Generative engine optimization spam: what the GEOHunter study measured

27 November 2026 · 4 min read · blogs.iictp.org/generative-engine-optimization-spam-network-study

Short answer: GEOHunter is a study, posted on arXiv on 19 November 2026 by Sean Tang and co-authors, that measured a paid generative engine optimization (GEO) operation publishing promotional articles to influence AI answers. Across nine generative engines, the authors attribute between 4.5% and 20% of retrievable cited pages to that one network. It matters because marketers need to know what coordinated content can do to AI recommendations, and what platforms may treat as manipulation.

TL;DR The short version

- GEOHunter, an arXiv study of 19 November 2026, measured a paid generative engine optimization network that publishes promotional articles to sway AI answers.
- Across nine generative engines, the authors attribute 4.5% to 20% of retrievable cited pages to that single network.
- Coordinated content can shift AI recommendations, but platforms may treat it as manipulation; invest in genuine, well-sourced content instead.

Read online, listen to it, or change the text size: https://blogs.iictp.org/generative-engine-optimization-spam-network-study

What is generative engine optimization?

Generative engines retrieve web pages and combine them into answers, recommendations and rankings. Generative engine optimization is the practice of trying to appear in those results. The paper looks at one form of it: services that sell visibility by placing promotional articles on platforms where users publish content.

How the study worked

The authors traced one paid-content operation across three major Chinese user-generated-content platforms. They identified accounts linked to the network, inferred which posts were promotional placements, and then checked how often pages from the network were cited by nine generative engines from the United States and China, and whether its clients were named in answers.

What is new compared with lab work on GEO

Much discussion of GEO rests on controlled experiments that rewrite a page and watch what a model does. This work observes a commercial network already operating, with many accounts acting over months and across platforms, and changing its list-style articles over time.

Results reported by the authors

- 1,985 accounts associated with the network.
- 466,357 inferred promotional placements.
- Between 4.5% and 20% of retrievable cited pages attributable to the network, depending on how strictly attribution is drawn.
- On eight of the nine engines, the network's clients were named in 52% of answers, compared with 68% for category leaders.
- The authors conclude that content-level defences are not enough and that account-level and platform-level signals are needed.

This is an arXiv preprint or workshop paper. Treat its numbers as the authors' own results, not as independently confirmed findings.

Limitations and open questions

- We read the abstract only. Query sampling, attribution rules and error rates are in the full paper.
- One network on Chinese platforms may not represent other languages or markets.
- Attribution is inferred, which is why the authors give a range.
- Appearing in answers is not the same as winning customers, and the abstract reports no sales effect.
- The authors raise particular concern for medical and financial topics, where misleading advice can cause harm.

Sources were read on 27 November 2026. Products, rules and data change, so check the linked originals before you rely on any detail.

How to apply it without crossing the line

1. Ask each GEO supplier where content will be published, under whose accounts and with what disclosure.
2. Reject offers built on many undisclosed accounts or invented comparisons.
3. Publish accurate first-party pages that answer comparison questions with sources.
4. Sample AI answers for your category on a fixed schedule and record which pages are cited.
5. Report coordinated false content to the platform that hosts it.
6. Check advertising disclosure rules in each market before paying for placements.

What it means for brand and search teams

Citation tracking should include a source quality check, not only a mention count. A rise in mentions driven by anonymous list articles is a reputational risk that can reverse when engines or platforms filter accounts. Our note on the Google back button hijacking policy is a reminder that tactics which work briefly can later count as spam.

Frequently asked questions

- **What is generative engine optimization?** Work intended to make a brand or page appear in AI-generated answers.
- **Did the study find GEO placement effective?** The authors report that the network's pages and clients appeared often in generative engine results.
- **Are all GEO services spam?** The paper studies one paid content network and does not judge every service.
- **How many engines were tested?** Nine, from the United States and China.

Related reading

- Search for Agents endpoint
- Google ranking principles and feedback
- Generative AI controls in Google Search
- AI Practitioner certification

> Related credential: AI Practitioner. Explore the exam facts and syllabus.

Sources and further reading

1. Tang, Wang, Li, Shivakumar, Meng, Tian. GEOHunter: Mapping a Generative Engine Optimization Network in the Wild. arXiv 2611.26564. 19 November 2026 (preprint). arxiv.org/abs/2611.26564

Published by IICTP, an independent technical credential body, on 27 November 2026. The current version of this post is at https://blogs.iictp.org/generative-engine-optimization-spam-network-study. More guides: blogs.iictp.org.