MarkenTIQ

DOCUMENT DE RÉFÉRENCE · MARKENTIQ

Recherche IA : ce qui est prouvé, ce qui ne l'est pas

Dix-sept idées fausses sur la visibilité dans les réponses des IA, chacune confrontée à sa source. Version française du manuel de Warren Hance, revérifiée, prolongée de cinq mesures.

Version MarkenTIQ : 14 septembre 2026, mise à jour le 24 septembre 2026
Travail d'origine : Warren Hance, AI Search Playbook, The Index Club, septembre 2026
Sources du manuel : 21, dont 20 rouvertes et vérifiées
Sources ajoutées par MarkenTIQ : 3
Écarts relevés et corrigés dans le texte : 6
Idées fausses ajoutées par MarkenTIQ, sur nos relevés : 5

Avant de commencer

Warren Hance est un praticien britannique du référencement. Il a publié son manuel sous le nom de **The Index Club**, gratuitement et sans formulaire. Nous l'avons lu en entier, puis nous lui avons écrit pour lui proposer une version française. Il a dit oui, à deux conditions que nous remplissons ici : **le nommer, et renvoyer vers son site**, The Index Club.

Ce que nous avons fait en plus de traduire. Le manuel s'appuie sur vingt et une sources ; nous en avons rouvert vingt, la vingt et unième est une synthèse d'Ahrefs que notre texte n'utilise pas. Une par une, nous avons vérifié que chaque affirmation y figure bien. **La plupart tiennent. Six ne tiennent pas tout à fait.** Nous les corrigeons dans le texte, à leur place, sans les effacer : une affirmation retirée n'apprend rien, une affirmation corrigée sous vos yeux vous montre comment vérifier la suivante.

Nous ajoutons aussi. Cinq constats que la vérification a fait apparaître, et **cinq idées fausses de plus, mesurées chez nous**, qui portent le compte de douze à dix-sept. Elles sont regroupées à part et signalées comme nôtres.

Cette version n'est donc pas son document en français. **C'est son document, relu et prolongé.**

1. La guerre des sigles est une perte de temps

AEO, GEO, LLMO. Trois sigles pour dire « être visible dans les réponses des intelligences artificielles ». Ils sont utiles quand ils désignent une surface précise. **Ils deviennent nuisibles quand ils laissent croire qu'il existe des systèmes de classement entièrement séparés, avec des tactiques magiques.**

Sur ce point, la documentation de Google est inhabituellement claire : pour apparaître dans ses fonctionnalités génératives, **une page doit être indexée et éligible à un extrait, et rien de plus**. Aucune exigence technique supplémentaire, aucun balisage spécial.

Nous avons écrit ailleurs ce que chacun des trois recouvre et ce qui les sépare. Ce qui a changé n'est pas la mécanique du classement. **C'est l'interface.** La question posée au moteur est plus longue, la réponse est rédigée, et l'utilisateur ne clique plus forcément. Le travail de fond, lui, ressemble beaucoup à ce qu'il était.

2. Ce qui se passe vraiment entre la question et la réponse

C'est rarement expliqué, et c'est pourtant ce qu'il faut avoir en tête pour lire la suite.

Une réponse d'assistant se fabrique en plusieurs étapes, **et votre page peut échouer à chacune**.

Étape 1. Le moteur décide s'il va chercher. Certains produits cherchent toujours. D'autres décident au cas par cas. Quand ils ne cherchent pas, ils répondent de mémoire, à partir de ce qu'ils ont appris à l'entraînement. **Votre site n'est alors même pas consulté.** Cette décision est la première porte, et c'est la plus méconnue.

Étape 2. Une question devient plusieurs recherches. Google appelle cela le déploiement en éventail. Une demande comme « quel logiciel de paie pour une agence de trente personnes avec des indépendants en Europe » peut déclencher des recherches séparées sur le logiciel de paie, la conformité, la gestion des indépendants, les prix et les avis. **Une page qui ne vise que l'expression principale peut rater la sous-question qui obtient la citation.**

Étape 3. Le moteur récupère et reclasse. Il rassemble des documents ou des passages, leur donne un score, les remet dans l'ordre, et n'en garde qu'une partie. **Le moteur ne peut lire qu'une quantité limitée de texte en une fois.** Ce qui n'y entre pas n'existe pas pour cette réponse.

Étape 4. Il rédige, et il cite, ou pas. La citation est une couche à part. Chez Anthropic, une étape dédiée associe les affirmations à leurs sources après coup.

D'où la phrase à retenir : **être indexé n'est pas être récupéré, et être récupéré n'est pas être cité.** Une page peut avoir toutes ses chances d'être citée une fois récupérée, et rester commercialement invisible parce qu'elle n'est jamais récupérée.

Un chiffre pour situer le placement dans la page : selon l'étude de Kevin Indig sur les citations de ChatGPT (Growth Memo, 16/02/2026), les 30 premiers pour cent du contenu produisent **44,2 % des citations**, le dernier tiers **24,7 %**. Ce que vous mettez en bas de page compte moins.

---

3. Trois niveaux de preuve, et il faut les tenir séparés

La plupart des conseils qui circulent mélangent trois choses très différentes.

Ce qui est établi. Les documentations des moteurs. Elles disent ce qui rend une page éligible, quels robots font quoi, ce qui est mesuré. C'est vérifiable et ça engage leur auteur.

Ce qui est corrélé. De grandes études d'observation. Elles montrent des liens statistiques, pas des causes. Exemple : sur 75 000 marques étudiées par Ahrefs (12/12/2025), **les mentions du nom sur le web sont bien plus corrélées à la visibilité dans les IA que les indicateurs classiques de liens**. C'est une direction stratégique utile. **Ce n'est pas un facteur de classement**, et les auteurs le disent eux-mêmes.

Ce qui est du folklore. Des recettes reprises d'un article à l'autre, sans source, ou avec une source qui ne dit pas ce qu'on lui fait dire. La deuxième moitié de ce document en donne dix-sept exemples.

Le test, quand vous lisez un conseil : demandez de quel niveau il relève. Si personne ne sait répondre, vous tenez du folklore.

---

4. Les sept portes

Warren Hance propose un modèle en sept étapes, depuis le moment où une machine peut atteindre votre page jusqu'à celui où quelqu'un agit. **Il ne sert pas à décrire, il sert à trouver quelle étape bloque.**

1. **Atteignable.** Le robot peut-il récupérer et afficher la page ?
2. **Pertinente.** Répond-elle à la question réellement posée ?
3. **Crédible.** Y a-t-il de quoi la croire, et qui d'autre peut le confirmer ?
4. **Extractible.** L'information est-elle lisible et isolable, ou noyée ?
5. **Sélectionnée.** Entre-t-elle dans ce que le moteur peut lire en une fois ?
6. **Représentée.** La réponse dit-elle de vous quelque chose de juste ?
7. **Prête pour l'action.** L'étape suivante est-elle claire pour un humain comme pour un agent ?

Pourquoi une note globale sur cent ne sert à rien. Une page très citable mais jamais récupérée vaut zéro. Une page très citée peut nuire si la réponse vous présente de travers. **Un score unique donne la même note aux deux, alors que le problème n'est pas le même.** Il faut savoir quelle porte est fermée.

---

5. Qui a le droit d'entrer chez vous

C'est la confusion la plus coûteuse du moment, et elle tient en une phrase : **bloquer l'entraînement d'un modèle n'est pas bloquer la découverte par un assistant.** Ce sont deux décisions différentes, qui se prennent avec des instructions différentes, adressées à des robots différents. Beaucoup d'entreprises ont fermé la mauvaise porte en croyant bien faire.

ROBOT	CE QU'IL FAIT	CE QUE VOUS DÉCIDEZ EN LE BLOQUANT
`Googlebot`	exploration et indexation pour la Recherche Google et ses fonctionnalités	vous sortez de Google, y compris des réponses génératives
`Google-Extended`	entraînement de certains modèles Gemini et ancrage des réponses	une décision de gouvernance, **sans effet sur votre présence ni votre classement dans la Recherche**, Google l'écrit noir sur blanc
`OAI-SearchBot`	découverte pour la recherche web de ChatGPT	vous sortez des réponses de ChatGPT qui s'appuient sur le web
`GPTBot`	collecte pour l'entraînement des modèles d'OpenAI	une décision de gouvernance, distincte de la précédente
`PerplexityBot`	fait apparaître et lie des sites dans Perplexity, **pas d'entraînement**	vous sortez de Perplexity
`Perplexity-User`	récupère une page parce qu'un utilisateur l'a demandé	presque rien : **Perplexity écrit que ces récupérations ignorent généralement le `robots.txt`**
`ClaudeBot`	collecte des contenus web susceptibles de servir à l'entraînement des modèles d'Anthropic (ajout MarkenTIQ)	une décision de gouvernance : vos contenus futurs sont exclus de l'entraînement, distincte des deux suivantes
`Claude-SearchBot`	indexe les contenus pour améliorer les résultats de recherche de Claude (ajout MarkenTIQ)	votre visibilité dans les réponses de Claude qui s'appuient sur le web peut baisser, écrit Anthropic
`Claude-User`	récupère une page quand un utilisateur de Claude pose une question (ajout MarkenTIQ)	Claude ne peut plus récupérer vos pages pour répondre à un utilisateur
`MistralAI-User`	visite une page quand un utilisateur pose une question à Vibe ; ni exploration automatique ni entraînement, écrit Mistral (ajout MarkenTIQ)	Vibe ne peut plus lire vos pages pour répondre à ses utilisateurs

Mistral publie deux autres robots sur la même page : `MistralAI-Index`, qui indexe le web pour sa recherche, et `MistralAI-Training`, qui collecte des contenus pour l'entraînement de ses modèles. Pages d'Anthropic et de Mistral relues le 24 septembre 2026.

Trois points pratiques que ces documentations donnent et que peu de sites appliquent.

Autoriser le robot ne suffit pas chez OpenAI. Il faut aussi que votre hébergeur ou votre réseau de diffusion laisse passer les adresses IP publiées. Un pare-feu trop zélé produit exactement le même résultat qu'un blocage volontaire, sans que personne ne l'ait décidé.

Bloquer `GPTBot` ne vous efface pas partout. OpenAI précise que si l'adresse d'une page interdite lui vient d'un fournisseur de recherche tiers et qu'elle semble pertinente, **le lien et le titre peuvent**

apparaître quand même. Pour l'empêcher, il faut poser une balise `noindex`, l'instruction qui demande de ne pas référencer la page. Ce qui suppose de laisser le robot la lire, pour qu'il voie l'instruction.

Il existe un paramètre pour compter les visites venues de ChatGPT. OpenAI ajoute automatiquement `utm_source=chatgpt.com` aux adresses de renvoi. C'est la réponse la plus simple à la question « comment savoir qu'une visite vient d'une IA ». Elle est partielle : elle ne dit rien de ceux qui vous lisent dans la réponse sans jamais cliquer, ni de ceux qui retapent votre nom dans un moteur ensuite.

---

6. Le contenu qui survit à la synthèse

Quand un assistant résume trois pages qui disent la même chose, aucune des trois n'est indispensable. **La question utile n'est donc pas « comment bien écrire », c'est « qu'est-ce qui, chez nous, ne peut pas être fabriqué en résumant les autres ».**

Ce sont les informations qu'un concurrent ne peut pas produire sans faire le même travail que vous. Des données que vous seul avez mesurées. Un test que vous avez conduit vous-même. Un exemple réel, avec ses chiffres. Une méthode décrite assez précisément pour être rejouée. Un document d'origine. Un jugement d'expert assumé. Une connaissance locale que personne d'ailleurs ne possède.

Ce qui ne survit pas : la définition que tout le monde donne, la liste des dix bonnes pratiques, la reformulation d'une documentation officielle. Ce sont précisément les pages que la réponse générée remplace.

Trois conséquences concrètes.

N'écrivez pas une page par sous-question. Quand on découvre qu'une question se déploie en dix recherches, la tentation est d'écrire dix pages. Google déconseille explicitement le contenu à faible valeur produit en masse, et écrit qu'il n'est pas nécessaire de viser chaque formulation rare d'une même demande.

Construisez un ensemble cohérent où chaque page apporte une information vraiment distincte.

Mettez l'essentiel en texte. Google recommande de rendre les informations importantes disponibles sous forme textuelle. Un fait qui n'existe que dans une image, un graphique interactif ou un contenu chargé après coup peut ne pas être vu.

Placez-le haut. Dans l'étude de Kevin Indig citée à la section 2, les 30 premiers pour cent du contenu d'une page produisent 44,2 % des citations de ChatGPT, le dernier tiers 24,7 %. Ce n'est pas une raison pour bâcler la fin, c'en est une pour ne pas garder votre meilleure information pour la conclusion.

Et une mise en garde qui vient de la recherche, pas de nous : réécrire une page uniquement pour être citée peut la rendre plus difficile à récupérer. La revue de littérature de juillet 2026 mesure environ **9 % de présence en moins dans les vingt premiers résultats** pour une optimisation limitée au corps du texte. **On peut se rendre moins visible en cherchant à être mieux cité.**

7. Des mots-clés aux questions réelles

Le référencement classique part d'une expression et de son volume de recherche. **Un assistant ne reçoit pas des expressions, il reçoit des situations.** « Quel logiciel de paie pour une agence de trente personnes avec des indépendants en Europe » n'a pas de volume de recherche mesurable, et c'est pourtant la demande réelle.

Trois changements d'habitude suffisent.

De l'expression à la tâche. Demandez-vous ce que la personne essaie d'obtenir, pas ce qu'elle tape. La tâche se formule en une phrase : choisir, comparer, vérifier, décider, remplacer.

De la question aux sous-questions. Une demande se déploie en plusieurs recherches. Pour la paie citée plus haut : le logiciel, la conformité, les indépendants, le transfrontalier, les prix, les avis. **Listez ces sous-questions avant d'écrire.** Elles vous disent ce que votre page doit contenir, et surtout ce qu'elle n'a pas besoin de contenir parce qu'une autre page à vous le fait mieux.

Des mots aux entités. Les moteurs raisonnent sur des choses nommées : votre marque, vos produits, vos catégories, vos concurrents, les normes de votre secteur. **Si votre nom n'est associé à rien de précis sur le web, aucune formulation de page ne le rattrapera.**

Et une conséquence de méthode. Pour mesurer votre visibilité, il faut une liste de questions fixe, qu'on repose à l'identique à chaque relevé. Notre méthode de scan décrit comment la construire. Cette liste doit contenir des reformulations et des variantes, parce que la réponse d'un assistant change avec la façon de demander. **Ne cherchez pas l'originalité dans cette liste** : elle doit être assez stable pour être rejouée, assez large pour représenter la demande, assez détaillée pour dire où les choses bougent.

---

8. Mesurer, avec ce qui existe vraiment

Quatre mots sont confondus en permanence, et ils ne veulent pas dire la même chose.

- **Mention** : votre nom apparaît dans la réponse.
- **Citation** : votre page est donnée comme source.
- **Recommandation** : la réponse conseille de vous choisir.
- **Renvoi** : quelqu'un clique et arrive chez vous.

Être cité n'est pas être recommandé. Votre page peut servir à étayer un fait pendant que la réponse recommande un concurrent. Ces quatre mesures se relèvent séparément, sinon on se raconte des histoires. La grille que nous utilisons pour les tenir séparées est publiée dans le Pack GEO.

Ce qui est disponible aujourd'hui, et ses limites :

OUTIL	CE QU'IL DONNE	CE QU'IL NE DONNE PAS
Search Console, rapport IA générative	des impressions dans les fonctionnalités génératives de Google, déployé partout depuis le 31/08/2026	**aucun clic**, et rien sur les expérimentations en cours
Bing Webmaster Tools, AI Performance	citations totales, pages citées, et les formulations que l'IA a réellement utilisées pour aller chercher votre contenu, sur Copilot et les résumés de Bing	ni classement ni score d'autorité, Microsoft le dit explicitement
Le paramètre `utm_source=chatgpt.com`	ChatGPT l'ajoute automatiquement aux adresses de renvoi	rien sur les gens qui vous lisent sans cliquer
Votre liste de questions, rejouée	l'évolution de vos mentions, citations et recommandations dans le temps	une vérité absolue, voir ci-dessous

La règle qui vaut plus que les outils. Une même question, posée deux fois, ne donne pas la même réponse. Cela dépend du moteur, de la date, du lieu, de la formulation, du fait que la recherche soit déclenchée ou non, et d'une part de hasard. **Une capture d'écran n'est pas une mesure.** Il faut plusieurs passages, une liste de questions stable, et des fourchettes plutôt qu'un résultat unique.

Ce que nous avons constaté de notre côté, sur cinq moteurs et cinquante-huit réponses en français, les 28 et 29 août 2026 : **les douze inventions d'identité que nous avons relevées sont toutes survenues quand le moteur ne cherchait pas**, et aucune sur les cinquante réponses avec recherche. **Mais vingt-deux des vingt-quatre reprises fautives se sont produites avec recherche.** Autrement dit : chercher empêche d'inventer, pas de mal recopier. Nous y revenons plus bas.

---

9. Douze idées fausses, revérifiées une par une

C'est la partie la plus utile du manuel d'origine, et celle que nous avons le plus travaillée. **Chaque affirmation a été confrontée à sa source.** Quand la source ne dit pas tout à fait ce qu'on lui fait dire, nous l'écrivons.

1. « Ajoutez un fichier `llms.txt` et ChatGPT vous classera »

Faux comme règle générale. Google écrit que ces fichiers texte pour IA ne servent pas à apparaître dans sa Recherche, y compris ses fonctionnalités génératives, parce qu'elle ne les utilise pas. Et sur un échantillon de 137 210 domaines, Ahrefs a trouvé 38 360 fichiers valides, dont **97 % n'ont reçu aucune requête du mois mesuré**.

Nuance omise dans l'original, et Google la donne dans la même phrase : maintenir ce fichier pour d'autres services **ne nuit pas**. Il est ignoré, pas pénalisé. Si vous en avez déjà un pour une documentation technique, gardez-le. N'en attendez pas de visibilité.

2. « Le balisage structuré fait que les IA vous citent »

Non démontré comme cause générale. Une étude d'Ahrefs (11/05/2026) a suivi **1 885 pages** ayant ajouté du balisage structuré (le JSON-LD, ce petit bloc de code qui décrit à une machine ce que contient la page) entre août 2025 et mars 2026, comparées à 4 000 pages témoins. Résultat : **−4,6 % sur les AI Overviews, +2,4 % sur AI Mode, +2,2 % sur ChatGPT.** Rien.

Le balisage reste utile pour les fonctionnalités de recherche qui s'en servent et pour la clarté de vos données. **Arrêtez seulement de promettre qu'il fait citer.**

3. « Chaque réponse doit tenir dans un bloc de 40 à 60 mots »

Google dit le contraire : il n'y a aucune obligation de découper son contenu en petits morceaux pour qu'une IA le comprenne mieux.

Précision que nous ajoutons : la fourchette de 40 à 60 mots ne vient pas de Google. **Aucun chiffre de ce genre ne figure dans sa documentation.** Il circule dans le métier, c'est tout. Écrivez des sections complètes et lisibles.

4. « L'étude de Princeton a prouvé 40 % de classement en plus sur ChatGPT »

Non, et l'étude n'est pas de Princeton.

Trois corrections plutôt qu'une. **Sur l'attribution** : le premier auteur est à l'Indian Institute of Technology Delhi, deux auteurs sont indépendants, et **trois des six seulement sont à Princeton**. L'appellation courante est commode, elle est inexacte, et nous la corrigeons y compris dans le manuel d'origine.

Sur le mot qui manque presque toujours : le résumé de l'étude annonce **« jusqu'à 40 % »**. Au mieux, donc. C'est le haut de la fourchette, obtenu par la meilleure des neuf variantes testées, là où elle marche le mieux. **Repris en français, le « jusqu'à » tombe, et un maximum devient une moyenne.** Le manuel d'origine, lui, le garde.

Sur ce que le chiffre mesure : dans l'expérience, **cinq sources étaient placées d'avance dans le contexte du modèle**, tirées une fois pour toutes de Google. Le gain mesure la part de texte attribuée à une source dans la réponse, pondérée par sa position. **Ce n'est ni un classement, ni une récupération sur le web vivant, ni du trafic, ni du chiffre d'affaires.**

Ce que l'étude montre de solide : des citations, des chiffres, des affirmations sourcées et une écriture fluide changent la part de texte reprise. **Le bourrage de mots-clés, lui, n'aide pas.** La lecture prudente n'est pas « mettez des statistiques partout », c'est **une information concrète et attribuable se réutilise plus facilement qu'une prose vague**.

5. « Si vous êtes premier sur Google, vous serez cité »

Non. Être visible dans la recherche classique aide : elle vous rend éligible et récupérable. Mais un produit d'IA fait autre chose ensuite. Il déploie la question en plusieurs recherches, va chercher d'autres sources, réordonne des passages et peut citer des adresses qui ne sont pas dans les dix premiers résultats. Ce qui se recoupe entre les deux varie selon le moteur et dans le temps.

Signalé pour ce qu'il est : le manuel ne donne aucune source sur ce point. **Nous en ajoutons une, mesurée** : selon Ahrefs (15/04/2026), sur 1,4 million d'invites analysées, **88 % des adresses finalement citées par ChatGPT viennent directement de la recherche**. Être trouvable par un moteur reste donc central, mais la place exacte ne garantit rien.

6. « Si vous êtes cité, vous êtes recommandé »

Non. Une page peut être citée pour étayer un fait pendant que la réponse recommande un concurrent. Mesurez la citation et la recommandation séparément. **Sans source dans l'original**, mais notre propre grille de relevé distingue ces deux états depuis septembre 2026, précisément parce que nous avons vu le cas se produire.

7. « Le contenu frais gagne toujours »

Non. La fraîcheur compte beaucoup pour les faits qui bougent. Pour le reste, la pertinence, l'autorité et les preuves peuvent rendre une source ancienne très utile. **Mettez à jour les faits, pas les dates pour faire joli.** Sans source dans l'original.

8. « Telle plateforme communautaire est le facteur de classement des IA »

Aucune plateforme unique n'est un facteur universel. Du contenu communautaire de première main peut être utile pour certaines questions et certains moteurs. **Fabriquer des messages est une manipulation, pas une stratégie**, et Google déconseille explicitement de rechercher des mentions inauthentiques.

Ce que les corrélations disent vraiment, et c'est plus fort que le manuel ne le dit. Selon l'étude d'Ahrefs du 12/12/2025, sur 75 000 marques, le facteur le plus corrélé à la visibilité dans les IA n'est pas un lien. **Ce sont les mentions sur YouTube, à environ 0,737**, devant les mentions du nom sur le web.

Deux précautions de lecture. **Corrélation n'est pas causalité**, les auteurs le répètent eux-mêmes : rien ne dit qu'ouvrir une chaîne vous rendra visible. Et le manuel d'origine confond ici deux tableaux de la même étude, ce que nous corrigeons.

9. « Avec l'IA, le trafic ne compte plus »

Le trafic n'est plus le seul résultat visible, il reste décisif quand il a lieu. Ajoutez des mesures de représentation et de considération. N'abandonnez pas la mesure commerciale. Sans source dans l'original.

10. « Il faut une équipe de contenu séparée pour l'IA »

Le plus souvent non. Le travail durable croise le référencement, le contenu, les relations presse, les données produit, la mesure, l'accessibilité et la marque. **Une équipe en silo produit surtout des pages en double et des vérités contradictoires.** Sans source dans l'original.

11. « Il faut écrire pour les machines, pas pour les humains »

Les recommandations des plateformes disent l'inverse. Google met explicitement en avant le contenu utile et fiable, pensé pour les personnes d'abord, et écrit qu'il n'est pas nécessaire d'écrire d'une façon particulière pour la recherche générative. Et ce qui rend un site lisible par un agent recoupe largement ce qui le rend accessible à un humain.

12. « Il existe un algorithme stable qu'on peut rétro-concevoir »

Non. Ces produits peuvent utiliser des modèles différents, des fournisseurs de recherche différents, des outils, des lieux, des contextes et des modes différents. OpenAI écrit que le placement en tête ne peut pas être garanti. Anthropic documente que le modèle décide lui-même quand chercher et peut enchaîner dix recherches ou plus.

Et la revue de littérature la plus récente est catégorique : sur 45 études publiées entre novembre 2023 et juillet 2026, **aucune technique examinée ne montre d'effet causal stable, durable et valable sur plusieurs plateformes** sur la découvrabilité organique. Elle relève même qu'une réécriture orientée citation peut **nuire** à la récupération : environ 9 % de présence en moins dans les vingt premiers résultats, 16 % dans les dix premiers après reclassement.

---

10. Cinq idées fausses de plus, que nous avons mesurées

Les douze précédentes sont de Warren Hance. **Les cinq suivantes sont de nous**, et chacune vient d'un relevé daté, le plus souvent d'une erreur que nous avons commise avant de la comprendre.

13. « Il suffit que le moteur cherche sur le web pour qu'il dise vrai »

Non, et la mesure est nette dans les deux sens. Sur cinq moteurs et cinquante-huit réponses en français, les 28 et 29 août 2026, dans le prolongement de notre premier relevé public : **les douze inventions d'identité relevées sont toutes survenues quand le moteur ne cherchait pas**, et aucune sur les cinquante réponses avec recherche. Densité de 1,75 énoncé inexact par observation sans recherche, contre 0,44 avec. **Un facteur quatre.**

Et le contre-résultat, qui interdit la conclusion facile : vingt-deux des vingt-quatre reprises fautives se sont produites avec recherche. Le moteur qui cherche cesse d'inventer votre identité. **Il se met à mal recopier ce qu'il lit.**

Conséquence pratique : deux défauts différents, deux corrections différentes. Une identité inventée ne se corrige pas à la source, puisqu'elle n'est mémorisée nulle part. Une reprise fautive, si.

14. « Un score de visibilité sur cent dit où vous en êtes »

Non. Notre propre score d'audit technique est passé de 91 à 100 en une journée. **Le même jour, notre visibilité dans les réponses des IA valait zéro.** Nous avons raconté cette journée en détail. Les deux chiffres étaient justes, et ils ne parlaient pas de la même chose.

Une seule porte fermée annule tout le reste, et un score global ne le montre pas. Reprenez les sept portes de la section 4 : une page parfaite techniquement mais jamais récupérée vaut zéro, et un score global lui donnera 100. **Demandez toujours quelle porte le chiffre mesure.**

15. « Mesurer sa propre marque depuis son compte habituel est neutre »

Non, et c'est notre erreur la plus embarrassante. Le 29 août, nous avons découvert qu'un compte Google permet de déclarer des sources préférées, et que ces sources reçoivent un marquage visible dans les réponses générées. **Si vous cochez votre propre marque, vos relevés sont faussés en votre faveur, et rien ne vous le signale sinon une petite pastille.**

Le même piège existe ailleurs : **la mémoire d'un compte contamine une conversation neuve.** Nous avons vu un moteur reporter une invention d'une conversation antérieure dans une conversation ouverte pour la mesure, sur le même compte.

Ce que nous faisons depuis : quatre vérifications avant la première question d'un relevé, dont l'état des sources préférées du compte. Vérifié à nouveau le 13 septembre 2026 : zéro source déclarée. **Une de ces vérifications ne se rattrape pas après coup**, et le relevé est alors perdu.

16. « Mon outil d'analyse d'audience voit les visites venues des IA »

Presque jamais. Trois raisons se cumulent, et aucune n'est un défaut de votre outil.

La personne ne clique pas forcément sur le lien. Elle lit votre nom dans la réponse, puis le retape dans un moteur, ou saisit votre adresse. La visite est alors comptée en « recherche » ou en « direct », et l'IA disparaît du tableau.

Certains assistants lisent votre page sans exécuter le moindre script. Un agent qui travaille en ligne de commande récupère le texte et s'en va. Votre marqueur de mesure ne se déclenche pas.

Notre propre relevé du 13 septembre 2026 : 105 visites, **aucune attribuée à une intelligence artificielle**. Nous savons pourtant que des lecteurs nous arrivent par là.

Ce qui marche un peu : le paramètre `utm_source=chatgpt.com` ajouté par OpenAI, décrit à la section 5. Et la question posée de vive voix, « comment nous avez-vous trouvés », qui reste l'instrument le plus fiable dont personne ne veut.

17. « Le vocabulaire de vos questions de test n'a pas d'importance »

Il peut détruire la mesure. Nous avons interrogé des moteurs sur le GEO sans expliquer le sigle. **Sur huit surfaces, cinq ont dérivé** vers la géographie ou vers un développement inventé du sigle. Reformulé lors du même relevé (27 au 29 août), même dérive, aux mêmes endroits.

Les mêmes questions, avec le sigle expliqué en une incise, n'ont produit aucune erreur.

Le manuel d'origine recommande d'inclure des reformulations dans votre liste de questions. **Il ne dit pas que le vocabulaire de cette liste peut invalider le relevé entier.** C'est le cas. Un sigle, un nom de produit ambigu, une marque homonyme suffisent.

---

11. Les dix étapes, à punaiser au mur

1. **Partez d'une tâche réelle**, une question qui a une valeur commerciale.
2. **Vérifiez l'accès** : le moteur peut-il récupérer, indexer et afficher la meilleure page ?
3. **Cartographiez l'éventail** : quelles sous-questions le moteur va-t-il poser séparément ?
4. **Possédez un fait unique**, que personne ne peut fabriquer en résumant vos concurrents.
5. **Rendez-le explicite** : dates, unités, méthode, limites, origine.
6. **Faites-vous corroborer** : donnez à des tiers une vraie raison d'en parler.
7. **Testez un panel** : les mêmes questions, sur plusieurs moteurs, plusieurs fois.
8. **Lisez vos propres signaux** : Search Console, Bing, journaux de robots, renvois, clients.
9. **Changez une chose à la fois**, avec des pages de comparaison et plusieurs passages.
10. **Rendez l'action possible** : l'étape suivante doit être claire et mesurable.

---

12. Par où commencer si vous partez de rien

L'ordre compte plus que la liste, et il tient en cinq mouvements.

1. Vérifiez que vous êtes joignable. Avant toute chose : vos pages sont-elles indexées, et vos règles laissent-elles passer les robots que vous voulez ? Regardez aussi votre pare-feu, pas seulement votre `robots.txt`. **Ce premier point élimine plus de cas qu'on ne le croit.**

2. Posez un relevé de départ. Une dizaine de questions que vos clients posent vraiment, jouées sur trois ou quatre moteurs, plusieurs fois, avec quatre colonnes : mention, citation, recommandation, exactitude. **Sans ce point zéro, vous ne saurez jamais si quelque chose a bougé.**

3. Corrigez ce qui vous représente mal avant d'aller chercher plus de visibilité. Si les moteurs racontent faux sur vous, être davantage cité aggrave le problème.

4. Publiez une chose que personne ne peut résumer à votre place. Une mesure, un test, une méthode complète. Une seule, faite sérieusement, vaut mieux que dix pages de synthèse.

5. Rejouez le relevé, à date fixe. Même panel, mêmes moteurs. **Comparez des fourchettes, pas des captures d'écran.**

Ce qui ne sert à rien tant que les cinq points ci-dessus ne sont pas faits : ajouter un fichier `llms.txt`, empiler du balisage, découper ses paragraphes en blocs courts. **Les trois premiers mythes de ce document portent précisément sur ces trois gestes.**

Ce que nous n'avons pas traduit

L'audit en 122 points et le plan de 90 jours de Warren Hance ne sont pas repris : ce sont ses outils de métier, et les traduire reviendrait à reprendre un dispositif commercial, ce qui n'est pas l'objet de

MarkenTIQ.

Les fiches par plateforme ne sont pas reprises non plus : elles reformulent des documentations qui changent tous les mois, et une traduction les figerait dans un état périmé.

Les sources, pour que vous puissiez vérifier

Documentation officielle, à ouvrir directement. Guide d'optimisation pour les fonctionnalités génératives de Google (mise à jour du 10/07/2026) · Fonctionnalités IA dans la Recherche (10/12/2025) · Robots courants de Google (14/07/2026) · Rapport IA générative dans la Search Console · ChatGPT et la recherche web, centre d'aide OpenAI · FAQ éditeurs et développeurs d'OpenAI · Robots de Perplexity · Outil de recherche web d'Anthropic · Comment Anthropic a construit son système de recherche multi-agents (13/06/2025) · Achat dans ChatGPT et Agentic Commerce Protocol, OpenAI (29/09/2025) · La découverte de produits dans ChatGPT, OpenAI (24/03/2026) · AI Performance dans Bing Webmaster Tools (10/02/2026).

Recherche. *GEO: Generative Engine Optimization*, Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan, Deshpande, novembre 2023, publié à KDD 2024 · *Optimizing Visibility in Generative Engines: A Critical Survey*, Olivier Martinez, 15 juillet 2026.

Études d'observation, Ahrefs. Étude `llms.txt` (15/06/2026) · Étude balisage et citations (11/05/2026) · Pourquoi ChatGPT cite des pages (15/04/2026) · Corrélations de visibilité de marque (12/12/2025) · Corrélations sur les AI Overviews (26/05/2025) · La génération augmentée par récupération expliquée (09/07/2026).

Ces vingt sources sont celles du manuel que nous avons rouvertes (la vingt et unième n'est pas utilisée ici). **Ajoutées par nous**, relues le 24 septembre 2026 : l'étude de Kevin Indig sur la place des citations dans une page (*The science of how AI pays attention*, Growth Memo, 16/02/2026, reprise par Search Engine Land le 18/02/2026) · Robots d'Anthropic, centre d'aide Claude (mise à jour du 07/04/2026) · Robots de Mistral, documentation Mistral (page non datée).

Le manuel d'origine. Warren Hance, *AI Search Playbook*, The Index Club, septembre 2026. www.theindexclub.co.uk

---

Version française établie par MarkenTIQ le 14 septembre 2026, avec l'accord de Warren Hance. La section 10 et les cinq constats ajoutés ailleurs reposent sur nos propres relevés, datés dans le texte. Le détail de la vérification est disponible sur demande. Mise à jour du 24 septembre 2026 : études attribuées dans le texte, liste des sources complétée, robots d'Anthropic et de Mistral ajoutés.
