

Benchfolk Index — dossier méthodologique

Benchfolk · presse@benchfolk.com

7 juillet 2026

CONTENTS

title: "Benchfolk Index — dossier méthodologique" subtitle: "Comment on mesure la visibilité des grandes maisons dans les réponses des IA" version: "v1"

Benchfolk Index — dossier méthodologique	
Ce que l'Index mesure	
Les 4 IA testées — pondérées par usage réel	
Corpus de prompts (60 questions publiques)	
Univers des marques suivies	
Comment on compte les citations — deux métriques	
1. Citations (couverture)	
2. Score IA (score 0-100)	
Cercle de tête (core club)	
Évolution mensuelle	
Fréquence et reproductibilité	
Ce que l'Index ne capture pas	
Reproduire, contester, approfondir	

title: “Benchfolk Index — dossier méthodologique” subtitle: “Comment on mesure la visibilité des grandes maisons dans les réponses des IA” version: “v1.1 · juillet 2026” author: “Benchfolk · presse@benchfolk.com” date: “2026-07-01”

Benchfolk Index — dossier méthodologique

Version v1.1 · juillet 2026 Contact : presse@benchfolk.com

Ce que l'Index mesure

Le Benchfolk Index mesure une seule chose : **quand un amateur pose à ChatGPT, Claude, Perplexity ou Gemini une question typique sur son univers de consommation premium, quelles marques sortent dans la réponse ?**

C'est une mesure de *share of voice* dans les réponses des grands modèles de langage — l'équivalent, pour les IA génératives, du rang organique dans Google.

Ce n'est **pas** une évaluation qualitative des cuvées, ni un baromètre d'appréciation critique. Un Champagne peut être une œuvre absolue de vinification et être totalement absent des recommandations IA ; l'inverse est aussi vrai. Les IA répondent à partir de leurs sources d'entraînement et du web indexé — pas d'une dégustation en aveugle.

Les 4 IA testées — pondérées par usage réel

Nous interrogeons les quatre grands assistants IA généralistes utilisés par le public francophone, avec la recherche web activée quand elle est disponible :

IA	Modèle	Mode web	Poids
ChatGPT	<code>gpt-4o</code>	Recherche web activée par défaut	55 %
Gemini	<code>gemini-2.5-flash</code> (Google)	<code>google_search</code> grounding activé	25 %
Perplexity	<code>sonar</code>	Recherche web native	10 %
Claude	<code>claude-haiku-4.5</code> (Anthropic)	Outil de recherche web activé	10 %

Pourquoi pondérer ? Un utilisateur français ne pose pas ses questions à « une IA moyenne » : ~55 % des requêtes IA généralistes en France passent par ChatGPT, ~25 % par Gemini (couplage Google + intégration Android), le reste se répartit entre Perplexity, Claude et niches. Une mesure équi-pondérée sur-représenterait les IA de niche. Le Score IA reflète ce que *les Français entendent vraiment*, pas ce que 4 modèles disent à parts égales.

Source des poids : estimation basée sur Similarweb et l'observation du marché IA généraliste FR (juillet 2026), révisable tous les 3 mois. Google AI Overviews n'est pas testé (pas d'API publique stable).

Corpus de prompts (60 questions publiques)

Pour chaque catégorie — **champagne, cognac, whisky, Bordeaux** — nous testons **15 prompts** en français, rédigés dans le vocabulaire de la catégorie :

- Recommandation générale (« Quelles sont les meilleures maisons de champagne pour célébrer en France ? »)
- Budget (« Quel champagne offre le meilleur rapport qualité-prix ? »)
- Occasion (mariage, cadeau, gastronomie)
- Style (blanc de blancs, single malt, rive droite)
- Comparaison entre grandes maisons (« Compare-moi Krug, Dom Pérignon et Salon — lequel choisir pour quel usage ? »)
- Découverte de vigneron artisans

Les 60 prompts sont téléchargeables :

- benchfolk.com/index-prompts/champagne.json
- benchfolk.com/index-prompts/cognac.json
- benchfolk.com/index-prompts/whisky.json
- benchfolk.com/index-prompts/bordeaux.json

Licence CC BY 4.0 (libre de réutilisation avec attribution).

Nous excluons volontairement les prompts qui visent des cavistes ou des services d'abonnement (« quelle cave en ligne... ? », « quel abonnement mensuel... ? ») : ces questions évaluent des distributeurs, pas des maisons.

Univers des marques suivies

Notre univers de suivi comporte entre **15 et 28 marques par catégorie**, soit une fraction volontairement restreinte du marché réel.

Catégorie	Nb marques suivies	Marché réel	Source
Champagne	24	~ 16 300 producteurs	Comité Champagne + Syndicat général des vignerons, croisés RVF / Decanter / Wine Advocate
Cognac	10	~ 250 maisons	BNIC (distribution internationale documentée)
Whisky	28	~ 150 distilleries écossaises seules	Whisky Bible + rayonnement international
Bordeaux	15	~ 4 600 châteaux	1855 (Médoc + Sauternes) + Pomerol + Saint-Émilion 1 ^{ers} grands crus classés A

Cet univers n'est **pas exhaustif**. Si les IA citent une marque hors panel (par exemple Fleury en biodynamie champenoise), nous l'observons dans les réponses brutes mais elle n'entre pas dans le classement du panel officiel. Nous élargissons le panel chaque trimestre selon les mentions récurrentes observées et les demandes presse.

Comment on compte les citations — deux métriques

Pour chaque combinaison (*prompt* × *IA* × *marque*), nous vérifions par pattern matching précis (nom canonique + alias déclarés, avec bornes de mot) si la marque apparaît dans la réponse générée — et, quand elle apparaît, à quel rang par rapport aux autres marques citées dans la même réponse.

Deux métriques exposées côte à côte :

1. CITATIONS (COUVERTURE)

Nombre de prompts distincts où la marque apparaît dans la réponse d'**au moins une** des quatre IA testées. Maximum théorique par catégorie : 15.

C'est la **couverture** — êtes-vous mentionnée ou pas ?

2. SCORE IA (SCORE 0-100)

Pondération par rang d'apparition dans chaque réponse LLM :

Rang dans la réponse	Points
1 ^{er} nommé	10
2 ^e	7
3 ^e	5
4-5 ^e	3
6 ^e et +	1
Absente	0

Le total est normalisé sur le maximum théorique ($\sum \text{prompt_weight} \times \sum \text{IA_weight} \times 10 \text{ pts}$) où le poids de chaque IA est celui du tableau ci-dessus et le poids de chaque prompt reflète son volume de recherche relatif dans la catégorie (structure en place, calibrage Ahrefs prévu édition d'août 2026 — poids uniforme = 1 en attendant).

C'est la **Score IA** — êtes-vous nommée en tête de l'esprit de l'IA, ou perdue au milieu d'un paragraphe ?

Le classement principal utilise la Score IA.

Pourquoi ces deux métriques ? Parce que « Krug est cité en 12^e position d'un paragraphe qui liste 15 champagnes » et « Veuve Clicquot est le premier nom qui sort quand l'IA recommande un champagne d'apéritif » ne signifient pas la même chose pour une maison qui veut être choisie par l'acheteur au moment de la décision.

CERCLE DE TÊTE (CORE CLUB)

Le plus petit sous-ensemble de marques qui capte 50 % ou plus de la Score IA totale de la catégorie. Un indicateur de concentration : plus le cercle est petit, plus le marché IA est verrouillé sur une poignée de noms.

Évolution mensuelle

Chaque édition mensuelle est comparée à la précédente. Le classement expose deux deltas par marque :

- **Δ rang** — mouvement dans le classement (↑ montée, ↓ descente, = stable). Une marque nouvellement dans le classement affiche « *New* ».
- **Δ Score IA** — variation du score /100. Filtre le bruit statistique LLM : ± 2 = bruit, ± 5 = signal, ± 10 = mouvement structurel.

Cette colonne apparaît à partir de l'édition d'**août 2026** — juillet 2026 étant la 1^{re} édition mensuelle publiée, il n'y a pas encore de mois précédent à comparer.

Fréquence et reproductibilité

L'Index est régénéré **une fois par mois**, le premier lundi à 06 h 00 Europe/Paris. Chaque édition est identifiée par sa semaine ISO (par exemple `2026-W27`) et conservée en base.

Pourquoi mensuel plutôt qu'hebdomadaire ? Parce que les IA génératives ont un rythme d'évolution mensuel (mises à jour de modèles, indexation web) et que le rythme éditorial de la presse vin est mensuel/bimestriel. Une lecture mensuelle lisse le bruit statistique naturel des LLMs (non-déterminisme) et donne un « Baromètre du mois » plus lisible qu'un « chiffre de la semaine ».

Les IA génératives ne sont **pas déterministes** : deux exécutions du même prompt peuvent produire des réponses différentes. Nous en tenons compte de deux manières :

1. **Le texte intégral de chaque réponse est conservé en base.** Un mouvement semaine-vs-semaine dans le classement s'explique en remontant à la donnée brute.
2. **Gate qualité** : si plus de 20 % des appels IA d'une édition échouent (rate limit, timeout, refus modéré), l'édition n'est pas publiée — les données brutes sont conservées pour diagnostic.

Ce que l'Index ne capture pas

- **La qualité intrinsèque du vin ou du spiritueux.** Une absence de l'Index n'est pas un jugement de qualité, c'est un déficit de présence dans les sources entraînant les IA (Wikipedia, presse indexée, Knowledge Graph, sites structurés).
- **Les marques hors panel.** Si les IA citent régulièrement une maison que nous ne suivons pas encore, elle n'apparaît pas dans le classement — mais elle est observée dans nos réponses brutes et prise en compte à l'itération suivante du panel.
- **La géolocalisation des utilisateurs.** Nous interrogeons les IA en français depuis nos serveurs, sans contexte utilisateur. Les IA peuvent personnaliser leurs réponses selon l'historique, la région IP, la version d'app — nos chiffres représentent la « ligne de base » sans personnalisation.
- **Le lien de causalité citation IA → vente.** Nous mesurons la visibilité ; nous ne prétendons pas mesurer directement l'impact commercial.
- **La performance sur les langues autres que le français.** L'Index actuel est FR-only. Extension EN prévue Q4 2026.

Reproduire, contester, approfondir

Nous fournissons sur demande :

- Le fichier complet des prompts (également téléchargeable sous licence CC BY 4.0 sur benchfolk.com/index-prompts/)
- Les réponses brutes agrégées de la dernière édition

- L'historique semaine par semaine
- L'accès aux critères de sélection des marques (grille détaillée par catégorie)

Contact presse dédié :

Martin Ohannessian, fondateur presse@benchfolk.com benchfolk.com

Contestation méthodologique bienvenue : si vous pensez que notre panel omet une maison essentielle, que nos prompts sont biaisés, ou que notre métrique passe à côté d'un phénomène important, écrivez-nous. Chaque évolution du protocole est datée et documentée publiquement sur benchfolk.com/methodologie-index.

Version v1.1 · juillet 2026 · Publiée sous CC BY 4.0