contexto y las celdas aportan datos concretos. Si un usuario pregunta «¿En qué se diferencia la plataforma A de la B?» y tu contenido tiene una tabla comparativa, el modelo puede extraer de ahí las diferencias clave directamente. Esto reduce la fricción cognitiva para la máquina: en vez de leer varios párrafos dispersos y sintetizar pros y contras, encuentra la información alineada lado a lado, lista para usar.

El error típico es usar capturas de pantalla de tablas en lugar de texto. Desde el punto de vista del LLM, una imagen con datos dentro es opaca: no puede leerla. También hay quien hace tablas demasiado complejas (anidadas,

con decenas de columnas, sin encabezados) que confunden tanto al modelo como al lector.

La regla que debes aplicar es: usa tablas para comparaciones directas o para condensar información repetitiva. Incluye siempre una fila de encabezados que explique las columnas. Y mantén tus tablas actualizadas: un dato erróneo en formato tabla es doble problema, porque la IA lo cita tal cual dándolo por cierto.

Densidad informativa vs. relleno: cada palabra debe ganarse su sitio

La lectura más pobre del mantra «content is king» es pensar que más texto significa más valor. Un LLM no se recrea en introducciones floridas. Examina el contenido para localizar la información útil que le permita responder a la pregunta que tiene delante.

Además, tiene límites físicos: la ventana de contexto es finita. Cada palabra irrelevante compite por ese espacio y puede empujar afuera el dato que contenía la respuesta. Si de 500 palabras solo 100 tienen sustancia, le estás dando muy poco material útil para trabajar.

He visto e-commerce españoles que abren una categoría de cámaras con tres párrafos sobre la historia de la fotografía. Ni el usuario, ni Google, ni ChatGPT necesitan ese desvío. También son habituales las introducciones genéricas del tipo «En la actualidad, las empresas buscan mejorar sus procesos...», fórmulas vacías que no

aportan nada concreto. Todo ese relleno ocupa un espacio valioso sin añadir información útil ni accionable.

La solución: escribe para la decisión, no para cumplir una cuota de palabras. Cada párrafo debe tener una misión clara: responder una pregunta, explicar un concepto, comparar opciones, dar un ejemplo concreto. Si detectas frases de compromiso tipo «Cabe señalar que...» o repites la misma idea por tercera vez con otras palabras, corta sin piedad.

Un ejemplo rápido. Antes (relleno): «En el mundo empresarial actual, el software ERP se ha convertido en una piedra angular para compañías que buscan eficiencia y eficacia en sus procesos.» Después (densidad): «ERP en la nube vs ERP local: un ERP en la nube elimina costes de servidores propios; un ERP local exige inversión en infraestructura y personal de mantenimiento.» La versión concisa no ha escatimado detalle relevante. Solo ha quitado el envoltorio vacío.

Densidad no significa telegráfico. A veces necesitas dos o tres frases para explicar bien un concepto complejo, y eso no es relleno, es rigor. Lo que hay que evitar son las repeticiones innecesarias, las introducciones de manual y las conclusiones obvias tipo «En conclusión, X es muy importante»: si ya lo demostraste, no hace falta decirlo.

Capas de contenido: resumen ejecutivo, detalle técnico, pruebas

Aunque los LLMs suelen prestar especial atención al inicio del texto y también pueden dar peso al cierre, no conviene reservar la idea principal solo para el final. La respuesta clave debería aparecer pronto y, si es posible, reforzarse o resumirse al cierre.

Los asistentes de IA ya funcionan así: dan un primer resumen conciso y, solo si se les pide, profundizan. Si tu página ya trae incorporados esos niveles, le estás haciendo el trabajo más fácil al modelo.

La primera capa es un resumen ejecutivo al inicio: un mini TL;DR con los puntos clave en tres o cuatro viñetas. Un asistente de IA ante una pregunta general podría citar tus bullet points, ofreciendo una respuesta rápida y segura.

La segunda capa es el detalle técnico: las secciones que desarrollan cada punto en profundidad. Explicaciones extensas, listas de pasos, casos prácticos. Este material es oro para cuando la IA formula preguntas más específicas.

La tercera capa son las pruebas y fuentes: datos concretos que respalden tus afirmaciones. Resultados de tests propios, cifras, citas a terceros. Un LLM entrenado para valorar confiabilidad detectará esas señales de calidad. Si en tu texto mencionas «según nuestra prueba con 5 clientes, el resultado fue X» o referencias un estudio re-

conocido, aumentas la probabilidad de que la IA te perciba como fuente confiable.

El beneficio práctico: si el LLM opera con retrieval, tener las capas separadas le permite recuperar exactamente la porción necesaria. El resumen para respuestas rápidas, el detalle para profundizar y las pruebas para verificar. Si todo está mezclado en un bloque enorme, la IA tiene más dificultad para identificar dónde está cada cosa.



Las tres capas de una página citable. Fuente: elaboración propia a partir del texto de este capítulo (Asier López Ruiz).

Optimización para extracción: ponle las cosas fáciles a la máquina

Cuando un LLM intenta responder usando tu contenido, busca y extrae piezas de información que le parezcan pertinentes. ¿En qué se apoya para hacerlo? En la estructura y los patrones de tu texto.

Un título descriptivo y semántico es fundamental. «Ventajas fiscales de la factura electrónica» le dice al modelo exactamente de qué va el bloque que sigue. «Ahorrando como hormiguita» no le dice nada. La IA no va a sentir curiosidad por tu título ingenioso; lo va a ignorar.

Las definiciones explícitas son otro patrón que los LLMs detectan inmediatamente. «Machine Learning es una disciplina de la IA que permite a las máquinas aprender de datos sin ser explícitamente programadas.» Esa línea es citable al 100%. Si la definición está diluida en un párrafo largo, el modelo tiene que trabajar más para extraerla y probablemente buscará otra fuente que se lo ponga más fácil.

Las listas y viñetas son igualmente potentes. Los modelos reconocen estructuras enumerativas. Si alguien pregunta «¿Cuáles son las ventajas de Z?» y tú tienes una lista con tres ventajas, es muy fácil para la IA extraer esos puntos.

Un glosario al final del contenido le da al LLM un mini diccionario sobre tu tema. Si usas jerga o acrónimos técnicos, un apartado con definiciones cortas hace que la IA pueda responder con tus palabras en vez de arriesgarse con otra fuente.

Hay un detalle técnico que perjudica en silencio: usar formatos visuales en lugar de estructuras semánticas. He visto autores que, en WordPress, hacen que una línea parezca un subtítulo solo cambiando su aspecto visual, aunque en el código siga siendo un párrafo normal. Para

la IA, eso rompe la estructura del contenido. Por eso es importante usar correctamente las herramientas de formato: encabezados reales, listas auténticas y tablas con sus etiquetas correspondientes. La IA no interpreta lo bonito, sino la estructura del código.

Optimizar para la extracción consiste en escribir pensando en cómo un sistema automático va a localizar, entender y reutilizar tu información. Eso implica usar títulos descriptivos con la jerarquía adecuada, definir con claridad los términos importantes, recurrir a listas cuando enumeras, a tablas cuando comparas, a un glosario cuando manejas terminología especializada y mantener siempre un HTML limpio como base de toda la estructura. Es el criterio con el que trabaja el AI Copywriter de Antropus: escribe o analiza cada texto pensando a la vez en SEO, en extractabilidad por IA y en conversión, para que ChatGPT, Claude o Gemini puedan citarlo.

Lo que controlas y lo que no

Controlas la estructura, la claridad, la calidad, el tono y la implementación técnica de tu contenido. Eso es mucho. Y es exactamente lo que separa al contenido que la IA elige del que descarta.

Influyes en cómo te perciben en términos de autoridad (aportando fuentes, mostrando experiencia) y en la relevancia respecto a la intención del usuario (cubriendo un tema a 360 grados). Pero no lo garantizas.

No controlas la pregunta específica que hará el usuario, los algoritmos internos de cada LLM, los sesgos de las plataformas, ni los errores de interpretación del modelo. Tampoco controlas que tu competidor haga exactamente lo mismo que tú pero con más información o más autoridad de dominio.

Lo que sí puedes garantizar es que, cuando la IA llegue a tu contenido, encuentre exactamente lo que necesita: información clara, bien estructurada y fácil de extraer. Todo lo demás es probabilidad. Pero, en un entorno así, facilitar al máximo la recuperación y la citación de tu contenido es una de las mejores estrategias que puedes seguir.

En el siguiente capítulo vamos a ir un nivel más profundo en la escritura: la optimización de claridad semántica. Cómo escribir menos y mejor, cómo cada token cuenta, y cómo evitar la dilución vectorial que hace que la IA te descarte, aunque tengas la respuesta correcta.

CAPÍTULO 14

Cada token cuenta

Optimización de claridad semántica

Tu artículo de 3.000 palabras no impresiona a la IA

Los modelos generativos no premian la longitud. Premian la densidad de significado. Cada frase de relleno, cada rodeo innecesario, es lastre que dispersa tu mensaje clave y diluye la señal entre ruido.

Un LLM procesa tu contenido transformándolo en representaciones vectoriales dentro de un espacio semántico. Si dos párrafos transmiten el mismo dato, uno de 100 palabras y otro de 30, la diferencia de longitud no implica necesariamente una ventaja. Al contrario: el más largo puede arrastrar tokens irrelevantes que no añaden información útil. Además, el contenido excesivo introduce otro problema: puede que ni siquiera quepa completo dentro de la ventana de contexto. Si entregas cinco páginas donde bastaba una, la IA tendrá que resumir, recortar o priorizar fragmentos. Cuanto más largo sea el texto, más fácil será que la idea clave quede diluida o se pierda.

La concisión semántica es una ventaja competitiva. No consiste en escribir de forma telegráfica ni en quedarse corto, sino en ser preciso: separar lo esencial de lo accesorio. Si al recortar descubres que no queda nada sustancial, el problema no era la longitud, sino la falta de contenido real. Pero cuando sí tienes algo valioso que aportar, expresarlo con claridad y sin rodeos refuerza mucho más el mensaje.

Cirugía de tokens: cómo escribir menos y mejor

La primera acción es la actuación radical. Revisa cada párrafo preguntando: ¿aporta algo nuevo que no se haya dicho ya? Si no, córtalo. Elimina frases introductorias vacías tipo «Cabe destacar que...». No repitas la misma idea con sinónimos. Si ya dices «nuestra solución optimiza rutas con IA», no necesitas luego «el sistema usa inteligencia artificial para mejorar la planificación de rutas». Tu mantra: una idea, una vez.

La segunda es maximizar la información por frase. Cada oración debe llevar un dato o un detalle importante.

Antes: «Nuestra plataforma es extremadamente poderosa y flexible, permitiendo una mejora significativa en eficiencia operativa para una variedad de sectores.»

Después: «Nuestra plataforma mejora la eficiencia operativa un 30% en promedio, y se adapta tanto a logística como a manufactura.»

En veinte palabras comunicas lo que antes no quedaba claro en treinta. Menos adjetivos vacíos, más sustancia

medible.

La tercera es el lenguaje preciso y directo. Frases cortas, sujeto-verbo-predicado. Corta muletillas corporativas del tipo «de alta calidad», «a nivel de soluciones integrales». No escribas «En el presente apartado procederemos a analizar las diferentes características» cuando puedes decir «Características del producto:» y listar las importantes.

Lenguaje natural predictivo: baja fricción o la IA te descarta

Un modelo generativo predice la siguiente palabra basándose en el contexto que le das. Si ese contexto es ambiguo o está mal estructurado, la predicción pierde precisión. La fricción cognitiva que siente un lector humano con una frase enrevesada es la misma que experimenta la IA.

«El servicio ofrece integración con sistemas legacy, cosa que mejora su rendimiento, lo cual demuestra nuestra dedicación a la innovación.» Esa oración encadenada marea a cualquiera. La IA prefiere: «El servicio se integra con sistemas legacy. Esto mejora su rendimiento.»

Los pronombres vagos son otra trampa habitual. En una frase como «Juan comentó a Pedro los resultados de su proyecto», no queda claro de quién es el proyecto. Un lector puede dudar y la IA también. Si el modelo extrae ese fragmento para incorporarlo a una respuesta, corre el riesgo de atribuir mal la referencia. Por eso, cuando

intervienen varios sujetos, conviene explicitar en cada frase a quién corresponde cada acción, dato o posesión.

La jerga excesiva sin definir es igualmente peligrosa. Palabras rebuscadas o siglas internas que el modelo quizá no asocie correctamente. Introduce las siglas definiéndolas la primera vez: «Nuestra Arquitectura de Microservicios (MSA) permite...». No asumas que la IA sabe a qué te refieres si no lo escribes.

Antes: «En virtud de la creciente demanda, nuestra empresa ha procedido a implementar mejoras sustanciales en el sistema.»

Después: «Debido a la alta demanda, implementamos mejoras sustanciales en el sistema.»

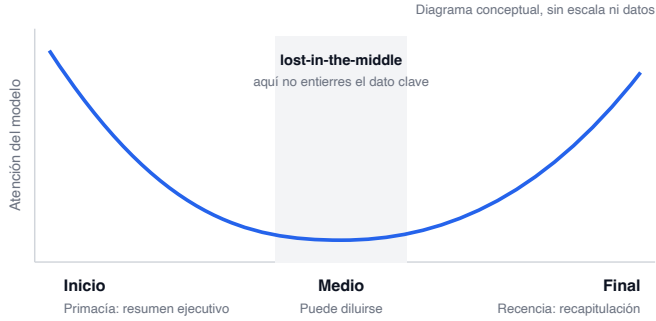
Mismo mensaje, menos palabras. Si una oración acumula demasiadas comas, probablemente puedas dividirla en dos y ganar claridad sin perder sentido.

Dónde poner la información para que entre

La posición de la información importa tanto como su calidad. Los modelos suelen prestar más atención a lo que aparece al principio del texto y, a menudo, también al final. Lo que queda enterrado en medio puede diluirse. Si colocas tu dato clave en el párrafo veinte de una página larga, aumentan mucho las probabilidades de que pase desapercibido.

Esto viene del efecto primacía/recencia: los tokens al inicio y al final de la ventana de contexto tienen más

peso que los del medio. Además, los sistemas de retrieval suelen truncar contenido largo, extrayendo fragmentos de donde encuentran señales contextuales fuertes, que suele ser en los primeros párrafos si el contenido está bien estructurado. Es el efecto lost-in-the-middle aplicado a tu página web.



Dónde poner la información para que entre. Fuente: elaboración propia a partir del texto de este capítulo y del capítulo 6 (Asier López Ruiz). Diagrama conceptual, sin datos.

La táctica principal es la pirámide invertida: lo más importante primero. Abre tu contenido con un resumen ejecutivo que contenga los principales hechos verificables. No empieces con florituras ni introducciones históricas. Si presentas un estudio con cinco hallazgos, que la primera frase los enumere. Si la IA solo toma ese primer párrafo, al menos lleva lo esencial.

En cada sección, la primera frase bajo el encabezado debe resumir la idea principal de ese apartado. «Las tres ventajas principales de X son: velocidad, coste y escalabilidad.» Luego ya desarrollas cada una. Si alguien esca-

nea tu contenido (humano o máquina), con solo leer el heading y esa primera línea sabe de qué va.

Conviene cerrar el contenido con una recapitulación breve que reafirme las ideas esenciales. Si la IA procesa también el final, ese cierre puede ayudar a fijar mejor los puntos clave. El patrón que se repite en auditorías GEO: dos empresas del mismo sector ofrecen las mismas garantías técnicas (por ejemplo, un 99,9% de uptime). La primera lo menciona en el penúltimo párrafo de su página de funcionalidades, enterrado entre características secundarias. La segunda lo coloca en la primera frase de su home. Cuando preguntas a ChatGPT, Claude o Perplexity «¿qué SaaS de [categoría] ofrece mejor garantía de disponibilidad?», el modelo suele citar a la segunda. No porque sea mejor, sino porque tenía el dato accesible al inicio del contenido, exactamente donde el modelo presta más atención (efecto de primacía, Capítulo 6). Esa comprobación no tiene por qué ser manual: con el Prompt Tracker de Antropus lanzas la misma pregunta a varios motores y ves a quién mencionan, a quién recomiendan y a quién citan.

Lo que NO debes hacer: el anti-capítulo

El relleno es veneno. Frases comodín que no añaden datos nuevos son prácticamente ignoradas en la vectorización, o peor, diluyen tu vector alejándolo de consultas concretas. Cada palabra vacía («excelente», «solución integral», «compromiso total») es como añadir ruido blanco a una canción: la melodía se pierde.

Los claims sin prueba son peor que el silencio. Cuando dices «somos los número uno del mercado» y ningún dato externo lo respalda, la IA lo toma con pinzas. Los LLMs saben distinguir entre «Empresa X lidera el sector según Gartner 2026» y «Empresa X dice ser líder». Lo primero es verificable. Lo segundo es promoción barata.

Los tecnicismos vacíos no te diferencian. «Contamos con algoritmos de machine learning avanzados» lo dicen miles de empresas. La IA ha visto esa frase genérica cientos de veces y no la va a asociar fuertemente contigo. En cambio: «Nuestro algoritmo propio predice la demanda con 95% de precisión», eso sí es un hecho.

La regla práctica: si una frase podría decirla cualquier empresa de tu sector, reformúlala hasta que sea algo que solo tú puedas decir por tus características únicas. Si no puedes respaldar un claim, no lo hagas. Es preferible describir tu empresa de forma neutra pero con hechos que soltar grandilocuencias sin fundamento.

Antes: «Somos la empresa líder en soluciones innovadoras para la transformación digital, ofreciendo un servicio excepcional y resultados sin precedentes gracias a nuestro equipo altamente calificado.»

Después: «Tenemos 10 años de experiencia en transformación digital. Nuestra plataforma ha implementado más de 200 proyectos en clientes Fortune 500. Equipo de 15 expertos certificados Azure y AWS. SLA de 99,9% uptime y soporte 24/7.»

El texto deficiente no aporta nada que pueda comprobarse. El buen texto, en cambio, retrata una empresa concreta a partir de métricas, credenciales y datos verificables. Eso le da a la IA algo sólido sobre lo que apoyarse: números, entidades y hechos. Del primer tipo de contenido, en cambio, ¿qué podría citar? ¿«Resultados sin precedentes»? La expresión suena grandilocuente, pero no dice nada preciso ni demostrable.

La dilución vectorial: por qué el relleno te aleja de las consultas que importan

Hay un concepto técnico que explica por qué el relleno es tan dañino más allá de la simple pérdida de atención. Cuando la IA convierte tu página en un vector para compararlo con la consulta del usuario, cada palabra contribuye a ese vector. Si tu página sobre «software de gestión de proyectos» incluye tres párrafos sobre la historia de la gestión empresarial desde los años 60, tu vector se desplaza ligeramente hacia «historia empresarial» y se aleja de «software de gestión de proyectos». Eso es dilución vectorial.

Cada edición fuera de tema, cada tangente innecesaria, cada párrafo de contexto que no aporta a la respuesta que busca el usuario, empuja tu embedding lejos del clúster semántico donde deberías estar. Es el vector drift aplicado a la propia página: no hace falta que cambies de tema entre publicaciones, basta con meter relleno dentro de una misma página para que tu posición en el espacio vectorial se degrade.

La solución es mantener cada página enfocada con láser en su tema central. Si necesitas dar contexto, encapsúlalo en una sección claramente separada para que el chunking lo trate aparte del contenido principal. Y si un párrafo no contribuye directamente a responder la pregunta que debería responder esa página, elimínalo. Si partes de cero, Content Architecture, en Antropus, convierte tus clusters de keywords en briefs con H1, esquema H2/H3 y bloques de respuesta pensados para que la IA los cite, de modo que cada página nazca enfocada en su tema.

Con la claridad semántica optimizada, cerramos las herramientas de ejecución de la Parte V. En el siguiente capítulo vamos al schema avanzado: el marcado estructurado que conecta todo lo que has construido con el idioma que la máquina entiende mejor.

CAPÍTULO 15

Schema avanzado para GEO

Más allá del JSON-LD genérico

Tu plugin SEO no entiende tu negocio

Tu plugin SEO añadió un par de etiquetas schema genéricas y te dijo «listo». Un Organization con el nombre y poco más. Eso no es schema avanzado. Eso es schema cosmético. Si tu marcado no describe con precisión quién eres, qué ofreces y cómo se relacionan tus contenidos, los sistemas que rastrean, indexan y recuperan información sobre tu web tendrán más dificultades para interpretarte correctamente. Eso puede debilitar tu visibilidad en entornos generativos con búsqueda o grounding.

Poner un JSON-LD genérico no hará, por sí solo, que una IA te mencione en sus respuestas, igual que tener un diccionario no te hace bilingüe. El schema sirve para aportar contexto y desambiguación, pero solo cuando está bien implementado y describe con precisión quién eres, qué ofreces y cómo se relacionan tus contenidos. Su valor no está en «activar» menciones de forma mágica, sino en ayudar a los sistemas que rastrean, indexan y recuperan información a entender mejor tu sitio. En en-

tornos generativos con búsqueda o grounding, ese contexto estructurado puede marcar la diferencia entre que tu empresa se interprete con claridad o se diluya como una entidad más.[4]

Los errores que encuentro una y otra vez: marcado minimalista que no dice nada útil, Person sin nombre real, servicios y productos sin marcar «porque no dan snippet visual», Article ignorado en los contenidos, schema duplicado o incoherente por copiar-pegar fragmentos sin integrarlos, y la «ceguera de plugin»: confiar ciegamente en que la herramienta lo hizo bien sin verificar.

Las seis tipologías que realmente importan

Seis tipos de schema cubren las entidades fundamentales de tu presencia online: quién eres (Organization), quiénes son tus personas clave (Person), qué haces (Service), qué vendes (Product), dónde operas (LocalBusiness) y qué publicas (Article).

Organization es el cimiento. Describe quién eres en términos corporativos. No escatimes propiedades: nombre legal y comercial, dirección completa, contacto, fundador, fecha de fundación, logo, y perfiles externos en sameAs. Le estás dando a Google un dossier completo de quiénes sois, mucho más allá del típico «nombre + URL». El @id que le asignes a tu organización será la referencia que todos los demás schemas usarán para conectarse.

Person representa a los individuos importantes: autores de contenido, CEO, fundadores, expertos. Cada persona debe tener nombre completo real, cargo, y vincularse a la organización con worksFor apuntando al @id de Organization. Los perfiles personales en sameAs desambiguan a la persona, especialmente si el nombre es común.

Service describe lo que haces por tus clientes, algo que la mayoría de sitios ignora por completo. Aunque no genere rich snippet visual, alimenta el Knowledge Graph con qué ofreces exactamente. En respuestas generativas, puede ser la diferencia entre «esta empresa habla de X» y «esta empresa ofrece el servicio X».

Product es probablemente el schema más cubierto por herramientas: nombre, imagen, descripción, SKU, marca, ofertas con precio y disponibilidad, y reviews si las tienes visibles en la página. Vincula cada producto a tu organización como seller o manufacturer.

LocalBusiness es para negocios con presencia física. Hereda de Organization y añade dirección completa, coordenadas geográficas, horarios y rango de precios. Lo que declares aquí debe coincidir exactamente con Google Business Profile: mismo nombre, misma dirección letra por letra, mismo teléfono. Cualquier discrepancia confunde.

Article cubre todo contenido publicado: posts, noticias, comunicados. Lo esencial: headline, imagen, fechas de publicación y modificación, y lo más importante author

y publisher referenciados por @id a Person y Organization respectivamente. Así el modelo sabe que este artículo lo escribió Jane Doe, CTO de Example Corp, y se publicó en el blog de Example Corp. Todo conectado.

Conecta las piezas: @id, @graph y anidación

Incorporar muchos datos estructurados es útil, pero si no están conectados entre sí de forma coherente, dejas que el sistema deduzca por su cuenta las relaciones. Y esa deducción puede ser incorrecta.

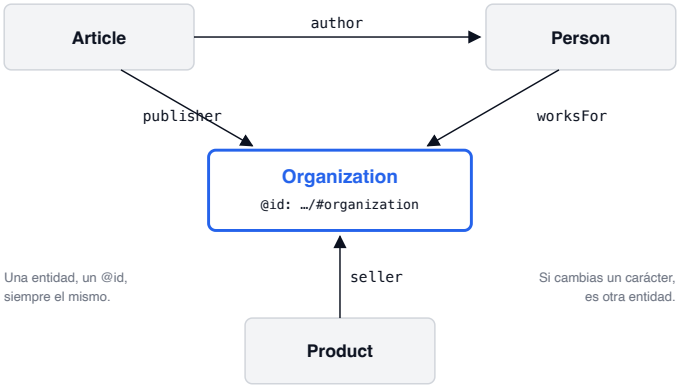
El @id es el mecanismo fundamental. Es un identificador único para cada entidad. Una URI con fragmento que permite que otros schemas la referencien sin repetir todos los datos. Defines tu organización una vez con `"@id": "https://tudominio.com/#organization"`, y luego en cada Article, Product o Service, simplemente apuntas a ese ID con "publisher": {`"@id": "https://tudominio.com/#organization"`}. Misma entidad, mismo @id, siempre. Si cambias un carácter, para la máquina es otra entidad.

El @graph es una forma de declarar múltiples entidades en un solo JSON-LD, agrupándolas en un array. No es obligatorio —Google procesa todos los scripts JSON-LD de una página en conjunto—, pero hace el código más limpio y reduce el riesgo de fragmentos aislados.

Los errores más comunes: no usar @id en absoluto (dejando fragmentos sueltos que parecen entidades distintas), usar IDs relativos que cambian según la URL de la

página, asignar IDs diferentes a la misma entidad en distintas páginas, y no conectar Person con Organization ni Article con Author.

La regla práctica es pensar el marcado como una base de datos relacional. Cada entidad tiene un identificador propio. Las relaciones entre ellas funcionan como vínculos explícitos: publisher apunta a Organization, author apunta a Person, worksFor conecta Person con Organization y seller enlaza Product con Organization. El resultado es un grafo coherente y bien conectado.



El marcado como base de datos relacional: un @id, muchas referencias. Fuente: elaboración propia a partir del texto de este capítulo (Asier López Ruiz).

Consolida tu identidad digital con sameAs

Implementar schema en tu web es un avance importante, pero resulta insuficiente si Google no encuentra señales externas que respalden la identidad que declaras. Para consolidarte como una entidad reconocible, nece-

sitas coherencia no solo dentro de tu sitio, sino también fuera de él, en los perfiles y referencias que te describen en la web.

SameAs es la propiedad que dice «esta entidad es la misma que esta otra en tal sitio». No es cantidad lo que importa, es calidad. Un sameAs a LinkedIn es altísimamente confiable. Uno a un directorio genérico donde cualquiera sube su ficha es irrelevante o peor.

Los perfiles que valen la pena: LinkedIn de empresa, X, Facebook (si es activo), Wikipedia si la tienes, Wikidata, Crunchbase, y perfiles sectoriales de alta confianza como G2 o Capterra. Todo lo demás probablemente sobra.

Consistencia de nombre e información en todos los perfiles es crítico. Si tu schema dice «Example Corp» pero tu Facebook pone «Example Corporation Inc.» y tu LinkedIn «ExampleCorp», estás creando ruido. Unifica la marca: mismo nombre, mismos datos, en todas partes. Y si un perfil externo tiene información desactualizada o errónea, corrígelo. En Antropus lo resolvemos con la Ficha de Entidad: genera y mantiene el JSON-LD canónico de tu marca, con desambiguación de homónimos y tus URLs oficiales publicadas como sameAs, listo para tu <head>.

Un error que veo repetidamente: mezclar perfiles personales con corporativos. El LinkedIn del CEO va en el schema Person del CEO, no en el sameAs de Organization. Cada entidad con sus propios perfiles.

Markup de credenciales: demuestra lo que vales

Si vas a poner en valor tus logros, exprésalos de forma verificable. Las IAs no se dejan convencer por un eslogan; necesitan hechos concretos, datos y señales que puedan respaldar.

Los premios pueden reflejarse en la propiedad award de Organization o Person, indicando el nombre del galardón y el año, siempre que exista una referencia verificable en fuentes externas. Si afirmas que fuiste «Mejor Startup Tech 2026» pero no hay ninguna fuente independiente que lo respalde, será difícil que la IA lo considere creíble.

Certificaciones profesionales van en hasCredential con el tipo EducationalOccupationalCredential: nombre de la certificación, organismo que la reconoce, fecha. En sectores YMYL (salud, finanzas, legal) esto suma puntos de confianza significativos.

Reviews deben estar visibles en la página para el usuario. Google penaliza markup de reseñas que no están en el contenido real. AggregateRating solo si es verdad. Jamás inventes reviews. La época de engañar con estrellas falsas terminó.

Para logros cuantitativos que no tienen propiedad directa («10.000 clientes en 35 países»), una FAQPage con la pregunta y respuesta explícitas es el vehículo más efectivo. Un LLM que rastree tu página encontrará esa cifra claramente estructurada.

Trata tu schema como un sistema vivo

Implementar schema una vez y no volver a revisarlo es una receta segura para que quede obsoleto en pocos meses. Tu negocio cambia, tu web evoluciona y Google actualiza sus criterios. Pero, si nadie lo mantiene, el JSON-LD permanece inmóvil y termina describiendo una realidad que ya no existe.

El mantenimiento tiene varios pilares. Validación en cada despliegue: incorpora la verificación de schema al flujo de CI/CD para que errores de sintaxis no lleguen a producción. Monitorización en Search Console: si la línea de elementos válidos cae de golpe, algo se rompió. Crawling periódico para detectar que ninguna página importante perdió su schema. Un proceso editorial interno donde cada cambio de negocio (nuevo CEO, nuevo producto, cambio de dirección, precio actualizado) dispare la actualización del marcado correspondiente.

El concepto clave es evitar el schema drift: la deriva entre tu contenido real y tu marcado. Si algo se elimina del frontend, elimínalo del schema. Si algo nuevo se agrega, agrégalo al schema. Sincronización constante.

Y un detalle que mata en silencio: desarrolladores que borran o modifican JSON-LD sin saber lo que hacen. Inserta los scripts vía código en plantillas, no como contenido editable. Pon comentarios claros en el código. Haz una sesión al equipo explicando que ese código es importante para Google.

Lo que controlas y lo que no

Controlas completamente la implementación, la calidad y la actualización de tu schema. Eso es mucho, y es lo que separa a los profesionales de los que hacen las cosas a medias.

Influyes en que Google incorpore tus datos al Knowledge Graph y en que los LLMs los utilicen. Cuanto más completo, consistente y verificable sea tu marcado, más probabilidades tienes. Pero el Knowledge Graph tiene sus propios criterios de notoriedad.

Tampoco controlas hasta qué punto un sistema generativo con búsqueda utilizará realmente tu JSON-LD en sus procesos de recuperación o interpretación. A día de hoy, no hay documentación oficial que lo garantice de forma explícita. Lo que sí puedes es observar el resultado: AI Visibility, en Antropus, te dice en qué búsquedas te cita el AI Overview de Google, con qué página y frente a qué competidores.

Cada snippet implementado correctamente es un cable más conectando tu sitio con la mente de la máquina. Mantener esos cables limpios y conectados es lo que separa la ingeniería de búsqueda seria del SEO superficial.

Con el schema construido, llegamos al final de la Parte V. Hemos recorrido el camino completo desde la priorización hasta la ejecución: filtrar ruido, escribir contenido que la IA quiera usar, optimizar la claridad semántica y construir el marcado estructurado que conecta todo. En la Parte VI vamos a dar el salto a la marca: cómo con-

solidar tu entidad, construir co-ocurrencia estratégica y gestionar la reputación que la IA percibe de ti.

NOTAS

4. No hay, a octubre de 2026, documentación oficial que confirme una lectura directa del schema por parte de los LLMs para mejorar sus respuestas. Donde sí está documentado su valor es en buscadores y sistemas de recuperación, que lo usan para entender mejor páginas, entidades y relaciones. En experiencias generativas con búsqueda o grounding, su impacto parece ser indirecto. (Google for Developers) ↩

PARTE VI

Entidad, reputación y autoridad

Construir la marca que la IA quiere recomendar.

CAPÍTULO 16

Consolida tu entidad

Desambiguación agresiva y coherencia categorial

Si tu nombre es genérico, para la IA eres un fantasma

Los motores de búsqueda piensan en entidades, no en palabras clave. Named Entity Disambiguation es el proceso por el cual distinguen entre entidades que se llaman igual o parecido. «Apple» puede ser la fruta o la empresa tecnológica, y solo entendiendo el contexto el modelo sabe cuál mostrar. En el Knowledge Graph cada significado vive en un nodo separado. Tu objetivo es que tu marca tenga su propio nodo único.

Si tu empresa se llama Jaguar, más te vale que en cada mención aclares de qué hablas.

Añade categoría, sector, ubicación u otro atributo diferenciador cada vez que cites la marca. Piensa que estás entrenando al modelo: a fuerza de ver tu marca junto a «software financiero B2B» o «restaurante mexicano en CDMX», asociará ese nombre con la ficha correcta.

Desambiguar significa lograr que tu marca sea tratada como una entidad canónica: única, distinguible y clara-

mente separada de cualquier otra con la que pueda confundirse. Pero no es un camino fácil. Si compartes nombre con alguien o con algo más conocido, empiezas con desventaja. Google tiende a reforzar la interpretación dominante, aquella que la web lleva más tiempo repitiendo y respaldando. Cuando un nombre ya está ampliamente asociado a otra entidad, cambiar ese patrón requiere tiempo, coherencia y muchas señales acumuladas.

La cirugía de precisión: seis pasos para consolidar

El primer paso es definir tu Entity Home, una página oficial en tu dominio dedicada 100% a describir quién eres. No frases de marketing vacías. Una descripción clara y desambiguada. «Acme es una empresa de ciberseguridad industrial en España, fundada en 2015, especializada en protección de infraestructuras críticas.» Más un bloque de datos estructurados completo con JSON-LD donde declares nombre, URL, logo, descripción, dirección, fundación, fundadores y sameAs con enlaces a todos tus perfiles oficiales. En Antropus ese bloque sale de la Ficha de Entidad: nombre canónico, descripción, desambiguación de homónimos y tus URLs oficiales como sameAs, con el JSON-LD listo para tu <head>.

El segundo paso consiste en aportar contexto en cada mención. No solo en tu web, sino en todos los canales: en el título de la página de inicio, en la meta description, en la bio de X, en la descripción de LinkedIn o en las notas de prensa. Tu nombre debe ir siempre acom-

pañado de un elemento que lo sitúe y lo defina. «Acme Cybersecurity» funciona mucho mejor que un simple «Acme». Puede parecer repetitivo, pero esa repetición contextual es justo lo que ayuda a entrenar al algoritmo.

El tercer paso es dejar claro quién no eres cuando la confusión es fuerte. Puedes hacerlo con una sección en tu web del tipo «No somos TalCosaParecida» o con una FAQ como esta: «¿Qué es Acme y en qué se diferencia de la plataforma XYZ, que tiene el mismo nombre?». En la práctica, es una forma de decirle a Google, de manera explícita, que esa otra entidad, no eres tú.

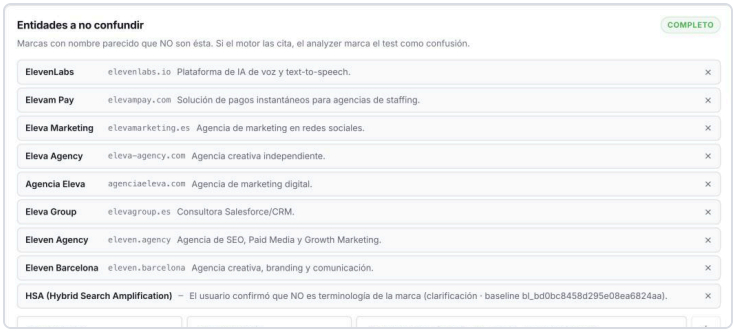
El cuarto paso es añadir un descriptor a tu marca si lo necesitas. Si tu nombre es demasiado genérico o poco distintivo, conviene acompañarlo temporalmente de una palabra que lo sitúe mejor. En lugar de presentarte solo como «Apollo», puede ser más útil usar «Apollo Broadband» durante una fase. Ese descriptor adicional ayuda a Google a entender de qué entidad se trata y a diferenciarla de otras con nombres parecidos. Cuando tu marca ya esté bien asociada a su contexto, podrás reducir ese apoyo y volver, si quieres, a usar solo «Apollo».

El quinto es entretejer enlaces de confirmación mutua. Desde tu Entity Home, enlaza a tus principales perfiles externos. Desde tus redes sociales, pon siempre el link a tu web oficial. Desde Wikipedia y Wikidata, asegúrate de que enlacen tu sitio oficial. Este loop de enlaces cruzados crea una red de confirmación: todas las fuentes se respaldan mutuamente apuntando al mismo centro.

El sexto paso es revisar periódicamente el efecto de lo que estás haciendo. Busca tu marca junto con tu sector o tu ciudad y analiza los resultados. ¿Ya aparece un Knowledge Panel? ¿Siguen mezclándose tus resultados con los de otra entidad? Si todavía aparece de todo menos tu marca, necesitas reforzar la estrategia.



Los seis pasos para consolidar tu entidad. Fuente: elaboración propia a partir del texto del capítulo 16.



Antropus · Identidad de marca, entidades a no confundir: si un motor cita alguna, el test se marca como confusión de entidad. Proyecto real de Elevam (elevam.es).

Consistencia total o no hay autoridad

Existe un viejo mito: «mientras más perfiles tengas por ahí, mejor». Error. No sirve de nada estar en mil sitios si en cada uno pareces una empresa distinta. Google recopila datos sobre tu negocio de múltiples fuentes: tu web, Wikipedia, Wikidata, Google Business Profile, Facebook, LinkedIn, prensa, directorios. La idea es que al juntarlo todo, pueda construir una sola entidad unificada. Tu objetivo debería ser facilitarle ese trabajo, no complicárselo.

Si cada fuente canta una versión distinta de tu nombre o tus detalles, el resultado es un Frankenstein semántico. Tu web dice «Acme Co.», en LinkedIn apareces como «Acme, Inc.», en un directorio estás como «Grupo Acme» y Wikidata muestra «Acme Corporation». Para el algoritmo pueden parecer entidades diferentes. El Knowledge Graph puede bifurcar el nodo de tu entidad, y tu autoridad se diluye entre nodos duplicados.

El remedio es una consistencia draconiana. Un solo nombre canónico usado igual en todos lados: NAP (Nombre, Dirección, Teléfono) idénticos hasta en la coma. Misma categoría o descripción principal y mismo logotipo. Si en un sitio pones «Consultora de software» y en otro «Proveedor de soluciones TI», estás dando mensajes mezclados.

La auditoría de entidad que vimos en el capítulo 11 cubre cuatro dimensiones: consistencia de marca, categoría, atributos y relaciones. Aquí la llevamos a la práctica. Haz un inventario exhaustivo de todas las variantes de tu marca y datos en la web. Busca tu nombre de todas las formas posibles. Documenta cada URL donde apareces y qué información muestra. Identifica discrepancias. Y después unifica: decide la verdad canónica para cada atributo y aplícala sin excepción en todo lo que controlas. Esa verdad canónica no debería vivir solo en una hoja de cálculo: en Antropus la guardamos en la Ficha de Entidad, con una puntuación de completitud que te dice cuál es la siguiente acción, y es la misma identidad que reutilizan el Prompt Tracker y Competitor Intelligence para medir.

El Knowledge Panel: cómo provocar que suceda

Un Knowledge Panel no aparece por arte de magia ni por hacer SEO on-page. Es el reconocimiento visible de que Google ya te tiene bien fichado y ha consolidado suficiente información confiable. Necesitas dos cosas: notabilidad (suficiente presencia e interés de usuarios) y

corroboración estructurada (datos fiables desde fuentes de alta autoridad que concuerdan entre sí).

El desencadenante clásico es tener un artículo en Wikipedia. Aunque no es la única vía. Google también se nutre de Wikidata, de tu Google Business Profile, del markup schema.org en tu web y de bases de datos sectoriales.

El schema markup por sí solo no crea un Knowledge Panel, pero sí ayuda a Google a entender y cruzar datos. Si en tu página incluyes un JSON-LD declarando Organization con tu información y haces sameAs a tus perfiles, y además tu Wikidata confirma esos mismos datos, y tu Google Business Profile tiene verificada la misma dirección, los planetas se alinean. Múltiples fuentes confiables dicen lo mismo, Google puede cuadrar el puzzle con confianza.



La red de confirmación mutua alrededor de tu Entity Home. Fuente: elaboración propia a partir del texto del capítulo 16.

Las señales que suman: presencia en Google News, volumen de búsquedas de tu marca, engagement en propiedades de Google (reseñas en Maps, actividad en YouTube), y menciones en fuentes de autoridad. Todo lo que indique que existes de forma activa en el ecosistema.

La acción concreta: refuerza tu Entity Home con schema completo. Crea o completa tu ficha en Wikidata con todos los campos posibles. Optimiza y verifica tu Google Business Profile. Cruza los streams de información: conecta todo con todo. Gana menciones en fuentes de autoridad y monitorea el progreso: la API del Knowledge Graph de Google te dice si ya tienes un entityId asignado aunque aún no veas panel.

El ecosistema de datos estructurados externos

No basta con predicar en tu web quién eres. Tienes que inscribir tu entidad en los registros y bases de datos que Google consulta regularmente.

Wikidata es la joya de la corona: pública, estructurada, multilingüe y Google puede hacer consultas directas para obtener datos como tu fecha de fundación o tu sector. Google Business Profile es ineludible si tienes presencia física. LinkedIn es casi universal. Crunchbase es imprescindible para startups y empresas tech. Y luego, están las bases sectoriales específicas de tu industria: G2 para software B2B, directorios de colegios profesionales para médicos o abogados, registros oficiales de tu sector.

No se trata de estar en todas partes. Calidad sobre cantidad, siempre. Mejor estar en cinco fuentes bien cuidadas que en cincuenta mediocres que no puedas mantener. Los primeros esfuerzos dan grandes saltos en claridad para Google. Del paso cero a uno, es decir, de no estar en Wikidata a estarlo, es un salto enorme. Pasar de ocho sitios a dieciocho probablemente ya no aporta tanto.

Mantén un documento maestro de control: todas las plataformas donde estás presente, con la fecha de la última actualización. Integra esto en tus procesos de negocio. Así cuando cambie algo, abres el documento y ves exactamente qué hay que actualizar y dónde.

Lo que controlas y lo que no

Controlas el contenido de tu web, la información en tus perfiles oficiales, y tus entradas en bases de datos que te permiten autogestión. También controlas la consistencia entre todas estas fuentes.

Influyes en Wikipedia (la comunidad decide, pero puedes aportar contenido neutral con referencias), en medios y prensa (no controlas qué dicen, pero puedes hacer PR inteligente), en reseñas de usuarios (no las dictas, pero puedes mejorar tu servicio), y en la notoriedad de tu marca (puedes generar búsquedas con marketing, pero es el público quien decide).

No controlas el algoritmo de Google ni sus decisiones sobre el Knowledge Panel. Tampoco controlas la interpretación dominante de tu nombre en la web, las reglas de las plataformas externas y los cambios que puedan introducir las plataformas de IA generativa o las tendencias de percepción pública de tu sector.

La desambiguación agresiva consolida tu entidad con el tiempo, pero no hace milagros instantáneos. Es una guerra de desgaste: persistir con señales claras hasta que ganas. Y a veces la decisión más pragmática (aunque dolorosa para el ego) es aceptar la derrota semántica. Si después de mucho batallar es imposible destronar a esa otra entidad ultra-famosa de la mente de Google, plantéate un rebranding. Un verdadero estratega de Search sabe cuándo insistir y cuándo pivotar.

En el siguiente capítulo vamos a salir de tu web y trabajar el terreno externo: cómo lograr que tu marca co-ocurra junto a los líderes de tu sector en contenidos de terceros, y por qué el Digital PR orientado a GEO ya no busca solo backlinks sino presencia en los contextos que la IA consume para construir sus respuestas.

CAPÍTULO 17

Off-Site

Co-ocurrencia y asociación de marca

Si no estás en la conversación, no existes para la IA

Los LLM no necesitan enlaces para aprender relaciones entre marcas. Les basta con detectar patrones de coocurrencia y mención conjunta. Si muchos artículos sobre «los mejores [tu producto]» mencionan de forma repetida a la marca líder junto a la tuya, el modelo acaba infiriendo que ambas compiten en la misma liga.

Pero no es lo mismo una mención contextual natural («empresas como X, Y y Z ofrecen esta función») que una mención forzada en cada párrafo. Las co-ocurrencias contextuales genuinas pesan mucho más que las metidas con calzador. Los algoritmos cazan patrones: si tu nombre aparece consistentemente cerca del del líder en contenidos de calidad, acabas relacionado con ese líder en la mente de la IA. Esa co-ocurrencia se puede medir: en Antropus, Competitor Intelligence cuenta cuántas veces sale tu marca frente a tus competidores cuando ChatGPT, Claude, Gemini y Perplexity hablan de tu

sector, comprobando antes que la mención sea de verdad tu marca.

Co-ocurrencia estratégica: cómo infiltrar tu marca

El primer paso es identificar las listas clave (los artículos de «Top 10» y «Mejores [tu producto]») que rankean más alto o que los chats de IA suelen citar. Esos son tus objetivos de co-ocurrencia. No lo adivines: con el Prompt Tracker de Antropus lanzas los prompts de tu categoría a ChatGPT, Claude, Gemini, Perplexity y la IA de Google y ves a quién mencionan y qué citan; con AI Visibility, qué páginas cita el AI Overview de Google en cada búsqueda, la tuya y las de tus competidores.

El segundo es ganarte un hueco. Contacta a los editores ofreciéndoles algo de valor: una prueba de tu producto, datos comparativos exclusivos, una funcionalidad única o métrica donde superes al líder. Tienes que darles razones para incorporarte, no esperar un favor altruista.

Si nadie te compara, hazlo tú con honestidad. Publica una comparativa imparcial en un medio neutral o un blog invitado, sin lenguaje de venta descarado. Un informe comparativo creíble tiene más impacto que mil posts en tu blog diciendo «somos los mejores».

Diversifica formatos y canales. Videos en YouTube comparando productos y preguntas en Quora tipo «¿Alternativas a [Marca Líder]?» con respuestas que aporten valor real. La idea es estar presente donde el público busca opciones.

La advertencia: co-ocurrencia mal hecha es un tiro en el pie. Si cada vez que te nombran en un foro es alguien diciendo «Sí, X es más barata, pero sale rana», estás generando una asociación perjudicial. Apunta a calidad de asociación, no cantidad ciega. Vale muchísimo más aparecer en dos listas respetadas junto al líder que en veinte blogs de dudosa fama.



Mapa de co-ocurrencia: los contextos donde la IA te ve junto al líder. Fuente: elaboración propia a partir del texto del capítulo 17.

Asociación de atributos: que la IA te defina con hechos, no con eslóganes

Da igual cuántas veces te proclames «líder del mercado» en tu web si afuera nadie lo apoya. La reputación de ser «fiable», «rápido» o «exclusivo» hay que ganársela fuera de tu sitio, en la voz colectiva de internet.

La IA aprende qué cualidades se ligan a tu marca leyendo y correlacionando información distribuida. Si en reviews, foros y artículos aparece repetidamente «Marca Z, entrega rapidísima», los modelos asocian Marca Z con rapidez. Si Wikipedia dice «Empresa Z es conocida por su excelente soporte», esa conexión estará presente en respuestas generativas. Tu propaganda unilateral pesa casi nada.

Los pasos concretos: elige un atributo en el que puedas sobresalir de verdad. Reúne pruebas tangibles: porcentajes de disponibilidad, certificados y casos de éxito con cifras. Siembra la conversación: colabora con medios para que publiquen benchmarks que destaquen tu fortaleza, incentiva reseñas que mencionen ese atributo y refuérzalo en eventos con datos concretos. En tus propios canales, apóyate en evidencia, no en eslóganes: «Índice de error del 0,01 % sobre 1 millón de transacciones» comunica mejor que «Líderes en confiabilidad».

La realidad brutal: si tu producto no sustenta el atributo que quieres, olvídalo. La asociación de atributos es fruto de reputación real. Y no intentes asociarte a cinco cosas a la vez — nadie te va a creer. Clava uno y cuando lo domines, amplías.

Digital PR para GEO: ya no buscas backlinks, buscas contexto

Una aparición de tu marca en un artículo de un medio líder le dice al algoritmo que eres relevante y digno de ser mencionado en ese tema. Incluso sin backlink, el

nombre apareciendo en un contexto de alto perfil alimenta la memoria del modelo de IA. Si TechCrunch publica algo sobre tu startup, esa información entra en el pool de datos que leen los LLMs.

La clave es crear contenido citable. Los periodistas aman los números frescos, los estudios comparativos, los informes anuales con datos. Un informe bien hecho titulado «Informe Sector X, 2026» se convierte en referencia que genera menciones citando tus datos. Esa es la mención ideal: con contexto rico que alimenta al modelo con hechos vinculados a tu nombre.

Las malas prácticas que la IA ignora: notas de prensa genéricas a mansalva que acaban en agregadores de tercera. Obsesión con el backlink por encima del contexto. Un párrafo valioso en un medio respetado vale más que cien links en sidebars. No aportar nada noticioso: si tu marca no genera datos interesantes, la única cobertura será publirreportaje encubierto que los modelos ignorarán. Y las acciones puntuales de PR, sin continuidad en el tiempo: un único impacto rara vez basta para consolidar una señal.

La hoja de ruta: apunta a los 5-10 medios clave de tu nicho, no a 50 menciones mediocres. Genera noticias reales (estudios propios, benchmarks, desafíos públicos). Dale a los periodistas el material masticado (dossieres visuales, gráficos, frases de expertos citables). Inyecta en tus declaraciones el atributo que te interesa ligar a tu marca, respaldado por datos. Y ten siempre assets citables públicos en tu web (whitepa-

pers, datasets, herramientas). Construye relaciones, no impactos sueltos.

Reddit, Quora y el consenso humano

Muchos ejecutivos ignoran lo que se dice en foros con el argumento «son cuatro frikis».

Error grave. Esas plataformas actúan como sensores del sentir popular. Lo que la gente comenta y repite allí, los modelos de IA lo captan para calibrar cómo perciben las personas a un producto.

Google en sus resultados generativos cita directamente foros para preguntas subjetivas. «¿Vale la pena [Producto]?». Los AI Overviews pueden contestar con pros y contras salidos de hilos de Reddit. Considera esos contenidos como la voz sincera del usuario, libre de marketing.

La IA no hace caso a un comentario suelto de 2 votos. Pero cuando ve volumen y consenso sobre un tema recurrente en muchos hilos, con muchos usuarios de acuerdo y montones de upvotes, se vuelve señal sólida. Las conclusiones colectivas se cuelan en las respuestas.

Las estrategias: monitoriza con herramientas para detectar patrones: la misma duda repetida, el mismo mito circulando, la misma queja. Participa con autenticidad cuando tenga sentido: un ingeniero tuyo respondiendo en Stack Overflow gana karma y respeto, y en su perfil aparece que trabaja en tu empresa. Usa Quora estratégicamente: busca preguntas con muchas visitas sobre tu

industria y responde con calidad, mencionando tu solución solo al final y con contexto. Gestiona la negatividad con cabeza fría: arregla el problema real y comunica la mejora en tus canales. Y amplifica los comentarios positivos genuinos: comparte en tus redes cuando alguien elogia tu producto espontáneamente.

Lo que la gente coincide en decir de ti suele ser cierto. La mejor forma de cambiar una narrativa negativa es solucionar el problema real y dejar que la comunidad se dé cuenta por sí misma.

Multimodalidad: más allá del texto web

Lo que se dice en vídeos, podcasts e imágenes también influye en la forma en que tu marca aparece en los resultados generativos. Google transcribe vídeos automáticamente, y la búsqueda ya funciona de manera multimodal, mostrando incluso fragmentos de vídeo relevantes. Además, el contenido audiovisual se propaga: se comenta en foros, se incrusta en blogs, se cita en redes sociales y deja huella en forma de texto, que es precisamente el material que la IA puede procesar con facilidad.

La trampa más común es separar los esfuerzos en compartimentos estancos: «lo de SEO va por aquí, lo de YouTube por allá». Muchas empresas invierten en webinars y emisiones en directo, pero luego no transcriben ni reaprovechan ese material. El resultado es que una par-

te valiosa de su conocimiento se diluye y deja de generar valor.

La acción es directa: transcribe y publica todo contenido audiovisual como artículo o resumen en tu web. Recicla entre canales (un informe técnico se convierte en video, un webinar en post). Ten presencia táctica en canales multimedia donde está tu público. Colabora con creadores que ya tienen la audiencia y optimiza la búsqueda dentro de cada plataforma (descripciones en YouTube con timestamps, notas extensas en podcasts).

Rompe los silos. Cada contenido importante debe existir en formato texto en alguna parte para ser indexado. Cada mensaje clave debería resonar en múltiples formatos.

Lo que controlas, lo que influyes y lo que aceptas

Controlas tu web, tus redes sociales, tus notas de prensa antes de enviarlas y tus perfiles oficiales. Úsalos para difundir mensajes clave con pruebas.

Influyes en artículos de prensa, videos de terceros, posts en comunidades, respuestas en Quora y fichas en plataformas de review. Puedes orientar el resultado proporcionando buen material y construyendo relaciones, aunque el resultado final no lo decides tú.

No controlas opiniones espontáneas, contenido generado por terceros sin tu intervención y ni la memoria de la IA una vez entrenada. Solo puedes escuchar, aprender, y

corregir la información en la web para que la próxima actualización del modelo refleje la realidad.

Tu marca será definida por lo que los demás dicen de ella. No por lo que tú digas de ti mismo. Trabaja para que digan algo bueno, y con razón.

En el siguiente capítulo vamos al otro lado de la moneda: qué pasa cuando ya apareces en la respuesta de la IA pero con mala fama. Reputación, sentimiento y cómo gestionar las crisis en la era donde un modelo de lenguaje puede destrozarte en silencio.

CAPÍTULO 18

Reputación y sentimiento

Cuando te recomiendan con asterisco

La IA te pone en tu sitio según tu fama

Un modelo de lenguaje no es un árbitro neutro. Es un espejo de la opinión colectiva. Si tu marca tiene mala fama, la IA lo refleja sin piedad. Los embeddings que representan tu marca incorporan el sentimiento subyacente de miles de textos: si en foros y reseñas abundan frases tipo «nunca más con TuEmpresa», el modelo internaliza ese patrón de negatividad. Cuando el sentimiento es bueno, la IA te pinta bien. Cuando es malo, o te ignora o te pone peros.

Un detalle que cambia el juego: Reddit es la fuente número uno que cita la IA generativa de Google, incluso por encima de páginas oficiales de empresas. El contenido generado por usuarios, ese hilo despellejándote o alabándote, puede pesar más que tu propia web corporativa en la mente de la IA.

Aparecer en la respuesta pero con mala fama es peor que no aparecer. Si la primera interacción de un cliente potencial vía IA planta una percepción negativa, perdiste antes del primer clic. Por eso no basta con contar si

apareces: el Prompt Tracker de Antropus mide si te mencionan, si te recomiendan y si te citan, que es justo donde se nota el asterisco.

Reseñas legítimas: el combustible de la recomendación

Las reseñas falsas huelen mal, y la IA tiene olfato fino para esa pestilencia. Los LLMs distinguen patrones de veracidad: valoran mucho más una reseña detallada en plataformas como G2 o Trustpilot, con usuarios verificados, campos estructurados y flujo constante en el tiempo, que veinte reseñas genéricas en un blog desconocido.

Los datos lo confirman: G2 es el sitio de reseñas de software más influyente en respuestas de IA, con un 22% de cuota de visibilidad. Y a más reseñas, más visibilidad: un 10% más de reviews que tu competidor puede traducirse en un 2% más de menciones en respuestas generativas. Parece poco, pero es la diferencia entre aparecer o no.

El sistema para conseguir reseñas legítimas: identifica los momentos «WOW» del cliente cuando acaba de lograr algo gracias a tu producto o cuando te envía un elogio espontáneo al soporte. Activa la solicitud en caliente, cuando el cliente está feliz, con un link directo a la plataforma relevante. Facilita el proceso al máximo. Incentiva de forma ética: acceso anticipado, descuento simbólico, mención de honor y sin manipular el contenido.

Diversifica las plataformas. La IA mira en todas partes. Fomenta también reseñas constructivas: esas 3 o 4 estrellas con alguna crítica. Muestran que no hay trampa. Una reseña que diga «No es perfecto en X, pero el soporte fue rápido en solucionarlo» te humaniza, y la IA al leerlo puede mencionarte con tono equilibrado pero positivo.

La calidad y la consistencia pesan más que la cantidad bruta. Cincuenta reseñas de cuatro estrellas, bien redactadas y sustanciosas, aportan más que quinientas de cinco estrellas clónicas y sin valor.

Gestión de crisis 2.0: cuando la IA te pinta mal

La IA no olvida ni se cansa. Antes, un rumor en X (antes Twitter) se perdía si la gente dejaba de hablar de él. Ahora, una vez la bola de nieve de mala reputación está en los datos de la IA, te persigue: cada vez que alguien pregunte por ti, el modelo añade el «pero» con el dato negativo.

Dos mecanismos generan esto: la alucinación, que es cuando la IA inventa datos que nunca existieron porque le faltan fuentes y rellena huecos, y la misinformación por sesgo de datos, que es cuando toma un hecho real pero lo saca de contexto o lo exagera. La segunda es más insidiosa porque suena creíble.

La IA tiende a amplificar: si tiene veinte artículos negativos con el mismo ángulo, aunque haya cien positivos, le da peso colectivo al negativo. Y prioriza fuentes co-

munitarias: Reddit suele superar a comunicados oficiales en las respuestas generativas.

La estrategia tiene dos fases. La prevención activa: tener tus activos de transparencia listos antes de que pase nada. Páginas de estatus con historial de incidencias. Secciones que anticipen tus puntos débiles. Borradores de respuesta pre-aprobados para escenarios de crisis, redactados con datos concretos y formato amigable para IA.

Y la reacción quirúrgica cuando estalla: actúa en horas, no días. Publica la explicación oficial inmediatamente con FAQs de crisis en formato Q&A. Siembra tu versión en los canales que la IA frecuenta (Reddit, medios, redes). Activa aliados que puedan avalar tu respuesta. Y monitoriza en tiempo real si la IA está integrando tu nueva información. En Antropus lo hacemos en dos frentes: Brand Monitoring recoge las menciones de tu marca en noticias, blogs y foros, las clasifica por sentimiento y detecta picos frente a tu media histórica; y el Prompt Tracker vuelve a lanzar los mismos prompts a los motores para ver si la respuesta ya recoge tu versión.

Si la metida de pata es real, la IA lo dirá con o sin tu intervención. Tu objetivo realista es lograr que la respuesta incluya tu versión junto a la acusación: «Sí, pasó X, pero la empresa hizo Y para solucionarlo.»



Gestión de crisis en dos fases. Fuente: elaboración propia a partir del texto del capítulo 18.

Transparencia proactiva: tu mejor escudo

La opacidad es suicidio digital. Si tú no hablas de tus objeciones y límites, la IA llenará ese silencio con contenido de terceros que no va a ser amable.

Las páginas de transparencia son secciones en tu web donde abordas de frente las posibles objeciones, las limitaciones de tu producto, cómo trabajas y las pruebas que lo demuestran. No son altruismo: son supervivencia. Las IAs premian la información clara, verificable y completa. Cuando un LLM necesita describir tu marca, preferirá sacar datos concretos de tu página oficial de preguntas difíciles antes que fiarse de un post anónimo.

Un portal de transparencia efectivo tiene cuatro secciones. Objeciones frecuentes: las pegas que has escuchado de clientes, con respuestas respaldadas por hechos. Lí-

mites y alcance: para quién NO es tu producto, dicho con honestidad radical. Metodología: cómo haces lo que haces, con métricas concretas. Y pruebas y evidencias: casos de éxito, certificaciones, resultados medibles.

El beneficio crítico: cuando admites tus límites, desactivas narrativas negativas potenciales. Si tú ya dices «No somos para empresas de más de 500 empleados», la IA responderá citando tu posición oficial en vez de especular «quizás no escala bien». Has enmarcado la limitación en tus términos.

Y hay un efecto colateral poderoso: estas páginas se convierten en imanes para fuentes externas. Periodistas, analistas y wikis las citan, amplificando su alcance y haciendo que la IA las encuentre aún más.

Lo que controlas, lo que influyes y lo que aceptas

Controlas tu contenido oficial, tus perfiles en plataformas reputacionales, y la calidad real de tu producto y servicio. Estas son las áreas donde puedes construir un pilar sólido de reputación.

Influyes en la opinión pública ofreciendo experiencia excepcional, motivando reseñas genuinas, participando en comunidades con valor, y reaccionando con inteligencia a las críticas. Puedes inclinar la balanza pero no la orquestas totalmente.

No controlas los comentarios espontáneos de terceros, los artículos negativos de prensa, los memes virales, ni la memoria histórica de la IA. Ese fiasco de 2019 puede

seguir apareciendo porque quedó grabado en los datos de entrenamiento.

La estrategia ante lo incontrolable: no puedes apagar cada incendio, pero puedes crear diez fuegos artificiales positivos a la vez. Mucho contenido útil, casos de éxito que se muevan y clientes felices hablando en público. Con el tiempo, eso diluye los ecos negativos.

Tu marca será definida por lo que los demás dicen de ella. La IA es un espejo amplificador: si la imagen que devuelve no te gusta, el problema no es el espejo. La única estrategia ganadora a largo plazo es la humildad y la transparencia radical, combinadas con un producto que de verdad merezca ser recomendado.

Con la reputación gestionada, cerramos la Parte VI. Hemos construido la marca que la IA quiere recomendar: entidad consolidada, presencia en las conversaciones correctas, y reputación que resiste el escrutinio de un modelo de lenguaje. En la Parte VII bajaremos a cada motor y a cada vertical, y en la Parte VIII pasaremos de la estrategia al sistema: roadmap de 90 días, automatización del pipeline de medición, y cómo empaquetar todo esto como servicio profesional.

PARTE VII

Estrategias por plataforma y vertical

Cada motor y cada sector tiene sus propias reglas.

CAPÍTULO 19

Domina los motores: Google, Perplexity, ChatGPT y Gemini

Un manual por motor, sin generalidades

Hasta ahora hemos trabajado principios que aplican a cualquier motor de IA. Pero cada plataforma tiene sus propias manías, sus propios sesgos y sus propios mecanismos de selección de fuentes. Lo que funciona en Google AI Overviews no necesariamente funciona en Perplexity. Lo que ChatGPT valora puede ser irrelevante para Gemini en local. Este capítulo es un manual de estrategias por motor, bajando al terreno práctico de cada uno.

Google AI Overviews: estar primero no te garantiza nada

Google, con AI Overviews, genera un overview con IA combinando múltiples fuentes. No se limita a un snippet de una página: el modelo busca en el índice de Google contenido relevante y lo sintetiza en una respuesta única. El detalle clave es que aplica un filtro de calidad antes de decidir a quién citar. Evalúa E-E-A-T de las posibles fuentes: si tu página no desprende experien-

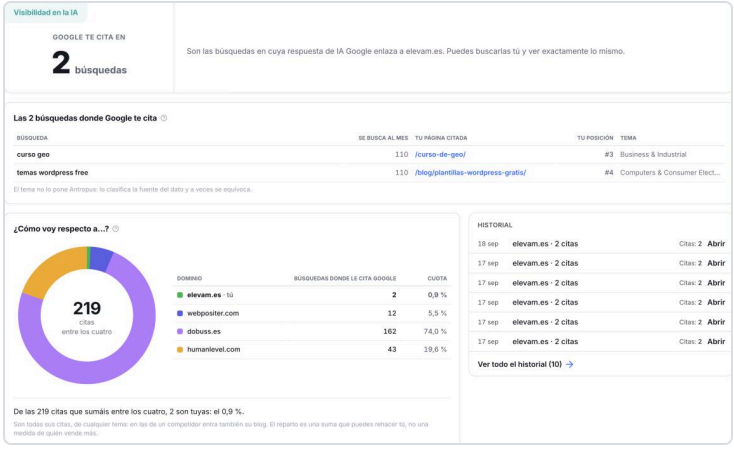
cia, autoridad y confianza, la descarta aunque tengas la respuesta perfecta en texto. Desde 2025 convive con AI Mode, la búsqueda conversacional con IA de Google, que llegó en español en septiembre de 2025 y a España en octubre de 2025.

AI Overviews prefiere contenido «answer-friendly»: bien estructurado, conciso y al grano. Puede ignorar al que rankea primero si otro resultado más abajo ofrece una explicación más confiable o completa. Y ha comprimido el espacio de visibilidad: ya no compites por top 10, compites por ser una de las 3-5 fuentes citadas en el overview.

Las estrategias específicas para AI Overviews: E-E-A-T al máximo nivel en cada contenido; autor con nombre y credenciales, casos prácticos y enlaces a fuentes externas confiables. Estructura answer-friendly: encabezados formulados como preguntas seguidos de respuesta concisa inmediata. Señales técnicas de calidad: schema adecuado, Core Web Vitals, contenido accesible sin login ni JavaScript pesado y contenido con information gain: datos originales, enfoque diferenciador y insights que tus competidores no ofrecen.

Lo que no controlas: AI Overviews puede ignorarte incluso si ocupas la primera posición, porque puede preferir combinar otra fuente con Wikipedia. Puede citarte un día sí y al siguiente no, sin que tú hayas cambiado nada. También puede favorecer a marcas muy conocidas por un sesgo de notoriedad. Y, en temas YMYL, puede no considerar fuentes que no tengan una autoridad

muy alta y ampliamente reconocida. Por eso conviene seguirlo búsqueda a búsqueda: el AI Overview Tracker de Antropus registra en cada chequeo si Google te cita, te menciona sin citarte, muestra el bloque sin ti o no lo muestra, y guarda el histórico; y AI Visibility te enseña en qué búsquedas te cita ya, con qué página y frente a qué competidores.



Antropus · AI Visibility: búsquedas en las que el AI Overview de Google cita tu web y reparto de citas frente a tres competidores. Proyecto real de Elevam (elevam.es), medición del 18 de septiembre de 2026.

Perplexity y motores con citas: sesgo hacia lo reciente y específico

Motores como Perplexity o Microsoft Copilot combinan un LLM con búsqueda en vivo. El sistema lanza consultas a un buscador (Bing, en el caso de Copilot; Perplexity usa su propio índice), abre los primeros resultados prometedores, lee trozos de cada página, y sintetiza una

respuesta con citas numeradas. Cada afirmación que hace la IA tiene su origen identificable.

Lo que hay que entender: estos motores no son bibliotecas neutrales. Tienen una lista restringida de sitios de confianza y prefieren citar contenido de ese conjunto. Si tu página no apareció en los resultados que el LLM consultó, no te citará, aunque tu contenido tuviera la respuesta perfecta. Y tienen sesgo hacia lo reciente y específico: contenido del año en curso y páginas que van directo al grano.

Las estrategias específicas: base técnica impecable. No bloquees a PerplexityBot ni a Bingbot en robots.txt, minimiza dependencia de JavaScript, optimiza velocidad. Respuesta primero, detalles después. La primera frase o párrafo debería dar la respuesta corta y directa. FAQ cubriendo preguntas relacionadas («La gente también pregunta» de Google y las related queries de Perplexity son oro puro para saber qué cubrir). Contenido chunkeable (cada párrafo autocontenido, longitud de 2-4 frases, viñetas y tablas cuando corresponda).

Lo que no controlas: el retardo de indexación puede ser significativo. El sesgo hacia dominios conocidos es real. La IA puede preferir un artículo decente de un dominio famoso sobre uno excelente de un sitio desconocido. Y la IA puede amalgamar información de dos fuentes y solo citar una.

ChatGPT con búsqueda: dos cerebros, una respuesta

ChatGPT funciona con dos componentes: un conocimiento estático base (su entrenamiento hasta cierta fecha) y un refresco dinámico limitado (lo que encuentra en búsquedas en vivo). La memoria estática suele pesar más de lo que quisiéramos.

Si tu marca nació después del corte de entrenamiento, simplemente no existes para el modelo. Incluso con la búsqueda activada, ChatGPT tiene sesgos formados por su entrenamiento. Si en sus memorias tu marca aparece poco, no te va a proponer por sí solo. Y tiene preferencias aprendidas por dominios: si hay un resultado de Wikipedia, casi siempre lo abre primero.

El concepto clave aquí es lo que podríamos llamar echo index: la representación que la IA tiene de ti en su memoria de entrenamiento. Para mejorar tu echo index necesitas aparecer en las fuentes canónicas del conocimiento online (Wikipedia, Wikidata, datasets abiertos, rankings sectoriales, medios reconocidos). Son como backlinks model-facing: no buscan PageRank sino presencia en los datos que la IA seguramente leerá.

Las estrategias específicas: alimentar la memoria del modelo para aparecer en fuentes canónicas que las próximas versiones del modelo incorporarán. Consistencia absoluta de entidad: una frase canónica de tu marca usada en todos lados para que la IA asocie correctamente nombre con producto y sector. Contenido echo-

friendly (datos concretos, sin jerga corporativa vacía). Señales dinámicas continuas (goteo constante de menciones en la web pública para que en sucesivos entrenamientos sigas presente). Y testear tu echo index regularmente (preguntarle a ChatGPT «¿Conoces la empresa X?» y «¿Cuáles son las principales empresas en [tu sector]?») para medir dónde estás. Y en el plano técnico: la búsqueda de ChatGPT rastrea la web con OAI-SearchBot; si lo bloqueas en robots.txt, tu web no se mostrará en sus respuestas con búsqueda. Es exactamente lo que automatizamos en Antropus con el Prompt Tracker: el mismo set de prompts, lanzado a ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview, con el histórico guardado para comparar mes a mes.

Lo que no controlas: no puedes re-entrenar ChatGPT tú mismo. Si tu sector es de nicho con poca información, el modelo puede seguir ciego. Los sesgos de idioma son reales: contenido mayoritariamente en español puede ser invisible para ChatGPT en inglés. Y el lag de entrenamiento significa que tus esfuerzos de hoy pueden tardar meses en reflejarse.

Google Gemini y el ecosistema local

Google está integrando Gemini en todo su ecosistema, especialmente en Maps y resultados locales. La IA local no necesita oír la historia de tu negocio desde tu web: le basta la ficha de Google Business Profile, las reseñas de usuarios, y los datos estructurados de Maps.

Gemini actúa como un concierge inteligente que conoce la ciudad al dedillo. Sabe qué ofrece cada negocio, qué dice la gente de ellos, cuál te queda cerca. La IA pondera en tiempo real: rating y número de reseñas, contenido de las reseñas (palabras clave como «romántico» o «familiar»), categorías y atributos de la ficha, proximidad y horario, y contexto del usuario.

Esto rompe el paradigma del SEO local clásico. Antes competías por top 3 en el pack local para una keyword. Con la IA, ella reorganiza sugerencias según lo que el usuario preguntó específicamente. Puedes ser tercero en el mapa para «pizzería cerca», pero si el usuario pregunta «pizzería con terraza», y tu competidor quinto tiene terraza y tú no, la IA lo mencionará a él.

Las estrategias específicas pasan por tener el perfil de Google Business Profile completamente trabajado: una descripción con palabras clave relevantes, todos los servicios bien listados, horarios actualizados, la categoría principal que mejor define el negocio acompañada de categorías secundarias relevantes, fotos recientes de calidad y publicaciones frecuentes. A eso se suma la consistencia NAP, con el nombre, la dirección y el teléfono exactamente iguales en todos los sitios; una gestión activa de las reseñas, incentivando opiniones genuinas y respondiendo siempre; los atributos bien completados (la sección de preguntas y respuestas de la ficha la ha ido retirando Google desde finales de 2025, en favor de respuestas generadas con IA a partir del perfil y las reseñas); y una web alineada con el perfil local, con páginas

por ubicación, mapa incrustado y schema LocalBusiness.

Lo que no controlas: la IA de Maps se alimenta de datos reales de usuarios. Si tu restaurante tiene 3,5 estrellas y muchas quejas, Gemini no va a mentir por ti. La geografía manda: si el usuario está lejos, la proximidad pesa más que tu optimización. Y Google hace experimentos A/B rotando sugerencias, así que tu posición en la respuesta puede variar sin que hayas cambiado nada.

Qué prioriza cada motor, según este capítulo

Motor	Fuentes	Qué prioriza	Lo que no controlas
Google AI Overviews	Índice de Google, combinando varias fuentes	E-E-A-T, estructura answer-friendly, señales técnicas, information gain	Citas que cambian de un día a otro; sesgo de notoriedad; YMYL
Perplexity y Copilot	Búsqueda en vivo; respuesta con citas numeradas	Lo reciente y específico; respuesta primero; contenido chunkeable	Retardo de indexación; sesgo hacia dominios conocidos
ChatGPT con búsqueda	Memoria de entrenamiento + búsqueda en vivo	Presencia en fuentes canónicas (Wikipedia, Wikidata, medios); entidad consistente	Corte y lag de entrenamiento; sesgo de idioma; nichos sin datos
Gemini en local	Business Profile, reseñas y datos de Maps	Rating y reseñas, atributos y categorías, proximidad y horario	Opiniones reales de usuarios; geografía; rotación de sugerencias

Resumen cualitativo del capítulo 19; no es una medición.

Qué prioriza cada motor, según este capítulo. Fuente: elaboración propia a partir del texto del capítulo 19.

En el siguiente capítulo vamos a ver cómo estas diferencias por plataforma se amplifican según la vertical: por qué la IA no «ve» igual la búsqueda de un restaurante

que la de un software ERP, y qué estrategias específicas necesitas según tu sector.

CAPÍTULO 20

GEO por verticales

Local, E-commerce, SaaS, B2B

Cada sector pelea una batalla distinta

La IA no se comporta igual en todas las industrias. La búsqueda de «restaurante romántico» y la de «mejor software ERP» activan mecanismos completamente distintos en el modelo: los datos que consulta, los sesgos aprendidos y las expectativas del usuario varían drásticamente. Necesitas estrategias calibradas según tu vertical porque la IA rankea, filtra y cita distinto según el contexto.

GEO por verticales: qué mira la IA en cada sector

Vertical	Qué consulta la IA	Señal dominante	Táctica clave
Local	Google Business Profile, Maps y reseñas	Proximidad y reputación de clientes	Ficha al 100 %, reseñas activas y NAP idéntico en todos los directorios
E-commerce	Atributos de producto, Merchant Center, Shopping Graph, reviews	Datos estructurados y diferenciación	Schema Product, contenido único en fichas, especialización en nicho
SaaS	Comparativas, contenido técnico, G2 o Capterra, señales de entidad	Consultas «alternativas a» y «vs»	Páginas comparativas, contenido BOFU con datos, FAQs técnicas
B2B y servicios	Rankings en portales, casos, testimonios, premios, certificaciones	Autoridad y pruebas	Casos públicos con cifras, testimonios verificables, marca personal

Resumen cualitativo del capítulo 20.

GEO por verticales: qué mira la IA en cada sector. Fuente: elaboración propia a partir del texto del capítulo 20.

GEO Local: reseñas como señal dominante

En local, la IA de Google suele apoyarse principalmente en Maps, Google Business Profile y las reseñas. Consulta tu ficha, interpreta las opiniones de clientes y tiene en cuenta detalles como horarios, categorías o fotos. Por eso, en muchas búsquedas locales, tu web queda en segundo plano. Si alguien pregunta «¿Cuál es la mejor pizzería por aquí?», la IA no necesita leer toda la historia de tu negocio en tu página: muchas veces le basta con tu ficha y con lo que dicen los clientes.

Las señales que mandan: proximidad geográfica y reputación de clientes. Un restaurante con 4,8 estrellas y 100 reseñas a un kilómetro saldrá antes que uno con 4,5 a cinco kilómetros, aunque este último tenga SEO on-page

impecable. La IA interpreta lenguaje coloquial («barato», «bueno para niños», «cerca de mí») y sabe distinguir matices locales. La misma búsqueda de «mejor bar de tapas» dará resultados distintos en Barcelona y Madrid, adaptando incluso la descripción.

Las tácticas: perfil de Google Business optimizado al 100% como si fuera tu nueva homepage, información NAP correcta y consistente, categoría apropiada, descripción clara, horarios actualizados, fotos atractivas y reales. Estrategia activa de reseñas: después de cada venta o servicio satisfactorio, invita al cliente a dejar su opinión; responde todas, buenas o malas, con profesionalismo. Y recuerda que más allá de la calificación numérica, los LLMs leen el texto de las reseñas. «La comida excelente y el dueño me recomendó un vino perfecto para nuestro aniversario» añade contexto que puede asomar en respuestas de IA.

Consistencia en todos los directorios: si en Google te llamas «Peluquería StilNovo» pero en Facebook sales como «Salon Stíl Novo» y en Yelp como «Stil Novo Peluqueros», confundes al algoritmo. Unifica en Google, Facebook, TripAdvisor, Yelp y páginas amarillas online. Y no abandones tu web: crea contenido orientado a búsquedas locales específicas que sirva como respaldo cuando la IA amplía detalles. Para saber en qué búsquedas el AI Overview de Google ya cita tu web, con qué página y frente a qué competidores, en Antropus usamos AI Visibility.

Para negocios con múltiples ubicaciones, el reto es doble: optimizar localmente cada sede mientras construyes la marca global. Que la IA entienda que tus franquicias en distintas ciudades son parte de la misma red con cierta reputación general.

GEO E-commerce: productos como entidades

En e-commerce, la IA conoce atributos estructurados de los productos (precio, marca, disponibilidad), confía en fuentes como Google Merchant Center y el Shopping Graph, y se informa en reviews agregadas tanto de expertos como de clientes. Para muchas consultas transaccionales, confecciona una especie de mini-vitrina con productos y fuentes.

El reto: cómo lograr que tus productos aparezcan en esa vitrina en lugar de los de Amazon. La respuesta está en los datos y la diferenciación.

Datos estructurados exhaustivos: implementa schema Product en todas tus fichas con nombre, descripción, precio, disponibilidad, marca, imágenes y AggregateRating si tienes reseñas. No hay confirmación oficial de que las IAs generativas utilicen este marcado directamente para redactar respuestas, pero sí ayuda a Google a entender mejor tu catálogo y a reforzar la información que puede aprovechar en Search, Shopping y otras superficies comerciales.

Contenido único en fichas, no solo marketing: si todos los sitios ponen la misma descripción genérica del fabri-

cante, la IA no tiene razón para elegirte. Añade valor: un párrafo de «Consejo del experto», comparaciones contra modelos similares, vídeos demostrativos. Ese contenido extra es el diferencial por el que el modelo te escoge.

Estrategia de reviews y UGC: facilita y muestra opiniones de clientes en tu web, y márcalas correctamente cuando proceda. Cuando varias reseñas coinciden en un detalle como «la talla va pequeña, pide una más», no solo ayudas al usuario: también añades contexto valioso sobre el producto. Ese tipo de señales prácticas puede reforzar la interpretación y la visibilidad de tu información en búsqueda.

Comparativas y guías en tu blog: «Mejores [categoría] de 2026» o «¿Qué [producto] comprar según [necesidad]?». Incluye tus productos si corresponde, pero sé objetivo. Si tu contenido es bueno, la IA lo usará.

Y especialización sobre generalismo: compitiendo con Amazon por «mejores zapatillas de running» lo tienes crudo. Pero quizá nadie ha cubierto «mejores zapatillas de trail para ultramaratón en alta montaña». En consultas nicho, las fuentes nicho brillan. Es preferible ser el rey de una categoría pequeña que invisible en la masiva.

GEO SaaS: la guerra de «alternativas a» y «vs»

En SaaS y B2B tech, las consultas suelen ser investigativas y comparativas. La IA responde listando 2-3 opciones con breve descripción de cada una, citando alguna

fuente notable para respaldar la recomendación, y señalando diferencias clave si la pregunta es comparativa.

Para hacer todo eso, necesita materia prima de calidad: listas comparativas actualizadas, contenido técnico de la empresa (blogs, documentación, whitepapers), reviews B2B en plataformas como G2 o Capterra y señales de entidad (Wikipedia, Crunchbase, menciones en prensa).

Prioridad número uno: páginas «Alternativas a [Líder]» y «[Tu Producto] vs [Competidor]». No basta con decir «somos la mejor alternativa». Estructura por funcionalidad, da tablas comparativas y menciona a todos los players relevantes. La IA podría usar tu página para responder a un usuario que pidió alternativas, y ahí te verás incluido porque tú mismo mencionaste al resto.

Contenido BOFU con datos reales: un caso de estudio concreto con «38% menos churn» supera a diez posts genéricos. Esa estadística podría salir en la respuesta con cita incluida hacia tu blog. FAQs técnicas exhaustivas: «¿Cómo se integra [Tu Software] con Salesforce?» respondido con todo detalle puede ser el texto recuperado y citado literalmente.

Fortalece tu perfil de empresa: Crunchbase como mínimo, Wikipedia si es viable y schema Organization en tu web. Cuanto más «real» parezcas para Google, más confianza para recomendarte. Y mantén monitoreo periódico de tu sector en IA. Pregunta por tu categoría y por tu empresa para saber cómo te ve el modelo. En Antropus

ese monitoreo es el Prompt Tracker: los prompts de tu categoría y los de tu marca, lanzados a varios motores, para ver si te mencionan, te recomiendan y te citan.

GEO B2B y servicios: autoridad personal como moneda

Empresas de servicios —agencias, consultoras, firmas legales y proveedores industriales— venden confianza, experiencia y resultados a medida. La IA evalúa a un proveedor de servicios buscando pruebas de autoridad y satisfacciones previas: rankings en portales especializados, casos de estudio publicados, testimonios, premios, certificaciones y menciones notables.

Casos de éxito públicos y detallados son el contenido estrella. No un párrafo genérico de «Cliente del sector retail aumentó sus ventas». Necesitas historias con todas las letras: nombra al cliente, explica el desafío, qué hiciste y los resultados en números concretos. Cuanto más eco tengan en medios o blogs, más huella para la IA.

Testimonios verificables con nombre, cargo y empresa. «Gracias a [Agencia X] incrementamos un 30% la conversión» —John Doe, CMO de Empresa Y. Si están en texto HTML en tu web, podrían aparecer en una respuesta de IA como respaldo. Plataformas como Clutch permiten reviews extensas que suelen ocupar posiciones prominentes en Google.

Credenciales y especializaciones visibles: certificaciones, premios, partners tecnológicos, todo claramente

mencionado online. Cada logro es un dato que la IA puede referenciar.

Y marca personal de líderes: si logras que tu CEO o expertos sean reconocidos en la industria (artículos de opinión, podcasts, eventos), ese prestigio se transfiere a la empresa. La IA asocia contenido técnico serio con credibilidad para recomendar.

El riesgo principal: invisibilidad por hermetismo. Si haces gran trabajo pero nunca publicas nada, la IA no lo sabrá. Encuentra material publicable sin violar acuerdos ni secretos.

Las realidades sectoriales que debes aceptar

Si en tu vertical falta data pública y nadie publica comparativas sobre «mejor proveedor de tornillos industriales», la IA no se va a mojar. Responderá diplomáticamente sin jugársela. Tu estrategia a largo plazo es convertirte en la fuente de información del sector.

En sectores saturados con narrativas dominantes, la IA refleja la tendencia general de los datos. Cambiar esa percepción es pelear contra años de contenido consolidado. Enfoca en nichos o en diferenciarte en algo muy específico.

Siguen existiendo dos límites importantes. Por un lado, una IA entrenada con una visión global puede infravalorar marcas locales si apenas aparecen en fuentes amplias o en contenidos en inglés. Por otro, las plataformas aplican reglas internas que condicionan qué fuentes

muestran o citan, sobre todo en temas sensibles como salud o finanzas.

La decisión pragmática: identifica las batallas que puedes ganar. Domina las consultas de nicho donde tu especialización te da ventaja. Es ahí donde el GEO por vertical marca la diferencia real.

Con las estrategias por plataforma y vertical cubiertas, estamos listos para la Parte VIII: pasar del plan al sistema. En los próximos capítulos veremos el roadmap operativo de 90 días, la automatización del pipeline de medición, y cómo empaquetar y vender todo esto como servicio profesional.

PARTE VIII

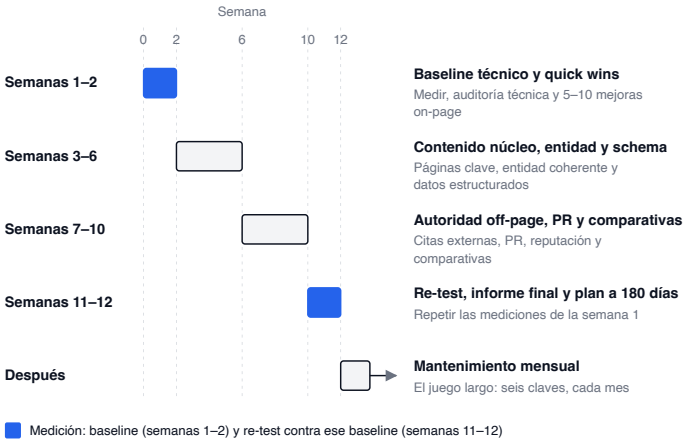
Ejecución, automatización y negocio

Del plan al sistema. Del sistema al servicio vendible.

CAPÍTULO 21

Roadmap operativo

90 días de GEO + mantenimiento



Los 90 días de GEO, fase a fase. Fuente: elaboración propia a partir del texto del capítulo 21.

Semanas 1-2: Baseline técnico y quick wins

El paso uno de cualquier plan GEO serio es obtener un baseline claro. Si no, vas volando a ciegas. Hay un patrón clásico en proyectos SEO que se repite con vergonzosa frecuencia: equipos puliendo el copywriting de la home mientras un metatag mal configurado bloquea la

indexación de un tercio del sitio. Sin auditoría previa, ese bloqueo puede pasar desapercibido durante meses. La acción menos glamurosa (revisar el robots.txt y los meta robots) es a menudo la que más mueve la aguja. Por eso, antes de invertir en estrategia, mide.

Antes de tocar nada, mide dónde estás: tráfico orgánico total, tráfico por página clave, posiciones de tus 10 keywords críticas, CTR orgánico, conversiones asistidas por orgánico. Incluye cualquier indicador de visibilidad en IA disponible, ¿tu marca aparece en alguna respuesta de ChatGPT o Copilot? Anótalo con contexto. Esta línea base te servirá para comparar en la semana 11-12 y demostrar avances con números reales, no sensaciones. Para la parte de IA, ese baseline es el que lanzamos con el Prompt Tracker de Antropus: el mismo set de prompts en ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview, midiendo si te mencionan, te recomiendan y te citan.

Después, la auditoría técnica sin piedad. Crawl completo del sitio identificando errores graves: páginas con status 4xx/5xx, huérfanas, bloqueadas por robots.txt o noindex accidental. Revisión de Search Console para cobertura e indexación. Examen de fundamentos: sitemap XML, Core Web Vitals, compatibilidad móvil. Todo documentado y priorizado en tres niveles: bloqueantes absolutos que requieren solución inmediata, problemas importantes a resolver pronto y mejoras menores para más adelante.

Con el diagnóstico en mano, los quick wins on-page: reescribir meta titles en páginas con CTR bajo, fusionar páginas que canibalizan, añadir contenido relevante que faltaba y corregir keywords mal elegidas en títulos. De 5 a 10 mejoras en producción que pueden dar pequeños saltos en días o semanas.

Semanas 3-6: Contenido núcleo, entidad y schema

Ni Google ni ChatGPT van a entender tu negocio por arte de magia. Tienes que dárselo mascado. Hay un error común entre startups con producto técnicamente sofisticado: asumir que la aplicación «se explica sola» por lo innovadora que es. La consecuencia previsible es no publicar nunca una página clara que cuente qué hace la plataforma, para quién y por qué. En las respuestas de los motores generativos sobre su categoría, esas empresas no aparecen, mientras la IA recomienda a competidores que sí construyeron contenido base.

La paradoja: cuanto más innovador es tu producto, más contenido base necesitas, no menos. La IA solo recomienda lo que entiende.

Identifica las 5 a 10 páginas más importantes que deberían existir en tu sitio. Si alguna no existe, créala. Si existe pero es floja, refuérzala. La página «Quiénes Somos» tiene que ser robusta, con historia, logros, equipo, credenciales y casos de éxito. No la típica misión/visión vacía. Páginas detalladas de productos o servicios con descripción clara, características técnicas, beneficios,

FAQs específicas y testimonios. Contenido pilar sobre los temas centrales de tu sector.

En paralelo, define tu entidad con consistencia: mismo nombre oficial en toda la web, detalles clave de la empresa (año, sede, tamaño, sector, clientes destacados, certificaciones), enlaces a fuentes externas fiables que hablen de ti, y contexto semántico que ubique tu marca en su mapa conceptual. En Antropus esto queda fijado en la ficha de entidad: nombre canónico, descripción, texto de desambiguación, alias y sameAs, para que todo el equipo publique lo mismo.

Implementa los datos estructurados esenciales: Organization en tu home o About, Product o Service en páginas de oferta, FAQPage donde corresponda, LocalBusiness si tienes ubicaciones físicas. Todo coherente con el contenido visible y validado sin errores. Estás hablando en el idioma de la máquina: en vez de que la IA tenga que inferir datos sobre ti, se los das directamente en formato que entiende sin ambigüedad.

Semanas 7-10: Autoridad off-page, PR y comparativas

Puedes tener la mejor web, pero si nadie fuera de ella habla de ti, las IAs seguirán sin considerarte relevante. Otro patrón habitual: empresas con producto excelente y documentación técnica impecable que, aun así, son invisibles en motores generativos. Sus potenciales clientes preguntan a ChatGPT por soluciones de su categoría y la IA menciona a competidores que tienen más ruido

externo (notas de prensa, rankings sectoriales, reviews públicas) aunque su producto sea técnicamente inferior. El producto no se recomienda solo. La calidad sin distribución es invisibilidad.

Las citas externas básicas: asegura que estás listado donde debes estar. Google Business Profile al día, directorios sectoriales, plataformas de reviews como G2 o Capterra si eres B2B, perfiles en asociaciones profesionales. Cada mención verificable de tu marca en una fuente externa de calidad es un punto en el mapa de internet diciendo «esta empresa existe y hace X».

El PR inteligente: convierte tus datos o expertise en noticias. Un dato interesante de tu negocio que indique una tendencia. Un estudio breve de tu sector con información pública más tu experiencia. Un caso de éxito con cifras. Prepara uno o dos ángulos de prensa, identifica periodistas que cubran tu sector, y acércate ofreciéndoles contenido de valor, no pidiendo un favor.

La reputación online: busca tu marca en Google y mira qué ves. Responde todas las reseñas con profesionalidad. Fomenta reseñas positivas pidiendo a clientes satisfechos que te evalúen. Si encontraste artículos con información errónea, contacta diplomáticamente para solicitar actualización.

Y las comparativas: crea en tu web contenido comparando tu producto con otros de forma honesta. Identifica quiénes publican las listas «Top 5» y acércate aportando información valiosa sobre tu solución.

Semanas 11-12: Re-test, informe final y plan a 180 días

El final del sprint inicial es, de hecho, el comienzo del ciclo continuo. Repite las evaluaciones clave de la semana 1 para ver el antes/después: re-crawl técnico comparando con el inicial, métricas SEO comparadas con tu baseline, y pruebas en IAs para medir si tu presencia mejoró.

El informe tiene dos niveles. Un resumen ejecutivo que el CEO entienda en cinco minutos: 3-5 logros cuantitativos clave, impacto en objetivos de negocio, y recomendación de ejecutar el plan semestral. Y un informe técnico detallado, con todas las implementaciones realizadas, métricas antes vs después, resultados GEO específicos, anomalías detectadas y conclusiones.

El plan a 180 días extiende y prioriza lo pendiente. Identifica áreas de oportunidad de tu análisis. Define 2-3 bloques estratégicos (contenido, autoridad, técnico). Lista iniciativas concretas con timing, responsables y recursos necesarios. Y marca 3-5 KPI objetivo para los próximos seis meses.

Ese plan es el que montamos con el GEO Strategy Builder de Antropus: un roadmap con las acciones priorizadas a 30, 60 y 90 días, 6 y 12 meses, en un documento editable que exportas en PDF o Word.

La reunión final busca un resultado concreto: compromiso de los decision-makers para seguir adelante. Deja claro qué pasa si se detiene el esfuerzo: el terreno gana-

do puede perderse en meses. Y que el plan no es garantía mágica, pero sin él seguro te quedas sin rumbo.

Mantenimiento mensual: el juego largo

Hay un patrón triste pero frecuente: empresas que tras tres meses de trabajo intenso consiguen resultados visibles, declaran victoria y sueltan el volante. En seis o doce meses, lo conseguido se desvanece, a veces de forma silenciosa, sin que nadie del equipo lo detecte hasta que el deterioro es evidente. Google cambia. Los competidores publican. La IA reentrena su modelo. Lo que un día te citaba, mañana puede ignorarte. Sin mantenimiento, el SEO y el GEO se descomponen.

El mantenimiento se estructura en seis claves. Revisión mensual de desempeño: una reunión al mes con responsable SEO, contenido y desarrollo para revisar KPIs, estado de indexación, cambios recientes en la web y tareas pendientes.

Control de cambios: ningún cambio relevante en el sitio se implementa sin revisión SEO previa. Una checklist breve que el equipo de desarrollo consulta antes de cada despliegue: ¿cambia URLs? ¿Elimina contenido? ¿Afecta títulos o headings? ¿Impacta velocidad? ¿Puede bloquear al crawler?

Monitoreo automatizado: alertas de Search Console, servicios de monitorización de cambios, seguimiento semanal de rankings clave, monitores de uptime y perfor-

mance. Herramientas que vigilen lo básico mientras tú te concentras en estrategia.

Iteración constante de contenido: cada mes al menos una pieza actualizada o mejorada. Revisiones trimestrales de las páginas que más tráfico traen. Keyword research ligero cada seis meses para detectar oportunidades emergentes. Y «prompt research», prestar atención a qué preguntan los usuarios a las IAs y crear contenido que responda a esas preguntas.

Capacitación y relevo: documenta los procesos clave, forma a un equipo transversal en los fundamentos y asegúrate de que exista un segundo responsable identificado. El SEO y el GEO no deberían depender de una sola persona ni vivirse como «el proyecto de fulanito», sino consolidarse como una disciplina integrada en la empresa.

Adaptabilidad constante: seguir fuentes de la industria, incluir un punto de novedades en alguna reunión periódica, y tener flexibilidad para ajustar el plan cuando las circunstancias cambien.



Las seis claves del mantenimiento mensual. Fuente: elaboración propia a partir del texto del capítulo 21.

En GEO, como en SEO, la única constante es el cambio. Por eso no conviene aferrarse ciegamente al plan inicial. Hace falta la humildad suficiente para recalibrar la brújula cuando sea necesario. El SEO y el GEO no son tareas que un día se acaban, sino disciplinas que deben administrarse y mantenerse en el tiempo.

En el siguiente capítulo vamos a automatizar todo esto: del Excel manual al pipeline de IA que ejecuta consultas, parsea respuestas, extrae métricas y genera informes reproducibles.

CAPÍTULO 22

Automatización y Tooling

Del Excel manual al pipeline IA

Adiós a las capturas sueltas

Si mides tu visibilidad en IA haciendo cuatro preguntas a ChatGPT un martes y cuatro más un jueves, estás a merced del azar. Los LLMs son probabilísticos: puedes preguntar lo mismo dos veces y que citen fuentes distintas. Si te conformas con muestras pequeñas, podrías autoengañarte con conclusiones erróneas.

La única forma de tener una imagen fiel es a escala y de forma continua: lanzar decenas de consultas relevantes a los modelos, recopilar sus respuestas y analizar patrones cuantitativos. Esto es lo que vamos a automatizar.

La arquitectura del pipeline

El sistema completo se apoya en una serie de piezas que trabajan de forma encadenada. La primera es la lista de preguntas que vas a monitorizar: las consultas clave que representan tu negocio, tu sector y tu nicho. Para construirla, puedes partir de tus palabras clave más importantes, de las preguntas frecuentes de tus clientes y de

las búsquedas de marca, tanto tuyas como de tus competidores. Lo importante es que cada consulta quede bien definida, clasificada y siempre formulada de la misma manera, para poder repetir el seguimiento con consistencia. Como punto de partida, suele bastar con unas cincuenta consultas estratégicas. Para Google hay un atajo: AI Visibility, en Antropus, te enseña en qué búsquedas ya te cita el AI Overview, con qué página y frente a qué competidores, y muchas son consultas que no habrías puesto en tu lista.

La segunda pieza es la automatización. En lugar de lanzar estas consultas a mano, conviene programarlas para que se ejecuten con una frecuencia fija. Así puedes observar la evolución en el tiempo sin depender de la memoria, la disponibilidad o la disciplina de una persona. En muchos casos, una revisión semanal es suficiente para detectar cambios relevantes sin generar un volumen innecesario de trabajo o de coste.

La tercera pieza es la recogida de respuestas. El sistema lanza cada consulta al motor elegido y guarda lo que este devuelve. Lo importante aquí no es tanto la vía técnica como el rigor del registro: cada respuesta debe conservarse con su fecha, el modelo utilizado, la consulta exacta y las fuentes citadas, si las hubiera. Así, aunque el proceso se interrumpa o falle en algún punto, no pierdes la información ya recogida y puedes reconstruir después qué ocurrió en cada medición.

La cuarta pieza es el análisis. Una vez recogidas las respuestas, toca convertirlas en información útil. Ahí es

donde se calcula qué dominios aparecen citados, con qué frecuencia, en qué categorías tienes más visibilidad y cuál es tu Share of Model, tanto de forma global como por tipo de consulta. El objetivo es pasar de una colección de respuestas sueltas a una lectura clara de patrones, tendencias y oportunidades.

Es exactamente lo que automatizamos en Antropus con el Prompt Tracker: el mismo set de prompts, lanzado a ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview, con cada respuesta parseada para saber si te mencionan, te recomiendan y te citan.

La quinta pieza es el informe. Los datos solo sirven de verdad cuando se convierten en una lectura útil para tomar decisiones. Por eso, el sistema debe terminar en un documento claro que resuma los principales hallazgos: tu Share of Model global, los dominios que más aparecen, las diferencias entre categorías, las coocurrencias con otras marcas, el tono de las menciones y las recomendaciones prioritarias. La clave es que ese informe pueda repetirse siempre del mismo modo, para que el seguimiento sea comparable y no dependa de interpretaciones cambiantes. En el Prompt Tracker de Antropus ese informe se exporta en Excel y en PDF con tu marca, listo para presentárselo al cliente.

Visto de forma simple, el proceso funciona así: primero defines las preguntas que quieres monitorizar; después las lanzas de forma periódica; a continuación guardas las respuestas; luego analizas qué marcas, fuentes y pa-

trones aparecen; y, por último, conviertes todo eso en un informe útil para decidir qué hacer.

Preguntas clave (prompts) → ejecución periódica → recogida de respuestas → análisis de patrones → informe final



Del Excel manual al pipeline. Fuente: elaboración propia a partir del texto del capítulo 22.

Lo que uso yo, sin filtros

Llevo años montando este tipo de pipelines para clientes, y la pregunta que me hacen siempre en eventos es la misma: «¿Qué stack usas tú?». Te lo cuento, porque después de 21 capítulos creo que te lo has ganado.

El 90% de mi operativa diaria pasa por Antropus. No solo la parte GEO (medición de Share of Model en los cuatro motores, auditoría de los cuatro pilares de visibilidad, generación metodológica de prompt sets, matriz

competitiva Google vs IA y reporting automatizado) sino también la operación de la firma: CRM con pipeline de clientes, gestión de proyectos con roles de Project Manager y asignaciones, control de tareas del equipo, plantillas de propuestas y reporting unificado. Antropus es la plataforma SaaS que desarrollamos en Elevam precisamente para resolver dos problemas a la vez: la falta de herramientas GEO serias en el mercado y la fragmentación operativa de las firmas y freelances que trabajan SEO y GEO. Antes pagábamos suscripciones a cuatro o cinco herramientas distintas y exportábamos datos entre ellas a mano. Ese problema se acabó.

Lo que sí uso fuera de Antropus: Google Search Console (gratuito, imprescindible para los datos crudos que Google solo expone ahí), Screaming Frog para auditorías técnicas puntuales que requieren crawling pesado, Google Workspace para correo y documentos compartidos, una herramienta de firma electrónica para contratos, y LastPass para gestión de contraseñas del equipo. Poco más.

No menciono Antropus porque sea mía. La menciono porque montar todo lo que verás en el resto del capítulo (orquestación multi-LLM, cache granular, observabilidad, scoring determinista, exportación de informes, además de un CRM con gestión de proyectos por encima) son doce meses de desarrollo serio si tu equipo tiene perfiles técnicos disponibles. Antropus es lo que usamos para que ese tiempo no se multiplique por cada

cliente que entra en cartera y para no tener al equipo saltando entre cinco pestañas para hacer su trabajo.

Logging estructurado: la base de todo

Cada entrada del log debe incluir como mínimo: identificador del prompt, texto de la consulta, modelo y versión usados, fecha y hora, texto completo de la respuesta, y lista de citas con URL y título de cada fuente referenciada. Guarda la mayor cantidad de información disponible aunque inicialmente solo analices las URLs. Tener el texto completo te permitirá después hacer análisis de sentimiento, posición de mención y co-mentions.

Además, el sistema debe ser resistente a los errores y a las limitaciones de uso. Si una consulta falla, hay que dejar constancia de ello, señalarla como fallida y seguir adelante con las demás. Un fallo aislado no debería interrumpir todo el flujo de seguimiento.

De los datos al informe reproducible

El informe automatizado elimina el copy-paste y reduce los errores de transcripción. A partir de los CSV de métricas, un script en Python puede rellenar una plantilla con los valores ya calculados: Share of Model global y por categoría, top 10 de dominios citados, variación respecto al periodo anterior, análisis de sentimiento y recomendaciones derivadas de los hallazgos.

También puede incorporar gráficos generados con matplotlib, como barras de dominios citados o la evolución

temporal del SoM. Conviene, además, versionar y documentar el proceso para que cualquier miembro del equipo pueda regenerar el mismo informe sin depender de la persona que lo montó. Y, si tiene sentido, puedes automatizar también la distribución: por ejemplo, con un bot que envíe cada lunes el informe a los interesados.

La automatización no sustituye al criterio. Antes de difundir cualquier informe, revísalo siempre con ojo crítico. Los valores pueden ser correctos en términos técnicos y, aun así, no tener sentido en su contexto de negocio.

Agentes de IA en auditorías: útiles pero no mágicos

Un agente de IA, ya sea una herramienta tipo AutoGPT o un framework como LangChain Agents puede tomar una lista de páginas de tu sitio, analizar su contenido con un LLM, buscar competidores, comparar resultados y consolidar hallazgos. La promesa es atractiva, porque permite automatizar muchas horas de auditoría.

El problema es que llevar todo eso a un entorno fiable resulta mucho más complejo de lo que parece. El agente puede alucinar problemas que no existen, entrar en bucles o generar informes vagos y poco accionables. Cada transición entre pasos añade un punto potencial de fallo. Cuando algo se rompe, depurarlo puede volverse muy difícil.

Si decides utilizarlos, conviene seguir algunas reglas básicas. Asigna a cada agente una función específica, en lugar de confiar en un agente generalista que intente hacerlo todo. Haz que intercambien resultados en formatos estructurados, como JSON, y no en párrafos narrativos. Limita el número de iteraciones para evitar bucles. Mantén una orquestación central con logging exhaustivo. Deja siempre la validación final en manos humanas: las sugerencias de la IA deben tratarse como hipótesis a verificar, no como verdades absolutas.

Donde mejor funcionan es en tareas acotadas, repetitivas y verificables. Por ejemplo, un agente especializado que audite elementos on-page de una URL: extrae el título, la meta description y el H1 con una función en Python, y después pasa esos datos al LLM para detectar problemas básicos. Ese tipo de uso tiene sentido. Lo que conviene evitar es el extremo opuesto: un agente generalista al que se le delega una estrategia SEO completa. Con esa lógica planteamos Antropus MCP: conectas Antropus a ChatGPT, Claude o Codex, pides el trabajo en la conversación y la plataforma pone los datos y las herramientas; la revisión final sigue siendo tuya.

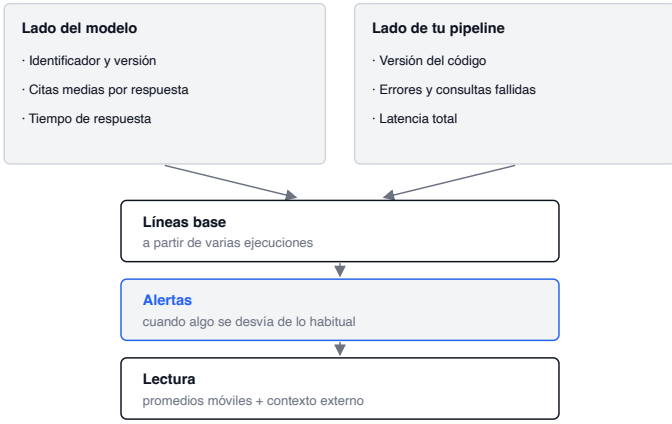
Observabilidad: saber por qué cambian tus métricas

Si tu SoM baja del 16 % al 10 %, necesitas entender qué ha cambiado. Puede deberse a una actualización del modelo, a que un competidor haya publicado contenido fuerte, a un fallo en tu propio sistema o, simplemente, a

variación aleatoria. Sin observabilidad, reaccionarás a ciegas y sacarás conclusiones equivocadas.

Por eso conviene registrar señales a ambos lados del proceso. Del lado del modelo, anota el identificador y la versión usada en cada consulta, el número medio de citas por respuesta y el tiempo de respuesta. Del lado de tu pipeline, registra la versión del código en cada ejecución, los errores, las consultas fallidas y la latencia total.

Después, establece líneas base a partir de varias ejecuciones. Por ejemplo: «este modelo suele dar unas dos citas por respuesta» o «mi SoM normalmente fluctúa entre el 10 % y el 15 % de una semana a otra sin cambios relevantes». A partir de ahí, configura alertas cuando algo se desvíe de ese comportamiento habitual. Si el promedio de citas cae de forma brusca, es posible que haya cambiado el comportamiento de citación del modelo, y no necesariamente que tu estrategia haya empeorado.



Observabilidad: de las señales a la alerta. Fuente: elaboración propia a partir del texto del capítulo 22.

Para no sobrerreaccionar al ruido, trabaja con promedios móviles y no con valores aislados. Añade siempre contexto externo a la lectura: si un proveedor como OpenAI lanza una nueva versión justo cuando cambian tus métricas, esa coincidencia puede explicar parte de la variación.

La observabilidad no te da control sobre las actualizaciones de la IA, te permite detectar cambios a tiempo y adaptar tu estrategia con más criterio.

Docker y versionado: que tu pipeline sobreviva

Si quieres que tu pipeline sea realmente reproducible, necesita poder ejecutarse de la misma forma en cualquier entorno. Empaquetarlo en un contenedor Docker ayuda precisamente a eso: evita diferencias entre má-

quinas y reduce los problemas de configuración. A la vez, conviene fijar las versiones de las dependencias para que el sistema no cambie de comportamiento sin que nadie lo haya tocado.

El código también debe estar versionado. Usar Git te permite registrar cada cambio, saber qué versión estaba activa en cada ejecución y mantener una trazabilidad clara de la evolución del pipeline.

Si trabajas con orquestadores como Airflow o Prefect, define el flujo en pasos bien separados. Primero ejecutar las consultas, después procesar los logs y, por último, generar el informe. Cuando cada fase queda aislada y registrada por separado, resulta mucho más fácil detectar dónde falló algo sin tener que revisar todo el sistema de punta a punta.

Y no olvides la documentación. Un buen README con instrucciones de instalación, ejecución y resolución de problemas puede marcar la diferencia entre un sistema reutilizable y uno que solo entiende quien lo creó. La regla es simple: cualquier persona del equipo debería poder levantarlo sin tener que adivinar configuraciones.

En el siguiente capítulo salimos de lo técnico para entrar en lo comercial: cómo empaquetar todo lo que has aprendido (auditoría, medición, ejecución, automatización) como un servicio profesional vendible a clientes que todavía no saben que lo necesitan.

CAPÍTULO 23

Packaging y venta de servicios GEO+SEO

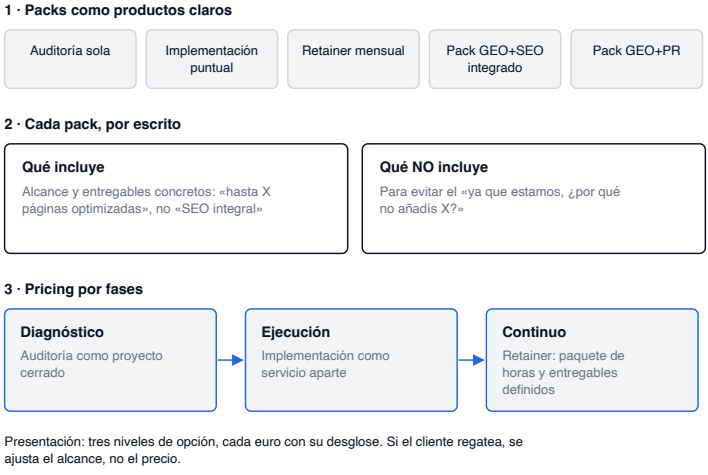
Vende estrategia, no humo

Uno de los errores más habituales es vender packs milagro y baratos que prometen de todo, pero no resuelven nada. La oferta debe presentarse como una solución con criterio, no como un combo mágico. Si un prospecto solo busca precio y promesas vacías, conviene educarlo desde el inicio o renunciar a esa oportunidad antes de arrastrar un proyecto condenado desde el principio.

Define y empaqueta tus servicios como productos claros con alcance y entregables delimitados. Ejemplos de packs bien articulados: auditoría sola (diagnóstico y roadmap sin implementación), implementación puntual (ejecutar recomendaciones en X semanas), retainer mensual (mantenimiento SEO+GEO continuo), pack GEO+SEO integrado (técnica clásica más optimización generativa), o pack GEO+PR (SEO generativo más acciones de PR digital para conseguir menciones que las IAs valoren).

Sé extremadamente explícito en qué incluye y qué no incluye cada oferta. Si vendes una auditoría GEO, detalla

qué analizarás y qué entregable obtendrá el cliente. Si ofreces implementación, define el alcance: «hasta X páginas optimizadas, integración de schema FAQ, creación de N contenidos LLM-friendly». Incluye también lo que NO harás para evitar el clásico «ya que estamos, ¿por qué no añadís X?». Evita términos vagos tipo «SEO integral». Mejor algo concreto como «Optimización de 5 contenidos existentes para que sean elegibles como fuentes en respuestas de IA».



Estructura de una propuesta de servicio GEO+SEO. Fuente: elaboración propia a partir del texto del capítulo 23.

No todos los clientes necesitan todo desde el día uno. Quizá una startup sin presencia web deba empezar por SEO básico antes de escalar a GEO. Ofrécele un roadmap evolutivo: primero lo fundamental y luego el componen-

te generativo. Eso demuestra honestidad y visión a largo plazo.

Pricing basado en valor, no en keywords

Cobrar «hasta X palabras clave por Y euros al mes» es un sinsentido en 2026. Hay miles de variantes de búsqueda. ¿De verdad dejarás de optimizar términos relevantes porque el cliente «pagó 10 keywords»? También huye de tarifas ridículas tipo «SEO básico por 299 al mes» sin definir alcance. Eso se traduce en una o dos horas reales de trabajo con cero impacto.

Conviene ser radicalmente honesto: un buen SEO o GEO requiere tiempo de perfiles cualificados, y eso tiene un coste. Si un cliente dice que otra agencia se lo hace por 500 euros, la respuesta debería apoyarse en hechos: por ese precio, lo más probable es que reciba cuatro acciones superficiales que no aporten nada o, peor aún, que terminen perjudicándolo.

Estructura tu pricing por fases separando diagnóstico de ejecución. La auditoría como proyecto cerrado, la implementación como servicio aparte. Para trabajo continuo, retainer mensual con paquete de horas y entregables definidos. Si consideras modelos híbridos con bonus por desempeño, que sea sobre métricas realistas bajo tu control razonable. Jamás prometas esquemas tipo «X euros por cada top 1 en Google».

La métrica que mejor ayuda a justificar el precio es el Share of Model: el equivalente al share of voice en moto-

res generativos. Puedes planteárselo al cliente de forma muy simple: «¿Cuánto pagarías por aparecer en 3 de cada 10 respuestas cuando alguien pregunta cuál es el mejor [tu producto]?» Esa presencia tiene un valor enorme. Conseguir 200 recomendaciones personalizadas de ese tipo mediante anuncios o PR tradicional probablemente costaría mucho más que la inversión en GEO.

Presenta tu propuesta con opciones claras, tres niveles por ejemplo y justifica cada euro con desglose de lo que obtiene. Si un cliente quiere regatear, defiende el valor ajustando el alcance, no tirando el precio.

Las matemáticas del retainer GEO (lo que nadie te cuenta)

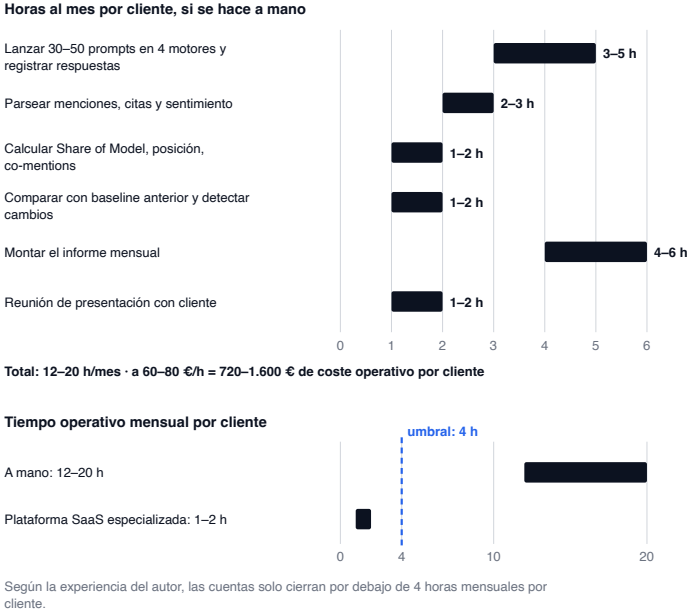
Hay una conversación que las agencias raramente tienen en voz alta: el coste real de operar GEO si lo haces manualmente. Vamos a hacerla aquí.

Un retainer GEO bien empaquetado puede facturar entre 1.500 y 4.000 € al mes según alcance. Una empresa mediana en un sector competitivo, con seguimiento mensual completo, suele cerrar en el rango medio-alto. Si pretendes ejecutarlo manualmente, la carga operativa por cliente está entre 12 y 20 horas mensuales, repartidas así:

- Lanzar 30-50 prompts en 4 motores y registrar respuestas: 3-5 horas
- Parsear menciones, citas y sentimiento: 2-3 horas
- Calcular Share of Model, posición, co-mentions: 1-2 horas

- Comparar con baseline anterior y detectar cambios: 1-2 horas
- Montar el informe mensual: 4-6 horas
- Reunión de presentación con cliente: 1-2 horas

A 60-80 € la hora de un técnico cualificado, estás hablando de 720 a 1.600 € de coste operativo por cliente y mes. El margen se queda en aire.



Las matemáticas del retainer GEO. Fuente: elaboración propia a partir del texto del capítulo 23.

Las agencias que están escalando GEO en 2026 no lo hacen con becarios y Excel. Lo hacen de una de estas dos formas:

- Pipelines automatizados propios (Capítulo 22) con un equipo técnico que los mantiene
- Plataformas SaaS especializadas que reducen ese tiempo operativo a 1-2 horas mensuales por cliente, dejando al equipo concentrarse en estrategia y acción

Es el problema para el que construimos Antropus: el Prompt Tracker lanza el set de prompts a los motores y devuelve el informe en Excel y PDF con la marca de tu cliente, para que las horas del equipo se vayan al análisis y no al copy-paste.

En mi experiencia con agencias que están escalando GEO, las cuentas solo cierran cuando el tiempo operativo por cliente baja de 4 horas mensuales. Por encima de eso, o subes precio o pierdes dinero. Por debajo, el modelo de negocio funciona y el equipo se concentra en lo que sí aporta valor: el criterio estratégico, la coordinación con el cliente y la mejora real de su visibilidad.

Es la única forma de que GEO sea un servicio rentable a escala. Lo demás es vender humo o vivir del margen del fundador.

Contrato y gestión de expectativas

En servicios GEO+SEO, manejar expectativas es tan importante como el trabajo técnico. Tu contrato debe blindarte contra malentendidos y contra promesas que tú nunca hiciste pero el cliente asume.

Incluye una cláusula explícita de no garantía de resultados: «No garantizamos rankings específicos ni menciones concretas. Trabajamos para aumentar probabilidades y visibilidad, pero los motores de búsqueda (especialmente los generativos) tienen factores fuera de nuestro control.» Las IAs son estocásticas: funcionan por probabilidad, no con reglas fijas. Un día el modelo te cita, al siguiente quizás no, incluso con el mismo prompt.

Las cláusulas GEO específicas a incluir: explicación del modelo no determinista para que el cliente entienda por qué no puede exigir respuestas fijas. Definición de Share of Attention como KPI de seguimiento, no como garantía contractual. Compromiso de trazabilidad con informes mensuales detallando acciones y resultados. Cláusula de ajustes continuos para adaptar el plan conforme evolucione el ecosistema sin renegociar todo el contrato. Y cláusula de white-hat: si el cliente exige prácticas contrarias a tu recomendación, puedes resolver el contrato sin responsabilidad.

El informe mensual que importa

Un director de tecnología (CTO) o un director de marketing (CMO) senior no quiere leer que este mes se han escrito 10.000 palabras ni revisar treinta páginas de métricas de vanidad. Si una métrica no sirve para tomar una decisión o activar una acción, sobra en el informe.

Un informe mensual efectivo debería concentrarse en cuatro bloques. Primero, el Share of Model, como indicador de presencia relativa en motores generativos, desglosado por plataforma y por grupos de preguntas, acompañado de una gráfica de tendencia y de una lectura en lenguaje de negocio. Segundo, las citas y menciones logradas, con una tabla que incluya la pregunta, el modelo, la fecha, el fragmento de respuesta y la fuente citada. Tercero, las acciones realizadas, seleccionando solo las que realmente explican o apoyan los resultados del mes. Y cuarto, el plan próximo, es decir, las acciones prioritarias del siguiente periodo en función de lo observado. La parte competitiva del Share of Model, con posición y tono frente a cada rival, es lo que mide Competitor Intelligence en Antropus en ChatGPT, Claude, Gemini y Perplexity.

La regla: si el cliente lee el informe en cinco minutos y sabe cómo vamos, qué hicimos y hacia dónde vamos, has hecho bien tu trabajo.

Cómo eliminar la percepción de humo

Cuando un prospect suelte «¿No será esto puro humo?», no te lo tomes a mal. Prepárate con hechos: «¿Sabías que casi el 60% de las búsquedas ya no generan clic? Si tu contenido no se adapta a eso, pierdes visibilidad aunque rankees primero.»[5]

Ofrece una prueba en vivo: «Preguntémosle ahora mismo a ChatGPT sobre tu producto, a ver qué dice.» Si la

IA no menciona su marca, acabas de generar urgencia en 30 segundos. Si responde con algo impreciso o citando a la competencia, mejor aún.

Cuando pregunten por qué deberían creer en GEO, explica con lógica de ingeniero: «Las IAs recuperan documentos relevantes usando embeddings vectoriales, luego generan una respuesta y citan fuentes. Nuestro trabajo es asegurarnos de que en esa fase de recuperación tus contenidos aparezcan y tengan la estructura suficiente para que la IA te cite.»

Muestra un baseline contundente: «De 50 preguntas relevantes, tu marca aparece en 2 respuestas: un 4 % de Share of Model. Tu competidor aparece en 10». Cuando un cliente ve números concretos de su posición relativa, la conversación pasa de «¿sirve para algo?» a «¿cómo lo mejoramos?»

Y reformula qué es éxito: solo el 12% de las URLs citadas por ChatGPT, Gemini y Copilot estaban en el Top 10 orgánico de Google. Dominar Google ya no implica dominar respuestas de IA. El nuevo éxito es que cuando alguien pregunte a una IA por algo de tu sector, tu nombre esté en la conversación.[6]

Con este capítulo cerramos la Parte VIII. Hemos pasado del plan al sistema, del sistema al servicio vendible. En la Parte IX abordaremos las fronteras: las líneas que no debes cruzar, y el sistema de investigación continua para no quedarte fuera del modelo.

NOTAS

5. SparkToro y Datos, «2024 Zero-Click Search Study»: el 58,5 % de las búsquedas de Google en EE. UU. y el 59,7 % en la UE terminaron sin clic (sparktoro.com). La actualización de SparkToro para enero–abril de 2026 sitúa la cifra de EE. UU. en el 68 %. ↩
6. Ahrefs, «Only 12% of AI Cited URLs Rank in Google's Top 10 for the Original Prompt», agosto de 2025: media de las URLs citadas por ChatGPT, Gemini y Copilot en 15.000 prompts (ahrefs.com/blog/ai-search-overlap/). ↩

PARTE IX

Seguridad, ética y fronteras

Lo que no debes hacer. Lo que nadie te cuenta. Y lo que viene.

CAPÍTULO 24

Seguridad, ética y lo prohibido

Las líneas que no debes cruzar

La trampa siempre sale cara

Los atajos sucios con la IA llevan directo al desastre. Nada destruye más rápido la reputación digital que cruzar las líneas rojas técnicas o éticas en motores generativos. Este capítulo desmonta cinco tácticas prohibidas: tentadoras pero letales.

Táctica	Señales y vectores	Qué hacer
Manipulación de sentimiento	Reseñas clónicas, ráfagas cortas, perfiles sin historial	Fomentar reseñas auténticas; si ya pasó, admitir y revertir
Inyección indirecta de prompts	Texto oculto en CSS, comentarios HTML, caracteres especiales	Sanitizar inputs, delimitar el contenido, red teaming
Granjas de entidades y redes falsas	Estilo casi idéntico, publicaciones coordinadas, dominios sin autoridad	Desmontar por fases y construir fuentes genuinas
Redes espejo (te clonan)	Tu contenido, estructura o marca copiados en varios dominios	DMCA, reporte a Google y comunicado de sitios oficiales

Las líneas rojas, de un vistazo. Fuente: elaboración propia a partir del texto del capítulo 24.

Manipulación de sentimiento: el espejismo de la narrativa perfecta

Forzar un sentimiento artificialmente positivo es un billete al desastre. La IA huele las narrativas «demasiado perfectas» a kilómetros. Un perfil de reseñas sospechosamente perfecto dispara filtros automatizados. Los LLMs, entrenados para buscar equilibrio, pueden descartar fuentes que suenen a publirreportaje.

Hay tres bombas de tiempo. La primera es la penalización algorítmica: los filtros antifraude pueden detectar con facilidad decenas de reseñas clónicas de cinco estrellas. La segunda es el ajuste de cuentas diferido: basta con que ocurra un hecho negativo real para que la IA recalibre bruscamente su tono y provoque un reputation drift, transformando una narrativa artificialmente positiva en una corrección severa hacia lo negativo. Y la tercera es la huella del engaño: muchas reseñas falsas repiten el mismo tono, se concentran en ráfagas cortas y proceden de perfiles sin historial. Un LLM moderno ha visto miles de variantes de este patrón durante su entrenamiento y puede aprender a identificarlo.

La mitigación pasa por fomentar reseñas auténticas, no por borrar las críticas. Conviene monitorizar el sentimiento en todos los canales para detectar desequilibrios sospechosos, respaldar los mensajes con fuentes independientes y establecer una política interna de honestidad radical: lo que ocultes, la IA lo reconstruirá o lo interpretará peor. Para la web tradicional usamos Brand Monitoring, en Antropus: rastrea noticias, blogs y foros,

clasifica cada mención como positiva, neutral o negativa y avisa cuando aparecen muchas más de lo habitual.

Si ya metiste la pata y sale a la luz: admite el error sin excusas, detén y revierte la manipulación, investiga la causa raíz, comunica los cambios públicamente, y reconstruye con contenido honesto. Las marcas que sobreviven son las que rinden cuentas y mejoran.

Inyección indirecta de prompts: el nuevo vector de ataque

Alguien esconde instrucciones maliciosas dentro de contenidos aparentemente inocentes para que cuando una IA procese ese contenido, ejecute esas instrucciones sin darse cuenta.

Texto invisible en CSS oculto, comentarios HTML que el usuario no ve pero el modelo sí lee, caracteres especiales que el LLM interpreta como comandos.

El OWASP Top 10 para aplicaciones con LLM de 2025 clasificó la Prompt Injection, que incluye la indirecta, como el riesgo número uno. No es ciencia ficción: en pruebas de 2023-2024 se comprobó la viabilidad de estos ataques.[7]

¿Por qué te importa como consultor GEO? Porque si terceros pueden manipular las fuentes de datos que tu marca presenta al mundo, pueden hacer que la IA cuente mentiras sobre ti sin que lo sepas. Es un vector de sabotaje que ataca la confianza que la IA tiene en tu contenido. Por eso conviene vigilar qué cuentan de ti los mo-

tores: el Prompt Tracker de Antropus guarda lo que responde cada uno, prompt a prompt, y una afirmación nueva sobre tu marca que no sale de ninguna de tus fuentes es una señal para investigar.

Cómo blindarte: sanitiza y filtra todo input de usuarios con la misma disciplina que aplicarías contra XSS o SQL injection. No permitas HTML arbitrario en contenido público.

Delimita claramente el contenido en integraciones IA con marcadores explícitos. Actualiza librerías y herramientas. Simula ataques en tus propios sistemas (red teaming) para encontrar las grietas antes que un mal actor.

Si ocurre un ataque: aísla y elimina la amenaza inmediatamente, investiga el vector de entrada, notifica a los proveedores de IA afectados, y fortalece defensas documentando todo el incidente.

Granjas de entidades y redes falsas: fama artificial con pies de barro

Crear un ecosistema artificial de personas ficticias, blogs fantasmas, medios inventados y perfiles sociales falsos que se referencian entre sí para inflar la prominencia de una marca. Es un castillo de naipes: en cuanto Google o cualquier IA analice con detalle la calidad e independencia de esas fuentes, todo se derrumbará. Y la caída te deja peor que al inicio: bajo tierra, con fama de tramposo.

Las señales que delatan este tipo de montaje suelen ser evidentes: un lenguaje y un estilo casi idénticos en múltiples sitios, fechas y publicaciones sospechosamente coordinadas, falta de menciones fuera de la red creada para inflarlas y una autoridad prácticamente inexistente de los dominios implicados. Cuando quedas asociado a una operación así, no regresas al punto de partida: caes por debajo de cero.

La tentación es comprensible cuando la presión por destacar es alta. La fama artificial es como esteroides para un atleta: es un empujón rápido que casi siempre termina en colapso. Los buscadores y las IA mejoran cada año su olfato anti-fraude, con décadas de aprendizaje enfrentando spam web. Una sola cita en un medio reputado pesa más que cincuenta artículos en blogs fantasma.

Si heredas una red falsa: desmóntala escalonadamente mientras construyes fuentes genuinas en paralelo. Usa la herramienta Disavow para los enlaces tóxicos. Si es un competidor haciéndote guerra sucia: recopila evidencia, reporta a los motores de búsqueda y considera acciones legales si hay difamación.

Redes espejo: cuando te clonan para perjudicarte

Son los competidores o spammers que copian sistemáticamente tu contenido, tu estructura web o tu identidad de marca. Lo despliegan en múltiples dominios para ro-

barte visibilidad, confundir usuarios o diluir tu autoridad con versiones distorsionadas.

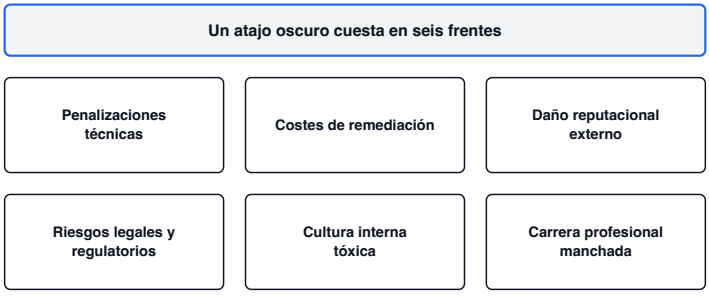
La protección proactiva pasa por adelantarte. Publica rápido y notifícalo en Search Console para ganar prioridad de indexación. Revisa con frecuencia si tu contenido está siendo copiado. Asegura dominios alternativos y nombres de usuario relevantes para que nadie se apropie de tu marca en otros canales. Y refuerza en tu web las señales técnicas de autoría, como las etiquetas canonical, el marcado del editor y las fechas visibles, para dejar claro qué contenido es original y cuál es la versión principal. Y mantén una ficha de entidad canónica (en Antropus es literalmente eso: nombre, descripción, desambiguación, alias y sameAs) que deje claro cuáles son tus perfiles oficiales.

Si ya te clonaron: envía avisos DMCA al hosting del sitio clon, reporta a Google para retirar resultados por infracción de copyright, publica un comunicado aclarando cuáles son tus sitios oficiales, y refuerza tu SEO original mientras tramitas las retiradas.

El verdadero coste del compliance: más allá de la penalización SEO

Cada atajo oscuro tiene un coste, a veces invisible al principio pero siempre real a medio plazo. Y no solo algorítmico: penalizaciones técnicas que hunden tu visibilidad. Costes de remediación que consumen muchísimo más tiempo y dinero que haber hecho las cosas bien. Daño reputacional externo con pérdida de confianza de

usuarios, clientes y socios. Riesgos legales y regulatorios con multas por publicidad engañosa, demandas por competencia desleal, y el AI Act europeo, cuyas obligaciones de transparencia se aplican desde el 2 de agosto de 2026. Efecto interno de cultura tóxica donde «vale todo mientras funcione». Y tu carrera profesional manchada en un mundo más pequeño de lo que parece.[8]



El verdadero coste de saltarse las reglas. Fuente: elaboración propia a partir del texto del capítulo 24.

Un e-commerce que inflaba contenido con IA y compraba enlaces perdió el 80% de su tráfico con un update de Google. Un consultor que fabricó su expertise con blogs falsos quedó en ridículo ante una audiencia técnica real. Una startup de reseñas falsas fue multada por la autoridad de consumo, salió en prensa como fraude, y cerró al año siguiente.

La única estrategia ganadora es entender cómo opera la IA y trabajar con ella, no contra ella. Todos los atajos tramposos chocan con la lógica de estos sistemas, que tienden a premiar el contenido útil, veraz y equilibrado. Puedes burlar un sistema durante un tiempo, pero la

banca siempre acaba ganando cuando juegas sucio en su mesa.

Si construyes tu estrategia sobre cimientos firmes (buenas prácticas, transparencia, valor real), obtendrás resultados que perduran y que podrás defender con la cabeza alta ante cualquier auditoría, sea de un algoritmo, un usuario inquisitivo o un regulador.

En el capítulo final estableceremos el sistema de investigación continua: benchmarks trimestrales, análisis de cambios por actualización de modelo, y el framework que garantiza que nunca te quedes fuera del modelo.

NOTAS

7. OWASP GenAI Security Project, «LLM01:2025 Prompt Injection» (genai.owasp.org/llmrisk/llm01-prompt-injection/). ↩
8. Reglamento (UE) 2024/1689 (AI Act), artículo 50, aplicable desde el 2 de agosto de 2026. ↩

CAPÍTULO 25

Labs GEO

Investigación continua para no quedarte fuera del modelo

Tu Top 1 orgánico no te protege si el modelo no te cita

Lo que hoy te cita, mañana puede dejar de hacerlo sin previo aviso. No por capricho, sino por la lógica del sistema. Un LLM no conserva una jerarquía fija sobre quién manda en tu sector, sino que reevalúa continuamente según las fuentes que recupera o tiene más a mano. En motores híbridos, como los apoyados en RAG, eso significa incorporar nuevas páginas, recalcular similitudes semánticas y, a veces, favorecer contenido más reciente o más relevante. Puedes destacar durante un mes y desaparecer al siguiente si un competidor publica una pieza más fuerte y el sistema la eleva a fuente de referencia.

GEO no es un proyecto que se termina. Es una disciplina que se mantiene.

Benchmarks trimestrales: las ligas de fuentes cambian sin pedirte permiso

Cada vertical tiene su propio ecosistema de fuentes. En B2B SaaS es típico ver citación de comparativas en G2 o TrustRadius. En salud, la IA prefiere Mayo Clinic o papers académicos. Si no sabes qué fuentes dominan tu sector para cada motor, estás optimizando a ciegas.

El ciclo trimestral: define un dataset fijo de 10 a 50 preguntas representativas del vertical, las que un usuario real haría cuando evalúa opciones, precios y mejores proveedores. Ejecútalas en varios motores generativos (ChatGPT, Copilot, Gemini, Perplexity) porque cada plataforma tiene sesgos y fuentes preferidas. Registra por cada pregunta y motor: dominios citados, dominios mencionados, timestamps. Construye una matriz consulta × fuente × motor.



El ciclo trimestral de benchmark de fuentes. Fuente: elaboración propia a partir del texto del capítulo 25.

A continuación, toca analizar los patrones. ¿Son siempre los mismos cinco dominios los que concentran la visibilidad? ¿Ha entrado alguna fuente nueva este trimestre? ¿Cuál gana presencia y cuál retrocede? Ese mapa solo cobra sentido cuando lo cruzas con tu propia posición: si en las consultas sobre «mejores X» aparecen diez fuentes y tú no figuras entre ellas, estás fuera del radar. Y si apareces, pero siempre en los últimos puestos, no deberías interpretarlo como una señal de fortaleza, sino de fragilidad. En Antropus esa lectura la hace Competitor Intelligence en ChatGPT, Claude, Gemini y Perplexity: qué competidores reconoce la IA, en qué posición y con qué tono, incluidos los emergentes que no tenías en el radar.

Y luego necesitas un informe trimestral con acciones concretas: «Durante este trimestre han entrado dos fuentes nuevas en respuestas sobre [tema] y nosotros seguimos sin aparecer. El plan: crear contenido similar al que impulsó esas fuentes y actualizar el contenido obsoleto». Si trabajas el plan con el GEO Strategy Builder de Antropus, cada estrategia queda versionada y puedes comparar un trimestre con el anterior.

La ventaja sostenible es de quien vigila y adapta, no del que se duerme en el ranking de ayer.

Cambios de comportamiento por actualización de modelo

Cada vez que un motor cambia su forma de responder, la reacción típica es infantil: «¿Nos penalizaron?». No

necesariamente. Puede ser una actualización del modelo base, un ajuste del algoritmo de recuperación, o un cambio en cómo prioriza fuentes. Los LLMs son probabilísticos: pequeñas afinaciones en pesos alteran resultados.

La clave para leer los updates con criterio es no reaccionar por impulso. Mantén una monitorización continua de tu batería de prompts y, cuando detectes un cambio, verifica si se repite en varios prompts y en varios motores. Si todo se altera al mismo tiempo, probablemente estés viendo un update general. Si el movimiento solo aparece en una plataforma, lo más razonable es pensar en un cambio específico de ese motor. A partir de ahí, conviene hacer pruebas deliberadas para confirmar hipótesis antes de modificar la estrategia. Y, sobre todo, evitar decisiones bruscas basadas en una única observación: hay que mirar el comportamiento durante dos o tres semanas para distinguir una señal real del simple ruido.



Cómo leer un cambio tras un update de modelo. Fuente: elaboración propia a partir del texto del capítulo 25.

Documenta todo en un update log: «En oct. 2026, ChatGPT dejó de citar blogs en recetas; ahora casi solo cita sitios oficiales. Impacto: nuestras recetas dejaron de salir.» Esa memoria histórica evita que rehagas el mismo diagnóstico cada mes.

A veces no habrá explicación clara. Te toca convivir con incertidumbre. Lo que sí haces: reforzar tus señales. Si haces bien lo básico, un update puede darte un susto pero no debería borrarte.

Biblioteca de ejemplos vivos: tu antídoto contra «yo creo»

No se comprende de verdad cómo responde una IA hasta que se observa un historial real y acumulado. La memoria humana es limitada, y los sesgos distorsionan la lectura. Por eso conviene construir una biblioteca viva: cientos de casos documentados de tu sector, con preguntas informativas, comparativas y transaccionales, la respuesta generativa obtenida, las fuentes citadas y el contexto necesario para interpretarlas con criterio.

Esa biblioteca sirve para detectar patrones. Si un prompt incluye la palabra «barato» y el modelo nunca menciona tu marca premium, ya tienes una señal útil: puedes decidir si te interesa crear contenido para quienes buscan opciones económicas o comparar precios, o si prefieres no competir en ese terreno. Si el 80 % de tus menciones dependen de una sola URL, tienes una vulnerabilidad clara. Y, además, te permite comparar la evolución en el tiempo: ¿ha cambiado el tono?, ¿cita menos fuentes?, ¿mejoraste después de una campaña?

El formato mínimo por caso: fecha, plataforma, prompt exacto, resumen de respuesta, fuentes citadas y observaciones. Si el prompt no está exacto, el caso no sirve. Centraliza en un repositorio único accesible al equipo. Actualiza periódicamente. Extrae patrones accionables.

Framework propio: métricas a medida para tu negocio

No existe un «PageRank de la IA» resumido en una sola cifra. Si quieres medir de verdad tu presencia en entornos generativos, necesitas un sistema propio que combine varias señales y las convierta en una lectura útil para el negocio.

En Elevam, esa lectura se apoya en una metodología de medición GEO centrada en la atribución. No se trata solo de comprobar si una marca aparece o no, sino de observar si el modelo la menciona, si la cita con una fuente visible, si la asocia a la URL correcta para esa consulta y en qué posición la sitúa cuando genera una shortlist. A partir de ahí, el SoM (Share of Model) permite calcular en qué porcentaje de prompts del dataset aparece tu marca. Es la misma lógica que aplicamos en Antropus: antes de contar una mención, la ficha de entidad comprueba que la IA habla de tu marca y no de un homónimo.

Lo importante es definir cada indicador con precisión y aplicarlo siempre del mismo modo. Hay que dejar claro qué cuenta como mención, qué se considera una cita válida, qué URL se da por correcta en cada tema y cómo se registra la posición cuando el modelo compara varias opciones. Solo así los resultados son comparables en el tiempo y entre plataformas.

La utilidad real del sistema no está en producir una cifra bonita, sino en orientar decisiones. Si la marca apenas

aparece, el problema es de presencia. Si aparece, pero no recibe cita ni URL atribuida correcta, el problema es de atribución. Si entra en la respuesta pero siempre en posiciones débiles, la visibilidad sigue siendo frágil. Y si el contexto es negativo o poco competitivo, el trabajo debe centrarse en reputación, claridad y refuerzo de fuentes. Ningún sistema de medición será perfecto ni definitivo, pero sí puede ser lo bastante consistente como para detectar patrones, priorizar acciones y seguir la evolución con criterio.



Del indicador al diagnóstico. Fuente: elaboración propia a partir del texto del capítulo 25.

Casos reales con evidencia reproducible

Si quieres convencer, deja de hablar: enseña pruebas. Un caso GEO bien documentado incluye situación concreta, medición inicial, cambios implementados, medición posterior y evidencia verificable.

Cada iniciativa GEO se trata como experimento: hipótesis, baseline, implementación, medición posterior, conclusiones. Sin baseline no hay resultado. Antes de tocar nada: qué responde la IA, qué fuentes cita y qué score tenías. Después: repites medición con mismas condiciones para aislar el efecto. Capturas o no pasó.

Conviene ocultar la identidad del cliente, pero manteniendo los detalles que hacen útil el aprendizaje. En lugar de ocultarlo todo, da el contexto suficiente para que el aprendizaje conserve valor: «Una fintech de préstamos P2P en fase Series B duplicó su presencia tras corregir la página de precios y lograr cinco menciones en blogs financieros.» Lo importante no es solo el caso, sino la lección transferible. Si aprendes, por ejemplo, que los PDFs sin texto útil ni contexto claro penalizan la visibilidad en GEO, esa conclusión debería aplicarse a todo tu portfolio.

También es importante documentar los fracasos. A veces, pese a haber optimizado bien, la IA seguirá sin responder como esperabas. Ese resultado no invalida el trabajo: te ayuda a entender los límites del sistema. Y como lo que hoy funciona puede quedar desfasado mañana tras un update, los casos deben tratarse como material vivo, con nuevas comprobaciones, notas de cambio y supuestos revisados con regularidad.

En GEO, la ventaja está en aprender más rápido que los demás. No se aprende sin experimentar, equivocarse, medir y volver a iterar.

Este capítulo cierra los 25 capítulos del libro. Pero como acabas de leer, GEO no tiene capítulo final, tiene iteración continua.

Apéndice diferencial: Anatomía de una auditoría GEO real

Cómo se ve, paso a paso, una auditoría completa con las seis capas del libro aplicadas a un caso real

Los 25 capítulos anteriores describen la metodología. Este apéndice la enseña en acción. Vas a ver, paso a paso y con capturas reales, cómo se hace una auditoría GEO completa sobre una empresa real, qué datos saca el sistema, cómo se interpretan y qué plan de acción produce.

El caso que sigue corresponde a un cliente real de Elevam cuya identidad se reserva por acuerdo de confidencialidad. Los datos numéricos son los reales del proyecto. Las capturas de pantalla son las que generamos durante la auditoría con Antropus, la plataforma que usamos para automatizar el sistema completo descrito en el Capítulo 22. Solo se han anonimizado nombre, URL y sector específico para proteger al cliente.

Lo que vas a ver no es una demo de producto. Es una auditoría real, con sus aciertos, sus sorpresas y sus puntos ciegos. Si después de este apéndice quieres replicar el ejercicio con tu propio dominio, al final encontrarás las instrucciones.

1. El cliente (NDA - Anónimo)

Empresa SaaS B2B española de tamaño medio. Ocho personas en plantilla. Facturación anual cercana a los dos

millones de euros. Operan en una categoría competitiva del software de gestión empresarial: cinco competidores directos, dos de ellos internacionales, tres nacionales.

Llevan cuatro años invirtiendo en SEO clásico con una agencia que les hace mantenimiento mensual. Los resultados en Google son decentes: rankean en las primeras posiciones para varios términos de su categoría, generan unas 4.000 visitas orgánicas al mes, y convierten un porcentaje razonable de ese tráfico a demos.

El problema lo plantea el CEO en la primera llamada: «Estamos en Google, pero cuando le preguntas a ChatGPT por las mejores herramientas de nuestra categoría, no aparecemos. Sale nuestra competencia. Y eso ya lo notamos en pipeline.»

La hipótesis del cliente: «Nos falta contenido». La hipótesis del equipo de marketing: «Es un problema técnico de la web». Las dos resultaron estar equivocadas.

2. El punto de partida

Antes de tocar nada, generamos el baseline. Este es el principio del Capítulo 21: medir antes de actuar. Sin foto del día cero, no hay forma de demostrar progreso después.

Lo primero que hicimos fue lanzar el conjunto de prompts representativos del sector en los cuatro motores principales. Las preguntas: «¿Cuáles son las mejores herramientas de [categoría]?», «Alternativas a [competidor líder]», «[Marca cliente] vs [competidor]», «Mejor soft-

ware de [categoría] para empresas medianas», y veintiséis variantes más segmentadas por intención.

Resultado en ChatGPT: el cliente aparecía en 2 respuestas de 30. Share of Model: 6,7%. Cuando aparecía, lo hacía en posición 4 o 5 de cinco opciones. Citation Coverage: 0%. Ninguna URL del cliente era citada como fuente.

Resultado en Perplexity, Claude y Gemini: peor. Aparecía en 1, 0 y 0 respuestas respectivamente. SoM agregado en los cuatro motores: 3,3%.

Para comparar, lanzamos los mismos 30 prompts buscando al competidor líder. SoM: 26,7%. Aparecía en primera o segunda posición de la shortlist en el 80% de los casos.

La hipótesis del cliente («nos falta contenido») y la del equipo («es un problema técnico») quedaban en cuarentena. Algo más profundo estaba pasando.

3. La auditoría con Antropus

Aquí es donde el sistema empieza a separar señal de ruido. Las seis capas del Capítulo 11 se ejecutan en paralelo, no en serie. El objetivo no es generar seis informes aislados, sino cruzar los hallazgos para encontrar la causa raíz.

3.1 Generación metodológica del prompt set

El primer paso es construir un prompt set que represente fielmente cómo busca el mercado del cliente. No vale

inventarse 30 preguntas. Tienen que estar segmentadas por intención, por idioma y por motor.

La diferencia entre un prompt set generado con criterio metodológico y uno improvisado a mano es la diferencia entre una medición seria y una colección de anécdotas.

3.2 Baseline GEO multi-motor

Con el prompt set definido, el sistema lanza las consultas en los cuatro motores y registra todas las respuestas. Por cada respuesta se extrae: marcas mencionadas, posición de cada marca, URLs citadas (si las hay), y el texto completo de la respuesta para análisis posterior de sentiment y co-mentions.

El dato más revelador de esta vista no es el porcentaje de apariciones. Es el patrón geográfico de los huecos. Donde no aparece el cliente, aparece consistentemente el mismo competidor.

3.3 Visibility Auditor por los cuatro pilares

Aquí es donde el sistema deja de medir y empieza a diagnosticar. Los cuatro pilares de visibilidad GEO (Entidad, Extractabilidad, Evidencia, Consenso) generan un scoring de 0 a 100 cada uno, basado en señales objetivas extraídas de la web del cliente, fuentes externas y el comportamiento observado en los motores.

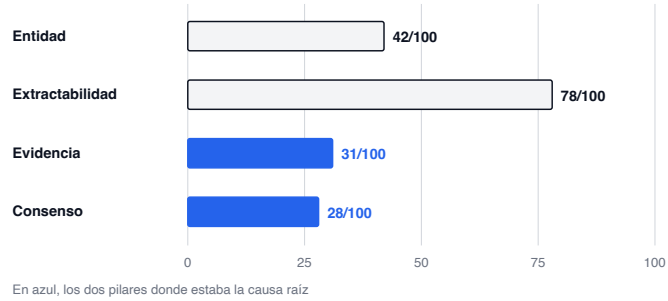
El diagnóstico se lee solo. La Extractabilidad técnica está bien (78/100): la web es accesible para crawlers, el

HTML está limpio, el contenido se puede parsear. Eso descarta la hipótesis del equipo de marketing.

La Entidad está mal (42/100). El nombre del cliente aparece de cuatro formas distintas entre web, LinkedIn, Wikidata y la ficha de Google Business: con guion, sin guion, con «SL» al final, sin «SL». Cada una de esas variantes podría estar siendo tratada como entidad distinta por los modelos.

La Evidencia está peor (31/100). El cliente no tiene datos propios publicados que la IA pueda citar. No hay informes anuales del sector, no hay benchmarks internos compartidos, no hay casos de éxito públicos con cifras verificables. Cuando la IA necesita una fuente con evidencia concreta, busca en otra parte.

Y el Consenso está crítico (28/100). El cliente no aparece en los directorios sectoriales que la IA consulta como fuente. No tiene perfil en las plataformas de reviews B2B donde sus competidores acumulan opiniones. No aparece en los rankings «Top 10» que dominan las respuestas comparativas.



Los cuatro pilares en el baseline del caso. Fuente: elaboración propia a partir del texto del apéndice.

El problema no era técnico ni de contenido. Era de Evidencia y Consenso. Eso era invisible sin los cuatro pilares medidos por separado.

3.4 Matriz competitiva Google vs IA

Una de las vistas más reveladoras del sistema es la que cruza posiciones de Google con apariciones en respuestas de IA. Sirve para ver de un vistazo qué competidores dominan cada terreno.

Hallazgos del caso: dos competidores del cliente rankean peor que él en Google pero aparecen mucho más en respuestas de IA. Tres competidores rankean similar a él en Google pero también aparecen más en IA. Ninguno de los cinco competidores principales tiene su mismo patrón de «fuerte en Google, débil en IA». Es un patrón inusual, casi siempre asociado a problemas de Evidencia y Consenso (los dos pilares en los que el cliente flojeaba).

3.5 Action Plan generado

Con el diagnóstico completo, el sistema genera un plan de acción priorizado por impacto y esfuerzo. No son sugerencias genéricas. Son acciones concretas con responsable, plazo y métrica de éxito asociada.

Las 7 acciones priorizadas, en orden:

1. Unificación de entidad (Pilar Entidad). Consolidar el nombre del cliente en una única forma canónica en web, LinkedIn, Wikidata, Google Business, prensa y directorios. Plazo: 2 semanas. Impacto: alto. Esfuerzo: bajo.
2. Crear página fuente canónica de marca (Pilar Entidad + Evidencia). Página «Acerca de [empresa]» con la estructura del Capítulo 8: definición, historia, datos, productos, FAQs, comparativa con alternativas. Plazo: 3 semanas. Impacto: alto. Esfuerzo: medio.
3. Publicar informe anual del sector con datos propios (Pilar Evidencia). El cliente tiene datos internos de uso de su plataforma agregables y anonimizables que ningún competidor publica. Convertirlos en un informe sectorial. Plazo: 6 semanas. Impacto: muy alto. Esfuerzo: alto.
4. Conseguir perfil en las dos plataformas B2B principales del sector (Pilar Consenso). Crear perfil, completar al 100%, activar campaña de captación de reviews entre clientes satisfechos. Plazo: 4 semanas. Impacto: alto. Esfuerzo: medio.

5. PR digital: 5 menciones objetivo en medios especializados del sector (Pilar Consenso). Acercamiento a periodistas con el informe del punto 3 como anzuelo. Plazo: 8 semanas. Impacto: alto. Esfuerzo: medio.
6. Optimización schema avanzado (Pilar Entidad + Extractabilidad). Aplicar la metodología del Capítulo 15: Organization, Person, Service, sameAs completo. Plazo: 2 semanas. Impacto: medio. Esfuerzo: bajo.
7. Crear 3 páginas comparativas honestas: «[Cliente] vs [competidor líder]», «Alternativas a [competidor líder]», «Mejor [categoría] para empresas medianas» (Pilar Evidencia + Consenso). Plazo: 4 semanas. Impacto: alto. Esfuerzo: medio.

Cada acción tiene asociada una métrica de éxito que se mide en la siguiente auditoría. Sin métricas, no hay forma de saber si funcionó.

4. La interpretación

Aquí entra el criterio humano, que no se automatiza. El sistema da el diagnóstico. El consultor decide qué hacer con él.

Lo que vimos en este caso era un patrón clásico que no se ve sin medir bien: cliente con producto sólido, SEO clásico decente, accesibilidad técnica perfecta, contenido razonable. Y al mismo tiempo, prácticamente invisible para la IA. La causa raíz no estaba en su web. Estaba en lo que el mercado decía (o más bien, no decía) sobre ellos.

Los modelos de lenguaje no aprenden quién eres por tu home. Aprenden quién eres por la suma de menciones, reviews, listados y referencias externas. Si el mercado guarda silencio sobre ti, la IA hereda ese silencio. No es un castigo, es una consecuencia matemática del entrenamiento.

El diagnóstico final que entregamos al cliente: «Vuestra invisibilidad en IA no es un problema de SEO ni de contenido. Es un problema de presencia distribuida. El plan de acción ataca primero la entidad (semanas 1-2), después la evidencia (semanas 3-8) y finalmente el consenso (semanas 4-12). Si lo ejecutáis disciplinadamente, esperamos ver movimiento de SoM medible a partir del mes 3.»

Ningún cliente quiere oír que el problema no estaba donde miraban. Pero los datos no dejan margen para la interpretación cómoda.

5. La ejecución a 90 días

El cliente aprobó el plan. La ejecución la coordinó su equipo de marketing interno con seguimiento semanal por parte de Elevam. Las acciones se ejecutaron en este orden y con estos hitos:

Semanas 1-2. Unificación de entidad ejecutada. Web, LinkedIn, Wikidata, Google Business y los seis directorios sectoriales donde aparecían: todos alineados con el nombre canónico. Schema avanzado implementado: Or-

ganization completo con sameAs apuntando a los seis perfiles unificados.

Semanas 3-6. Página fuente canónica publicada. Estructura completa siguiendo el Capítulo 8: 4.500 palabras, índice navegable, datos verificables, FAQs, comparativas con competidores. En paralelo, recopilación interna de datos para el informe del sector.

Semanas 7-10. Informe sectorial publicado. 28 páginas, datos anonimizados de 1.200 usuarios de la plataforma, gráficos, conclusiones accionables. Distribución a 40 medios y periodistas especializados con dossier de prensa adjunto. Resultado: 7 menciones en medios sectoriales en las dos semanas siguientes, dos de ellas en cabeceras de referencia del sector.

Semanas 8-12. Perfiles en plataformas B2B activos y con primeras 30 reviews acumuladas (campaña interna a clientes satisfechos con incentivo simbólico). Tres páginas comparativas publicadas con la estructura recomendada en el Capítulo 20.

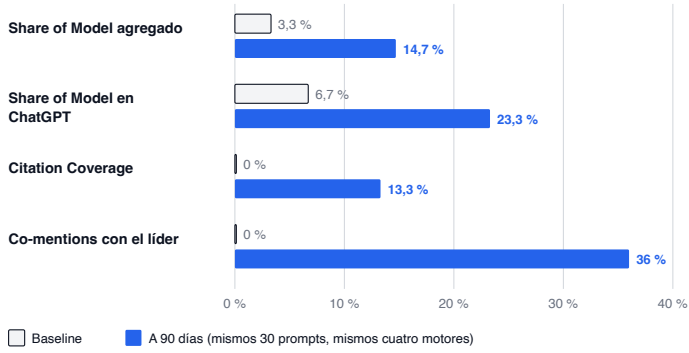
Hito clave: en la semana 6, el cliente recibió el primer correo de un periodista pidiendo entrevista a partir del informe sectorial. Es la primera vez en la historia de la empresa que un medio les buscaba en lugar de ser ellos buscando al medio. Las acciones del pilar Consenso empiezan a tener efectos en cadena que no se ven en métricas todavía pero sí en operativa real.

6. El resultado a 90 días

Al cumplirse los 90 días, repetimos la auditoría exactamente igual que el baseline. Mismos 30 prompts, mismos cuatro motores, misma metodología.

Resultados a 90 días:

- Share of Model agregado: del 3,3% al 14,7%. En ChatGPT, del 6,7% al 23,3%.
- Shortlist Position media en ChatGPT: de 4,3 (cuando aparecían) a 2,8.
- Citation Coverage: del 0% al 13,3%. La página fuente canónica y el informe sectorial empezaron a aparecer como fuentes citadas.
- Co-Mentions con el competidor líder: del 0% al 36%. La IA empezó a verlos como categoría comparable, no como dos cosas distintas.
- Sentiment Score: pasó de neutro mayoritario a 41% positivo, 52% neutro, 7% negativo. Los adjetivos asociados a la marca en las respuestas: «completo», «robusto», «buena documentación», «alternativa sólida a [competidor líder]».



Antes y después: el caso a 90 días. Fuente: elaboración propia a partir del texto del apéndice.

Lo más importante no son los porcentajes. Es lo que esos porcentajes significan para el negocio. El equipo comercial empezó a recibir llamadas de prospectos que mencionaban explícitamente «lo vi en ChatGPT». Tres deals que estaban estancados se reactivaron. Pipeline de demos creció un 28% en el trimestre, atribuible solo parcialmente a GEO pero coincidiendo con las acciones del plan.

El cliente no dominaba la categoría. Tampoco era el objetivo a 90 días. Pero había pasado de invisible a presente, y eso ya cambia la conversación comercial.

7. Lo que aprendimos

Cinco lecciones de este caso, transferibles a cualquier proyecto GEO serio.

Primera. El problema casi nunca está donde el cliente piensa. La hipótesis inicial («nos falta contenido») y la

del equipo («es técnico») eran las dos equivocadas. Sin medición rigurosa de las seis capas en paralelo, ese diagnóstico es imposible.

Segunda. Los pilares Evidencia y Consenso son los más infravalorados por los equipos de marketing y los más determinantes en GEO. Casi todo el mundo trabaja Entidad y Extractabilidad porque son los que se sienten más «técnicos». Los otros dos requieren PR, partnerships y trabajo con el mercado externo, que es menos cómodo de planificar pero es donde se gana o se pierde la batalla.

Tercera. Sin baseline, no hay resultado defendible. Si este cliente hubiera ejecutado las mismas acciones sin medir antes, en 90 días estaría diciendo «creo que estamos mejor pero no estoy seguro». Ahora puede mostrar cifras concretas a su consejo. Esa diferencia, en términos comerciales, vale el coste entero de la auditoría.

Cuarta. La medición automatizada cambia el qué se puede prometer. Con 50 prompts ejecutados manualmente en cuatro motores, una vez por trimestre, no puedes detectar movimiento real. Con 30-50 prompts ejecutados semanalmente de forma automatizada, sí. La diferencia entre un servicio GEO que vende humo y uno que vende resultados es esa capacidad de medir a escala.

Quinta. GEO es un servicio de paciencia. El cliente vio resultados medibles a partir de la semana 6 y resultados de negocio a partir del mes 3. Cualquier proveedor que

prometa resultados en 30 días en GEO está mintiendo o midiendo mal.

¿Quieres ver Antropus aplicado a tu propio dominio?

Puedes crear tu cuenta de Antropus gratis, sin tarjeta, y lanzar tu primer baseline con el Prompt Tracker: el mismo set de prompts en ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview, con el plan de acción priorizado que sale de cada informe.

→ **antropus.io**

Epílogo: Tu siguiente movimiento

Has llegado al final de 25 capítulos sobre RAG, vectores, embeddings, auditorías, pipelines, schema avanzado, gestión de entidad, reputación, automatización y ética. Si has llegado hasta aquí leyendo de verdad, ya has hecho algo que el 95 % de los profesionales de tu sector probablemente no hará. Y eso, por sí solo, te sitúa en una minoría que ya ha entendido que las reglas del juego han cambiado.

Ahora sabes algo que la mayoría de tu competencia todavía no sabe. Eso solo tiene valor si lo conviertes en decisiones, foco y ejecución. La verdadera pregunta es qué vas a hacer con ello.

Porque el conocimiento sin ejecución es solo entretenimiento caro. Un libro técnico leído y archivado no mueve la aguja. Lo que mueve la aguja es la primera auditoría que hagas con la metodología que acabas de aprender. Es el primer prompt set que ejecutes. Es la primera página que reestructures pensando en recuperabilidad. La primera vez que midas tu Share of Model y veas el número real y decidas que vas a cambiarlo. Si no quieres montar esa medición a mano, es exactamente lo que

automatizamos en Antropus con el Prompt Tracker: el mismo set de prompts, lanzado a ChatGPT, Claude, Gemini, Perplexity y Google, con el histórico guardado para comparar mes a mes.

La ventana de oportunidad existe ahora. El 90 % del mercado sigue optimizando para un mundo que ya no existe: peleando por posiciones en un ranking que la IA a veces ni mira, midiendo clics que cada vez representan menos del total de interacciones, ignorando que las respuestas generativas están decidiendo quién merece ser recomendado y quién no.

Tú ya no eres parte de ese 90 %.

Tienes el mapa completo. Los 25 capítulos que acabas de leer son los mismos planos que usamos en Elevam cada semana con nuestros clientes. Sin versión light, sin secretos guardados, sin el típico «para más información contáctanos». Todo está aquí.

Puedes construir el barco tú solo. Tienes todos los planos, las herramientas y las métricas para hacerlo. Si tienes equipo, tiempo y ganas de ensuciarte las manos, adelante. Este libro se escribió para que eso sea posible.

Si después de leer este libro has llegado a la conclusión de que necesitas ayuda para aplicar todo esto en tu caso, la conversación está abierta. En Elevam llevamos años trabajando precisamente en ese terreno: convertir conocimiento técnico en sistemas reales de visibilidad.

A estas alturas, no se trata de vender, sino de coherencia. Después de 25 capítulos de método, ejemplos y criterio, el lector ya sabe si quiere recorrer este camino por su cuenta o apoyarse en quien lleva años haciéndolo.

→ **elevam.es/contacto**

El GEO no es un proyecto que se termina. Es una disciplina que se mantiene. Los modelos cambian. Las fuentes rotan. Los competidores despiertan. Lo que hoy te cita, mañana puede omitirte sin pedirte permiso.

Pero si has interiorizado lo que contiene este libro, no como teoría sino como sistema, tienes algo que la mayoría no tiene: la capacidad de ver lo que se mueve bajo la superficie, medir lo que importa y adaptar antes de que sea tarde.

Eso no caduca.

Ahora elige tu camino. Y sea cual sea, ejecútalo hoy.

Glosario GEO de Elevam

Diccionario técnico de la A a la Z — referencia rápida de todos los conceptos del libro

Alucinación (Hallucination)
Cuando un LLM genera información que suena convincente pero es falsa o inventada, por falta de datos fiables en su entrenamiento o contexto. → Cap. 2

Arquitectura híbrida de búsqueda
Sistema que combina búsqueda determinista (tipo Google clásico) con generación probabilística (LLM), como Copilot o Perplexity. → Cap. 2

Atribución difusa
Fenómeno por el cual tu contenido influye en la respuesta de una IA sin que el usuario sepa que tú fuiste la fuente, porque la IA no siempre cita explícitamente. → Cap. 1

Auditoría de entidad
Análisis de la consistencia de tu marca, categoría, atributos y relaciones en todas las fuentes digitales donde apareces. → Cap. 8

Auditoría de fuentes citadas
Identificación de qué URLs y dominios domina la IA en las respuestas de tu sector. → Cap. 11

Auditoría de respuesta generativa

Análisis de qué dice la IA sobre tu marca y qué se inventa. → Cap. 11

Auditoría reputacional IA

Evaluación del sentimiento y percepción que los modelos de IA proyectan sobre tu marca. → Cap. 11

Auditoría técnica de rastreabilidad (bots IA)

Verificación de que los crawlers de IA (GPTBot, PerplexityBot, Bingbot, etc.) pueden acceder y leer tu contenido. → Cap. 11

Baja fricción semántica

Escribir con claridad y desambiguación para que el LLM no tenga que «adivinar» qué quieres decir. → Cap. 14

Baseline técnico GEO

La medición inicial de tu situación antes de implementar cambios, que sirve como punto de comparación. → Cap. 21

Benchmarks trimestrales GEO

Ciclo de medición periódica de qué fuentes citan los motores de IA en tu vertical, para detectar cambios y oportunidades. → Cap. 25

Brand Association (asociación de marca en IA)

Los atributos y cualidades que la IA asocia a tu marca basándose en evidencias distribuidas en la web. → Cap. 17

Búsqueda determinista vs. probabilística

Determinista: un algoritmo con reglas fijas (Google clásico). Probabilística: un modelo que predice la respuesta más probable (LLM). → Cap. 2

Capas de contenido

Estructura de resumen ejecutivo / detalle técnico / pruebas y datos que facilita la extracción por parte de la IA. → Cap. 13

ChatGPT browsing vs. índice base

La diferencia entre el conocimiento estático del entrenamiento de ChatGPT y la información que obtiene en búsquedas en vivo. → Cap. 19

Chunking de contenido

Cómo los sistemas RAG dividen tu página en fragmentos para procesarlos y compararlos con la consulta del usuario. → Cap. 7

Citable snippets (datos citables)

Fragmentos de tu contenido estructurados para ser fácilmente extraídos y citados por una IA. → Cap. 13

Citation bias

Sesgo de los modelos a citar ciertos tipos de fuentes (dominios conocidos, contenido reciente) sobre otros igualmente válidos. → Cap. 9

Citation Coverage / Cobertura de Citas

Métrica que mide si la IA enlaza tu URL o solo te nombra de pasada. → Cap. 9

Clúster semántico
Grupo de conceptos relacionados en el espacio vectorial donde tu marca debería estar posicionada. → Cap. 5

Co-Mentions (co-menciones)
Marcas que la IA menciona junto a la tuya, revelando con quién te asocia. → Cap. 9

Co-ocurrencia estratégica
Lograr que tu marca aparezca junto a los líderes de tu sector en contenidos de terceros. → Cap. 17

Coherencia categorial
Que tu marca esté clasificada de forma consistente en la misma categoría en todas las fuentes. → Cap. 16

Compliance GEO
Cumplimiento de directrices éticas, legales y de plataformas en la optimización para motores generativos. → Cap. 24

Concisión semántica
Economizar tokens sin perder significado; decir más con menos palabras para que cada frase aporte valor. → Cap. 14

Consistencia de entidad
Que el nombre, datos y descripción de tu marca sean idénticos en todas las fuentes digitales. → Cap. 16

Consistencia de marca en IA
Que la identidad digital de tu marca sea coherente para que la IA construya una representación unificada. → Cap. 8

Consultas de Conocimiento vs. Decisión vs. Acción
Taxonomía de intenciones de búsqueda: informativas, comparativas/recomendación y transaccionales. → Cap. 6

Consultas «money» (contexto GEO)
Las preguntas que generan negocio: «mejores», «alternativas», «vs», «opiniones». → Cap. 11

Contenido «citable»
Contenido que ofrece información verificable, estructurada y con fuentes que la IA puede extraer y citar. → Cap. 2

Contenido «echo-friendly»
Contenido diseñado para que la IA lo absorba y replique fácilmente: claro, con datos concretos, sin jerga vacía. → Cap. 19

Context window (ventana de contexto)
La cantidad máxima de tokens que un LLM puede procesar a la vez, limitando cuánto de tu contenido «ve». → Cap. 6

Contrato GEO (cláusulas específicas)
Cláusulas contractuales adaptadas a servicios de optimización generativa, incluyendo no-garantía y modelo probabilístico. → Cap. 23

Control de cambios en GEO
Protocolo para verificar que ningún cambio en la web rompa la visibilidad en IA. → Cap. 21

Correlación vs. causalidad en GEO

La trampa de atribuir mejoras en visibilidad IA a una sola acción sin verificar otras variables. → Cap. 10

Decisiones sin clic

Cuando un usuario obtiene la respuesta directamente de la IA y nunca visita ningún sitio web. → Cap. 1

Definiciones embebidas

Explicaciones de conceptos integradas en el texto de forma que la IA las identifique sin buscar en otra fuente. → Cap. 13

Densidad informativa vs. relleno

La diferencia entre contenido que aporta datos nuevos en cada frase y contenido que repite o divaga. → Cap. 13

Digital PR para GEO

Relaciones públicas digitales orientadas a generar menciones en fuentes que los modelos de IA consumen. → Cap. 17

Dilución vectorial por relleno

Cuando el texto irrelevante desplaza el embedding de tu página lejos de las consultas objetivo. → Cap. 14

Diseño orientado a verificación

Estructura de página pensada para que la IA pueda verificar datos (fuentes, fechas, cifras) de forma directa. → Cap. 8

Diseño RAG-friendly

Formato, chunking y estructura de página optimizados para que el sistema RAG te encuentre y cite. → Cap. 7

Distancia coseno
Medida matemática de similitud entre dos vectores en el espacio semántico; cuanto menor, más relacionados. → Cap. 5

Docker para pipelines GEO
Uso de contenedores Docker para empaquetar y ejecutar pipelines de medición de forma reproducible. → Cap. 22

Echo index
La representación que un LLM tiene de tu marca en su memoria de entrenamiento. → Cap. 19

Economía de influencia sin visita
Valor que genera tu contenido cuando la IA lo usa para responder sin que el usuario visite tu web. → Cap. 1

Efecto primacía
Los tokens al inicio del contexto tienen más peso en la atención del modelo. → Cap. 6

Efecto recencia
Los tokens al final del contexto también reciben más peso que los del medio. → Cap. 6

Embeddings
Representaciones numéricas (vectores) que codifican el significado de palabras, frases o documentos en un espacio multidimensional. → Cap. 5

Entity Home (página central de entidad)
La página oficial dedicada a describir quién eres, que actúa como referencia central para la IA. → Cap. 16

Entity SEO
Optimización orientada a que los buscadores y modelos de IA reconozcan tu marca como entidad única y definida. → Cap. 16

Espacio semántico
El espacio multidimensional donde los embeddings representan significados y las distancias indican relación. → Cap. 5

Fact retrieval (recuperación de hechos)
El mecanismo por el cual un LLM busca datos verificables para construir su respuesta. → Cap. 2

Fan-out de sub-consultas
Cuando un motor de IA descompone una pregunta compleja en varias sub-preguntas para buscar información. → Cap. 11

Fases del Pipeline GEO
Las cinco fases: Señales → Recuperación → Citación → Síntesis → Shortlist. → Cap. 3

Fenómeno zero-click
El 60% de las búsquedas terminan sin que el usuario haga clic en ningún resultado. → Cap. 1

Fragmentos recuperados (chunks)
Los trozos de texto que un sistema RAG extrae de las fuentes para pasarlos al modelo generativo. → Cap. 4

Framework HSA (Elevam)
Holistic Semantic Authority: sistema de scoring propietario que combina SoM, posición, sentimiento, cobertu-

ra de citas y comparativa competitiva. → Cap. 25

«Frontlinks» (citas en IA)
Las menciones y citas que tu marca recibe en respuestas de IA, equivalente moderno de los backlinks. → Cap. 9

GEO (Generative Engine Optimization)
Disciplina de optimización para que los motores de búsqueda con IA te encuentren, te citen y te recomienden. → Cap. 1

GEO B2B
Estrategias GEO específicas para servicios profesionales: autoridad personal, casos de éxito verificables, reputación. → Cap. 20

GEO E-commerce
Estrategias GEO para tiendas online: productos como entidades, Shopping Graph, reviews estructuradas. → Cap. 20

GEO Local
Estrategias GEO para negocios físicos: Google Business Profile, reseñas, consistencia NAP, atributos geolocalizados. → Cap. 20

GEO SaaS
Estrategias GEO para software: batalla «alternativas a», «vs», contenido técnico, reviews B2B. → Cap. 20

Gestión de crisis 2.0
Protocolo de respuesta cuando la IA difunde información negativa o errónea sobre tu marca. → Cap. 18

Google AI Overviews (antes SGE)
La experiencia generativa de Google que sintetiza respuestas con IA en la página de resultados. → Cap. 19

Granjas de entidades
Redes artificiales de sitios, perfiles y personas ficticias creadas para inflar la prominencia de una marca. → Cap. 24

Índice vectorial
Base de datos que almacena embeddings y permite buscar por similitud semántica en vez de por palabras exactas. → Cap. 4

Information gain (ganancia informativa)
Lo que tu contenido aporta que la IA no obtendría de ninguna otra fuente. → Cap. 7

Inyección indirecta de prompts
Instrucciones maliciosas ocultas en contenido aparentemente normal que manipulan la respuesta de la IA. → Cap. 24

Inventario de hechos
Lista completa de datos verificables sobre tu marca que debe contener tu Source of Truth. → Cap. 8

JSON-LD avanzado para GEO
Marcado estructurado con relaciones complejas (@id, @graph, anidación) para comunicar datos a la IA. → Cap. 15

Knowledge Graph de Google

La base de datos de entidades y relaciones que Google usa para entender el mundo. → Cap. 16

Knowledge Panel

El recuadro informativo que Google muestra cuando reconoce una entidad; señal de que estás fichado. → Cap. 16

Lag de entrenamiento

El desfase temporal entre la realidad y lo que un LLM «sabe» por haber sido entrenado con datos anteriores. → Cap. 19

Lenguaje natural predictivo

Escribir con claridad y baja ambigüedad para que el LLM prediga correctamente el significado. → Cap. 14

Logging estructurado

Registro sistemático de resultados en formato consistente (JSON/CSV) para garantizar trazabilidad. → Cap. 10, 22

Lost-in-the-middle

Efecto por el cual los tokens en el centro de la ventana de contexto reciben menos atención del modelo. → Cap. 6

Manipulación de sentimiento en IA

Intentar forzar una narrativa artificialmente positiva o negativa sobre una marca en los datos que consume la IA. → Cap. 24

Mapa de co-ocurrencia

Herramienta visual que mapea dónde quieres que tu marca aparezca junto a líderes del sector. → Cap. 17

Marco de controlabilidad GEO

Framework que distingue lo que controlas, lo que influyes y lo que no controlas en tu visibilidad IA. → Cap. 12

Mención primaria vs. secundaria

Primaria: la IA te nombra como opción principal. Secundaria: te menciona de pasada o al final. → Cap. 10

Métricas vanidosas en GEO

Indicadores que parecen relevantes pero no miden visibilidad real en IA. → Cap. 12

Modelo mental operativo GEO

El pipeline de cinco fases como brújula para diagnosticar y mejorar tu visibilidad. → Cap. 3

Named Entity Disambiguation

El proceso por el cual los motores distinguen entre entidades que se llaman igual o parecido. → Cap. 16

Observabilidad del pipeline

Instrumentar tu sistema de medición para distinguir cambios del modelo, bugs propios y variación aleatoria. → Cap. 22

Optimización de tokens

Reducir la cantidad de tokens necesarios para comunicar un mensaje sin perder información. → Cap. 14

Optimización para extracción
Estructurar contenido (títulos semánticos, definiciones, datos citables) para facilitar que la IA lo extraiga. → Cap. 13

Página canónica
La página definitiva que reúne toda la verdad verificable sobre tu marca para que la IA la use como referencia. → Cap. 8

Páginas fuente vs. páginas satélite
Fuente: páginas maestras de alto valor. Satélite: páginas de apoyo que refuerzan y enlazan a las fuente. → Cap. 7

Patrones de citación por vertical
Qué tipo de contenido cita la IA en cada sector (reviews en local, comparativas en SaaS, etc.). → Cap. 7

Pipeline GEO automatizado
Sistema que ejecuta consultas a IAs, registra respuestas, extrae métricas y genera informes de forma automática. → Cap. 22

Precision@3
Porcentaje de los tres primeros resultados/fuentes citadas que son relevantes para tu marca. → Cap. 12

Pricing basado en valor
Modelo de precios que cobra por el proceso y aumento de visibilidad, no por keywords posicionadas. → Cap. 23

Productización de servicios GEO
Empaquetar servicios GEO como productos claros con alcance, entregables y precio definidos. → Cap. 23

Prompt Set
Conjunto estructurado de consultas de prueba que se ejecutan periódicamente para medir visibilidad en IA. → Cap. 10

Proximidad marca-intención
La distancia en el espacio vectorial entre tu marca y la intención del usuario cuando busca una solución. → Cap. 5

QA de datos estructurados
Control de calidad continuo del marcado schema para garantizar que esté correcto y actualizado. → Cap. 15

Quick wins GEO
Mejoras rápidas de alto impacto que se implementan en las primeras semanas del roadmap. → Cap. 21

RAG (Retrieval-Augmented Generation)
Técnica donde la IA primero recupera documentos relevantes y luego genera su respuesta basándose en ellos. → Cap. 2 (conceptual), Cap. 4 (técnico)

Redes espejo (mirror sites)
Sitios que copian sistemáticamente tu contenido o identidad para perjudicarte o aprovecharse. → Cap. 24

Redes falsas de autoridad
Conjuntos de perfiles y sitios creados artificialmente para simular reconocimiento. → Cap. 24

Reference Rate
Frecuencia con la que una IA cita tu contenido en relación al total de respuestas relevantes. → Cap. 1

Relevancia semántica vs. autoridad de dominio
En RAG, la similitud de significado pesa más que el PageRank o Domain Authority tradicional. → Cap. 4

Reporting automatizado GEO
Generación de informes de visibilidad en IA mediante scripts que eliminan procesos manuales. → Cap. 22

Retrievability / recuperabilidad
La capacidad de tu contenido de ser encontrado y seleccionado por el sistema RAG. → Cap. 4

Retrievability cost
El esfuerzo necesario para que un sistema RAG encuentre tu contenido entre todas las fuentes disponibles. → Cap. 7

sameAs (consolidación de identidad)
Propiedad de schema.org que conecta tu web con tus perfiles externos para crear identidad digital unificada. → Cap. 15

Schema como sistema vivo
Tratar tu marcado estructurado como algo que requiere mantenimiento continuo, no configuración única. → Cap. 15

Scheduler (cron/Airflow/Prefect)
Herramienta que ejecuta automáticamente tu pipeline de medición según frecuencia definida. → Cap. 22

Scorecard LLM
Plantilla que consolida las cinco dimensiones de visibilidad en IA en un panel único de seguimiento. → Cap. 9

Sentiment Score (contexto LLM)
Puntuación que mide si la IA habla bien, neutro o mal de tu marca cuando te menciona. → Cap. 9

Share of Model (SoM)
Porcentaje de respuestas de IA en tu sector que mencionan o citan tu contenido. → Cap. 9

Share of Attention (SoA)
Variante comercial de SoM usada en propuestas de servicios para comunicar el valor en lenguaje de negocio. → Cap. 23

Shopping Graph de Google
Base de datos que conecta productos, vendedores, reviews y precios, usada por AI Overviews y AI Mode para recomendaciones de compra. → Cap. 20

Shortlist Position
Tu posición en la lista de opciones que la IA presenta al usuario (primera mención, segunda, última...). → Cap. 9

Significancia estadística (mediciones LLM)
El umbral a partir del cual un cambio en tus métricas GEO puede considerarse real y no aleatorio. → Cap. 10

Source of Truth (fuente canónica)
La página definitiva interna que reúne toda la verdad verificable para que la IA la use como referencia. → Cap. 8

Tablas comparativas para IA
Formato de contenido que facilita que la IA extraiga comparaciones estructuradas entre opciones. → Cap. 13

Test A/B en contexto GEO
Experimentación controlada donde cambias una variable y comparas resultados de visibilidad IA antes y después. → Cap. 10

Token prediction
El mecanismo fundamental de un LLM: predecir la siguiente palabra basándose en el contexto previo. → Cap. 2

Training Memory de modelos
El conocimiento que un LLM retiene de su entrenamiento, que puede estar desactualizado. → Cap. 19

Transparencia proactiva
Estrategia de publicar tus limitaciones y objeciones antes de que otros las descubran, para controlar la narrativa. → Cap. 18

Truncamiento de contenido en RAG
Cuando un sistema RAG corta tu contenido por exceder la ventana de contexto, perdiendo información. → Cap. 6

Vector drift
Desplazamiento de la posición de tu marca en el espacio semántico por cambios en tu contenido o en el contexto del modelo. → Cap. 5

Vectores multidimensionales
Las coordenadas numéricas que representan significados en el espacio de embeddings. → Cap. 5

Wikidata (como herramienta GEO)

Base de conocimiento estructurada que alimenta a Google y otros modelos; inscribir tu entidad aquí refuerza tu visibilidad. → Cap. 16

Este glosario es un recurso vivo. Los términos y definiciones se actualizarán conforme evolucione la disciplina GEO. Para la versión más reciente, visita elevam.es/glosario-geo.

Bibliografía y referencias

Fuentes citadas a lo largo del libro, organizadas por capítulo para facilitar la consulta. Los enlaces estaban activos en el momento de la edición; algunos pueden haber cambiado.

Estudios y datos citados

- **Ahrefs (2024). «AI Search Overlap: How Often Do AI Citations Match Google's Top 10?».** Estudio que documenta que solo el 12 % de las URLs citadas por asistentes de IA coinciden con el Top 10 orgánico de Google. Citado en Capítulo 1. → https://ahrefs.com/blog/ai-search-overlap/
- **Bain & Company (2025). «Goodbye Clicks, Hello AI: Zero-Click Search Redefines Marketing».** Encuesta Bain–Dynata, diciembre 2024, n=1.117 consumidores. Documenta el 60 % de búsquedas sin clic y la caída del 15-25 % en tráfico orgánico. Citado en Capítulo 1. → https://www.bain.com/insights/goodbye-clicks-hello-ai-zero-click-search-redefines-marketing/
- **Liu, N. et al. (2023). «Lost in the Middle: How Language Models Use Long Contexts».** Investigación de Stanford que documenta el sesgo de los LLMs hacia

el principio y final del contexto, ignorando información del medio. Citado en Capítulo 6.

- **OWASP (2025). «OWASP Top 10 for Large Language Model Applications».** Clasificación de riesgos en aplicaciones con IA generativa. La inyección de prompts (Prompt Injection, directa e indirecta) figura como riesgo número uno. Citado en Capítulo 24. → https://owasp.org/www-project-top-10-for-large-language-model-applications/
- **Lewis, P. et al. (2020). «Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks».** Paper fundacional de RAG, Meta AI Research. Citado en Capítulos 2 y 4.
- **Karpukhin, V. et al. (2020). «Dense Passage Retrieval for Open-Domain Question Answering».** Base técnica de la recuperación densa que usan los sistemas modernos. Citado en Capítulo 4.

Documentación técnica de referencia

- **Google for Developers.** Documentación oficial sobre datos estructurados, AI Overviews y AI Mode. → https://developers.google.com/search/docs
- **Schema.org.** Vocabulario estructurado para marcado de entidades, productos, organizaciones y artículos. Referencia obligada para implementación de JSON-LD. → https://schema.org
- **Wikidata.** Base de conocimiento estructurada de libre acceso, citada en los capítulos sobre consolidación de entidad. → https://www.wikidata.org

- **Anthropic Research.** Publicaciones técnicas sobre el comportamiento de modelos de lenguaje, incluido el funcionamiento interno de Claude. → https://www.anthropic.com/research
- **OpenAI Research.** Documentación pública sobre sus modelos GPT, embeddings y la API de búsqueda web. → https://openai.com/research

Plataformas y herramientas mencionadas

A lo largo del libro se citan plataformas comerciales (Semrush, Ahrefs, G2, Capterra, Crunchbase, Trustpilot, entre otras) y motores generativos (ChatGPT, Claude, Gemini, Perplexity, Microsoft Copilot, Google AI Overviews y AI Mode). Las menciones tienen carácter ilustrativo. Las dos herramientas vinculadas a este libro son Elevam (elevam.es), la firma de crecimiento del autor, y Antropus (antropus.io), la plataforma SaaS de SEO + GEO desarrollada como parte del ecosistema Elevam.

Lectura complementaria recomendada

Para profundizar más allá de este libro, el autor recomienda seguir las publicaciones técnicas de Anthropic, OpenAI, Google DeepMind y Perplexity, que documentan periódicamente cambios en sus modelos. También las actas de conferencias académicas EMNLP, ACL y NeurIPS, donde aparece la investigación más rigurosa sobre recuperación y generación aumentada.

Nota: las URLs citadas en notas al pie a lo largo del libro forman parte de esta bibliografía y no se repiten aquí

para evitar duplicación.

Índice analítico

Referencia rápida de términos y conceptos por número de capítulo. Para definiciones detalladas, consulta el Glosario GEO de Elevam.

A Ahrefs, 1 · Alucinación de marca, 2, 11 · AI Overviews, 19 · Anti-IA-detection, 13 · Antropus (plataforma), 22 · Asociación de marca, 17 · Atribución difusa, 1 · Auditoría de entidad, 8, 11, 16 · Auditoría reputacional, 11, 18 · Auditoría técnica de extractabilidad, 11 · Autoridad de dominio, 4, 7

B Backlog unificado, 12 · Bain & Company, 1 · Baseline GEO, 10, 21 · Benchmarks trimestrales, 25 · Bing Chat (hoy Copilot), 2, 19 · Brand Association, 17 · Browsing (ChatGPT), 19 · Búsqueda determinista, 2 · Búsqueda probabilística, 2

C Capas de contenido, 13 · ChatGPT, 1, 2, 19 · Chunking, 7, 13 · Citation Coverage, 9 · Claude, 1, 2 · Clúster semántico, 5 · Co-Mentions, 9, 17 · Co-ocurrencia estratégica, 17 · Coherencia categorial, 16 · Compliance GEO, 24 · Concisión semántica, 14 · Consistencia NAP, 19, 20 · Context window, 6 · Core Web Vitals, 19, 21 · Crunchbase, 16, 20

D DataForSEO, 22 · Decisiones sin clic, 1 · Densidad informativa, 13 · Digital PR, 17 · Dilución vectorial, 14 ·

Docker (pipelines), 22

E E-E-A-T, 19 · Echo index, 19 · Elevam, introducción, 22, sobre el autor · Embeddings, 5 · Entity Home, 16 · Espacio semántico, 5 · Extractabilidad, 11

F Fan-out de subconsultas, 11 · FAQ estructurado, 7, 13, 15 · Featured snippets, 1 · Fragmentación de contenido, 7 · Frontlinks, 9

G G2, 18, 20 · Gemini, 1, 19 · Google AI Overviews, 1, 19 · Google Business Profile, 16, 19, 20 · GPTBot, 11 · Granjas de entidades, 24

H Hallucination, ver Alucinación · HTML semántico, 11, 13

I ICE / RICE (priorización), 12 · Information gain, 7, 19 · Intención híbrida, 6 · Inyección indirecta de prompts, 24

J JSON-LD, 15 · @id, 15 · @graph, 15

K Keyword (densidad), 5 · Knowledge Graph, 16 · Knowledge Panel, 16

L LangChain Agents, 22 · LinkedIn, 16, 17 · LocalBusiness schema, 15, 20 · Lost-in-the-middle, 6, 14

M Mantenimiento mensual, 21 · Manipulación de sentimiento, 24 · Memoria estática (LLM), 19 · Multimodalidad, 17

N NAP (Nombre, Dirección, Teléfono), 19, 20 · Named Entity Disambiguation, 16

O Observabilidad, 22 · Off-page, 17 · OpenAI, 2, 4, 19 · Optimización para extracción, 13 · OWASP, 24

P Página canónica, 8 · Páginas fuente, 7, 8 · PageRank, 1, 5 · Perplexity, 1, 2, 19 · Pillar pages, 6 · Pipeline GEO, 22 · Plantilla de informe, 23 · Precision@3, 12 · Pricing basado en valor, 23 · Primacía (efecto), 6, 14 · Prompt Set, 10, recursos · Prompts mágicos (mito), 3

Q Q&A (estructura), 13 · Quick wins, 21

R RAG (Retrieval-Augmented Generation), 2, 4, 7 · Recencia (efecto), 6, 14 · Reddit, 17, 18 · Redes espejo, 24 · Redes falsas, 24 · Reference Rate, 1 · Retrievability, 4 · Reviews (gestión), 18

S sameAs, 15, 16 · Schema avanzado, 15 · Schema drift, 15 · Scorecard LLM, 9 · Semrush, 22 · Sentiment Score, 9 · Share of Attention, 23 · Share of Model, 9, 21, 22, 23 · Shopping Graph, 20 · Shortlist Position, 3, 9 · Significancia estadística, 10 · Source of Truth, 8

T Tablas comparativas, 13 · Test A/B en GEO, 10 · TL;DR, 7, 13 · Token prediction, 2 · Tono comercial (evitar), 6 · Trustpilot, 18

V Vector drift, 5, 12 · Vectores multidimensionales, 5 · Visibility Auditor (Antropus), 22

W Wikidata, 16 · Wikipedia, 4, 16, 17 · WordPress, 13

Z Zero-click, 1

Sobre el autor

Asier López Ruiz es fundador y CEO de Elevam (elevam.es), una firma de crecimiento especializada en optimización para motores de búsqueda (SEO) y para motores generativos (GEO), con sede en Reus (Tarragona) y proyectos en mercados de Europa y Latinoamérica.

Elevam —antes FullSEO y FullWeb— fue nº 1 en el primer ranking de agencias GEO de España de MarketingDirecto (marzo de 2026), con 86/100. En 2023, Forbes España ya había seleccionado a FullSEO y FullWeb entre las 50 mejores agencias SEO de España.

Su trabajo se centra en convertir el conocimiento técnico en sistemas reales de visibilidad: empresas que pasan de ser invisibles para la IA a aparecer como primera recomendación en su categoría, sin atajos ni promesas vacías. Entre sus clientes hay empresas que facturan miles de millones al año en sectores como SaaS, servicios financieros, e-commerce y consultoría.

Antes de fundar Elevam, Asier atravesó la etapa más dura de su carrera profesional: proyectos mal elegidos, deslealtades empresariales, bloqueo cognitivo y una depresión real, no de LinkedIn. Salió reconstruyendo desde cero, con un compromiso claro: no más parches rápi-

dos, no más clientes que no encajan, no más servicios que no se puedan respaldar con resultados. Esa decisión es el origen de la metodología que aparece en este libro.

Es ponente habitual en eventos del sector, escribe sobre la trastienda real del marketing digital sin filtros corporativos, y es también fundador y CEO de Antropus (antropus.io), la plataforma SaaS de SEO + GEO que automatiza buena parte de los procesos descritos en este libro y que ya usan agencias, consultores y equipos in-house para medir su visibilidad real en ChatGPT, Claude, Gemini y Perplexity.

Vive en Reus (Tarragona), donde compatibiliza Elevam con la vida fuera del trabajo: familia, deporte y muchas horas leyendo sobre los temas que acabas de leer aquí.

Contacto profesional: asier@elevam.es

Firma de crecimiento (SEO + GEO consultoría): → https://elevam.es/

Plataforma SaaS de SEO + GEO: → https://antropus.io/

LinkedIn: → linkedin.com/in/asier-lopez-ruiz

Tu siguiente movimiento

Has terminado el libro. Ahora vienen las decisiones.

Si quieres aplicar lo que has leído o profundizar más, tienes dos puertas abiertas. Elige la que encaje con tu momento. Ninguna te compromete a nada.

1. Prueba Antropus

Si después de leer el Capítulo 22 has pensado «esto es exactamente lo que necesito pero no quiero construirlo desde cero», Antropus lo tiene resuelto.

Con el Prompt Tracker mides tu Share of Model en ChatGPT, Claude, Gemini, Perplexity, Google AI Mode y AI Overview; con Prompt Intelligence generas prompt sets metodológicos; y cada informe termina en un plan de acción priorizado. Puedes crear la cuenta gratis y sin tarjeta.

Antropus → antropus.io

2. Trabaja con Elevam

Si después de leer el libro tienes claro que necesitas que alguien lo aplique en tu caso (porque no tienes tiempo, equipo o ganas de ensuciarte las manos), Elevam trabaja con un número limitado de clientes al año. No es para

todo el mundo, ni hace falta que lo sea. Si encaja, encaja.

Contacto → elevam.es/contacto

Gracias por haber llegado hasta el final.

— *Asier*