

Caribbean Hubspot User Group · August 2026

How to track and measure AI impact on your website

We will show you the tools, explain you the benefits, the challenges and give you a scorecard you can run every month.

Vuk Ilić · SEO specialist, CRO:NYX Digital

hello

Quick intro, then we get into it

- I'm Vuk. I've been doing SEO for 15+ years.
- I've done this job from most of the angles it has. Small businesses, agencies, in-house, enterprise.
- I am a genuine SEO geek. I nerd out on this stuff for fun!
- Outside of work, I watch and play sports (basketball, soccer, padel, squash). Mostly badly, but enthusiastically.

50%+

I turned 29 a couple of weeks ago. Which means, officially, I have now spent more than half my life doing SEO.

Two things before we get into the weeds



GEO, AEO, AI search - same thing

People are still arguing about which acronym wins. I'm going to use them interchangeably and you should too, because none of them change what you actually do. If someone corrects you on LinkedIn, they're procrastinating.



And this lands on the SEO team

In almost every company I've seen, whoever owned organic ends up owning AI search. It is not a new department but it's a new surface on a job SEOs already do.



Also worth saying up front: nobody in this space has it fully figured out. I'll tell you where the data is solid and where I'm guessing.

Here's where we're going

1	What's the setup?	What actually changed, and why your analytics are quieter
2	AI search 101	How these things pick sources and why the numbers wobble
3	The AI impact stack	Four layers with free native tools first
4	The scorecard to use	Eight numbers you should be tracking
?	Questions?	The best part!

part one

What AI search actually brings us

Before we measure anything, let's agree on what changed

the setup

the numbers everyone is quoting right now

-27%

organic traffic, year over year, across
HubSpot's customer base

~3x

growth in AI referral traffic over the
same period

4.4x

better conversion rate from LLM-
referred visitors than from organic
search



This is one company's customer base, heavily B2B, heavily marketing-software. Your numbers will be different.

the setup

So, is SEO dying?

No. The playing field got bigger.



People are searching more, not less

There is no shortage of demand. Conversational search made research faster, so people do more of it.



More of your category gets discovered

Buyers now surface brands, solutions and competitors they would never have scrolled to on page 4 of Google.



The visibility just moved

And it moved somewhere your current reporting doesn't have a column for.

Great for users. Harder for marketers.



What got better for the buyer

- Instant answer instead of multiple links to go through
- Conversational refinement which means they can argue with the answer
- Comparison across many brands in a single question
- A shortlist in minutes instead of an afternoon



What got worse for us

- Fewer clicks because the answer absorbs them
- You no longer control how you're framed
- Selection criteria are a black box
- The occasional confidently-wrong claim about you



People who do click through have already read about you. They arrive researched and pre-qualified.

part two

AI search 101

What it is, how it actually picks sources, and the two problems that make it hard to measure.

Three different things with the same name

1

Standalone assistants

ChatGPT, Gemini, Claude, Copilot, Perplexity. Someone asks, the assistant answers, and sometimes (not always) it links out to you.

This is the bucket you can actually measure today.

2

AI inside traditional search

AI Overviews and AI Mode. Same google.com URL, completely different game.

This is almost certainly your biggest AI surface. It is also the one that hides inside "Organic Search".

3

AI browsers and agents

Software that browses on a human's behalf and compares options, fills forms, books things.

Real category. Very unsettled. We'll deal with it honestly in Layer 3.



Bucket 2 is where most of your AI exposure actually is, and it is the hardest one to see. Almost every "AI traffic" dashboard you've been shown only measures bucket 1.

How it all works?



What the model already knows

Baked in at training time. Months old, frozen, and you cannot influence it between now and the next training run.



What it goes and fetches right now?

Live retrieval or "grounding". It runs real searches against a live index and reads what comes back. This is the door your site walks through.



One synthesised answer

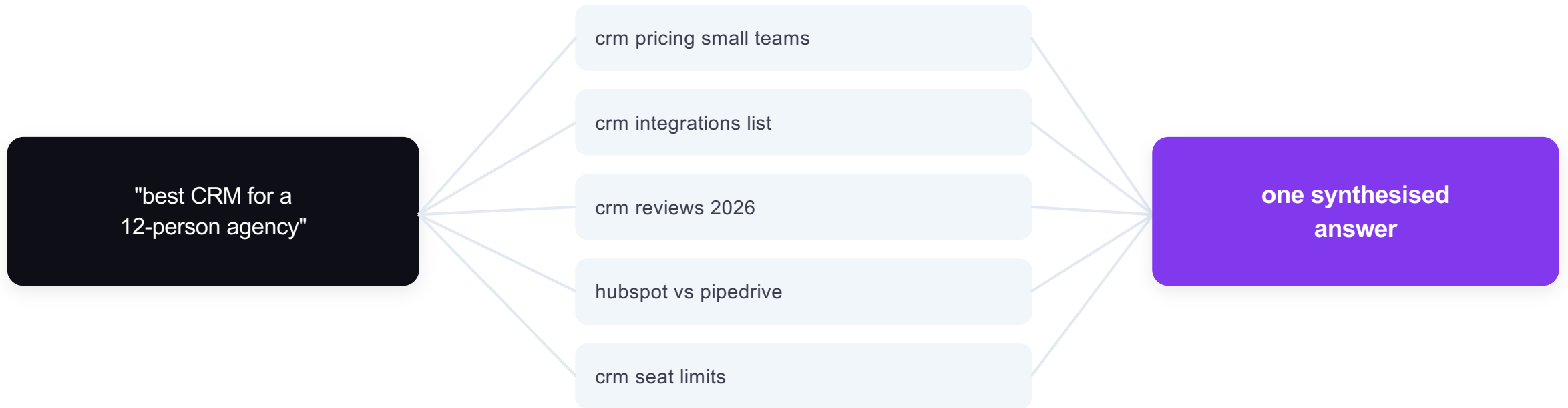
It merges both into a single response. Citations are optional and inconsistent so you can be used and not credited.



Step 2 is the one that matters for us. Grounding means you don't have to wait for a retrain to become visible, and you can publish something useful today and it can be pulled into an answer this week. It also means your visibility can move without you touching anything.

One question becomes twenty searches

They call it query fan-out, and it is the single biggest reason keyword rank stopped being a useful number.



You never ranked for the question. You ranked for one of the twenty which is why "what's my position for X" is now the wrong question to bring to a client call.

Ask the same question twice, get two different answers

- These systems are **non-deterministic** by design. You can use the same prompt on the same day and get a different answer.
- Add personalisation, location, chat history and A/B tests on the platform side, and "your position" stops being a thing that exists.
- So any tool that shows you a single, confident "you rank #3 in ChatGPT" is showing you one roll of the dice, presented as a fact.

The same prompt, three times

- 1 you're cited 2nd
- 2 you're not mentioned
- 3 you're cited 1st

Nothing changed on the site. Nothing changed in the prompt.



if the answer changes every time, what are we even measuring?

Stop treating it like a rank tracker

Every number in the AI search should serve as directional. Once you accept that, the whole thing gets much easier and much more useful.



Read direction, not position

Up or down over eight weeks is a signal. "We're number three" is noise.



Look at who gets cited, and why

The real insight is the pattern. Which sources keep appearing, and what they cover that you don't.



Same prompts, same schedule

Fix your prompt set. Run it monthly. Change the prompts and you've thrown away your baseline.



Sample size beats precision

Thirty prompts across four weeks tells you more than one perfect check on a Tuesday. Aggregate, then look.



Directional data still tells you what to do. It tells you which topics you're invisible for, which competitors keep getting pulled in instead of you, and whether last quarter's work moved anything. That's enough to build a plan on.

problem two

The clicks you can see are the tip of it

Monday

**She asks ChatGPT
which tools to shortlist.
You're in the answer.**

Thursday

**She googles your
brand name to check
you're real.**

The following week

**She converts on a
branded search.**

Your report says

**"Organic
search/branded." AI is
nowhere in it.**



Attribution rarely says "AI". It says "organic". Or "direct". The AI did the persuading and traditional search got the credit which means the channel that is actually growing can look flat in your reporting, and the channel that is actually shrinking looks like it's holding up fine.

When the data goes dark, go and ask people

We are losing a real chunk of the tracking. The only way to fill that in is to talk to the humans on the other end of it.

- Add one question to the form and one to the sales script: "how did you first hear about us?" with an AI assistant option in the list.
- Then actually read the answers. Tag them. Most companies collect this field and nobody opens it or analyzes it.
- Do five customer calls/interviews a quarter. Ask what they searched, what they read, who else they looked at, and what nearly stopped them.
- Reverse-engineer it: the questions they describe asking become your prompt set in Layer 2. Real buyer language, not keyword-tool language.

"what made you buy from us?"

Five answers to that question will tell you more about your AI visibility than any tool in this deck because the buyer remembers the journey your analytics lost.



This isn't a consolation prize for lost tracking. Done properly it is better data than the tracking ever was.

part three

The AI impact stack

Four layers. We go one at a time, free native tools first!

Four layers, four different questions

1	Traffic	Who actually clicked through to you?	GA4 · HubSpot
2	Visibility	How often does AI show and cite you?	Search Console · Bing WMT · prompt trackers
3	Consumption	Which AI systems are reading your site?	server logs · your CDN
4	Outcomes & influence	What does it all turn into? leads, revenue, demand?	HubSpot · branded search · self-reported



GA4 finally has a channel for this

Rolled out across accounts in May and June of 2026. If you haven't looked since spring, go and look.

Where it lives?	Reports → Acquisition → Traffic acquisition
What to look for?	the "AI Assistant" default channel
How it decides?	medium = ai-assistant, or campaign = (ai-assistant)
What is the cost?	free, on by default, nothing to configure



what it gets right

- You get a named line in the report instead of a pile of referrals
- It works retroactively for anyone who reports on channel
- Your CFO can find it without you



If you do one thing from this whole section: open GA4 after this call and check whether that channel is showing data.

Now, before you put that number in a board deck...

1

The list moves without telling you

Which assistants count has already changed since launch. Claude came off the list. Perplexity was never on it. Your month-over-month can move because Google changed its mind.

2

It's a black box

Google does not publish the full matching list, so you cannot audit it. You cannot tell a client exactly what is and isn't in that number.

3

Nothing is reclassified backwards

When a new assistant gets added, your old data stays wrong. So a year-on-year comparison is comparing two different definitions.

4

And the copy-paste regex is broken

Google's own documented example uses `^.*ai` which happily matches gmail. If you pasted that in, some of your email traffic has been sitting in your AI bucket.



None of this makes the channel useless. It makes it a starting point rather than a source of truth. Use it for direction, and build your own version underneath it so you know exactly what you counted.

Build your own channel group anyway

You should do this because:

- Custom channel groups **apply retroactively** to your historical data. That means you will get a consistent definition across last year too, not just from today.
- **You own the list** - you decide which assistants count, you can add the next one the day it launches and you know everything that goes into it.

Channel details

Channel name

AI Traffic

Channel conditions

Add conditions to define your custom channel

! At least one condition is required and every channel condition requires a dimension and a value. A channel can have across all condition groups.

Match AT LEAST ONE rule in this group

Source

+ Add condition group

Conditions*

matches regex

^(?:chatgpt\.com|chat-gpt\.o

Cancel Apply

HubSpot has this too and it goes a bit further

It's called AI Referrals, it sits alongside Organic and Paid as its own traffic source, and it follows the person into the CRM.



the bit GA4 can't do

- Break it down by individual assistant, not one blended bucket
- It stamps the contact record so you can see which deals started in an AI answer
- Those contacts can be segmented and reported on like any other source

Summarized Data		Unsummarized Dataset	
☐	SOURCE	NEW VISITOR SESSIONS T1	RETURNING VISITOR SESSIONS T1
☐	Organic search	11,593	576
☐	Direct traffic	1,881	605
☐	Referrals	137	71
☐	AI Referrals	108	16
☐	Other campaigns	41	3
☐	Organic social	18	1
☐	Email marketing	2	6
☐	Paid search	1	0
☐	Offline sources	-	0
Report Total		13,781	1,278

< Prev 1 Next > 10 rows per page



This is the closest thing anyone has to a straight line from "an assistant recommended us" to "that became pipeline". If you're on HubSpot, this is your Layer 1.

And here's what layer one will never show you...

This is just how this works.



Mobile apps don't pass a referrer

When someone taps a link inside the ChatGPT or Gemini app, the referrer is frequently stripped before it reaches you. That session lands in Direct, forever. Given how much assistant usage is on a phone, this is not a rounding error.



AI Overviews and AI Mode look like organic

Both live on google.com and both pass a normal google/organic referrer. Your single largest AI surface is blended into a channel and you can't separate your AI search efforts from your SEO efforts in an actual report.



This is exactly why the stack has four layers. If clicks were the whole story, one GA4 report would do it. Clicks are the smallest, most visible slice so we go to Layer 2 and start counting the times you showed up without a click at all.

2

layer two

Visibility

How often you show up in an AI answer at all, including every single time nobody clicked.

Search Console added an AI report in June

New, limited and...somewhat useful, I guess?



What you get

- A generative-AI breakdown inside the Performance report
- Impressions, how often your pages surfaced in AI experiences
- Trended over time, so you can watch a direction



What you don't get

- No clicks. No CTR. No position.
- No queries so you cannot see what triggered you
- It's rolling out in stages, so your account may not have it yet



This exists largely because a UK regulator pushed for, which tells you something about how fast the rest of the transparency is likely to arrive.

Bing is ahead of Google here

Bing Webmaster Tools has had an AI Performance report in public preview since February, and it got a serious upgrade in June.

- Clicks and impressions from Copilot - split out from regular Bing search so you can see the AI surface on its own rather than blended into total.
- Intents and topics - Added in June of 2026. What people were actually trying to do when you appeared, grouped into themes rather than raw strings.
- Citation share - how often you get cited relative to the rest of the field, plus a compare view. This is the closest thing to a share of voice number that comes free.

Grounding Query	Intent	Topic	Citations ↓	Citation Share
hawaii flooding 2026	Informational	Flooding & Flood ...	40.2K	22.45%
cuaca hari ini	Local	Weather Forecasts & Radar	39.7K	14.98%
El Nino 2026	Live Event	Weather Science & Meteorology	33.3K	7.98%
unterwasser-meereshöhlen	Informational	Oceans & Seas	30.6K	18.24%
weather today	Local	Weather Forecasts & Radar	29.1K	17.06%
lunar eclipse	Informational	Severe Weather & Storm Tracki...	28.9K	1.82%
weather channel	Navigational	Severe Weather & Storm Tracki...	24.6K	35.75%
the weather channel	Navigational	Severe Weather & Storm Tracki...	21.5K	35.34%
clima saltillo	Local	Weather Science & Meteorology	19.1K	35.94%
what is el nino	Informational	Weather Science & Meteorology	19.0K	11.67%
tornado watch vs warning	Comparison	Weather Alerts & Notifications	18.9K	15.07%
الطقس اليوم	Local	Weather Forecasts & Radar	18.5K	11.68%
weather	Local	Weather Forecasts & Radar	18.0K	16.65%
cleveland ohio meteorite	Informational	Weather Science & Meteorology	17.1K	10.87%



Yes, Bing's absolute volume is small. But you should look at it anyway, as it is a live window into how a real AI system chooses and cites sources, and that behaviour generalises even when the traffic numbers don't.

Grounding queries are the closest thing to AI keyword data

Remember the fan-out diagram? These are the middle column. And Bing will show them to you.



What they are

The searches the AI generated for itself while building an answer. Machine-written retrieval strings. They tell you what the system thought it needed to know, and which of those needs you turned out to satisfy.



What they are not

They are not what the human typed. Do not present them to a client as "here are the prompts people used to find you". That is a genuinely wrong claim and someone in the room will know it.



Used properly they're a content brief generator. If the same retrieval need keeps appearing and you keep not being the answer to it, you've just found your next piece of content and you found it from machine behaviour rather than a keyword tool's guess.

Before you get excited about impressions

48,000

AI citations over the measured period



14

clicks that came back from them

0.03%

the click-through rate, if you want to call it that

- Impressions are not traffic, and citations are not clicks.
- Report AI impressions as a growth number without this context and you will be asked where the revenue is.
- Visibility is real value. It just cashes out somewhere other than your sessions report.

So why measure it at all?

Because being in the answer is how you get into the consideration set. You just have to stop expecting the click to prove it.

And if you want to watch the assistants directly

This is where most of the investment in AI search reporting happens and where most of the conversation is at today.

HubSpot AEO	\$50/mo standalone. Tracks ChatGPT, Gemini, Perplexity only. No Claude.
Ahrefs Brand Radar	Mentions and citations across assistants, with share-of-voice trending.
Semrush AI toolkit	Prompt-level visibility and sentiment, sold as an add-on.
Roll your own	Hit the APIs on a schedule with a fixed prompt set and log the answers.



what actually matters

- A fixed prompt set you never change
- Prompts written in buyer language
- The same schedule every month
- Who else got cited, every time



Honestly, the DIY option is underrated. Thirty prompts, an API key and a scheduled script gets you 80% of what these tools sell, you own the data, and you know exactly how it was collected.

3

layer three

consumption

Which AI systems are actually reading your website and what they took while they were there.

Your server logs already know all of this

Every crawling request is being logged. It is the most raw dataset you own and a lot of people will miss it.



GPTBot

OpenAI · training



OAI-SearchBot

OpenAI · search index



ChatGPT-User

Fetching live, for one person's question



ClaudeBot

Anthropic



PerplexityBot

Perplexity



Google-Extended

Google's AI opt-out control



A training crawler took your content for a future model. A live-fetch agent like ChatGPT-User went and got your page because a real person asked a real question in that moment.

You are the one who decides who gets in

Once you can see who is reading you, someone will ask whether you should let them. That is a business decision, not a technical one.



The case for letting them in

If you block the crawlers, you cannot be cited. For most B2B companies, being in the answer is worth more than the traffic you are protecting because you were never getting that traffic anyway.



The case for keeping them out

If your content is the product such as publishing, research, anything behind a subscription... giving it away to be summarised for free is a real commercial loss, and citations don't really make up for it.

robots.txt

The actual control - allow or disallow per bot, and remember training and search bots are separate

llms.txt

A proposed convention for pointing AI systems at your best content. Low cost, unproven, do it if you are curious

4

layer four

Outcomes - Leads, sales and revenue

And the much larger thing you will never get credit for.

First, the part that is actually straightforward

If you're on HubSpot, AI Referrals follows the person all the way through. That gives you a real, defensible number.



An assistant sends someone over

The session is tagged AI Referrals at the source level.



They convert into a contact

Original Source is stamped on the record and stays there.



The contact becomes a deal

Now you can report revenue by original source, like any other channel.



Two numbers to pull every month: how many contacts had AI Referrals as their original source, and what those contacts were worth. That is a straight answer to "is any of this making money".

The influence gap

The number you can attribute and the number you actually influenced are not remotely the same number.

what you can attribute



the sliver with a referrer attached

what actually influenced the sale



every answer they read, every comparison, every mention that never sent a click



This is not new and brand marketing has lived with it forever. What is new is how much of the buying journey just moved into that unmeasurable space, and how quickly. If you only defend the blue bar, you will lose the budget that produced the orange one.

Branded search is your cheapest AI proxy

If an assistant recommends you, a lot of people go and google your name to check you're real. That leaves an obvious trace.

- Pull branded query impressions and clicks from Search Console, monthly, on a fixed date.
- Watch the trend against your AI visibility trend from Layer 2 and 3. If both climb, something is working.
- Also watch branded search volume against your own paid and PR activity, so you can rule out the obvious other causes.
- A rise in branded search with no campaign behind it is the single clearest sign that something is recommending you somewhere you cannot see.



Be careful with it

This is correlation, and it will never be anything else. Branded search moves for a dozen reasons. It can be a conference, a funding round, a competitor's bad week etc. Present it as a supporting signal but never as proof.

Just ask your customers. Seriously.

Self-reported attribution is unfashionable, unglamorous, and currently the best tool you have for the dark part of the funnel.

30-50%

of pipeline routinely comes back attributed to sources your click tracking never sees. So don't be lazy and ask these questions.

- One open field on the demo form: "how did you first hear about us?"
- One question in the sales discovery script, asked the same way every time.
- Tag the answers into a handful of buckets and report them monthly next to your attributed numbers.
- Add "an AI assistant" as an explicit option. People will not volunteer it, but they will tick it.



The reason this works is boring: the buyer remembers the journey your analytics lost. They know they asked ChatGPT. They just have no reason to tell you unless you ask, and no way to tell you unless there is a box.

Look for things moving together

Three correlation checks that take an afternoon to set up and give you a story you can defend in a QBR.



visibility vs direct traffic

If AI citations climb and unexplained Direct sessions climb with them, you are probably watching the Layer 1 referrer loss show up in the only place it can.



AI-sourced leads vs lead quality

Compare close rate and sales-cycle length for AI-sourced contacts against everything else. In most places these people convert better because they arrive pre-researched.



citations vs branded search

Chart them on the same axis over eight to twelve weeks. Two lines moving together is not proof, but it is a far better argument than either line alone.



Say the word "correlation" out loud when you present these, every single time. The moment you let a stakeholder hear "AI drove this revenue" from a correlation chart, you have written a cheque that the next quarter's data will bounce.

8 numbers to track every month

If you build nothing else from this talk, build this. Same day every month, same definitions, no exceptions.

1

Traffic

sessions from the AI Assistant channel

contacts with AI Referrals as original source

2

Visibility

AI impressions - Search Console + Bing

citation share across your fixed prompt set

3

Consumption

AI crawler hits, split live-fetch vs training

the pages they keep coming back for

4

Outcomes

pipeline from AI-sourced contacts

branded search + self-reported AI mentions

and just as important

3 things I'd ignore for now



chasing a position in an AI answer

The answer changes between two people asking the same question ten seconds apart. Optimising toward a number that unstable is a way to look busy while learning nothing.



schema as an AI-ranking lever

Structured data is good practice, and you should have it. But the evidence that more of it makes an assistant cite you is thin. Careful testers keep failing to find it.



anyone selling you certainty

If a vendor demo shows a precise AI rank with no error bars and no methodology, ask how it was collected. The good tools will tell you. The rest change the subject.



The stuff that does still work is unglamorous and familiar: be genuinely useful on the questions your buyers actually ask, be quotable, be technically accessible to the crawlers, and be mentioned in the places the models read. That is most of the job. It just has a new acronym on it now.

questions.

Feel free to ask anything that interests you on this topic - and any adjacent ones!

Vuk Ilić · SEO lead, CRO:NYX Digital