Technical SEO for AI search

Optimizing for non-human visitors

Jan-Willem Bobbink / Friends of Search / 2026


We live in a retro game,
again.

For nearly three decades,
we optimized for one
gatekeeper.

Google.

Google

jan-willem bobbink

Jan-Willem Bobbink

Seo-specialist

11-internet.nl

11 Internet - Jan-Willem Bobbink: Freelance SEO consultant

Jan-Willem Bobbink: Freelance SEO consultant & 11 Internet: Wij verzorgen voor u het complete online visitekaartje!

Leeftijd

78 jaar

13 okt 1947

Opleiding

Radboud Universiteit

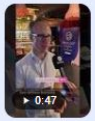
2007–2009

YouTube • Chantal…

Jan-Willem Bobbink over hoe LLM &…

Jan-Willem Bobbink over hoe LLM & GEO…

20 okt 2025



LinkedIn · Jan-Willem Bobbink

9,1K+ volgers

Jan-Willem Bobbink - Glippy

I'm specialised in: - (Enterprise) SEO Strategy and alignment - International… · Ervaring: Glippy · Opleiding: Hogeschool van Arnhem en Nijmegen · Locatie: …

11-internet.nl

https://www.11-internet.nl

11 Internet - Jan-Willem Bobbink: Freelance SEO consultant

11 Internet - Jan-Willem Bobbink. Freelance SEO Consultant Piet Heinstraat 79. 2518CD Den Haag, Nederland KVK: 09190645. Algemene leveringsvoorwaarden

jwbobbink.nl

http://www.jwbobbink.nl

Jan-Willem Bobbink - Portfolio - Webdesign - Skeeleren ...

Website Jan-Willem Bobbink: op deze website vind u een portfolio, informatie over scripting en webdesign, zoekmachine optimalisatie en een skeeler weblog.

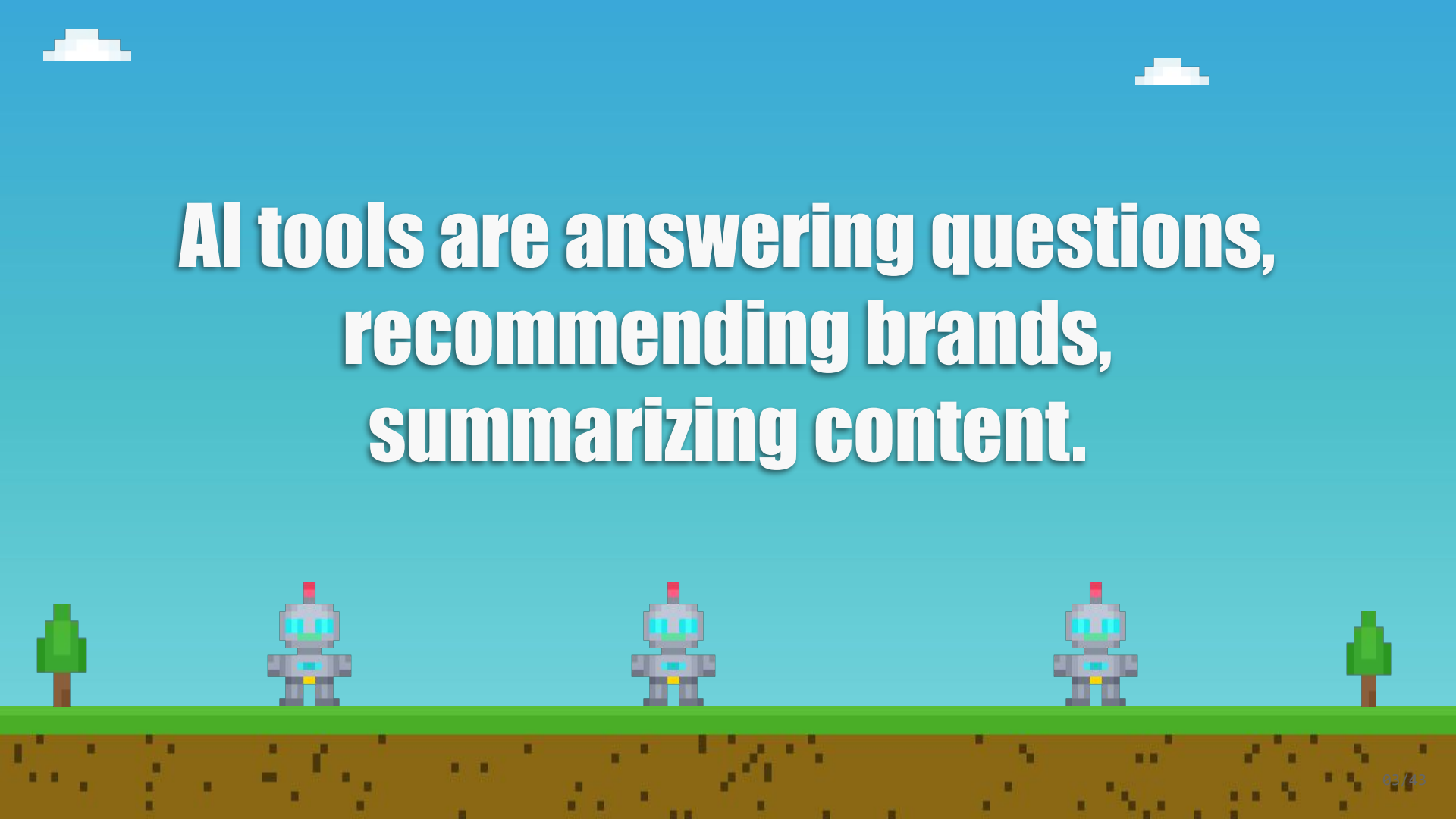
Over

Geboren	13 oktober 1947 (78 jaar), Warnsveld
Opleiding	Radboud Universiteit (2007–2009)

Feedback

Profielen

LinkedIn X (Twitter) Facebook


AI tools are answering questions,
recommending brands,
summarizing content.

EXAMPLE

The Zero-Click Answer

Traditional Search

```
User searches: "best CRM for startups"

Google returns:
  1. hubspot.com/crm-guide
  2. salesforce.com/small-biz
  3. pcmag.com/best-crm
  4. g2.com/categories/crm

User clicks 3 links, compares, decides
```

AI Answer

```
User asks: "best CRM for startups"

ChatGPT responds:
  "For early-stage startups, I'd
  recommend HubSpot's free tier
  for its ease of use..."

No click. No visit. Just an answer.
```

The question is no longer just

"Do we rank?"

It's

"Do AI systems understand us, trust us, and choose us?"

THE SHIFT

Two Different Readers

Search engine Crawler

- > Crawls and indexes pages
- > Matches keywords to queries
- > Ranks by links and signals
- > Returns 10 blue links (and some ads!)
- > User clicks, reads, decides

VS

LLM / AI Agent

- > Reads and understands content
- > Grasps meaning and context
- > Evaluates trust and accuracy
- > Synthesizes a direct answer
- > User gets the answer. No click.

EXAMPLE

What a search engine sees

Googlebot's View

```
<title>Best CRM Software 2026 | Brand.com</title>
<meta name="description" content="Compare top CRM...">
<h1>Best CRM Software for Startups</h1>
<h2>Top 10 CRM Tools Compared</h2>
<a href="...">Read our full review</a>

// Googlebot sees: tags, links, meta, keywords

// It INDEXES. It doesn't UNDERSTAND.
```

EXAMPLE

What an LLM sees

LLM Context Window

```
"This page compares CRM tools for startups.
The author recommends HubSpot for teams
under 10 people due to its free tier.

The content seems credible: named author,
published date, cited sources from Gartner.

However, the comparison table is rendered
via JavaScript and I cannot parse it."

// The LLM READS. It REASONS. It JUDGES.
```

Strip away the CSS.
Remove the images.
Kill the JavaScript.

That's your site to an AI agent.

HOW AI FINDS YOU

Three Pathways to Your Content

TRAINING DATA

Pre-baked knowledge

- > Learned during training
- > Static, potentially outdated
- > No attribution
- > You can't control it

RAG / RETRIEVAL

Real-time lookup

- > Pulls live web content
- > Cites sources
- > You CAN influence this

AGENT BROWSING

Live web access

- > AI navigates your site
- > Reads, clicks, evaluates
- > Emerging and growing
- > The future is here

PATHWAY 1

Training Data

Ask an LLM: "What does [Brand X] do?"

```
// Sometimes accurate:
"HubSpot is a CRM platform for marketing,
sales, and customer service..."  OK

// Sometimes outdated:
"They offer three pricing tiers starting
at $45/month..."  WRONG (changed 6mo ago)

// Sometimes hallucinated:
"They recently acquired DataSync
for $2B..."  WRONG (never happened)
```

PATHWAY 2

RAG / Retrieval

Perplexity answering "best CRM for startups"

```
"Based on recent reviews, HubSpot and
Pipedrive lead for startups in 2026..."

Sources cited:
  [1] hubspot.com/startup-crm      CITED
  [2] pipedrive.com/pricing        CITED
  [3] g2.com/crm-comparison        CITED
  [4] yourbrand.com/crm-guide      NOT CITED

// Why was your page skipped?
// Poor structure? No schema? Thin content?
```

PATHWAY 3

Agent Browsing

AI Agent navigating your website

```
Agent task: "Find the best CRM under $50/mo"

Step 1: Navigate to yourbrand.com
Step 2: Read pricing page
Step 3: ERROR: pricing loaded via JS widget
Step 4: ERROR: comparison table is an image
Step 5: Leave site. Try competitor.

// The agent tried. Your site failed the test.
// It moved on in 2 seconds.
```

Three pathways.
Three optimization challenges.

If your content isn't structured for machine comprehension, you're invisible in all three.

EXAMPLE

Same Query. Different Outcomes.

Brand A: Invisible to AI

- > Content behind JS tabs
- > Generic, keyword-stuffed copy
- > No author attribution
- > No 3rd party references
- > Blocks AI crawlers

VS

Brand B: Cited by AI

- > Content in initial HTML
- > Semantic HTML + Structured data
- > Clear, specific language
- > Named experts and brands
- > External websites write about the brand
- > Welcomes all crawlers


What Actually
Matters Now?

BAISOM

Bobbinks ~~Basic~~ AI Search Optimization Model

FRAMEWORK

The 7 Layers of BAISOM

Get in. Render. Navigate. Comprehend. Annotate. Write well. Earn trust.

> Layer 1: Crawlability & AI Access
> Layer 2: Performance & Renderability
> Layer 3: Accessibility & Machine Overlap
> Layer 4: Content Architecture
> Layer 5: Structured Data & Schema
> Layer 6: Clarity Over Keywords
> Layer 7: Authority & Trust Signals

LAYER 1

Crawlability & AI Access

New crawlers, new rules

> Googlebot is no longer the only bot that matters
> GPTBot (OpenAI), ClaudeBot (Anthropic), PerplexityBot are all visiting
> Your robots.txt now controls AI access, not just search access
> Blocking AI crawlers = choosing invisibility in AI answers
> Clean internal linking, render speed, and clean HTML matter more than ever

THE SWARM

226 AI Crawlers and Counting

Cloudflare 2025: The crawl volumes tell a stark story

> GPTBot grew 305% YoY. ChatGPT-User surged 2,825%.

> PerplexityBot: +157,490% in raw requests

> Crawl-to-refer ratio: Google 3:1 vs OpenAI 400:1 vs Anthropic 100,000:1

> 30% of AI scrapes now bypass explicit robots.txt blocks (TollBit Q4 2025)

> AI crawlers are becoming indistinguishable from human visitors in logs

EXAMPLE

Your Server Logs Right Now, Use it!

New visitors you didn't invite

```
66.249.x.x  Googlebot/2.1        -- the usual
20.15.x.x   bingbot/2.0          -- old friend
52.230.x.x  GPTBot/1.0           -- OpenAI
18.238.x.x  ClaudeBot/1.0        -- Anthropic
44.228.x.x  PerplexityBot/1.0    -- Perplexity
35.196.x.x  Bytespider           -- ByteDance
52.167.x.x  cohere-ai            -- Cohere

// They're already here.
// Are you letting them in?
```

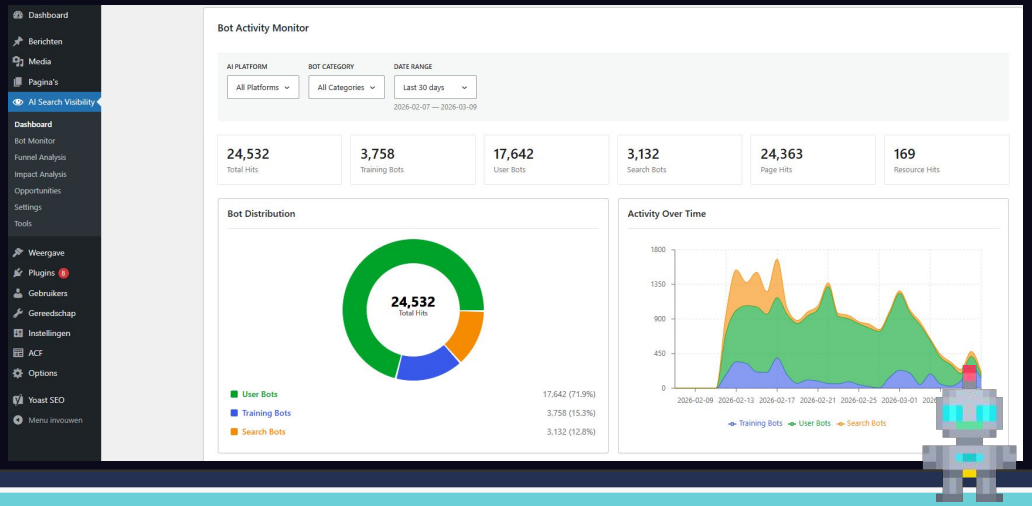
EXAMPLE

Monitor AI bots behaviour

Are pages being picked up and generating clicks?

1. Check for plugins for your platform

1. I built dashboards for WordPress and Shopify



EXAMPLE

robots.txt Decisions

Blocking AI

```
User-agent: GPTBot
Disallow: /

User-agent: ClaudeBot
Disallow: /

User-agent: PerplexityBot
Disallow: /

// "Safe" but invisible.
// AI won't cite what
// AI can't read.
```

Welcoming AI

```
User-agent: GPTBot
Allow: /
Disallow: /admin/

User-agent: ClaudeBot
Allow: /

User-agent: PerplexityBot
Allow: /

// Open for discovery.
// Strategic about what you share.
```

LAYER 2

Performance & Renderability

AI agents won't wait for your page to load

> Bloated DOM = more tokens to parse. Keep your HTML lean
> Async/defer scripts prevent render-blocking that hides content from agents
> Images without dimensions cause layout shifts that confuse agent screenshots
> Lazy loading is fine for humans, but AI agents need content in the initial HTML
> If your content needs JavaScript to appear, it doesn't exist for most AI.
> Use HTML5 <header> <main> <footer>

LAYER 3

Accessibility & Machine Overlap

What helps humans helps AI agents too

> Image alt text is how AI sees your visuals. No alt = invisible content
> ARIA landmarks and lang attributes give AI structural understanding
> Descriptive link text tells agents where a link goes before clicking
> Form labels and keyboard navigation make your site agent-browsable
> Accessibility was always the right thing. Now it is the smart thing too.

LAYER 4

Content Architecture

Making content navigable for machines

> Clear H1 > H2 > H3 hierarchy tells AI what matters most
> Descriptive headings, not clever ones: "Pricing table" beats "Let's Talk Numbers"
> Topic clustering creates connected knowledge, not isolated pages
> Summary paragraphs at the top give AI the answer immediately
> AI doesn't scroll. It reads top-down and decides in milliseconds.

EXAMPLE

Flat vs. Structured Content

Wall of Text

```
Our CRM is great for startups
and enterprises alike. With many
features including contact mgmt
and pipeline tracking and email
automation and reporting tools
you can do everything you need.

// AI: What does this page answer?
// AI: I can't extract specifics.
// AI: Moving on.
```

Clear Structure

```
<h1>CRM for Startups</h1>
<p>Summary: A CRM built for
teams under 50, from $29/mo.</p>

<h2>Key Features</h2>
<h2>Pricing Plans</h2>
<h2>Customer Reviews</h2>

// AI: Clear topic. Clear structure.
// AI: I can cite this confidently.
```

EXAMPLE

Topic Clustering Done Right

Connected content = comprehensive AI answers

```
Pillar: /crm-for-startups/
  |
  |-- /crm-for-startups/pricing/
  |-- /crm-for-startups/features/
  |-- /crm-for-startups/vs-hubspot/
  |-- /crm-for-startups/setup-guide/
  |-- /crm-for-startups/case-studies/

Result: AI pulls from the entire cluster
to build a comprehensive, trustworthy answer.
Your brand becomes THE source on this topic.
```

LAYER 5

Structured Data & Schema

Why machines need explicit context?

> Entity relationships connect your brand to the knowledge graph

> Without proper HTML like tables, AI has to guess. And it guesses wrong.

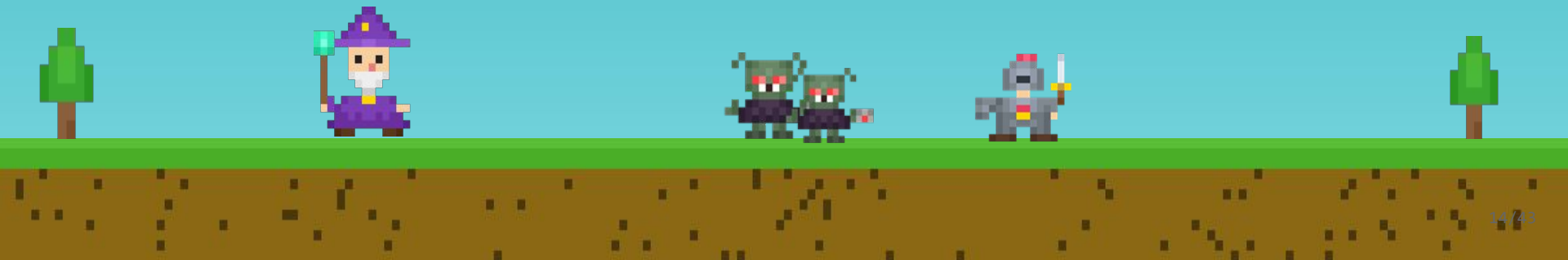
> LLMs have been trained on HTML.

> Schema.org seems to make an impact, sometimes

Schema.org markup?

"Schema markup helps Microsoft's LLMs understand content."

Fabrice Canel (Bing)



WARNING

Do not use LLMs for Schema.org

Built it yourself, learn something useful

The "LLM4Schema.org" study (University of Nantes/Inria, 2024-2025) found:

> 40-50% of markup produced by GPT-3.5 and GPT-4 contains errors
> After error filtering GPT-4 actually outperforms human annotators
> This demonstrates that LLMs have been trained on substantial Schema.org data

GOOD EXAMPLE

HTML Done Right

Proper HTML table

```
<table>
	<thead>
		<tr><th>Offer Type</th><th>Offer Name</th><th>Offer Price</th><th>Offer Currency</th></tr>
	</thead>

	<tbody>
		<tr><td>Product</td><td>Starter Plan</td><td>29.00</td><td>USD</td></tr>
	</tbody>
</table>

// AI now knows: who, what, how much
```

BAD EXAMPLE

Price in image, non structured text

What AI has to work with

```
<!-- No JSON-LD. No schema. Just HTML. -->

<div class="hero">
  <h1>Welcome to Brand X</h1>
  <p>Brand X is a CRM platform built for improving your workflow. It is founded by Jane Smith. We offer a product called the AI Starter Plan</p>
  <img src=”pricing-offer.jpg” alt=”Pricing table” />
</div>

// AI: For whom is the CRM?
// AI: What do they cost? What is included in the plan?
// AI: I don't know. Skip this result.
```

LAYER 6

Clarity Over Keywords

LLMs understand meaning, not keywords

> Keywords are irrelevant to language models
> LLMs understand synonyms, context, and intent natively
> Clear, specific, unambiguous language gets cited
> Write like you're explaining to a smart colleague
> The best content for AI is also the best content for humans

TESTING

GEO: What Actually Moves the Needle

Princeton KDD 2024 tested 9 optimization tactics

> Adding statistics to content: up to +41% AI visibility
> Expert quotations: +35-40% visibility improvement
> Citing credible sources: +115% for lower-ranked sites
> Keyword stuffing performed WORSE than baseline
> Lower-ranked sites benefit most. GEO is the great equalizer

Optimized

Be very explicit

AI needs facts and proper context

```
<div class="hero">
  <h1>Brand X: AI-Powered CRM for B2B SaaS Teams</h1>
  <p>Brand X helps B2B SaaS sales teams of 10-200 people close deals faster with AI-driven pipeline management,
  automated follow-ups, and real-time forecasting. Founded in 2019 by Jane Smith, former VP of Sales at
Salesforce.</p>
</div>

<section class="pricing">
  <h2>AI Starter Plan</h2>
  <p>$49 per user per month, billed annually. Includes AI lead scoring, automated email sequences, pipeline
analytics, and CRM integrations with HubSpot, Slack, and Google Workspace.</p>
 </section>
```

LAYER 7

Authority & Trust Signals

How LLMs assess credibility

> Named authors with verifiable expertise
> Citations and references to primary sources
> Publication dates and update signals
> Consistent entity presence across the web
> E-E-A-T isn't just a Google thing. LLMs evaluate trust too.

EXAMPLE

The Knowledge Graph Effect

Google Knowledge Panel

- > Company name, logo, description
- > Founded date, headquarters
- > Key people, social profiles
- > Products and services
- > Powered by schema + Wikidata

LLM Entity Understanding

- > Same data feeds LLM knowledge
- > Schema creates the "entity card"
- > Richer content = more accurate AI
- > Connected entities = stronger trust
- > You're building for both at once

EXAMPLE

Cited vs. Not Cited

Skipped by AI

- No author name, no bio
- "Admin" as author
- No publish or update date
- No sources or references
- Generic stock photo

VS

Cited by AI

- Named expert with credentials
- Author page with schema markup
- Clear publish date, last updated
- Links to Gartner, Forrester data
- Real author photo, LinkedIn linked

EXAMPLE

Authorship That Signals Trust

Building entity authority

```
{
  "@type": "Article",
  "author": {
    "@type": "Person",
    "name": "Jane Smith",
    "jobTitle": "VP of Product",
    "url": "brand.com/team/jane-smith",
    "sameAs": ["linkedin.com/in/janesmith"]
  },
  "datePublished": "2026-01-15",
  "dateModified": "2026-02-28"
}
// Person + Role + Links = verifiable expertise
```

PARADIGMS

Brand Mentions Are the New Backlinks

Ahrefs, December 2025: 75,000 brands analyzed

> Brand mentions correlate 3x more with AI citations than backlinks
> YouTube mentions: strongest correlation (0.737) across all AI platforms
> Domain Rating correlation dropped to just 0.266 for ChatGPT
> 76.1% of AI-cited URLs also rank in Google's top 10
> “Topics are the new keywords. Topical authority is the new PageRank.”

BAISOM

Crawl. Comprehend. Convince.

CRAWL

Can AI access it?

- > Allow AI crawlers
- > Clean sitemap structure
- > Fast, renderable pages
- > No JS-only content

COMPREHEND

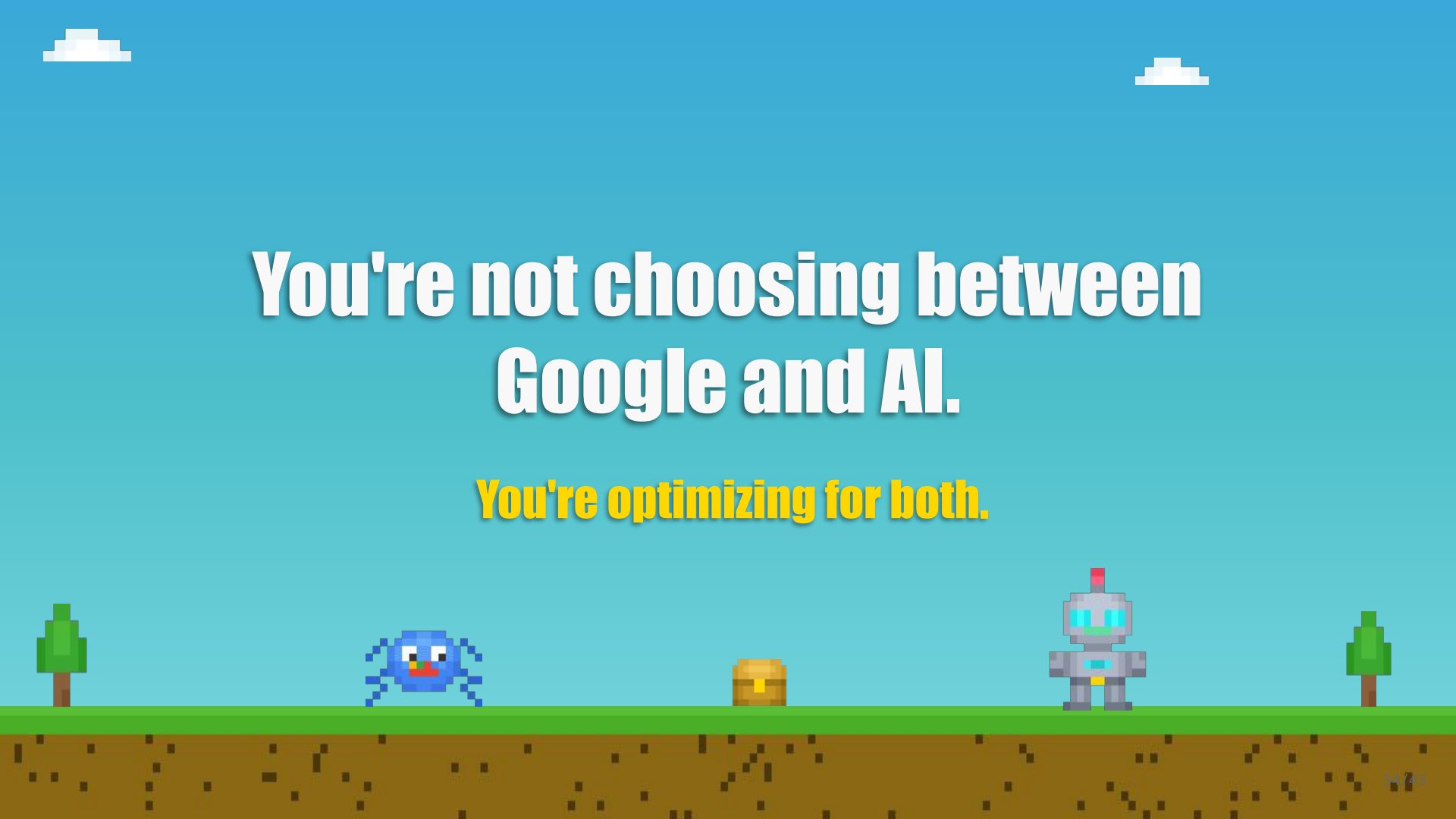
Can AI understand it?

- > Structured data / JSON-LD
- > Clear heading hierarchy
- > Specific, descriptive copy
- > Entity-rich content

CONVINCE

Does AI trust it?

- > Named expert authors
- > Citations and sources
- > Fresh, accurate info
- > Consistent brand entity


You're not choosing between Google and AI.
You're optimizing for both.

OVERLAP

Google + AI: Shared Ground

GOOGLE ONLY

> PageRank / backlinks
> Core Web Vitals weight
> SERP feature targeting
> Local pack optimization

SHARED

> Quality content
> Structured data
> Site architecture
> Authority / E-E-A-T

AI ONLY

> Machine readability
> AI crawler access
> Entity clarity
> Factual specificity

EXAMPLE

Scoring a Real Site

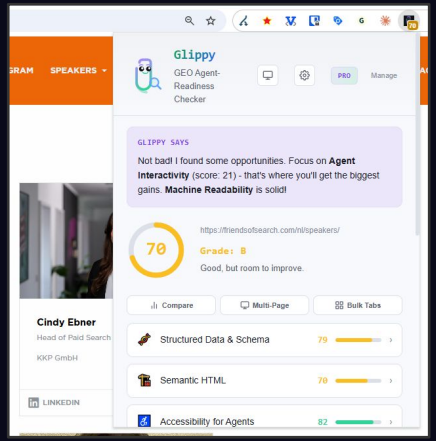
Audit: brand-example.com

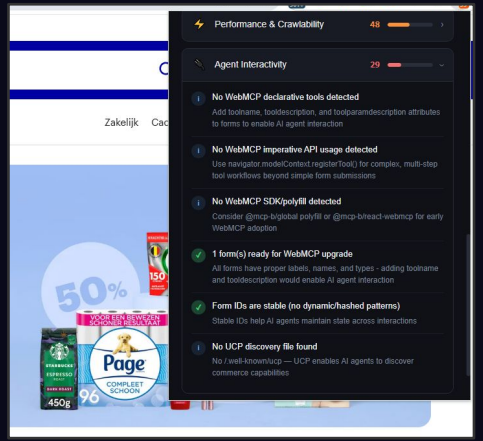
```
CRAWL:
  robots.txt allows AI bots      OK
  Sitemap is current              OK
  JS rendering required           FAIL
  Score: 2/3

COMPREHEND:
  JSON-LD on key pages            OK
  Clear heading structure         FAIL
  Topic clusters exist            OK
  Score: 2/3

CONVINCE:
  Named authors                   FAIL
  Sources cited                   FAIL
  Fresh content                   OK
  Score: 1/3
```

EXAMPLE

Check out tools like Glippy.dev

Audit: friendsofsearch.com and bol.com





STRATEGY

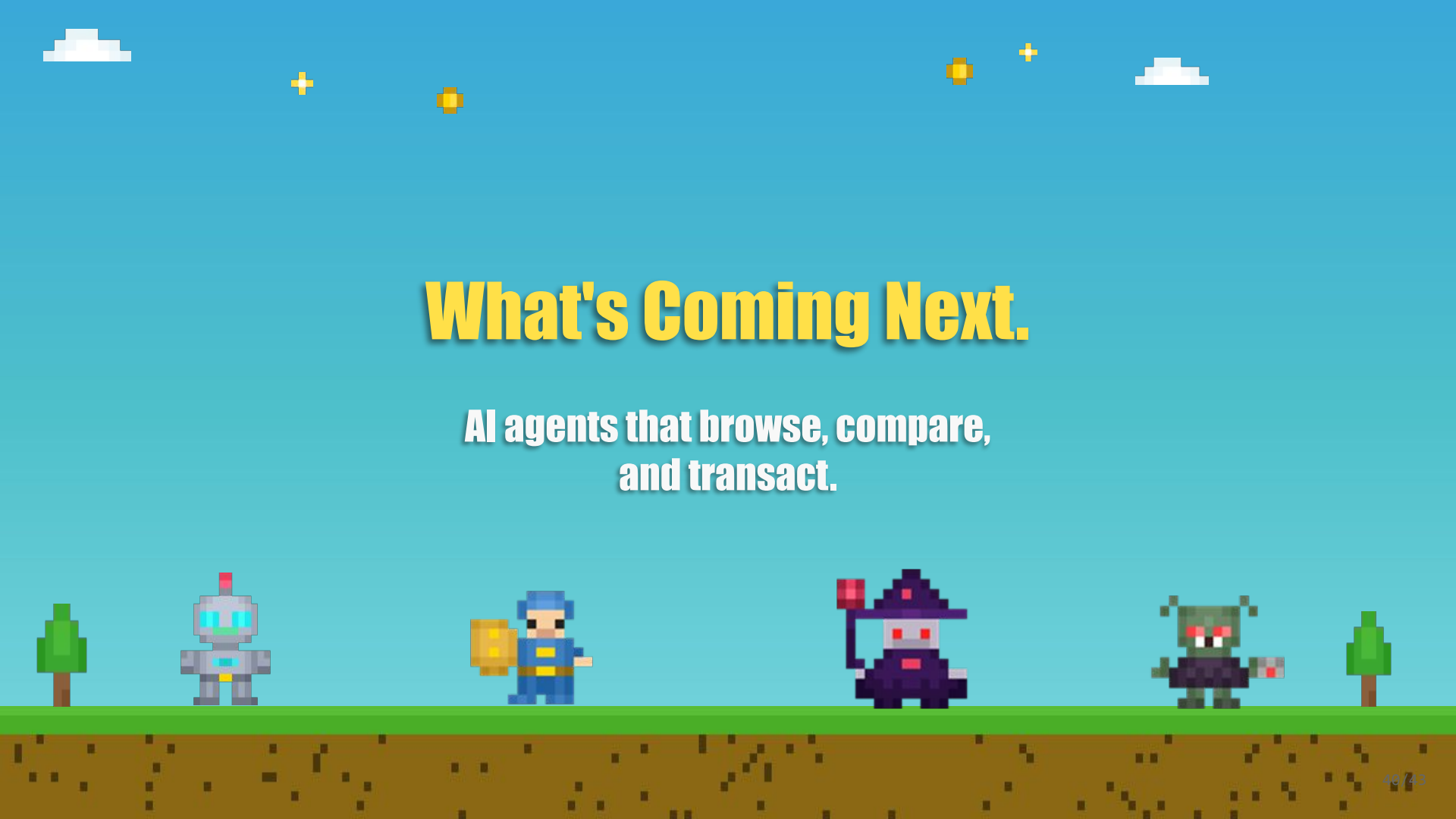
Where to Start

Quick Wins (This Week)

- Update robots.txt for AI bots
- Add Organization schema
- Add author names to content
- Check JS rendering issues
- Add publish dates to articles

Strategic Bets (This Quarter)

- Build topic cluster architecture
- Work on external references
- Create author and brand entity pages
- Audit and fix JS-dependent content
- Develop AI-specific content strategy


What's Coming Next.
AI agents that browse, compare,
and transact.

THE FUTURE


```
{
  "ucp": {
    "version": "2026-01-11",
    "services": {
      "dev.ucp.shopping": {
        "version": "2026-01-11",
        "rest": {
          "endpoint": "https://myshop.example.com/"
        }
      }
    },
    "capabilities": [
      {
        "name": "dev.ucp.shopping.checkout",
        "version": "2026-01-11"
      },
      {
        "name": "dev.ucp.shopping.discount",
        "extends": "dev.ucp.shopping.checkout"
      },
      {
        "name": "dev.ucp.shopping.fulfillment",
        "extends": "dev.ucp.shopping.checkout"
      }
    ]
  }
}
```

README Contributing License

WebMCP 🧪

Enabling web apps to provide JavaScript-based tools that can be accessed by AI agents and assistive technologies to create collaborative, human-in-the-loop workflows.

First published August 13, 2025

Brandon Walderman <brwalder@microsoft.com>
Leo Lee <leo.lee@microsoft.com>
Andrew Nolan <annolan@microsoft.com>
David Bokan <bokan@google.com>
Khushal Sagar <khushalsagar@google.com>
Hannah Van Opstal <hvanopstal@google.com>

THE FUTURE

An AI Agent Books a Flight

Your site is now a product page for a robot customer

```
User: "Book me a flight to Barcelona under $500"

Agent:
  Step 1: Browse airline-a.com     $489 found
  Step 2: Browse airline-b.com     ERROR: JS widget
  Step 3: Browse airline-c.com     $520 found
  Step 4: Compare and recommend    airline-a wins
  Step 5: Complete booking

Result: Airline B lost the sale.
Not because of price. Because of accessibility.
```


The companies that
figure this out first win.
The rest become invisible.

AI search is not a content problem: It is a brand problem

Jan-Willem Bobbink / Friend of Search / 2026



Thank You!

Level Complete.

Jan-Willem Bobbink

https://www.linkedin.com/in/jbobbink/

GSCwizard.com | Glippy.dev | noinfo@notprovided.eu