

Improve Your SEO Strategy with Cosine Similarity and Topic Clustering Tool

Metehan Yesilyurt

AI Search & Information Retrieval Researcher · metehan.ai

Published: November 28, 2024 ·

Canonical: <https://metehan.ai/blog/improve-your-seo-strategy-with-cosine-similarity-and-topic-clustering-tool/>

SEO professionals and content marketers know that identifying internal link opportunities and creating optimized topic clusters are essential for driving organic traffic. However, performing these tasks manually can be tedious, time-consuming, and prone to errors. To address these challenges, I used AI and developed a new **Cosine Similarity & Content Analysis Tool**.

Access here: <https://metehan.ai/cosine-similarity/>^[1]

This tool is specifically designed to extract content from up to **100 URLs simultaneously**, group the extracted data into **topic clusters**, and calculate **cosine similarity scores** to pinpoint the best opportunities for internal linking.

Tech stack; -Backend: Flask, Python -Backend platform: Replit + Streamlit -UI HTML -Topic Cluster: OpenAI - Embeddings: OpenAI

Key Features of the SEO Content Analysis Tool

1. Bulk Content Extraction from Up to 100 URLs

Gone are the days of manually copying and pasting website content for analysis. Our tool allows you to:

- Input a list of up to 100 URLs.
- Automatically extract content, metadata, and semantic details from each page.
- Generate a consolidated report for easy review.

This bulk extraction feature saves time and ensures you have all the necessary data in one place.

2. Automatic Topic Clustering for Content Optimization

Topic clusters are crucial for building a well-structured internal linking strategy. The tool uses **natural language processing (NLP)** to:

- Analyze the content from all URLs.
- Group similar topics together into clusters based on relevance and semantic relationships.
- Highlight content gaps and overlapping areas within your site.

With this feature, you can visualize how your existing content fits into broader topic clusters, improving your site's topical authority.

3. Cosine Similarity Scores for Internal Linking Opportunities

Finding the right pages to link to is critical for improving user navigation and distributing link equity. Our tool calculates **cosine similarity scores**, a mathematical measure that:

- Compares the semantic similarity between pieces of content.
- Identifies pages with related topics that can be linked for enhanced relevance.
- Suggests link anchors to use based on keyword context.

This data-driven approach ensures that your internal links are strategic, valuable, and impactful for SEO.

Why Cosine Similarity Matters in SEO

Cosine similarity is a robust metric in content analysis because it quantifies how closely two texts align in meaning. For SEO, this means:

- Avoiding irrelevant or forced links that could harm user experience.
- Creating a logical internal linking structure that search engines can easily crawl and understand.
- Enhancing contextual relevance, which is increasingly important for ranking algorithms like Google's BERT and MUM.

By incorporating cosine similarity, the tool not only boosts your SEO strategy but also ensures a better user experience for your audience.

Benefits of Using the Cosine Similarity Content Analysis Tool

You can develop a better version. Here's how this tool can transform your SEO efforts:

- **Efficiency:** Save hours of manual work by automating content extraction and clustering.
- **Data-Driven Insights:** Base your internal linking and content decisions on measurable similarity scores.
- **Improved Topical Authority:** Build well-structured topic clusters that signal expertise and relevance to search engines.
- **Enhanced User Experience:** Provide users with a logical site structure and valuable internal links that guide them through your content.

Figure

How to Get Started

Getting started with our SEO Content Analysis Tool is simple:

1. **Upload Your URL List:** Copy and paste up to 100 URLs into the tool.
2. **Analyze and Cluster:** Let the tool process your content and generate topic clusters.
3. **Review Cosine Similarity Results:** Use the similarity scores to identify internal linking opportunities.
4. **Optimize Your Site:** Implement the tool's recommendations to enhance your SEO performance.

Cosine similarity and topic clustering are the math behind modern AEO auditing. AI citations in ChatGPT, Perplexity and Gemini concentrate on pages that cluster cleanly with the query. Answer engine optimization inherits that geometry.

REFERENCES & SOURCES

[1] <https://metehan.ai/cosine-similarity/> — <https://metehan.ai/cosine-similarity/>