

GEO Casa · deck 4 of 4

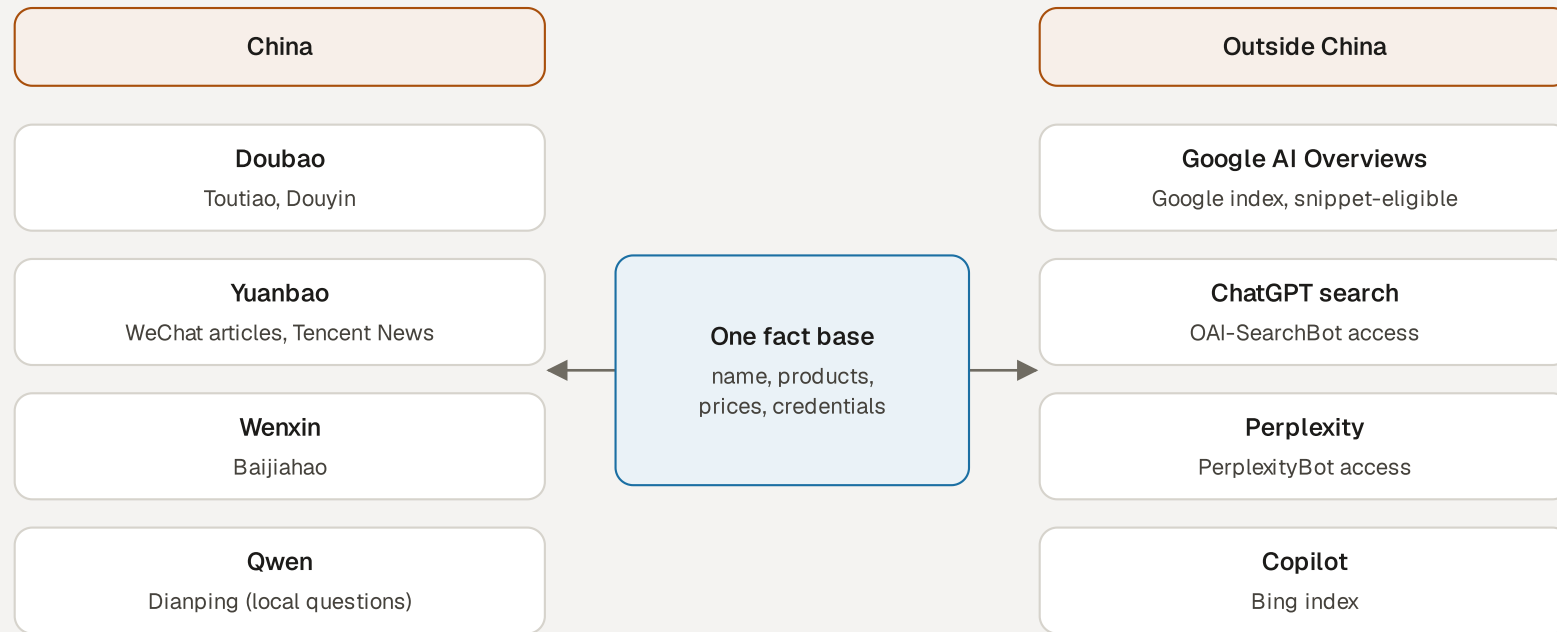
GEO by market

China platforms · global platforms · cross-border brands

GEO



Do China and abroad need separate work?



Observed source preferences (2025–2026 industry studies) · documented access rules

China's source preferences are dated industry observations; the global column is documented access rules.

[Read the full guide](#)



China: why can't one article serve every assistant?

In a Hefei study (3,469 citations), Doubao cited classifieds, Douyin and Toutiao most; Yuanbao Toutiao and Tencent; Qwen Dianping.

● Observation · 2026-08 · 合肥图灵象人工智能科技有限公司

Of 51,000 sources cited across seven platforms, only 9.34% were cited by four or more.

● Observation · 2025-10 · 增长黑盒 · 百分点科技

In China you optimise for ByteDance, Tencent, Baidu and Alibaba separately.



Where each assistant looks

Different assistants, different sources. The basis column says how sure we are.

Assistant	Reads mainly	Your lever	Basis
Doubao	Toutiao, Douyin, local classifieds	Publish inside ByteDance's ecosystem	Observation
Yuanbao	WeChat articles, Tencent News	Official account articles	Observation
Qwen	Dianping for local questions	Complete local listings	Observation
Google AI Overviews	Google's search index	Indexed, snippet-eligible pages	Official
ChatGPT search	Pages OAI-SearchBot may reach	Allow OAI-SearchBot	Official
Perplexity	Pages PerplexityBot may reach	Allow PerplexityBot	Official

China rows: industry studies from 2026, which change over time. Global rows: platform documentation.



Doubao: ByteDance's own channels first

In the Hefei study, Doubao's top sources were local classifieds (854), Douyin (244) and Toutiao (111).

● Observation · 2026-08 · 合肥图灵象人工智能科技有限公司

On pharma topics, Doubao leaned to ByteDance's Toutiao accounts, Douyin and Kuaidong Baike.

● Observation · 2026-06 · 艾昆纬 (IQVIA) · 上海零假设信息科技

A 2026 guide reports Doubao's citations were more concentrated and stable than Qwen's, DeepSeek's or Yuanbao's.

● Observation · 2026-09 · GDMS全球数字营销峰会 · 秒针营销科学院

For Doubao, a Douyin account and a Toutiao account do more than another page on your site.



Yuanbao: WeChat and Tencent, tested in the app

On pharma topics, Yuanbao leaned to WeChat official accounts and Tencent News.

● Observation · 2026-06 · 艾昆纬 (IQVIA) · 上海零假设信息科技

In the Hefei study, Yuanbao recommended the studio in 3% of API runs but 99% of app runs.

● Observation · 2026-08 · 腾讯新闻 (报道合肥图灵象普拉提行业GEO白皮书)

WeChat's AI search mostly drew on 5–10 official-account posts plus 1–3 pages from Tencent's ecosystem.

● Observation · 2026-06 · 艾昆纬 (IQVIA) · 上海零假设信息科技

An API test can tell you the opposite of what your customers see.



Qwen and ERNIE: thin evidence so far

We found only a few dated observations. Treat them as things to test.

Qwen, local questions



Dianping was its top source in the Hefei study (207 citations).

Qwen, over time



Its versions differed widely in what they cited. Retest after updates.

ERNIE (Wenxin)



On pharma topics it leaned to Baijiahao and Baidu's medical Q&A sites.

Baidu's crawler



Baidu's FAQ names no separate robots.txt token for its AI products.

A pharma study found Wenxin leaning to Baijiahao and Baidu's medical Q&A sites.

● Observation · 2026-06 · 艾昆纬 (IQVIA) · 上海零假设信息科技



DeepSeek: it changes, and it reads complaints

DeepSeek narrowed to academic and government sites in late 2025, then widened to portals such as Sohu and NetEase from Q2 2026.

- Observation · 2026-06 · 艾昆纬 (IQVIA) · 上海零假设信息科技

In the Hefei study, DeepSeek recommended the studio in 69% of API runs but 1% of app runs.

- Observation · 2026-08 · 腾讯新闻 (报道合肥图灵象普拉提行业GEO白皮书)

DeepSeek cited complaint and consumer-rights platforms 75 times, 23.9% of its citations, and often cited business registry lookups.

- Observation · 2026-08 · 腾讯新闻 (报道合肥图灵象普拉提行业GEO白皮书)

DeepSeek's sources change over time, so re-test regularly.



Kimi: Zhihu, Bilibili and more overseas sites

Two studies, two categories. Little else is published.

Kimi favoured Zhihu, Bilibili and mainstream portals, with more overseas sites than the others, especially in Agent mode.

● Observation · 2026-06 · 艾昆纬 (IQVIA) · 上海零假设信息科技

On infant formula questions in June 2025, Kimi relied on Zhihu (24.5%) and general portals.

● Observation · 2025-06 · Xsignal (奇异因子)

If your category is discussed on Zhihu, Kimi is likely to hear about it there.



Teams selling in China

China: the playbook

- 1 Measure in the consumer apps, not only APIs
- 2 Map each assistant's ecosystem for your category and city
- 3 Publish in each ecosystem; don't copy one article everywhere
- 4 Keep Baidu, 360, Sogou and Toutiao crawlers able to reach you
- 5 Look after complaint and registry pages about you
- 6 Stay inside the rules on AI data poisoning

Check what others publish about you regularly, especially complaint and registry pages.

[Read the full guide](#)



China: keep the search crawlers in

Chinese assistants lean on Chinese search and content platforms. Each crawler has its own rules.

Crawler	What the docs say	Watch out for	Basis
Baiduspider	Follows robots.txt	Blocking it removes you from Baidu search	Official
360Spider	Supports user-agent, allow, disallow	No reverse DNS check; match its IP ranges	Official
Sogou	Supports robots.txt	Changes can take several weeks	Official
Bytespider	Toutiao Search crawler, IP ranges from 2019	Its page doesn't mention robots.txt	Official

From each search engine's own help pages, as listed in the China platforms guide.



China: do assistants read complaint and registry pages?

In the Hefei study, DeepSeek cited complaint and consumer-rights platforms 75 times, 23.9% of all its citations.

● Observation · 2026-08 · 腾讯新闻 (报道合肥图灵象普拉提行业GEO白皮书)

Across six Chinese platforms, cited content was mostly ranking lists (22.8%–72.0%); user reviews were under 3% in every category.

● Observation · 2026-08 · 数说故事

Answer every open complaint and check your registry record before you write another article.



Studio A Pilates

★★★★★ 4.8

Address **12 Example Road** [same everywhere](#)

Phone **0551-0000 0000** [same everywhere](#)

Hours **09:00–21:00** [same everywhere](#)

Illustration; made-up shop.



China: the regulator is watching GEO

The CAC's 2026 campaign targets AI data poisoning, including forging authoritative data and using GEO for malicious marketing.

- Official · 新华网 (中央网信办通知)

Since 1 September 2025, AI-generated content must carry labels, and users who publish it must declare it.

- Official · 国家互联网信息办公室

CCTV's 3·15 show reported a GEO tool getting a fitness band that did not exist recommended near the top by two major AI models.

- Observation · 2026-03 · 第一财经

What worked on 3-15 is exactly what the regulator now targets.



Myth: seed mentions everywhere

? **Mass-publishing articles that mention us is the fastest way into AI answers.**

No evidence for it

It has moved answers in tests, which is why China now targets it as data poisoning. Google says such mentions help less than they seem, because its AI features rely on systems that block spam.

Google: seeking inauthentic mentions across the web "isn't as helpful as it might seem".

● Official · Google Search Central

[Read the full entry](#)



How does a page get into each AI search?

Google: a supporting link "must be indexed and eligible to be shown in Google Search with a snippet".

- Official · Google Search Central

OpenAI: sites opted out of OAI-SearchBot "will not be shown in ChatGPT search answers".

- Official · OpenAI

Open to Google but blocked for OAI-SearchBot means invisible in ChatGPT search, however good the page.



Google AI Overviews: built on Search

AI Overviews show only when Google decides they add something beyond classic Search, and so "often don't trigger".

• Official · Google Search Central

Google grounds answers in pages its core ranking systems retrieve, and fans out into related sub-queries to find more.

• Official · Google Search Central

A site "must be included in Search generative AI features in Search Console" to be eligible.

• Official · Google Search Central

Rank in Google first. AI Overviews draw on the same index.



AI Overviews and AI Mode: are you eligible?

Walk one important page through Google's documented conditions.

1 Can Googlebot crawl the page under robots.txt?

⚠ Allow Googlebot. Google lists robots.txt rules for Googlebot as a control on AI features.

2 Is the page indexed in Google Search?

⚠ Check it in Search Console and fix noindex tags or indexing problems first.

3 Can it show with a snippet (no nosnippet)?

⚠ Remove nosnippet. It keeps content out of AI Overviews and AI Mode; max-snippet limits how much is used.

4 Is the site included in Search generative AI?

⚠ In Search Console, Settings > Search generative AI. Exclusion removes the site from both features.

✓ **Eligible. No special markup or AI text file is required.**

[Read the full guide](#)



ChatGPT search: one bot decides

OpenAI: sites opted out of OAI-SearchBot "will not be shown in ChatGPT search answers, though can still appear as navigational links."

- Official · OpenAI

GPTBot is a separate setting for training and does not affect search. Systems take about 24 hours to adjust.

- Official · OpenAI

Blocking training and staying in ChatGPT search are two separate switches.



Perplexity: which visit is which

Click a line. Perplexity runs two fetchers with different rules.

```
203.0.113.10 - - [12/Oct/2026:09:14:02] "GET /pricing HTTP/1.1" 200 "PerplexityBot/1.0"
```

Needs checking The search crawler, if the IP matches Perplexity's published list. It follows robots.txt.

```
198.51.100.7 - - [12/Oct/2026:09:20:45] "GET /blog/guide HTTP/1.1" 200 "Perplexity-User/1.0"
```

Needs checking A fetch for a user's question. Per Perplexity it generally ignores robots.txt.

```
192.0.2.55 - - [12/Oct/2026:09:21:03] "GET /wp-login.php HTTP/1.1" 404 "PerplexityBot/1.0"
```

Needs checking Probing a login page doesn't look like search crawling, and the name in the user agent proves nothing. Check the IP before trusting it.

Simulated log lines with documentation-range IPs. Check real visits against each vendor's published IP list.

Perplexity-User fetches pages when users ask: "this fetcher generally ignores robots.txt rules."

● Official · Perplexity



Copilot: Bing's index and product feeds

Microsoft: Copilot breaks a shopping query down using crawled web data and product feeds; "traditional SEO remains essential".

- Official · Microsoft Advertising

Bing's AI Performance report counts citations in Copilot and Bing AI answers. "It simply shows which content was cited."

- Official · Microsoft Bing

Submit your product feed and verify in Bing Webmaster Tools.



Gemini: Google-Extended, not Search

One robots.txt token for Gemini training and grounding, separate from Search.

⚠️ What it covers



Training future Gemini models, plus grounding in Gemini Apps and on Vertex AI.

⚠️ What it doesn't



Inclusion or ranking in Google Search.

⚠️ In your logs



Nothing. It has no user agent of its own; it is read only from robots.txt.

⚠️ AI Overviews?



Use the Search generative AI control. It doesn't affect training.

"Google-Extended does not impact a site's inclusion in Google Search nor is it used as a ranking signal in Google Search."

● Official · Google



Preferred sources and Discover

Readers who pick you as a preferred source may see your content with a "preferred" badge in Top Stories, AI Mode and AI Overviews.

• Official · Google Search Central

Only domain-level and subdomain-level sites are eligible as preferred sources; a /blog folder is not.

• Official · Google Search Central

"Traffic from Discover is less predictable or dependable when compared to keyword-driven search visits."

• Official · Google Search Central

Ask loyal readers to add you. Don't plan around Discover traffic.



Brands selling in China and abroad

Cross-border: split the work

- 1 Two programmes, one shared fact sheet
- 2 One URL per language, linked with hreflang
- 3 Open each market's crawlers
- 4 Write natively for each market
- 5 Measure each market in its own apps

Reviews earned in one market don't carry over. Build them again in the next.

[Read the full guide](#)



Languages, domains and moves

Google recommends a different URL for each language version, not cookies or browser settings, and no automatic redirects.

• Official · Google Search Central

If two pages don't both point to each other, hreflang tags will be ignored. Each version lists itself and all others.

• Official · Google Search Central

Moving domains: submit Change of Address for all verified variants of the old domain, and keep redirects at least a year.

• Official · Google Search Central

Settle the URL structure before entering a market. Changing it later is harder.



Measure each market on its own

Different tools for each side. Never add the numbers up into one score.

Where	What to use	Keep in mind	Basis
China	Prompts run in the consumer apps	API and app can give opposite results	Observation
China	30+ runs per prompt	Low-visibility brands need 60+	Observation
Google	Search Console Generative AI report	Group by country	Official
Bing / Copilot	Bing AI Performance report	Shows citations, not causes	Official
Any vendor tool	Prompt sampling	No access to Google's internal AI systems	Official

China rows: Hefei study (2026-08) and Miaozen (2025-09). Other rows: platform documentation.



Launching in a second market

A phased plan suggested by one industry report. It hasn't been tested.



Meltwater suggests a central team plus regional teams; the report shows no results for it.

[Read the full guide](#)



Know which assistants your customers use

Each slide condenses a full guide. Read the guide for the evidence behind it.

China platforms

Global platforms

Cross-border brands

China AI platform sources

GEO regulation in China

Search generative AI control

Google-Extended

Preferred sources

hreflang

Site moves

Back to the overview deck