



AI Visibility Is Math

A Data Science Approach to Making Brands Visible in LLMs

March 2026

Workshop plan

- About us
- State of the market
- How models work
- Tracking & first steps
- Case study
- What can you do
- Q&A

Who we are



AI visibility lab for data-driven brands

metaturbine.com



Daria, Co-founder, Marketing Lead

8 years in marketing, including Publicis Groupe (Coca-Cola). Led teams for fintech scale-ups, business tools, and apps. Founded a content marketing agency and an HR-tech startup.



Cyril, Co-founder, Data Science Lead

10 years in applied data science and machine learning. Built scoring models for fintech startups (Series A led by SBI) and predictive models (Burger King, McCann London, Century21).



Denis F., Fullstack Developer

4 years in AI-powered services, fintech, and end-to-end development across backend and frontend.



Denis A., Fullstack Developer

6 years in development across enterprise, fintech, edtech, and multimedia projects. Main stack: React.js and Node.js.

State of the market

Chatbots vs. Google

ChatGPT

18B messages

July 2025, [Open AI](#)

Google

96B searches per week

Estimation from 5T+ searches per year

December 2025, [Google](#)

ChatGPT: 700M users/week (July 2025, [Open AI](#))

Gemini app: 750M monthly active users (Feb 2026, [Google](#))

Google AI Overviews: 2B users/mo (July 2025, [Google](#))

Copilot: 100M MAU (October 2025, [Microsoft](#))

Grok: 600 MAU across the X and Grok apps (January 2026, [xAI](#))

Claude: not disclosed

Perplexity: not disclosed

→ People phrase queries differently in chatbots.

→ Search systems work differently.

→ The results are different too.

The image shows a Google search results page for the query "email marketing tool for startups". The search bar at the top contains the query and the Google logo. Below the search bar, there are navigation tabs for "AI Mode", "All", "Images", "Videos", "News", "Forums", "Short videos", "More", and "Tools". The main content area is divided into several sections:

- Sponsored results:** This section contains four sponsored ads. The first is from Pipedrive, titled "Easy Email Marketing", with a sub-headline "Try Pipedrive™ For Free — Get Full Access for 14 Days Free. No Credit Card Required. Try Any Feature During Trial. Over...". The second is from MailerLite, titled "11 Best Email Marketing Platforms in 2026 - MailerLite", with a sub-headline "Grow Your Business with MailerLite! Free marketing resources, tools to help you excel in email marketing. Unlimited Emails. No Contract. Try Premium for Free. Sell Digital Products.". The third is from SendPulse, titled "Send 15,000 Free Emails/Month - Unlimited Emails from \$10/mo", with a sub-headline "#1 Email Marketing Platform When Price and Features Matter. Unlimited Emails from \$10/mo. Smarter Marketing Starts Here: Explore SendPulse's Email Features.". The fourth is from Attio, titled "Attio | CRM of the Next Era", with a sub-headline "Attio | Next Generation CRM — Attio is the CRM of the future - powerful, flexible, and built for scale."
- AI Overview:** This section provides a summary of the best email marketing tools for startups. It mentions that the best tools offer free tiers, automation, and user-friendly interfaces. It lists top choices: MailerLite for ease of use, Brevo for all-in-one automation (SMS/email), Loops for SaaS-focused marketing, and HubSpot for robust CRM integration. It also includes a "Top Email Marketing Tools for Startups" section with a bullet point for MailerLite, noting it is highly regarded for a clean interface, great automation features, and a generous free tier (typically up to 1,000 subscribers and 12,000 emails per month).
- Search results:** This section contains several organic search results. The first is from Mailtrap, titled "11 Best Email Marketing Platforms for Startups [2026]", with a sub-headline "I tested and compared 7 best email marketing platforms for startups on deliverability, pricing, scaling, email design, and more. Click here.". The second is from Get AI Perks, titled "Best Email Marketing Software for Startups: Tools That ...", with a sub-headline "13 Feb 2026 — A practical guide to the best email marketing software for startups, covering tools that help teams automate campaigns, grow lists, ...". The third is from EmailTooltester.com, titled "The 17 Best Free Email Marketing Services for 2026", with a sub-headline "9 Mar 2026 — Building up your mailing list, but not quite ready to make the jump to a paid service? We look at the 17 best free newsletter services."
- Videos:** This section contains a video result from YouTube, titled "My Top 5 Email Marketing Software for Small Business in 2026", with a sub-headline "YouTube - Stewart Gauld".

**What we know about how chatbots work
(and what to do with that)**

How models work

They rely on vector embeddings

Embeddings are numerical representations of concepts. They convert meaning into number sequences, which helps computers understand relationships between concepts.

- synonyms can behave almost like the same word.
- concepts matter more than exact keywords.
- semantic similarity is the key.

How models work

They use pre-trained knowledge + the web

The model may rewrite the prompt internally.

```
gpt_chat_output_responses("Best digital wallet for kids", "General")
```

```
'queries': ['best debit cards and digital wallet apps for kids 2026',  
'Greenlight debit card for kids features and fees 2026',  
'GoHenry debit card for kids US features and pricing 2026',  
'Apple Cash family children features'],  
'query': 'best debit cards and digital wallet apps for kids 2026',
```

It uses pre-trained knowledge first.

It may turn to web search when:


- the initial answer does not meet its reliability threshold,
- the prompt requires fresher information,
- personal or product settings enable search.

→ We do not know the exact weighting of each factor.

→ Semantic content similarity is likely one of the main drivers.

Sources

 surferstack.com
Best Email Marketing Platforms in 2026: Complete Comparison of Mailchimp, Brevo, ActiveCampaign & More – Surferstack
13 Feb 2026

 LinkedIn
11 Best Email Marketing Software for Startups in 2026
17 Jan 2026

 vibegrowthstack.io
Best Email Marketing Tools for Startups 2026: Complete Comparison
20 Feb 2026

 Startup Email Marketing
15+ Best Email Marketing Tools for Startups in 2026 | Startup Email Marketing

 Reddit
Best email marketing tools in 2026? My honest breakdown
13 Dec 2025

How models work

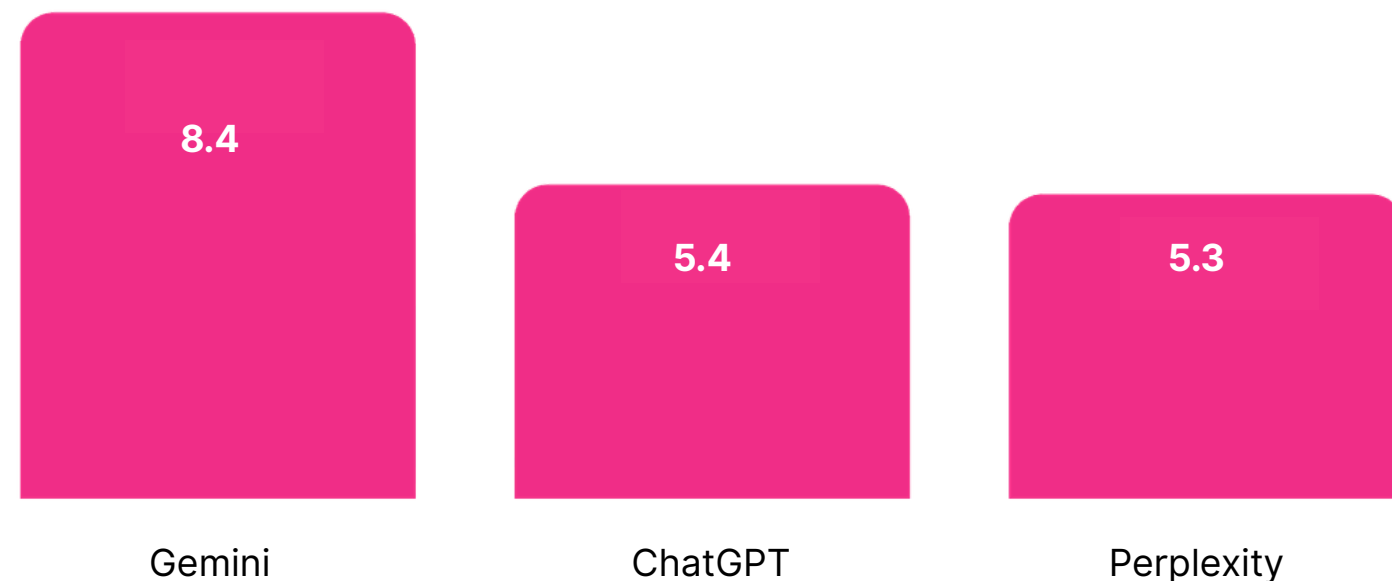
Answers differ across models

Because models are trained differently, use different weights, and make different ranking decisions.

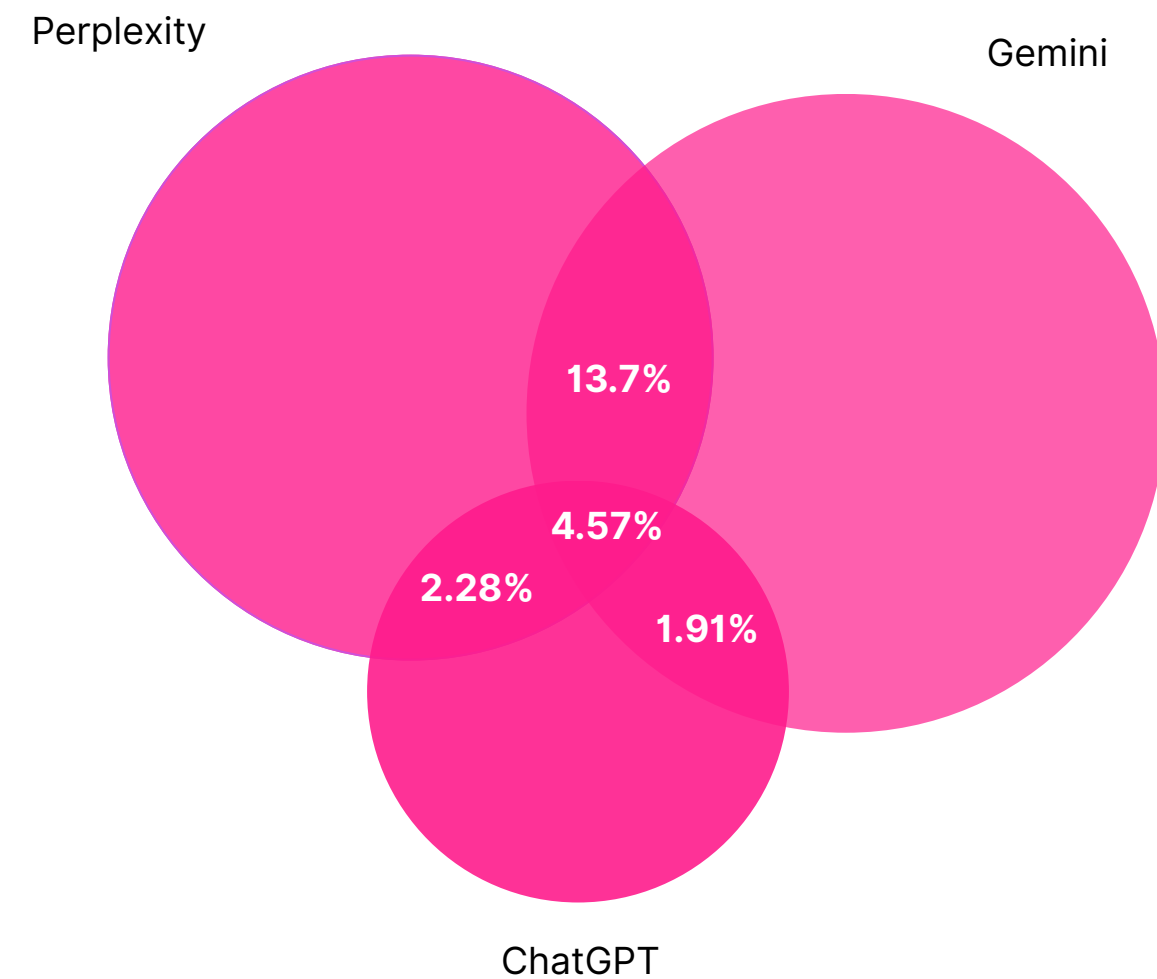
→ **Each chatbot requires its own optimization approach**



How many brands do chatbots recommend on average?



How similar are chatbot sources?

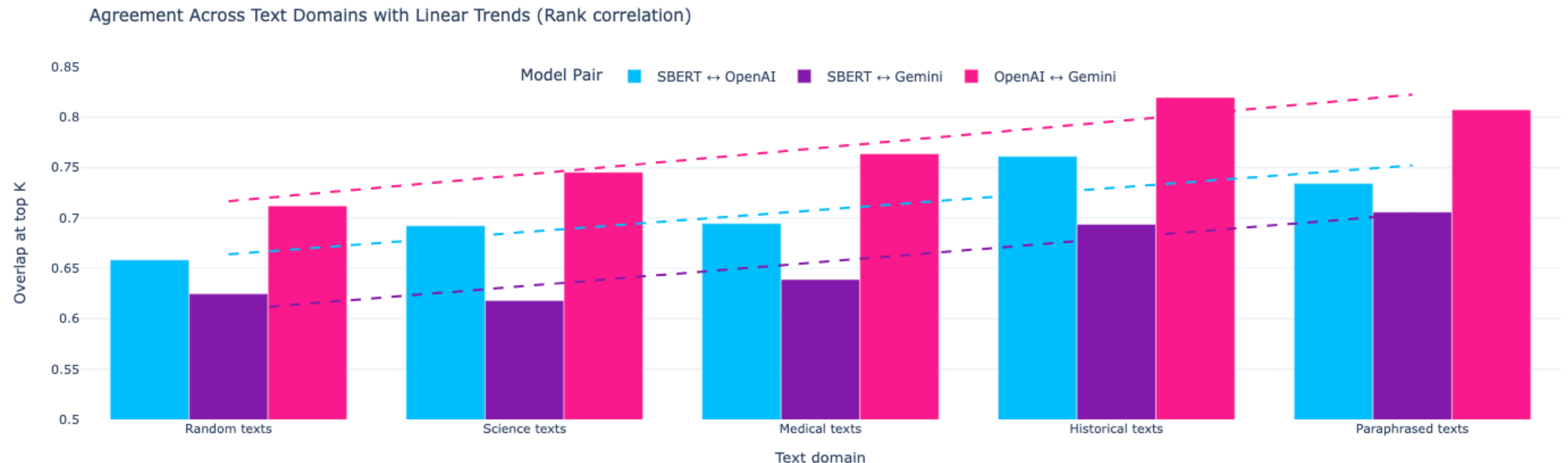


How models work

Experiment: OpenAI vs. Gemini vs. SBERT

- 1,000 texts from different domains
- OpenAI, Gemini, and SBERT embeddings
- We measured how similarly they ranked texts.

Result: OpenAI and Gemini are similar, but not identical with semantic agreement ranging from ~71% to ~82%.



How models work

What we do not know

- No one outside LLM developers knows exactly how users phrase their prompts.
- Responses are also shaped by user context.
- Hallucinations still occur.

What we do know

- We can see chatbot traffic in Google Analytics.
- We also hear from self-reported users who found us through LLMs.
- And there is likely additional LLM-driven traffic that we cannot fully attribute.

What we can do about it

- Track performance regularly across a broad enough set of prompts and topics.
- Use mathematical methods to measure and predict as much as possible.

Tracking and first steps

Tracking & first steps

Define the tracking set

Build a test set

Check whether the brand appears at all.

If visibility is zero

The first thing to check is whether chatbot crawlers can technically access the site. Once that is verified, review the results again.

Update the prompt set

Not every prompt is worth tracking. Choose the ones that:

- matter for the business,
- you want to be visible in,
- have relatively low competition.

Other inputs: Google, Bing, SEO.

→ Build a solid tracking set of 50+ prompts

Tracking & first steps

Check logs for crawlers

Platform	Crawler Name (User-Agent)	Purpose
OpenAI (ChatGPT)	GPTBot, OAI-SearchBot, ChatGPT-User	Model training, indexing, real-time content fetching
Google Gemini	Googlebot (standard crawling), Google-Extended (AI content use)	Search indexing and Gemini AI features
Perplexity AI	PerplexityBot, Perplexity-User	AI search indexing and on-demand citation fetching
Microsoft Copilot	Bingbot	Search + AI integration
Anthropic Claude	anthropic-ai, ClaudeBot, claude-web	AI model training and live answers

Traffic controls

On | Allow Search Engine Crawlers

On | Allow Artificial Intelligence Bots

Tracking & first steps

Tracking set

Prompts [How it works](#)

Show less ^

<input checked="" type="checkbox"/> All topics and prompts	Visibility	Best rank	Sources share	Competition index	Response
<input checked="" type="checkbox"/> Workforce visibility	Yes	Ranked	1/28	●●●●	
<input checked="" type="checkbox"/> improving real time visibility into global payroll costs	Yes	#1	1/13	●●●●	📄
<input checked="" type="checkbox"/> how to centralize reporting for international workforce data	None	#0	0/15	●●●●	📄
<input checked="" type="checkbox"/> Cross-border payments	Yes	Ranked	5/30	●●●●	
<input checked="" type="checkbox"/> safest methods to send cross border salary payments	Yes	#2	5/16	●●●●	📄
<input checked="" type="checkbox"/> how to optimize international payroll payments to employees	None	#0	0/14	●●●●	📄
<input checked="" type="checkbox"/> Compliance management	Yes	Ranked	1/17	●●●●	
<input checked="" type="checkbox"/> tools to manage multi country payroll tax compliance	Yes	#3	1/9	●●●●	📄
<input checked="" type="checkbox"/> how to stay compliant with international payroll regulations	None	#0	0/8	●●●●	📄
<input checked="" type="checkbox"/> Employer of record	None	Not ranked	0/23	●●●●	
<input checked="" type="checkbox"/> when to use an employer of record for global hiring	None	#0	0/12	●●●●	📄
<input checked="" type="checkbox"/> employer of record vs opening a local entity comparison	None	#0	0/11	●●●●	📄
<input checked="" type="checkbox"/> Global payroll automation	Yes	Ranked	5/21	●●●●	
<input checked="" type="checkbox"/> best way to automate multi country payroll processes	Yes	#3	3/10	●●●●	📄
<input checked="" type="checkbox"/> how to reduce manual work in global payroll operations	Yes	#2	2/11	●●●●	📄

Test report for Papaya Global

Tracking & first steps

Analyze each topic and prompt

Identify which competitors already appear in chatbot answers.

These are often not only the competitors the product team sees as direct competitors.

Then ask: what are they doing to show up there?

Look at the sources behind those answers:

media, Wikipedia, Reddit, LinkedIn, G2, review platforms, and company websites.

Also look at the formats:

articles, category pages, comparison pages, or other types of content.

Tracking & first steps

Build a plan

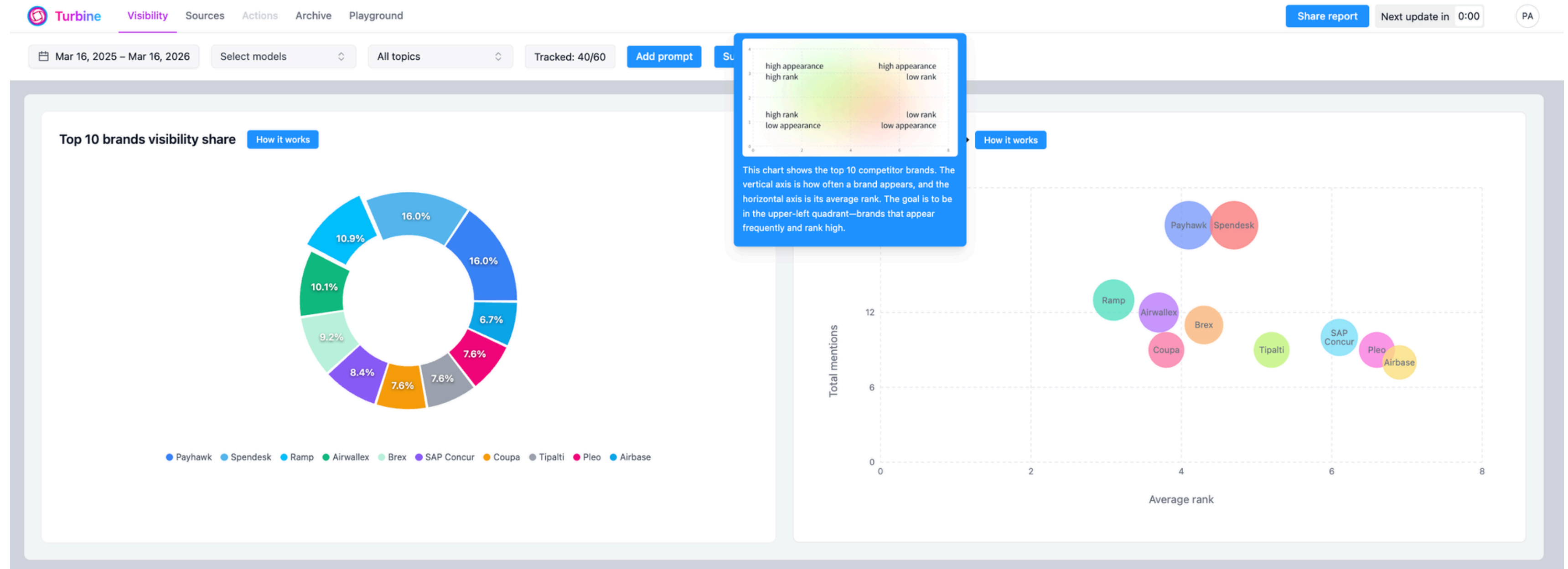
Based on what is realistically possible for your brand.

In many cases, this means:

- strengthening media presence,
- improving visibility on review platforms and social media,
- updating existing content or adding new content to the site.

Tracking & first steps

Top competitors



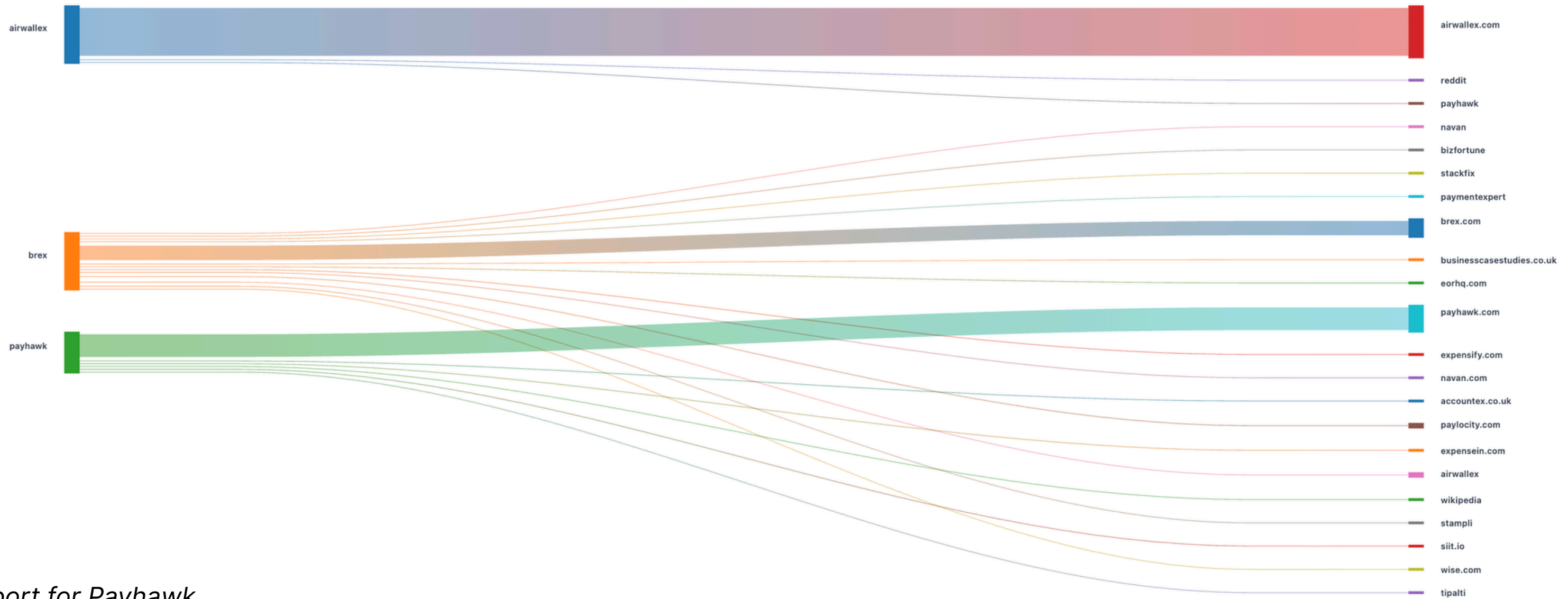
Test report for Payhawk

Tracking & first steps

Top sources

Source connections

Flow view Network view How it works



Test report for Payhawk

How to create content for LLMs

How to create content

Simplified workflow

- Start with a prompt.
- Identify the key terms and concepts as the LLM interprets them.
- Create an agent and provide context: company information, competitors, sources, and target keywords.
- Ask it to generate a draft.
- Validate it with a content validator.
- Review the text.
- Publish.
- Monitor the result.

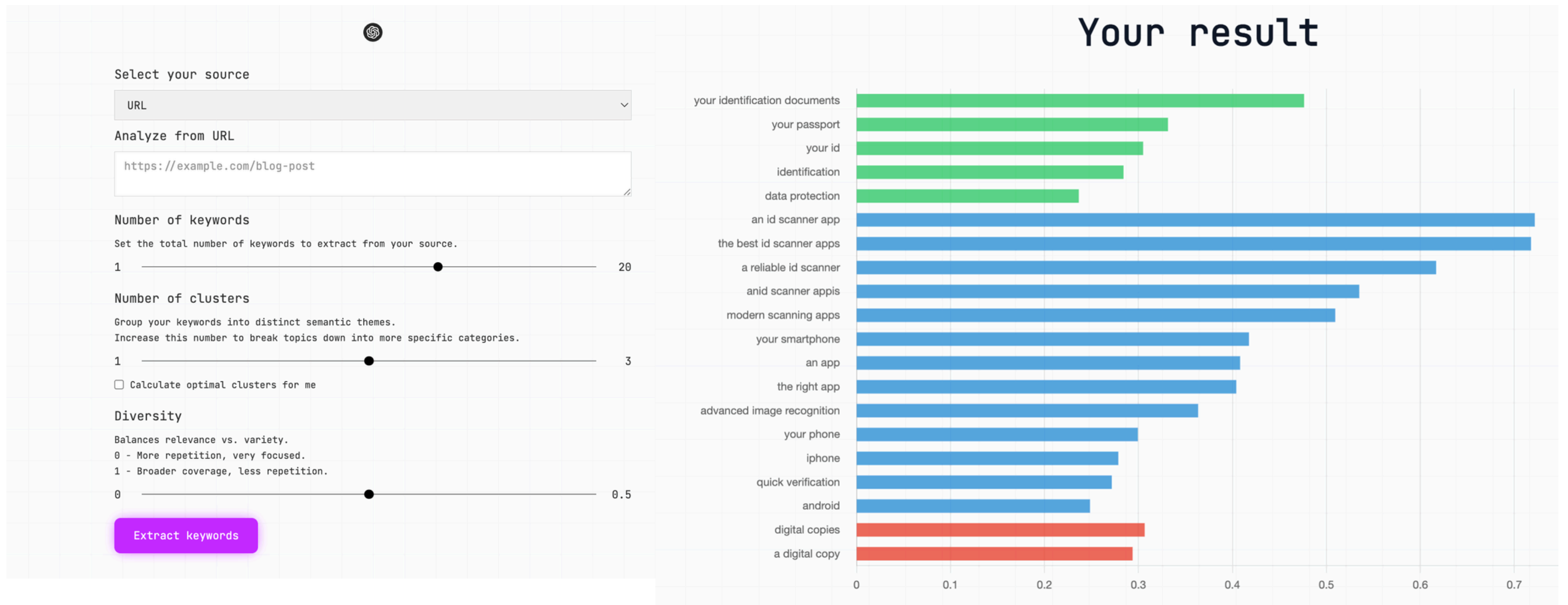
***SEO fundamentals still work:**

good page structure, overall site trust, a table of contents, FAQ sections, etc.

How to create content

Keyword extractor

metaturbine.com/tools/semantic-keyword-extractor



How to create content

Content validator

metaturbine.com/tools/validator

Measure the semantic proximity between a prompt and your text.
High scores improve your chances of being cited by chatbots.

0 = unrelated, 1 = identical
≥ 0.6 safe to publish, ≥ 0.8 high alignment (rare)



Select model

ChatGPT/OpenAI

Insert your prompt

Target prompt

Select your source

URL

Analyze from URL

<https://example.com/blog-post>

Calculate

Your result

Semantic proximity score

0.78

Meaning

very similar / strong paraphrase

Explanation

This is the sweet spot for visibility. Your content covers the exact answer the AI expects while maintaining your unique brand voice. You have high semantic relevance and a great chance of ranking in AI snapshots.

Model

Gemini/Google

Semantic proximity score

0.682

Meaning

very similar / strong paraphrase

Explanation

This is the sweet spot for visibility. Your content covers the exact answer the AI expects while maintaining your unique brand voice. You have high semantic relevance and a great chance of ranking in AI snapshots.

Model

ChatGPT/OpenAI

How to create content for LLMs (like a pro)

Case study

Case study

Client: Folio Wallet

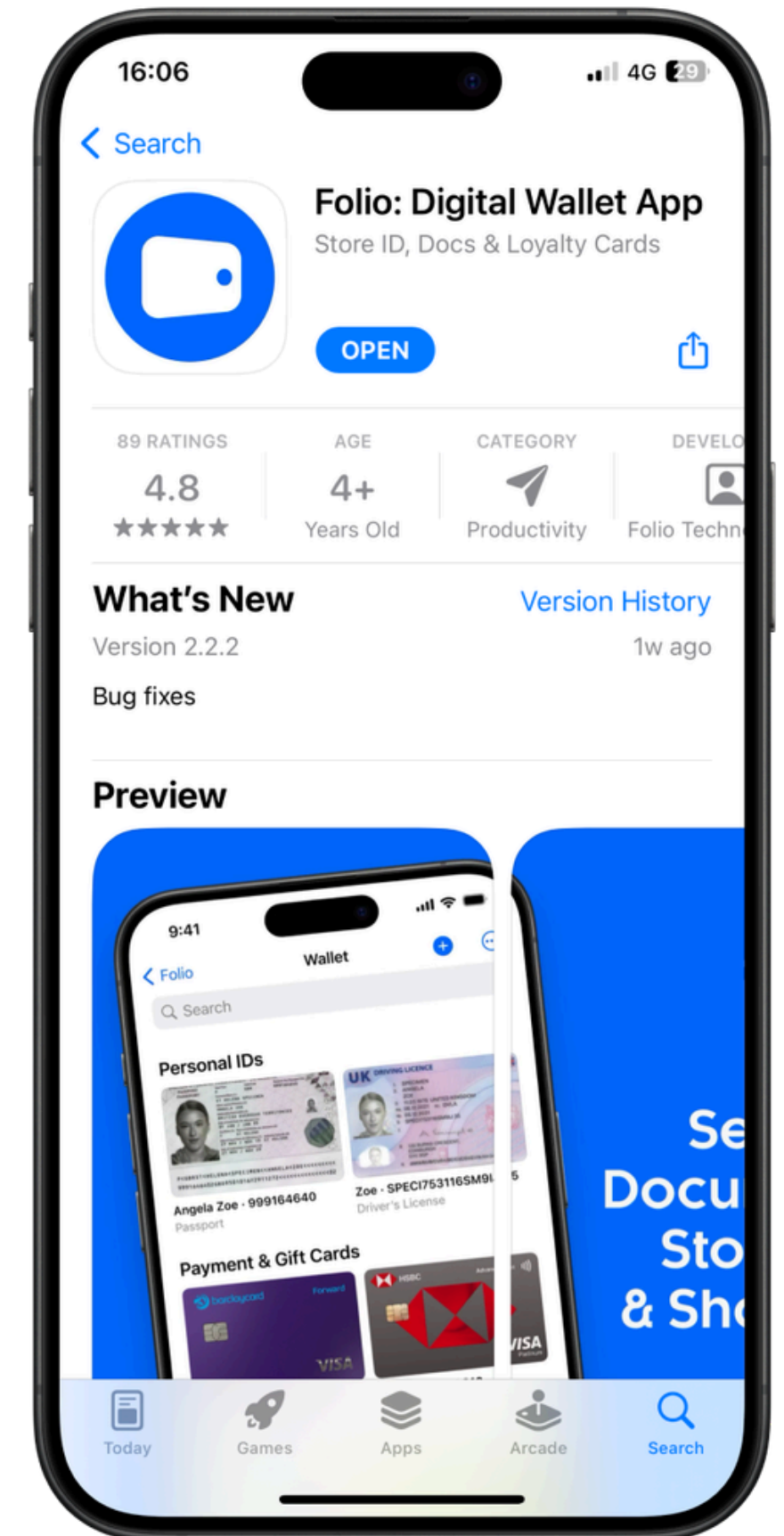
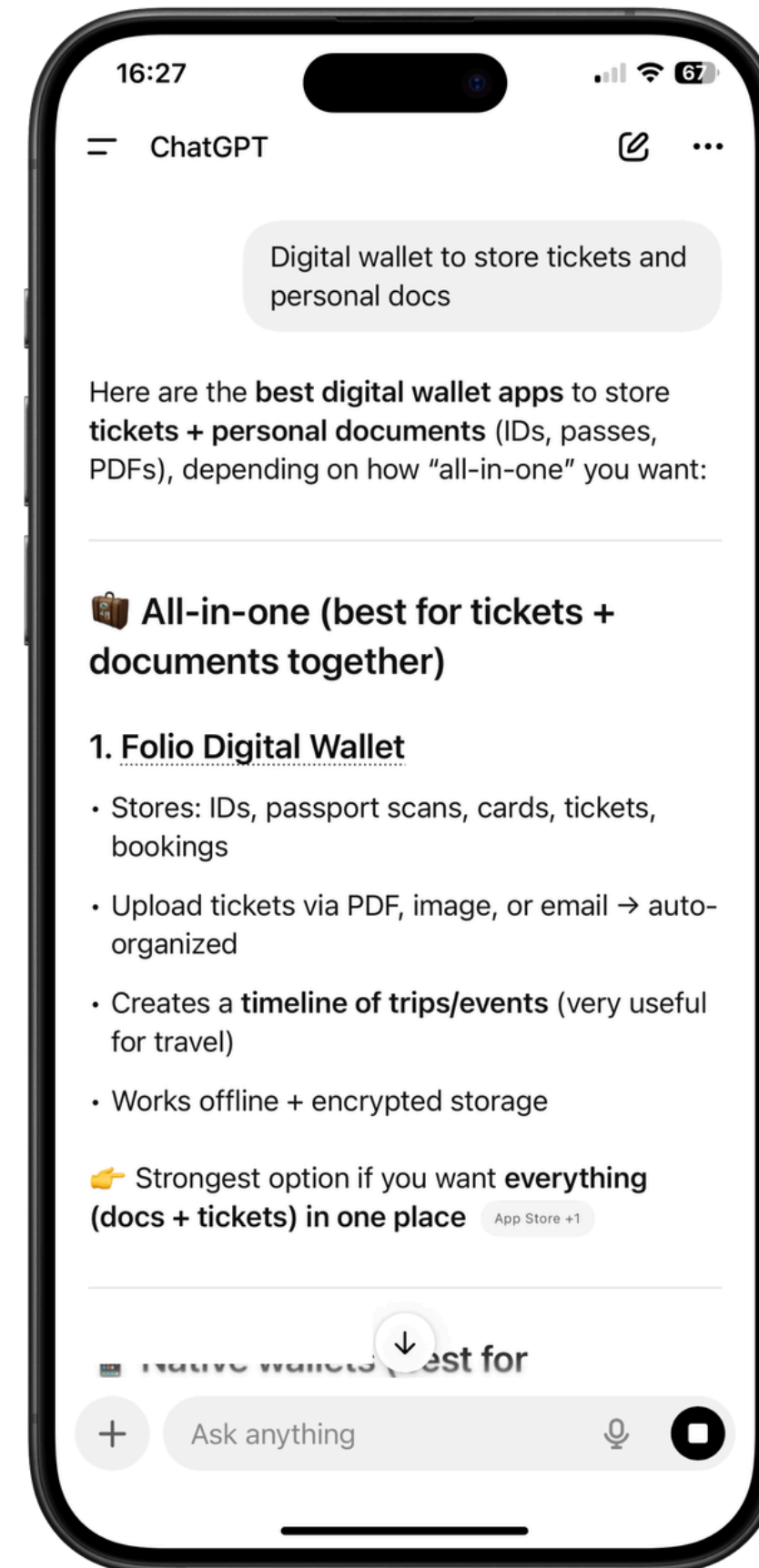
Service Provider: Turbine

Timeframe:

- 10 weeks of optimization
- 3 weeks of stabilization

Primary Model Measured: ChatGPT

Key Metric: AI search visibility share



Case study

Phase 1: Baseline & Testing

Aug 5 – Sep 10

We ran initial tests across a wide range of prompts. It confirmed two things:

- There was no accidental or residual visibility.
- Any future movement would be attributable to the work done.

Result: 0% visibility. Folio did not appear in model outputs.

Folio Wallet results					
Mentions	?	Visibility share	?	Average rank	?
0	–	0.0 %	–	0.0	–
Overall results					
Total / Unique brands	?	Average visibility share	?	Average brands per response	?
512 / 206	↑ 188	2.9 %	–	6.0	↑ 2.1

Case study

Phase 2: Setup & Prompt Strategy

Sep 11 – Sep 17

We started with two fundamentals:

- **Technical accesibility.** Ensured the site was properly accessible for LLMs.
- **Prompt set discovery.** We built and validated a prompt set that could be tracked consistently over time.

We deliberately looked for prompts at the intersection of:

- Folio's real product features
- Topics valuable for the brand narrative
- Areas where competitors appeared, but Folio didn't (yet)

→ That intersection became the core prompt set we tracked going forward.

Result: The initial setup and technical fixes resulted in the first 18% of visibility share.

Case study

Phase 3: Core Optimization

Sep 18 – Nov 22

This was the main execution phase.

We focused on three core thematic clusters: security, general document storage, and new niche Folio wanted to win: storing tickets and personal documents in one place

What we did:

- Updated existing content (metadata, page structure, FAQ blocks)
- Created new content to close clearly identified semantic gaps

Results: Week 1: 20% → Week 2: 37% → Week 3: 50% → Following month: stabilized at 43% visibility share

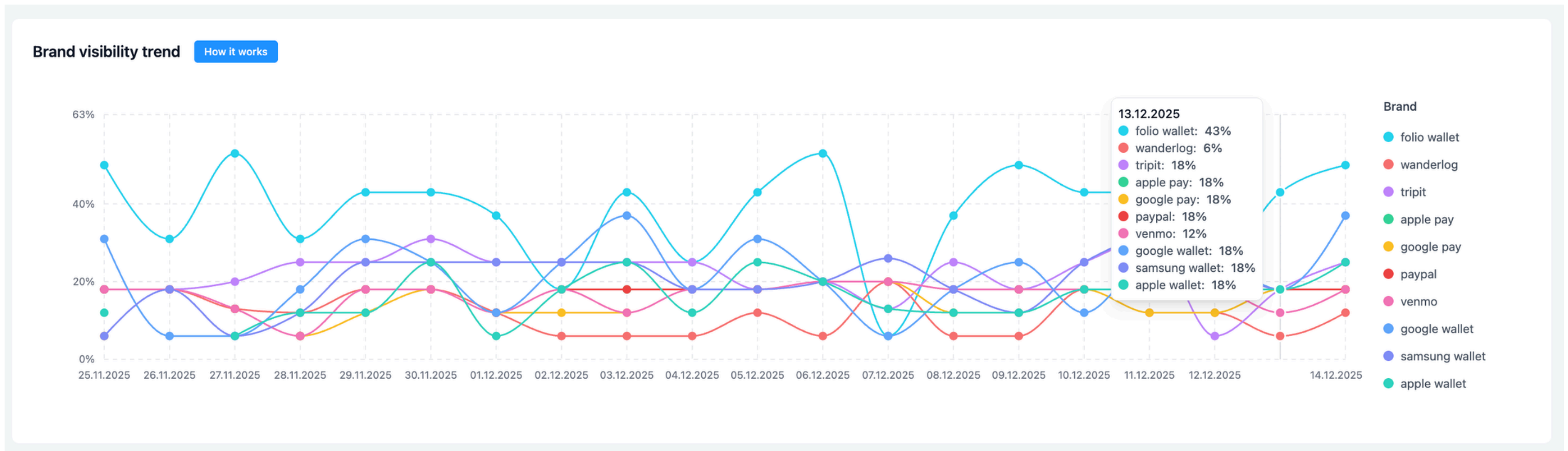
Case study

Phase 4: Stabilization check

Nov 25 – Dec 14

For three weeks, we made no changes at all.
We simply observed whether visibility would hold.

Result: it did. The gains were not short-term fluctuations.



What we did in practice

Initial assessment

At the start, Folio Wallet and the brands already appearing in responses were located in different semantic clusters.

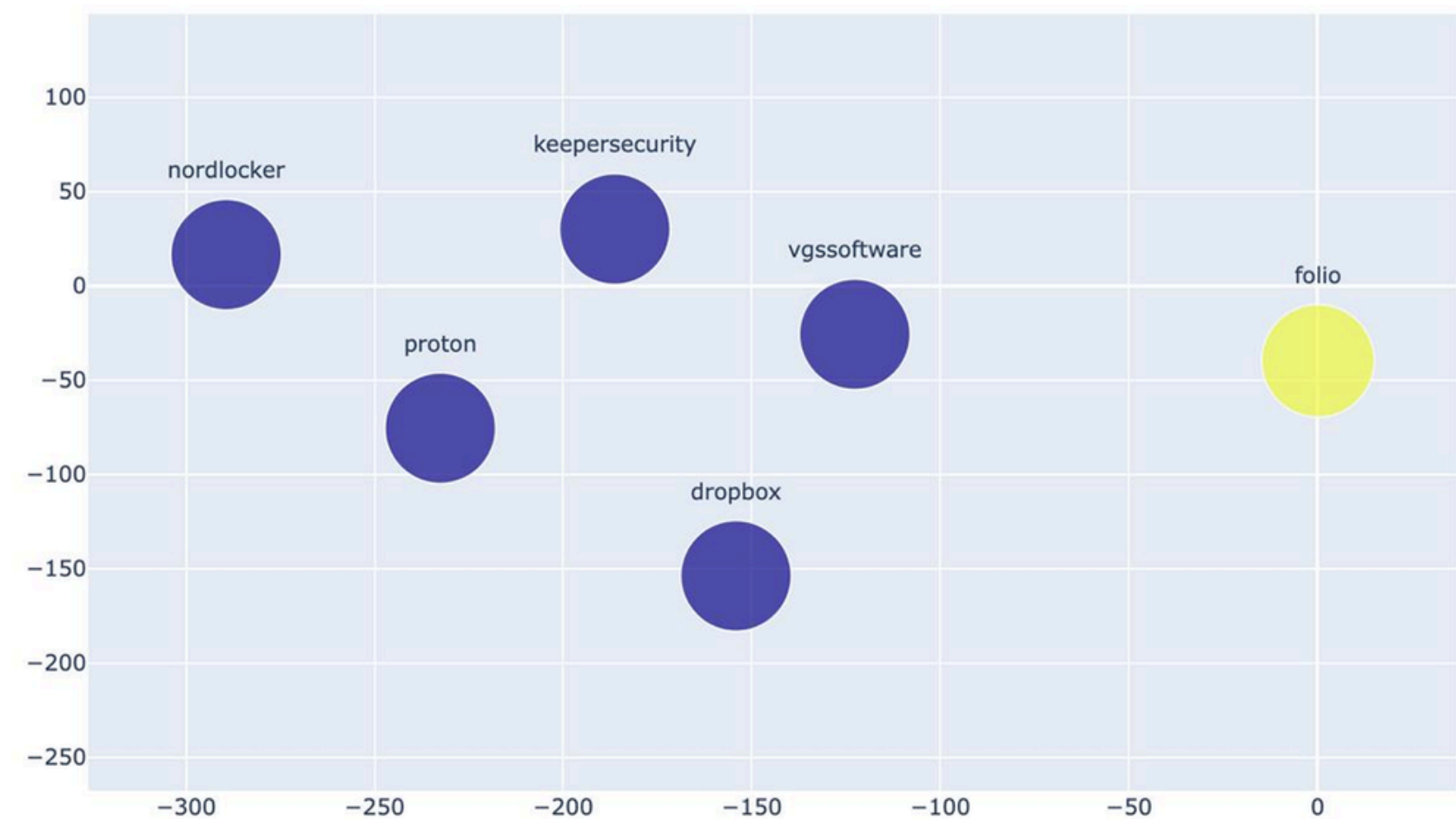
Existing brands occupied one cluster (blue), while Folio sat in another (yellow). In this state, Folio had a low probability of being shown as an answer.

Brand: Folio Wallet

Prompt: encrypted vault for personal documents

Current content

Two different clusters, brand is not visible



What we did in practice

Content creation

We use standard content guidelines (structure, relevance, length, etc.), but the most important input is semantics as seen by the model itself.

Even in [Turbine's free reports](#), you can see a preview of this approach: model answers are annotated with semantic entities, showing how the LLM interprets the text.

The same logic can be explored using Turbine's semantic [keyword extractor](#).

Response history

Model: ChatGPT

← Prev

Oct 16, 2025

Report 1 of 1

Next →

Response ?

Brand mentioned

Highlights indicate terms with the highest semantic importance to the model.

If you're looking for free **apps** to **securely store** and manage your **documents** and **cards**, here are some options to consider:

Doc Store r

Doc Store r allows you to create a **protected digital archive** of **personal files**, **including documents**, notes, **cards**, and receipts. You can capture photos of paper **files** or upload **digital** ones, organizing them into folders for convenience. The **app** **offers PIN security** to protect your information. (docstor.com)

Folio Wallet

Folio Wallet enables you to scan and **securely store** various **cards** and **documents**. It **offers** features like **PIN** and biometric authentication, ensuring your sensitive information remains protected. The **app** also provides alerts and notifications for any suspicious activity. (folio.id)

Show more ▾

What we did in practice

Content evaluation

We built a heatmap: the brighter the color, the closer the text is to those already being surfaced for that prompt (shown in the first five rows).

Brand: Folio Wallet

Prompt: encrypted vault for personal documents

Brands already visible for the prompt

Current content

Low scores (darker colours), not visible for the prompt

New content

Higher scores (brighter colours), high chances to get visible



What we did in practice

Content evaluation

The final step was empirical validation: publishing the content and monitoring AI visibility for the same prompt. We observed that it was in fact cited by the model, and Folio began appearing consistently in the results.

We then verified this again using semantic clustering. When this content was taken into account, the model placed Folio in the same cluster as the brands already visible for that prompt (highlighted in pink).

Brand: Folio Wallet

Prompt: encrypted vault for personal documents

New content

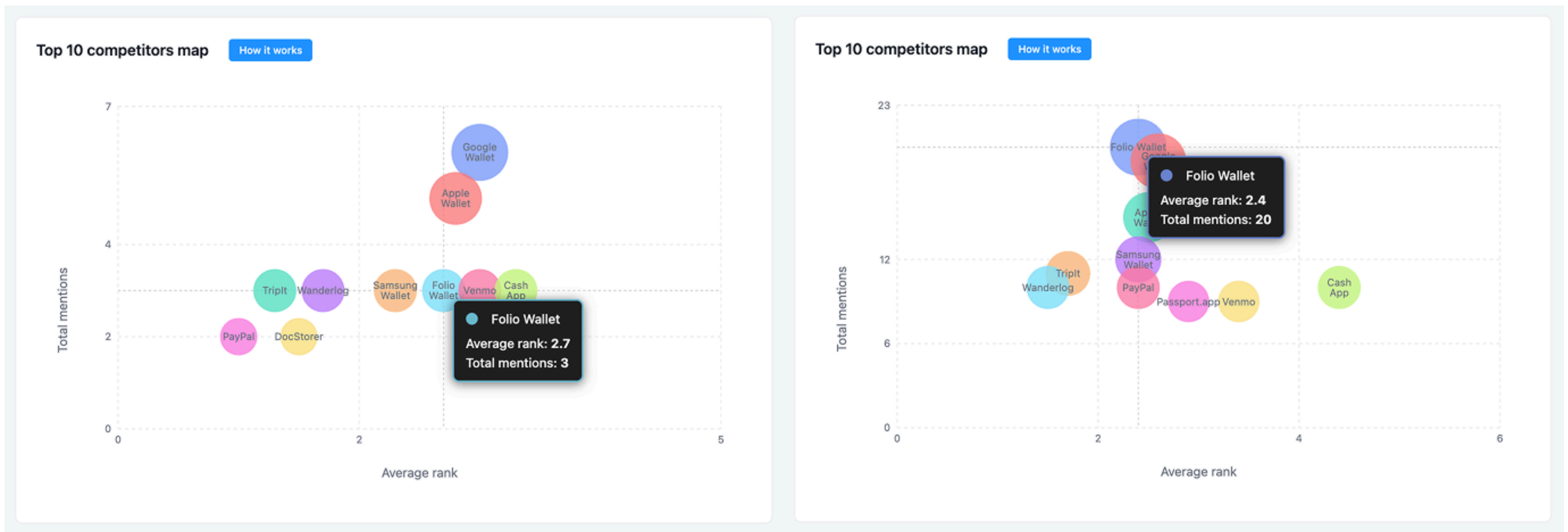
One cluster, high chances to get visible



Case study

Overall results

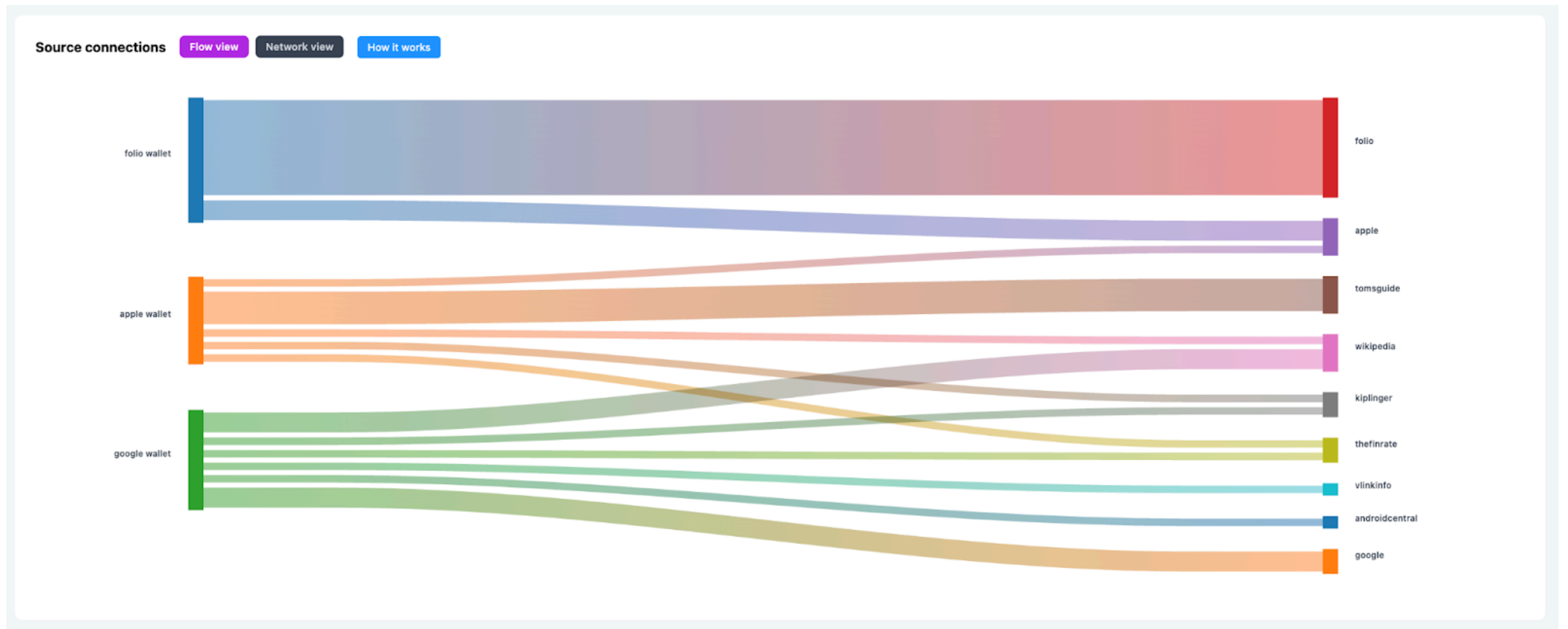
Folio moved from low visibility to a leading position, a clear narrative was established and reinforced.



Case study

Overall results

The Folio website became a primary citation source when models talk about the brand.

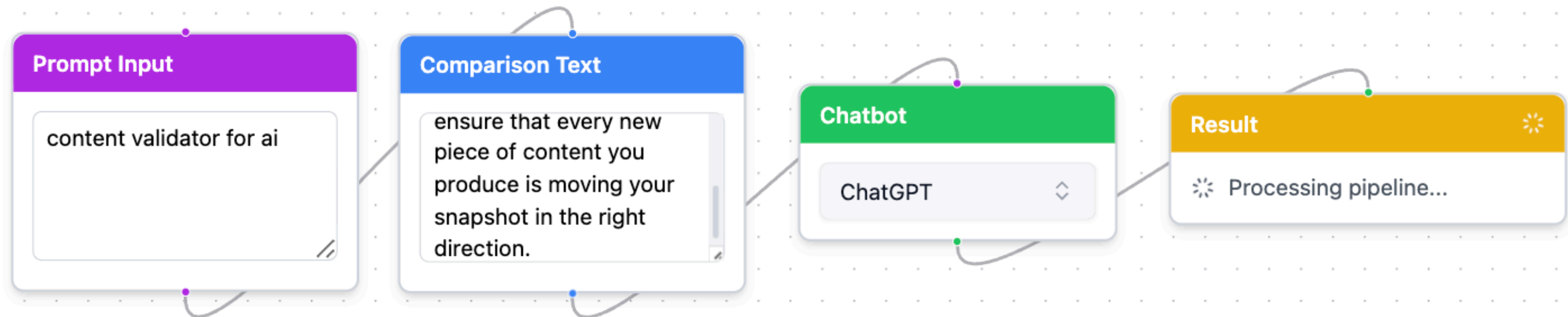


What can you do

Content development

- All of this can be built with APIs, so it is possible to create a workflow like this in-house.
- If you work with external vendors, check whether they actually use this kind of approach.
- We build content development pipelines to automate it. If you want to test it, come talk to us about a pilot.

+ vector databases



Key takeaways

- It's a new channel
- AI visibility is measurable (to some extent)
- Different models behave differently
- Track first
- Then create content
- Create content using maths

Let's stay in touch

 **Turbine AI Visibility Lab**



metaturbine.com



Daria, Co-founder, Marketing Lead

[linkedin.com/daria-ofitserova](https://www.linkedin.com/daria-ofitserova)



Cyril, Co-founder, Data Science Lead

[linkedin.com/cyril-kotrikov](https://www.linkedin.com/cyril-kotrikov)