

ÉTUDE STRATÉGIQUE · AVRIL 2026

↑ Reddit dans les *réponses des LLM.*

Ce que 9 286 citations nous apprennent sur la visibilité des marques dans l'AI Search.

Étude · Insights stratégiques

Pourquoi *cette étude*.

15 questions opérationnelles posées par les marques et les agences. Une base de données de citations réelles pour y répondre.

Les 4 thèmes couverts

- A** **Stratégie de présence**
Q1 · Q2 · Q3
- B** **Signaux d'engagement**
Q4 · Q5 · Q6
- C** **Structure du contenu**
Q7 · Q8 · Q9 · Q10
- D** **Sentiment, subreddits & mécanique LLM**
Q11 · Q12 · Q13 · Q14 · Q15

Le jeu de données

- 9 286** citations Reddit analysées (sur 15 783 récupérées)
- 291** projets clients couverts
- 1 854** subreddits uniques
- 2 066** threads Reddit uniques
- 04/04 – 20/04 / 2026** fenêtre temporelle

Le point *clé*, en une phrase.

La visibilité Reddit dans l'AI Search est un jeu de « *présence dans les bonnes conversations* », pas de « publication et promotion de son propre thread ».

Quand un LLM cite Reddit, il pointe bien plus souvent vers une discussion organique tierce que vers un thread créé par une marque. Le contenu cité est typiquement âgé de **1 à 3 ans**, court, au format Q&R, et provient plus souvent d'un **commentaire** que du post d'origine. Les LLM **reformulent plutôt qu'ils ne citent** (80 % de paraphrases), et puisent dans l'ensemble de la distribution des upvotes — pas uniquement dans les contenus très upvotés.

Cinq enseignements *à retenir.*

22 %

des citations pointent vers un thread dédié à la marque.

Les 78 % restants sont des threads organiques où la marque apparaît par infiltration. Publier son propre thread est 3x moins productif que s'insérer dans une conversation existante.

62 %

des citations proviennent d'un commentaire, 38 % du post.

Le titre du post ne fait que la moitié du travail. Celui qui rédige le premier commentaire substantiel capte l'essentiel de la valeur.

82 %

des commentaires cités ont entre 1 et 5 upvotes.

Les LLM lisent toute la plage d'upvotes — pas seulement les commentaires viraux. Couverture et volume comptent plus que l'engagement par commentaire.

280 j

c'est l'âge médian du contenu Reddit cité (~9 mois).

Le contenu récent n'a pas de prime ; la tranche 1–3 ans concentre 32 % des citations. L'horizon pertinent d'une stratégie Reddit est de 6 à 24 mois.

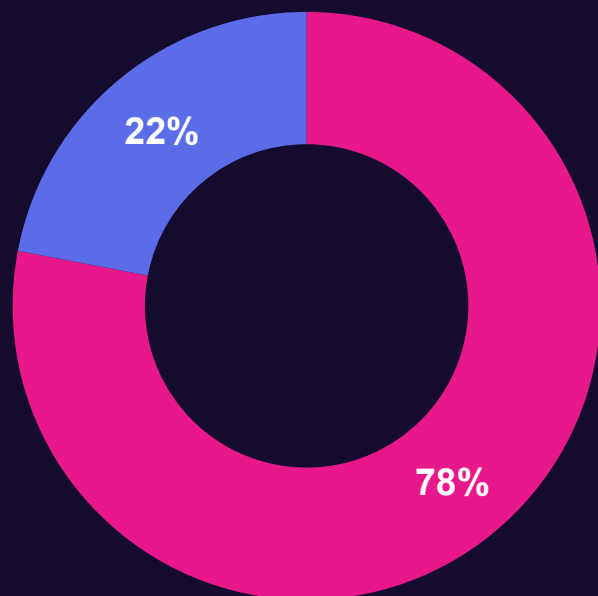
80 %

des citations sont des paraphrases, pas des citations textuelles.

Similarité cosinus médiane de 0,62. Les LLM utilisent Reddit comme matière première conceptuelle, pas comme source à citer.

Thread dédié ou *infiltration* ?

Question 1 — Vaut-il mieux créer un thread autour de sa marque, ou intervenir dans des threads existants ?



■ Infiltration (thread organique) ■ Thread dédié (posté par la marque)

L'infiltration l'emporte largement.

Sur 6 248 citations traçables à une stratégie :

4 874 threads organiques cités (infiltration)

1 374 threads dédiés cités (marque)

Implication stratégique

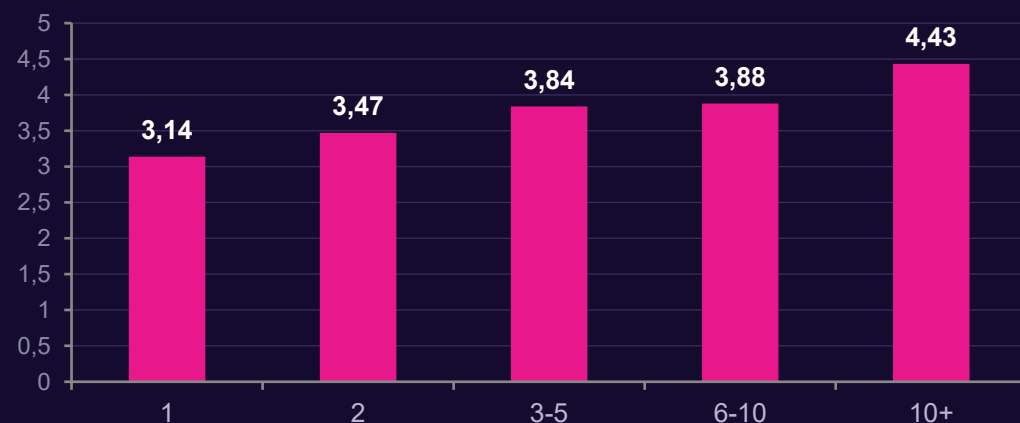
Les AMA orchestrés et les posts auto-publiés sous-performent d'un facteur 3. Le vrai gisement est la participation dans les discussions qui existent déjà.

Densité de mention & *co-présence*.

Questions 2 & 3 — Répéter une marque dans un thread : effet réel mais modeste. Cohabiter avec des concurrents : effet positif, pas dilutif.

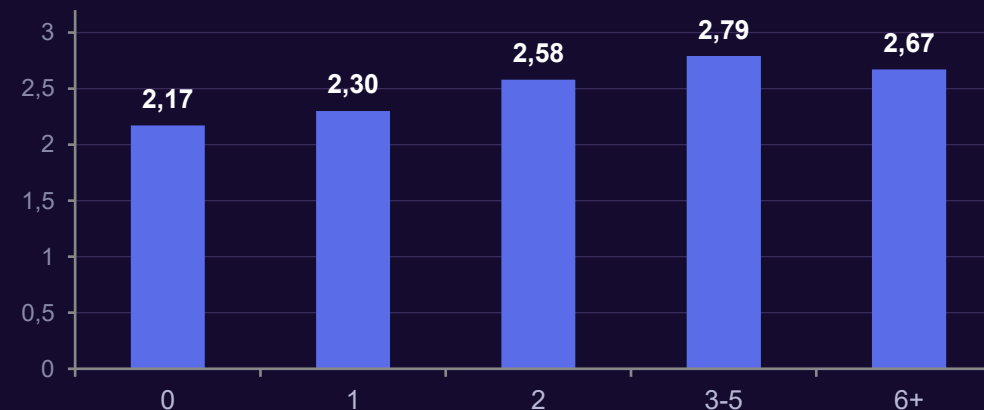
Mentions de la marque dans un thread

Citations LLM moyennes du thread



Concurrents co-cités dans la réponse

Citations Reddit moyennes par réponse LLM



L'evergreen *gagne*.

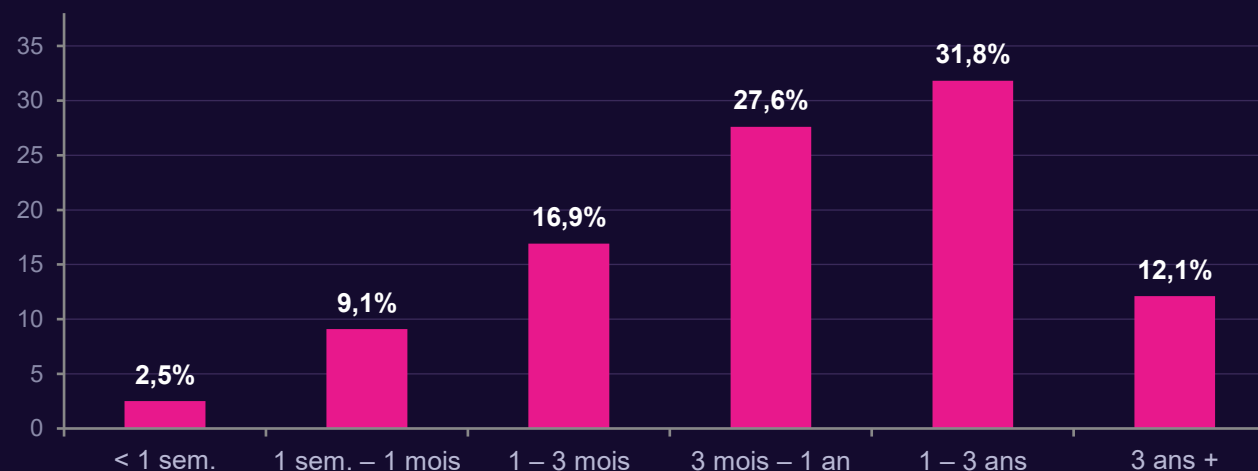
Question 6 — Les LLM favorisent-ils le récent ou l'evergreen ? Existe-t-il un âge optimal ?

280 j

âge médian du contenu Reddit cité (~9 mois)

La tranche 1–3 ans concentre à elle seule 31,8 % des citations. Les 10 premières semaines ne représentent que 11,6 %.

Répartition des citations par âge du post (%)

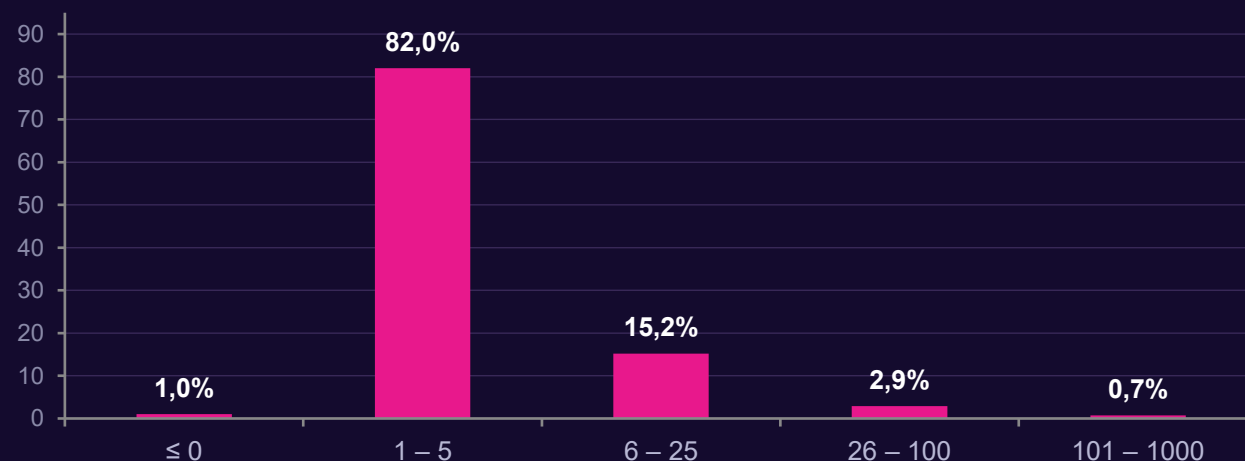


Implication · l'horizon pertinent d'une stratégie Reddit est de 6 à 24 mois. Les KPI trimestriels sous-récompenseront le travail de fond.

Les upvotes *ne sont pas un filtre.*

Questions 4 & 5 — Upvotes du commentaire cité et karma de l'auteur : signaux faibles, très loin des idées reçues.

Upvotes du commentaire le plus proche (%)



82 %

des commentaires cités ont 1 à 5 upvotes.

Les LLM ingèrent tout le thread, pas seulement les commentaires « top ».

Karma auteur

Pas de prime pour le fort karma.

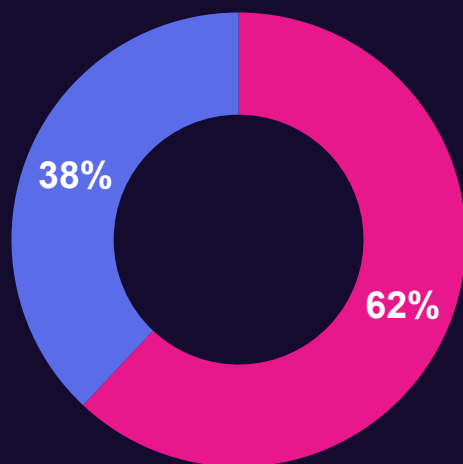
Être reconnu sur Reddit ne rend pas plus citable par un LLM.

Contre-intuitif · la chasse à l'upvote et le recrutement de contributeurs à fort karma sont mal alignés avec le fonctionnement réel des LLM.

D'où le LLM *tire-t-il l'extrait* ?

Questions 7 & 9 — Post vs commentaires, et longueur optimale des extraits cités.

Origine de la citation



■ Commentaire ■ Corps du post

Longueur optimale · court et dense

522

caractères

médiane du selftext des posts cités (~ 1 paragraphe)

150 – 400

caractères

longueur médiane des commentaires cités dans le thread

0,6 %

seulement

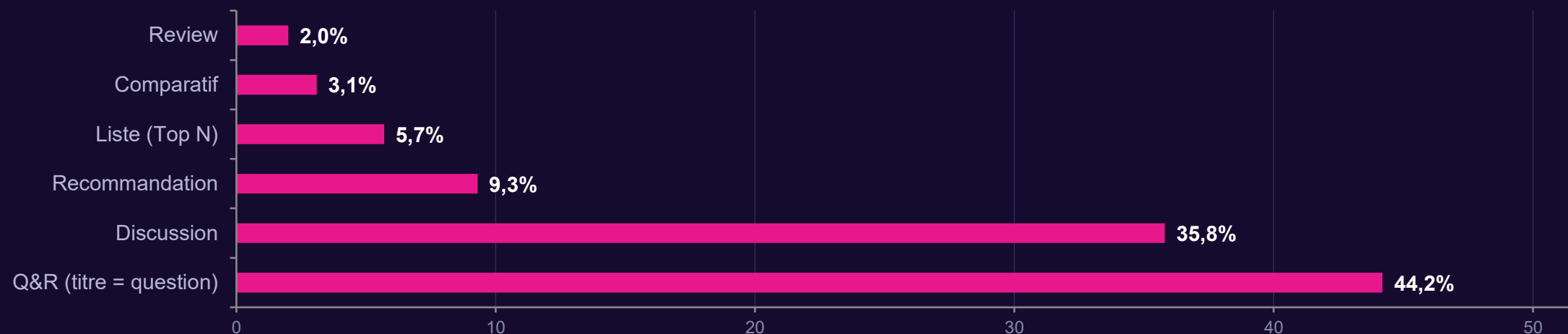
des citations viennent de commentaires > 1 000 caractères

Implication · viser le premier commentaire clair, de longueur moyenne, en phase avec la question — plutôt que des posts longs et exhaustifs.

Les LLM adorent *le Q&R.*

Question 8 — Le format du post compte : question-réponse devant tout le reste, avec près du double de la discussion classique.

Répartition des citations par format de post (%)



Implication · viser les threads dont le titre est une question. Écrire des réponses structurées façon Q&R, même dans des threads de discussion ouverte.

Une citation Reddit est *traçable*.

Question 10 — Quand un LLM cite Reddit, inclut-il une URL ?

65 %

des citations incluent l'URL Reddit directement dans la réponse LLM.

Une mention de marque sans URL Reddit visible est le schéma minoritaire. La citation IA est typiquement vérifiable en un clic.

Avec URL vs sans URL

URL Reddit incluse

65 %

Mention textuelle seule

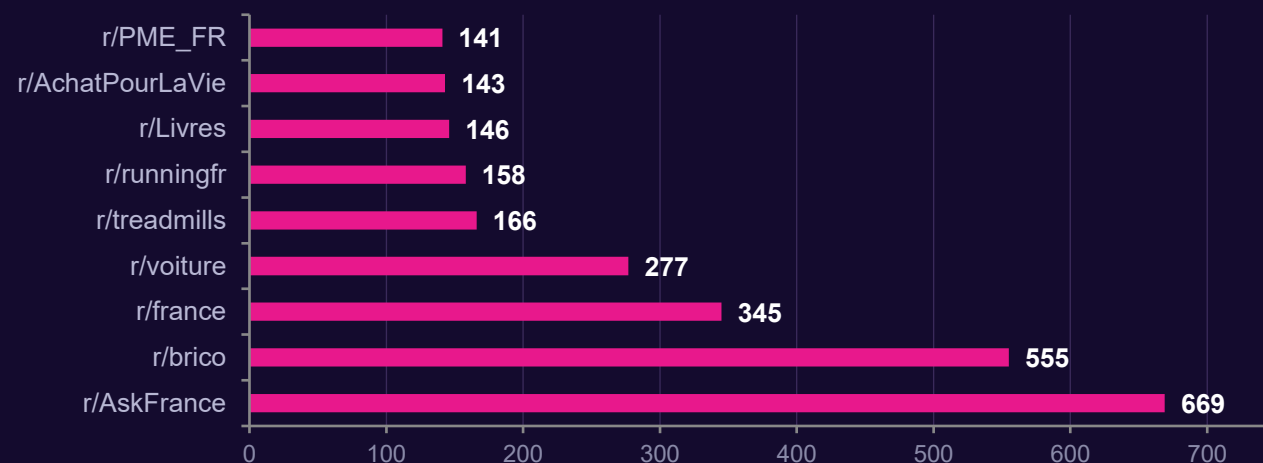
35 %

Implication · le trafic de retour via les réponses IA est un KPI réel et mesurable, pas une hypothèse.

Où vivent *les conversations.*

Questions 12 & 13 — 9 subreddits concentrent l'essentiel du volume. r/AskFrance est le levier universel francophone.

Top 9 subreddits — nombre de citations



Note · la cartographie complète 944 industries × 1 007 subreddits est disponible pour un travail stratégique client par client.

51

projets bénéficient de r/AskFrance.

Sur 291, c'est le plus gros levier unitaire de visibilité IA sur le marché francophone.

Verticaux

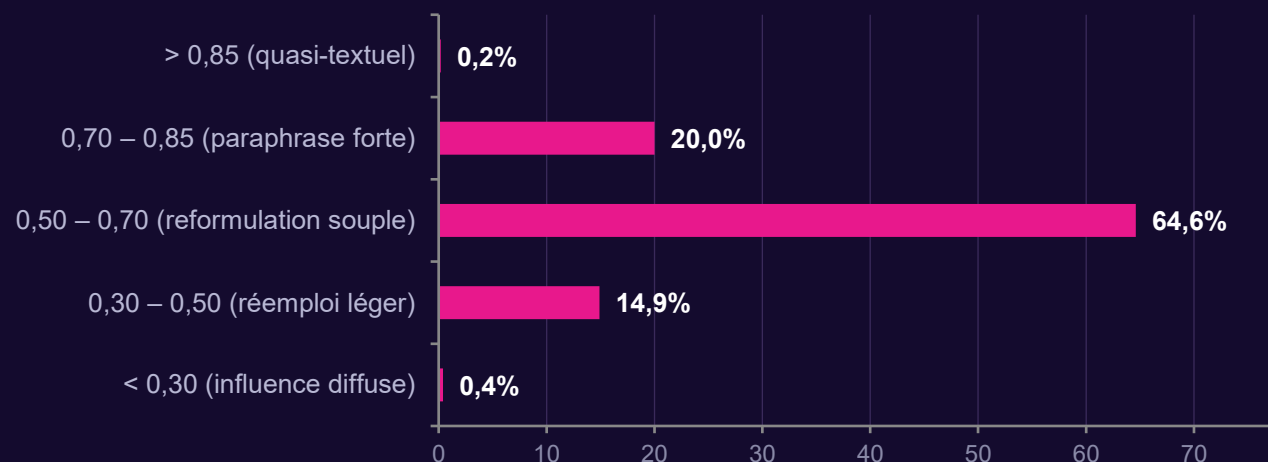
Chaque secteur a son hub.

Bricolage → r/brico · Automobile → r/voiture · Running → r/runningfr · Fitness indoor → r/treadmills. Cartographie industrie × subreddit disponible en livrable.

Les LLM *paraphrasent*, ils ne copient pas.

Questions 14 & 15 — Similarité cosinus entre la réponse LLM et le contenu Reddit cité.

Distribution de la similarité sémantique (%)



Ce que ça implique

0,62

similarité médiane

Reformulation souple, le mode dominant.

80 %

paraphrases, pas des citations

La citation textuelle n'est quasiment jamais le mode.

0,2 %

de quasi-textuel

Les LLM réécrivent dans leur propre voix 99 % du temps.

Implication · écrire pour la paraphrase, pas pour la citation. Optimiser pour « la manière dont un utilisateur Reddit réel expliquerait cela ».

Cinq *actions* pour les marques et agences.

01

Passer de « créer du contenu » à « semer des conversations ».

Identifier 5–10 threads en cours pertinents par mois plutôt que de publier de nouveaux threads. Gain : ×3 sur la productivité de visibilité LLM.

02

Prioriser le commentaire stratégique plutôt que le post principal.

62 % de la valeur se joue dans les commentaires. Viser le premier commentaire clair, de longueur moyenne, en phase avec la question, indépendamment des upvotes obtenus.

03

Écrire pour la paraphrase, pas pour la citation.

Les formulations commerciales sont perdues. Structurer pour répondre à une vraie question, dans un langage communautaire naturel.

04

Raisonner sur un horizon 6–24 mois.

Le contenu Reddit placé aujourd'hui n'apparaîtra pas dans les réponses LLM avant plusieurs mois. Adapter les KPI trimestriels en conséquence.

05

Cartographier chaque client sur ses 3–5 subreddits productifs.

r/AskFrance est universel pour les marques francophones. Les verticaux ont leurs hubs concentrés. La cartographie industrie × subreddit est prête en livrable.

Méthode & *périmètre.*

Pipeline d'analyse

1

Enrichissement

30 colonnes dérivées par citation depuis le JSONL brut.

2

Tier A — local (quelques sec.)

Q1, Q2, Q5, Q6, Q8, Q9, Q10, Q12, Q14.

3

Tier B — joins DB (quelques min.)

Q3 (concurrents), Q11 (sentiment), Q13 (cartographie industrie).

4

Tier C — similarité sémantique

~20 min CPU · Q4, Q7, Q15 · sentence-transformers multilingue.

Limites acceptées

- **Analyse descriptive, pas causale**
- Heuristique de format à 89 % d'accord
- Karma lu au moment du fetch, pas historique
- Top-30 commentaires par thread pour la similarité

Parlons de votre cas.

La cartographie industrie × subreddit est disponible immédiatement pour tout client. Nous pouvons déployer cette méthode sur votre marque sous 2 semaines.