

Être référencé par TOUS les moteurs — et la recherche intelligente

SOS Disque · sos-disque.fr — complète le dossier SEO classique. Google + Bing + ChatGPT + Gemini + Perplexity + Copilot + Claude.

La vérité contre-intuitive : le « référencement IA » repose à ~80 % sur le SEO classique. ChatGPT lit l'index **Bing** ; Gemini/AI Overviews lisent l'index **Google** ; Perplexity a son index. **Si tu n'es pas indexé, l'IA ne peut pas t'inventer.** Donc : SEO propre + 3 couches spécifiques.

Les 5 chantiers (par impact)

#	CHANTIER	POURQUOI
1	robots.txt : autoriser les bots de RECHERCHE	Tu bloques GPTBot/Perplexity/Google-Extended → tu t'interdis d'être cité
2	Être indexé partout (Bing Webmaster + IndexNow)	Bing nourrit ChatGPT/Copilot/DuckDuckGo ; IndexNow = 1 clic Cloudflare
3	Contenu « answer-first » + stats + citations	Étude Princeton : +28 à +41 % de visibilité IA
4	Entité + réputation tierce (fiches locales, Reddit, avis Trustpilot )	Les IA recommandent ce qui est corroboré ailleurs
5	Schema d'entité (Organization + sameAs)	Aide l'IA à savoir « qui » tu es

1 Par moteur : source · robot à autoriser solide

MOTEUR	SOURCE	ROBOT(S) À AUTORISER
ChatGPT Search	index propre + Bing	OAI-SearchBot · ChatGPT-User
Perplexity	index propre + tiers	PerplexityBot · Perplexity-User
Google AI Overviews / Gemini	index Google	Googlebot (<i>Google-Extended</i> ≠ <i>AI Overviews</i>)

Microsoft Copilot	index Bing	Bingbot
Claude (web)	crawl Anthropic	Claude-SearchBot · Claude-User
DuckDuckGo	index Bing	(Bingbot)
Brave	index propre	Bravebot
Apple (Siri/Spotlight)	index propre	Applebot
Ecosia/Qwant (FR)	Bing + index « Staan » émergent	(Bing)

2 Chantier 1 — le robots.txt cible prioritaire

Séparer les bots de **RECHERCHE** (à autoriser = être cité) des bots d'**ENTRAÎNEMENT** (choix éditorial, blocables sans perdre la citation). ⚠ Le blocage actuel vient du bloc **Cloudflare Managed** → à désactiver d'abord.

```
# RECHERCHE – AUTORISER
User-agent: Googlebot
Allow: /
User-agent: Bingbot
Allow: /
User-agent: OAI-SearchBot
Allow: /
User-agent: ChatGPT-User
Allow: /
User-agent: PerplexityBot
Allow: /
User-agent: Perplexity-User
Allow: /
User-agent: Claude-SearchBot
Allow: /
User-agent: Claude-User
Allow: /
User-agent: Applebot
Allow: /
User-agent: Bravebot
Allow: /

# ENTRAÎNEMENT – reco : autoriser aussi (commerce local : être recommandé >
protéger le texte)
User-agent: GPTBot
Allow: /
User-agent: ClaudeBot
Allow: /
User-agent: Google-Extended
Allow: /
```

```
User-agent: *
Allow: /
Sitemap: https://sos-disque.fr/sitemap-index.xml
```

3 Chantier 2 — être indexé partout

- **Bing Webmaster Tools** (gratuit) : vérifier + soumettre le sitemap. Couvre Bing + ChatGPT Search + Copilot + DuckDuckGo. *(le plus rentable après le robots.txt)*
- **IndexNow** : Bing/Yandex/Seznam/Naver (⚠ pas Google). Activation **1 clic via Cloudflare → Cache → Crawler Hints** (gratuit).
- **Google** : Search Console (déjà). **Apple Business Connect** : fiche Plans/Siri.

4 Chantier 3 — le contenu qui se fait citer étude Princeton

Gains mesurés de visibilité dans les moteurs génératifs (KDD 2024) :

ACTION	GAIN
Ajouter des citations / verbatims	+41 %
Ajouter des statistiques chiffrées	+33 %
Améliorer la fluidité	+29 %
Citer des sources crédibles	+28 %

⚠ Le « +115 % en gras » qui circule = mythe, absent du papier.

Appliqué à SOS Disque, chaque page : réponse en tête (answer-first) · 2 stats sourcées (Backblaze / tes cas) · FAQ Q→R (vraies questions : « disque clac-clac ? », « prix récup Paris ? », « SSD mort ? ») · tableau symptôme→cause→récupérable · bloc « 💡 À retenir ».

5 Chantier 4 — entité & réputation tierce

Tu es une entreprise de 6 mois : les IA ne te « connaissent » pas. Multiplier les sources cohérentes :

- **3 fiches locales**, NAP (nom/adresse/tél) **identique partout** : Google Business + Bing Places + Apple Business Connect.
- **Reddit r/datarecovery** + forums FR (très cités par les IA), **avis Trustpilot** ✅ (collecter du réel après chaque dossier) + Google avis.

- **Contenu citable** : l'étude « taux de panne 2026 » (data Backblaze + tes cas) → les autres la citent → les IA te nomment comme **source**.
- **E-E-A-T** : page À propos factuelle, mentions légales, auteur identifié.

6 Schema 2026 & mesure

JSON-LD only. Donne encore des étoiles/rich results : **LocalBusiness, AggregateRating, BreadcrumbList. Déprécié** (plus d'affichage Google) : FAQPage (rich result supprimé 07/05/2026), HowTo (2023) — mais **FAQ reste utile pour les IA**. Pour l'entité : **Organization + sameAs** (relie Google Business, Trustpilot, LinkedIn, Wikidata).

Mesure : GA4 → groupe « AI Traffic » (regex `perplexity\.ai|chatgpt\.com|gemini\.google\.com|copilot\.microsoft\.com`) — ⚠ ~70 % du trafic IA tombe en « Direct » (sous-estimé). + **tester les prompts** chaque mois dans les 4 IA.

⚡ Igor (serveur GPU) — ligne rouge vs usages légitimes

🚫 **Ne jamais faire** : générer en masse de **faux avis / faux témoignages** (« grâce à SOS Disque j'ai sauvé mes photos » ×100 000). = **astroturfing** : illégal (faux avis = pratique trompeuse, DGCCRF / AI Act), **détecté et puni** (SpamBrain, bannissement Trustpilot/Google), et **contre-productif** (les IA cherchent un consensus de vraies sources). **On ne force pas le message, on le mérite.**

✅ Ce qu'Igor peut faire (puissant ET légal) :

1. **Brouillons de contenu en masse** (LLM local) → relus/validés par toi (sceau humain). Beaucoup de vrai contenu, vite, gratuit.
2. **L'étude data citable** : Backblaze + tes ~2000 cas → « Taux de panne 2026 » → source que les IA reprennent.
3. **Observatoire GEO** : interroger automatiquement ChatGPT/Perplexity/Gemini sur des centaines de prompts → mesurer si/où tu es cité.
4. **Analyse sémantique** des vraies questions → quelles pages écrire.
5. **Version honnête de ton souhait** : récolter et mettre en avant les **vrais témoignages** de tes vrais clients → sentiment positif RÉEL capté par les IA.

Igor accélère la **production de vrai** et la **mesure**, pas la triche.

7 Plan 90 jours & prompts Largo

Mois 1 : robots.txt · Bing Webmaster + IndexNow · schema Organization+sameAs · GA4 AI Traffic. **Mois 2** : pages answer-first + 3 fiches locales + collecte avis. **Mois 3** : étude data citable · Reddit/forums · presse · test mensuel des prompts.

▼ Prompts prêts à coller pour Largo (Claude Code)

Détaillés dans le .md (§11) : P1 robots.txt · P2 Bing Webmaster + IndexNow (Cloudflare Crawler Hints) · P3 schema Organization+sameAs · P4 gabarit answer-first · P5 GA4 AI Traffic.

▼ Sources (vérifiées 06/2026) & mythes corrigés

OpenAI/Perplexity/Anthropic docs bots · Google-Extended ≠ AI Overviews (doc Google) · ChatGPT Search = Bing (openai.com) · étude GEO Princeton arXiv 2311.09735 · Profound/Peec (citations Reddit) · IndexNow.org (Google NON) · Cloudflare Crawler Hints (gratuit) · Google : FAQ rich result supprimé 07/05/2026 · Applebot (Apple) · Staan/EUSP (Ecosia/TechCrunch 08/2025) · llms.txt rejeté par Google (Illyes 2025) → à ignorer. Mythes écartés : « +115 % gras », « Google supporte IndexNow », « 46,7 % Reddit » (périmètre différent).

Dossier GEO · SOS Disque · 10/06/2026 · faits sourcés, mythes corrigés · complète le dossier SEO classique.