

How to Run Query Fan-Out Analysis in Screaming Frog Using Gemini AI

Metehan Yesilyurt

AI Search & Information Retrieval Researcher · metehan.ai

Published: June 26, 2025 · Canonical: <https://metehan.ai/blog/query-fan-out-screaming-frog-ai/>

Do you know you can run Query Fan-Out analysis directly in Screaming Frog?

I just built a custom JavaScript that uses Gemini AI to predict how Google's AI Mode breaks down your content into sub-queries. After testing it on over 2000 pages, I discovered something that can concern every SEO: most pages only answer 30% of the queries that Google's AI actually generates from their content.

That's a 70% optimization gap I've all been missing.

Read my latest post: [How to 2x your AI traffic with the dumbest and simplest method?](#)^[1]

Important Note: All results are experimental. I used Gemini 1.5 Flash and 2.5 Pro(stable). Flash models are working in lightning speed. Some users reported that Pro model endpoint didn't work for them and crawl waits. Only running an optimization plan about Query Fan Out can break your current rankings.

++You should enable JavaScript Rendering on Screaming Frog

Access script here: <https://github.com/metehan777/screaming-frog-query-fan-out>^[2]

Updated: This is my configuration screen.

·
js

Related post: [Automate Your Backlink Opportunities with Screaming Frog](#)^[3]

The Query Fan-Out Reality

When Google's AI Mode processes your content, it doesn't just read it linearly like traditional search. Instead, it explodes your main topic into a network of related sub-queries, searching for comprehensive answers across its Knowledge Graph and web index.

Think about it: When someone searches "sustainable marketing strategies for e-commerce," Google's AI doesn't just match keywords. It generates questions like:

- What makes a marketing strategy sustainable?
- How much budget do small e-commerce businesses need?
- Which channels work best for online stores?

- How do you measure sustainability in campaigns?
- What are real-world success stories?

The problem? Most content only answers 2-3 of these sub-queries, leaving massive gaps that competitors can fill.

Related, great posts to read:

Emina Demiri: <https://www.womenintechseo.com/knowledge/deep-dive-into-ai-overviews/>^[4]

Mike King: <https://ipullrank.com/how-ai-mode-works>^[5]

Marie Haynes: <https://www.mariehaynes.com/ai-mode-is-the-future-of-search-how-will-you-adapt/>^[6]

Enter the Query Fan-Out Detector

I've created a Screaming Frog custom extraction script that leverages Gemini AI to analyze your pages the way Google's AI Mode does. You can easily play with the prompt. Here's what makes it powerful:

What It Does

1. **Extracts Semantic Chunks:** Uses layout-aware chunking to segment your content (just like Google's Document AI)
2. **Identifies Primary Entity:** Determines the main ontological topic of your page
3. **Predicts Fan-Out Queries:** Generates 8-10 sub-queries Google would likely create
4. **Scores Coverage:** Evaluates whether your content answers each query (Yes/Partial/No)
5. **Suggests Follow-Ups:** Predicts what users would ask next

The Script

TEXT

```
// Gemini AI Query Fan-Out Detection for Screaming Frog
// Replace with your actual Gemini API key
const apiKey = 'YOUR_GEMINI_API_KEY';

// Extract semantic chunks from page (layout-aware chunking)
function extractSemanticChunks() {
  const chunks = [];

  // Extract title and main heading
  const title = document.title || '';
  const h1 = document.querySelector('h1')?.textContent || '';
  if (title || h1) {
    chunks.push({
      type: 'primary_topic',
      content: `${title} ${h1}`.trim()
    });
  }

  // Extract headings and their content (layout-aware chunking)
  const headings = document.querySelectorAll('h2, h3');
  headings.forEach(heading => {
    let content = heading.textContent;
    let sibling = heading.nextElementSibling;
    let sectionContent = '';

    while (sibling && !['H1', 'H2', 'H3'].includes(sibling.tagName)) {
      if (sibling.textContent) {
        sectionContent += ' ' + sibling.textContent;
      }
      sibling = sibling.nextElementSibling;
    }

    if (sectionContent.trim()) {
      chunks.push({
        type: 'section',
        heading: content,
        content: sectionContent.trim().substring(0, 500)
      });
    }
  });

  // [Rest of the script continues...]
}
```

Real-World Results That Should Wake You Up

I tested this on various types of content, and the results were eye-opening:

Case Study 1: Technical SEO Guide

- Semantic Chunks Found: 24
- Query Coverage: 7/10
- Missing Queries:
- Implementation costs

- Tool comparisons
- API alternatives
- Key Insight: Even well-structured content missed critical buying-intent queries

Case Study 2: E-commerce Product Page

- Semantic Chunks Found: 12
- Query Coverage: 3/10
- Missing Queries:
 - Sustainability information
 - Comparison with alternatives
 - Use case scenarios
 - Maintenance guides
- Key Insight: Product pages focus on features but miss the “why” and “how” queries

Case Study 3: Local Service Page

- Semantic Chunks Found: 8
- Query Coverage: 4/10
- Missing Queries:
 - Pricing structures
 - Service area specifics
 - Booking process
 - Qualification requirements
- Key Insight: Local pages assume too much prior knowledge

The Pattern That Changes Everything

After analyzing 2000+ pages, I discovered a clear pattern:

- Pages with Custom > Custom Javascript
- Paste the custom JS code
- Run the test script

Step 2: Analyze Your Key Pages Run the crawler on your most important pages first. Look for:

- Coverage scores below 5/10
- Patterns in missing queries
- Common gaps across similar pages

Step 3: Fill the Gaps Strategically For each missing query type:

- Related Queries → Add context sections
- Implicit Queries → Address unstated needs in FAQ format
- Comparative Queries → Create comparison tables
- Procedural Queries → Add step-by-step guides
- Contextual Refinements → Include specific scenarios

Example Output Analysis Here's what the script found on a recent AI Mode article:

TEXT

=== GOOGLE AI MODE QUERY FAN-OUT ANALYSIS ===

PRIMARY ENTITY: Google Document AI Layout Parser and its implications for SEO

FAN-OUT QUERIES:

- How does Google's AI Mode use webpage structure information? - Coverage: Partial
- What are the best practices for HTML structure to optimize for Google's AI Mode? - Coverage: Yes
- How does Google Document AI Layout Parser differ from Google Search's actual parser? - Coverage:
- Comparison with other website analysis tools - Coverage: Partial
- How can I use Google Document AI Layout Parser to improve my website's SEO? - Coverage: Yes
- What are the Layout Parser limitations? - Coverage: Yes
- How does layout affect rankings? - Coverage: Partial
- Can I use Layout Parser for free? - Coverage: No
- What content types does it identify? - Coverage: Yes
- How to access the Layout Parser API? - Coverage: Yes

COVERAGE SCORE: 7/10 queries covered

RECOMMENDATIONS:

- Expand on how Google's AI Mode utilizes webpage structure information
- Add cost and accessibility information
- Include comparison table with other tools
- Provide ranking impact case studies

The New Optimization Framework

For AI Mode, an SEO should also optimize for query networks. Here's your new framework:

1. Entity-First Approach Start with your primary entity and map all possible query branches
2. Coverage Over Keywords Measure success by query coverage percentage, not keyword density
3. Semantic Chunking Structure content for AI parsing
4. Follow-Up Anticipation Every section should answer current queries AND seed the next question

What This Means for Your Strategy

If your pages average 30% query coverage, you're leaving 70% of potential AI visibility on the table. While competitors scramble to understand AI search, you can systematically improve coverage using data.

Start with your money pages. Run the analysis. Fill the gaps. Watch your AI Overview visibility grow.

Have you run the analysis on your site? Share your coverage scores on Twitter/X@ [metehan777](#)^[7]. I'm collecting data on industry benchmarks and would love to see what patterns emerge. Connect me on [LinkedIn](#)^[8]

If you're applying fan-out analysis to a real site, these help: [an AI visibility audit guide for e-commerce brands](#)^[9], [AI search optimization platforms for global marketing teams](#)^[10], and [how agencies can report AI visibility to clients](#)^[11].

Running query fan-out analysis in Screaming Frog turns AEO from theory into pipeline. The pages that survive fan-out are the ones that earn AI citations in ChatGPT, Perplexity and Gemini. Answer engine optimization at scale needs this workflow.

REFERENCES & SOURCES

- [1] How to 2x your AI traffic with the dumbest and simplest method? — <https://metehan.ai/blog/citemet-ai-share-buttons-growth-hack-for-llms/>
- [2] <https://github.com/metehan777/screaming-frog-query-fan-out> — <https://github.com/metehan777/screaming-frog-query-fan-out>
- [3] Automate Your Backlink Opportunities with Screaming Frog — <https://metehan.ai/blog/automate-your-backlink-opportunities-with-screaming-frog-and-openai/>
- [4] <https://www.womenintechseo.com/knowledge/deep-dive-into-ai-overviews/> — <https://www.womenintechseo.com/knowledge/deep-dive-into-ai-overviews/>
- [5] <https://ipullrank.com/how-ai-mode-works> — <https://ipullrank.com/how-ai-mode-works>
- [6] <https://www.mariehaynes.com/ai-mode-is-the-future-of-search-how-will-you-adapt/> — <https://www.mariehaynes.com/ai-mode-is-the-future-of-search-how-will-you-adapt/>
- [7] metehan777 — <https://x.com/metehan777>
- [8] LinkedIn — <https://www.linkedin.com/feed/update/urn:li:activity:7343991127873597440/>
- [9] an AI visibility audit guide for e-commerce brands — <https://metehan.ai/articles/ai-visibility-audit-ecommerce-brands-complete-guide/>
- [10] AI search optimization platforms for global marketing teams — <https://metehan.ai/articles/ai-search-optimization-platform-for-global-marketing-teams/>
- [11] how agencies can report AI visibility to clients — <https://metehan.ai/articles/how-agencies-can-report-on-ai-visibility-for-their-clients/>