

The 2026 AI Citation Visibility Study

How 50 crypto protocols appear in ChatGPT, Perplexity, and Google AI Overviews

Author	David Wood — CryptoContent.dev
ORCID	0009-0007-3353-7188
Published	10 May 2026
DOI	10.5281/zenodo.20146677
Sample	50 protocols, 7 categories, 40 prompts
Platforms	ChatGPT, Perplexity, Google AI Overviews

Key findings at a glance

Schema does not predict citations	Protocols with the best schema scores are not the most cited. Pendle, DeFi's only schema protocol, scored zero across all three platforms.
Ethereum — 81.9/100	Highest overall AI citation visibility score
1%	Official pages: share of Perplexity citations
4%	Official pages: share of Google AI Overviews citations
0%	ChatGPT responses that include source URLs
YouTube — 111 citations	Top cited domain in Google AI Overviews
CoinGecko — 6 citations (9%)	Top cited domain in Perplexity
74% (37 of 50)	Crypto protocols with no JSON-LD schema
1 of 10	DeFi protocols with any structured schema
6 protocols	Invisible across all three AI platforms
1,016 records	Total citations collected (Perplexity + Google AIO)

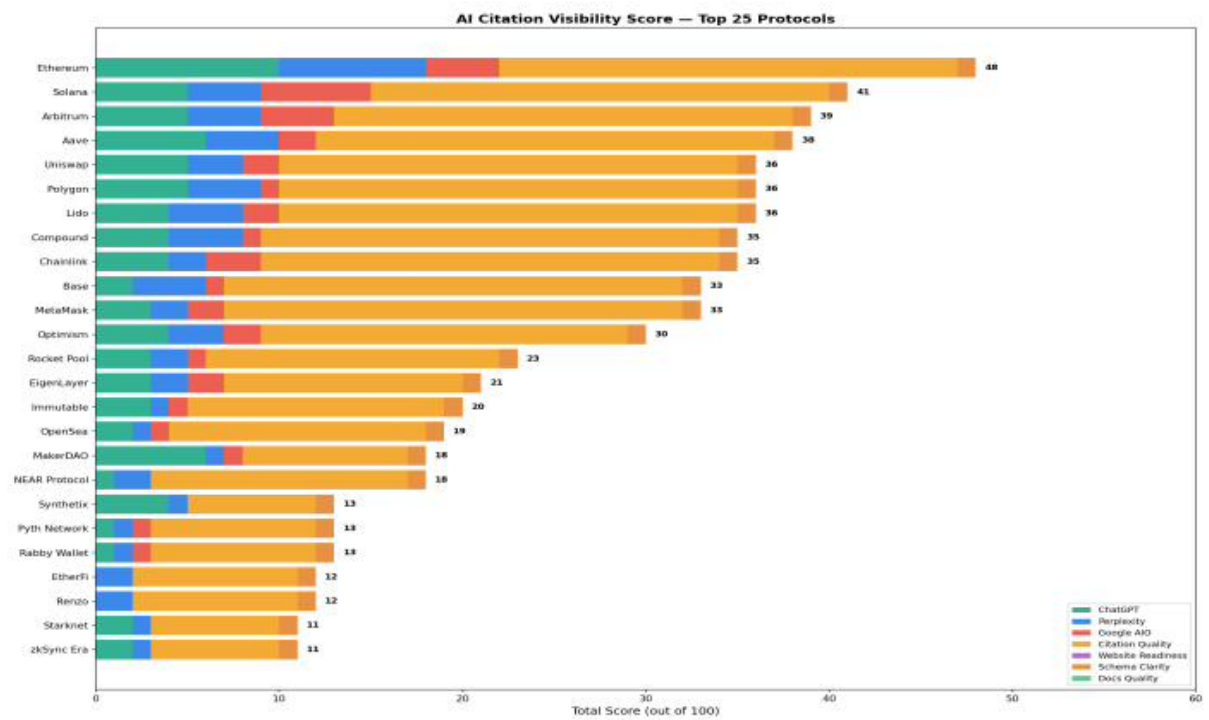


Fig. 1 — AI Citation Visibility Score: top 25 protocols (out of 100 points)

Schema markup does not predict AI citation visibility in crypto. The data says so directly.

The 2026 AI Citation Visibility Study audited 50 crypto protocols across ChatGPT, Perplexity, and Google AI Overviews, scoring each against a 100-point framework covering AI answer presence, citation quality, website structure, schema implementation, and documentation quality. The results contradict one of the most repeated claims in generative engine optimisation: that structured data improves how AI systems find and cite your content.

Pendle is the only DeFi protocol in the study with any JSON-LD schema on its homepage. It scored zero across all three platforms for AI presence. Zero mentions in ChatGPT. Zero in Perplexity. Zero in Google AI Overviews. Its entire score of 29.3 out of 100 comes from structural audit points that no user will ever see translated into a citation.

Starknet has the joint-highest schema quality score of any protocol in the 50-protocol sample. It ranks 15th overall. Aave has no schema at all. It ranks 2nd overall, with a citation quality score of 24 out of 25.

The protocols that appear most consistently in AI-generated answers are not the ones with the best technical infrastructure. They are the ones with the deepest content corpus: protocols that journalists and researchers have written about most extensively over the longest period. AI systems are drawing on that accumulated third-party record, not on what protocols have declared about themselves in structured data.

For the past year, a significant strand of GEO and AEO advice has positioned schema markup as a near-direct path to AI citation. This study does not support that claim. Structured data may help AI systems resolve entities and understand page relationships, but in this sample it was not a reliable predictor of which protocols actually appeared in AI-generated answers. The protocols that did appear had something schema cannot provide: years of third-party coverage that AI systems had already indexed.

Official protocol pages are losing the citation layer entirely

Across 1,016 citation records from Perplexity and Google AI Overviews, official protocol pages accounted for 1% of Perplexity citations and 4% of Google AIO citations. The top cited domain on Perplexity is CoinGecko. The top cited sources on Google AI Overviews are YouTube at 111 citations and Reddit at 94. Protocol teams are being described by AI systems almost entirely through third-party content.

Study methodology

The 2026 AI Citation Visibility Study scores 50 crypto protocols across a 100-point framework measuring AI answer presence, citation quality, website retrieval readiness, schema and entity clarity, and documentation quality. The full dataset, including raw audit data and scoring formulas, is available in the downloadable spreadsheet.

Protocol selection

The 50-protocol sample spans seven categories, selected using publicly available rankings from DefiLlama, L2Beat, and CoinGecko. Protocols were chosen to represent a range of sizes, ages, and technical approaches within each category rather than to favour high-scoring outcomes.

Category	Protocols
DeFi	10
Layer 1	8
Layer 2	8
Infrastructure	6
Wallets	6
Staking	6
Gaming/NFT	6
Total	50

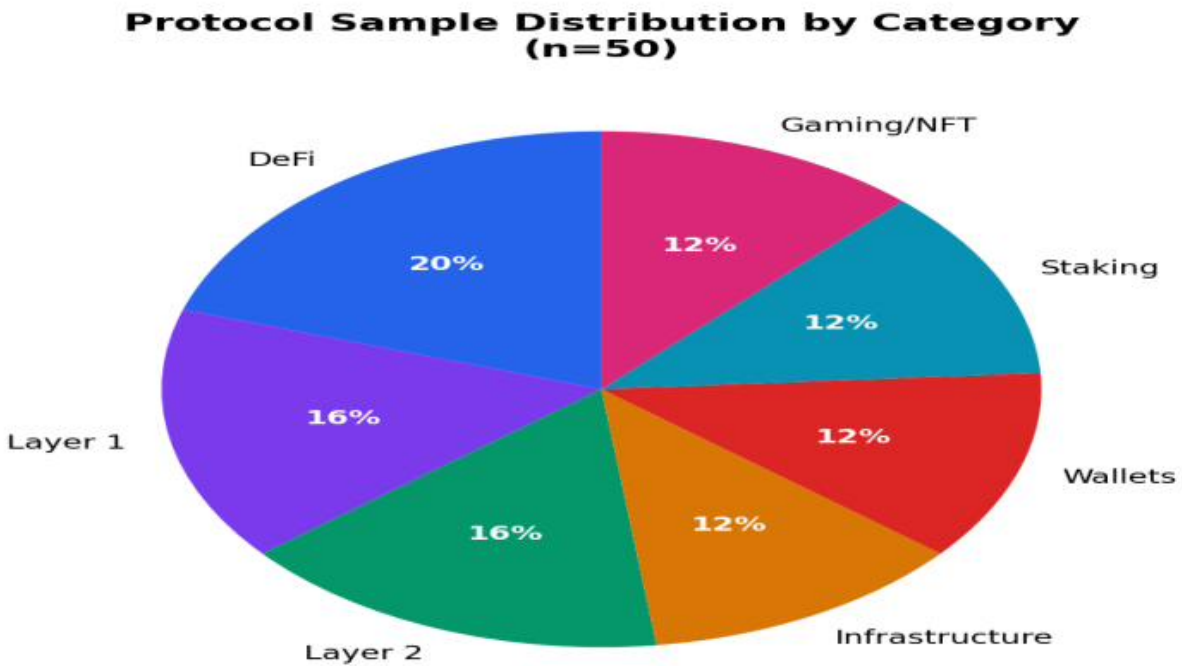


Fig. 2 — Protocol sample distribution by category (n=50)

Prompt set

Forty prompts were tested across five types, written in natural user language rather than keyword-style queries.

Prompt type	Count	Purpose
Category	10	Discovery queries for protocol categories
Comparison	10	Head-to-head and multi-protocol comparisons
Branded	10	Protocol-specific named queries
Risk / Security	5	Queries about protocol safety and audits
Developer / Docs	5	Technical and integration queries
Total	40	

Testing conditions

Prompts were submitted to ChatGPT, Perplexity, and Google AI Overviews in May 2026 via API and browser automation from a US-based endpoint. ChatGPT returned no source URLs across any tested prompt. Perplexity citations total 69 real entries after removing 38 Cloudflare CAPTCHA and privacy-footer URLs scraped during automated sessions. Google AI Overviews citations total 947 entries.

Scoring model

Layer	Points	What it measures
AI answer presence	30	Protocol appears by name across ChatGPT (0-10), Perplexity (0-10), and Google AIO (0-10)
Citation quality	25	Official site cited, docs cited, high-intent prompt citations, answer rank position, factual accuracy
Website retrieval readiness	20	H1 clarity, heading structure, extractable facts, internal linking, indexable HTML, FAQ and security page presence
Schema and entity clarity	15	JSON-LD presence, schema types used, sameAs links, validation errors, schema-to-content match
Documentation quality	10	Indexable docs, conceptual and architecture coverage, developer guides, security information, retrieval-ready structure
Total	100	

Limitations

Sample size. Fifty protocols is sufficient to identify patterns but not to make claims about the sector as a whole. The sample skews toward established, well-resourced protocols.

Single time period. All testing was conducted in May 2026. AI platform behaviour changes as models update. Findings reflect conditions at the time of testing.

Automated testing. Prompts were submitted via API and browser automation rather than manual entry. Results may differ from interactive sessions.

ChatGPT source opacity. ChatGPT returned no source URLs. Citation quality for ChatGPT reflects mention data only.

Geographic scope. Testing used a US-based endpoint. Google AI Overviews results vary by geography.

Causation. Statistical patterns reflect correlation within the sample. The study design does not establish causation.

AI presence findings: which protocols appeared and where

Across 40 prompts run on ChatGPT, Perplexity, and Google AI Overviews, 44 of the 50 audited protocols appeared in at least one AI-generated answer on at least one platform. Six protocols were invisible across all three: Pendle, Ethena, Sui, The Graph, Filecoin, and Arweave. Appearing at least once is a low bar. On measures of consistency and citation, the field narrows sharply.

Protocol	ChatGPT prompts	Perplexity prompts	AIO prompts	Score
Ethereum	30	12	4	81.9
Aave	8	4	2	63.7
Chainlink	5	2	3	62.4
Polygon	6	4	1	62.0
Solana	6	4	6	61.8
Arbitrum	6	5	4	58.3
Lido	4	4	2	55.7
Optimism	5	3	2	51.0
MetaMask	3	2	2	49.6
Safe (Gnosis Safe)	1	0	3	48.5

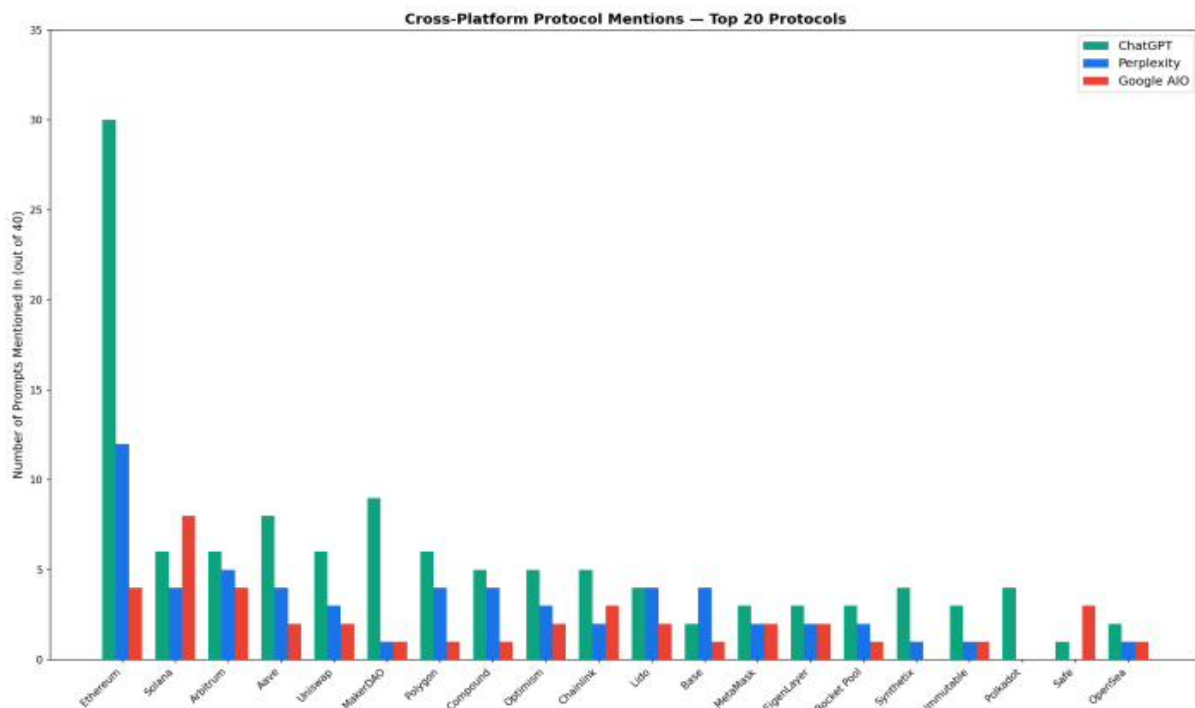


Fig. 3 — Cross-platform protocol mentions: top 20 protocols by prompt count

Ethereum is not competing in the same category as any other protocol

Ethereum appeared in 30 of 40 ChatGPT prompts. The next closest protocol appeared in 9. On Perplexity, Ethereum was mentioned across 12 prompts with 43 total mentions. The nearest competitor had 17. Its total score of 81.9 is 18 points ahead of Aave in second place, reflecting the depth and age of its third-party content record accumulated over a decade. Ethereum functions as a different data type in this study and should not be read as a representative data point.

Schema markup was not a reliable predictor of AI citation visibility

Thirteen of the 50 audited protocols have JSON-LD schema on their homepage. The two with the highest schema quality scores are Starknet (12/15) and Ledger (12/15). **Starknet ranks 15th overall. Ledger ranks 14th. Both sit below ten protocols with no schema at all. Aave has no schema and ranks 2nd. Solana has no schema and ranks 5th.**

Protocol	Schema quality (0-15)	Overall rank	AI presence (0-30)
Starknet	12.0	15th	3
Ledger	12.0	14th	3
Ethereum	10.5	1st	22
Chainlink	10.5	3rd	9
Polygon	10.5	4th	10
Safe	10.5	10th	4
Pendle	9.0	25th	0
Aave	0	2nd	12
Solana	0	5th	15
Lido	0	7th	10

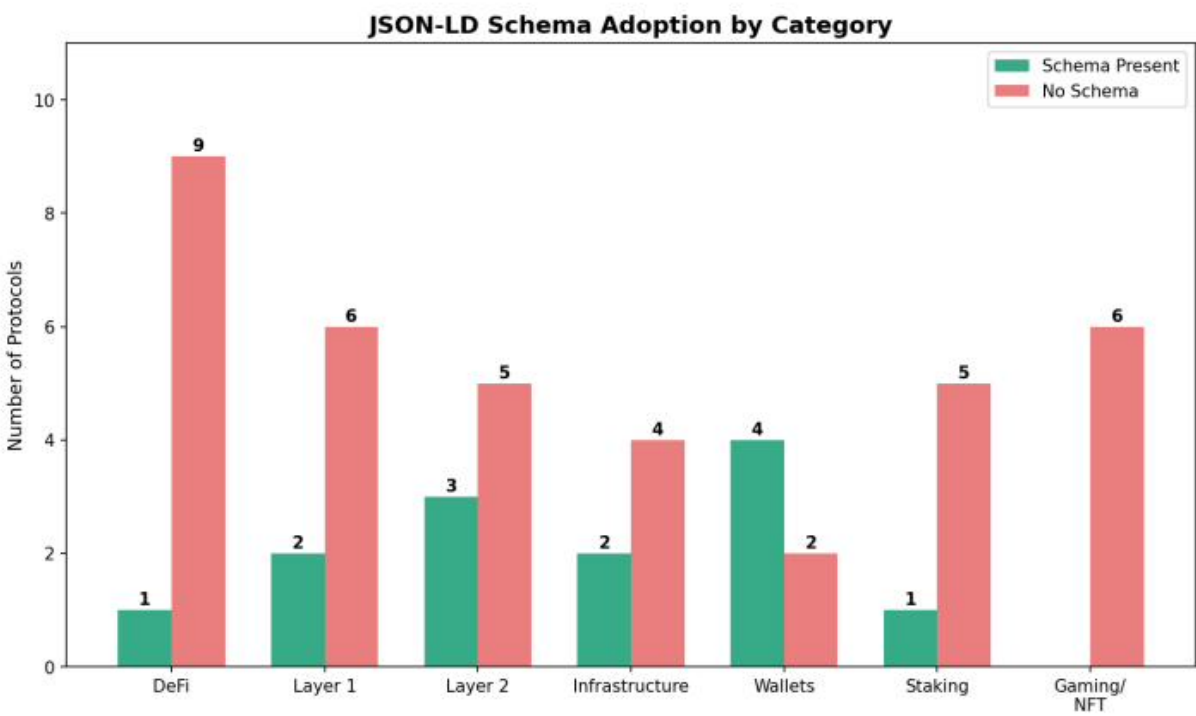


Fig. 4 — JSON-LD schema adoption by category: 74% of audited protocols have no schema

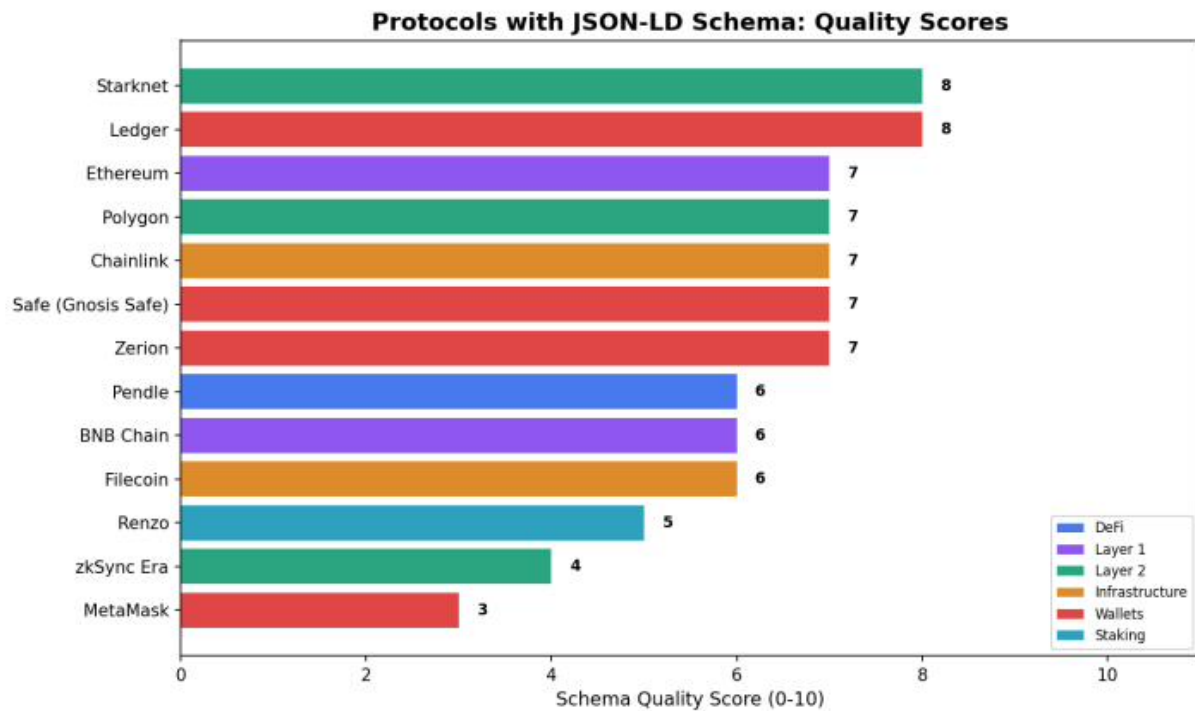


Fig. 5 — Schema quality scores for the 13 protocols with JSON-LD

A note on the correlation between schema and mention rates

Protocols with schema averaged 7.0 AI mentions across the prompt set. Protocols without schema averaged 4.6. That gap does not mean schema caused the higher mention rate. The 13 schema-adopting protocols tend to be among the larger and more established names in the sample. They would likely score higher on AI presence regardless of schema. The Graph is the clearest counter-case: it has JSON-LD schema and registered zero AI presence across all three platforms. **Pendle has schema and scored zero for AI presence while being the only DeFi protocol in the study with any schema at all.** The 7.0 vs 4.6 gap reflects protocol maturity, not schema effectiveness.

Website readiness vs schema quality: four protocol profiles

Plotting readiness scores against schema quality reveals four distinct groups in the sample. The table below describes each quadrant.

Quadrant	Readiness	Schema	Example protocols	What it means
High readiness high schema	>16	>8	Ethereum, Starknet, Chainlink, Polygon	Best structural position. AI presence depends on corpus depth.
High readiness low/no schema	>16	<4	Aave, Solana, Cardano, Lido, Avalanche	Retrievable content. No entity declaration. Largest group in the sample.

Quadrant	Readiness	Schema	Example protocols	What it means
Low readiness some schema	<12	>5	Filecoin, Renzo, Pendle	Structured but harder to find. Schema has not compensated for weak readiness.
Low readiness no schema	<12	0	Immutable, Jito, Arweave, Curve Finance	Invisible tier. No structural foundation. All score below 20 overall.

Protocol pages are losing the citation layer to third parties

When AI systems cited sources in this study, they almost never cited official protocol pages. **Across 1,016 citation records from Perplexity and Google AI Overviews, official protocol pages accounted for 1% of Perplexity citations and 4% of Google AIO citations.** ChatGPT provided no source URLs in any of its 40 responses.



Fig. 6 — Who controls the AI narrative: domain citation breakdown by platform

Source type	Perplexity (69 citations)	Google AIO (947 citations)
Third-party	80%	62%
Exchange	10%	7%
Aggregator	9%	4%
Media	n/a	14%
Dev / Docs	n/a	6%
Official protocol	1%	4%
Educational	n/a	2%
Forum / Video	n/a	5%

The ChatGPT finding deserves attention. It does not mean ChatGPT is not drawing on sources. It means protocol teams have no visibility into which sources are shaping its answers. A response naming Aave as a leading DeFi protocol could be drawing from Aave's own documentation, a CoinGecko aggregator page, a Reddit thread from 2021, or all three simultaneously. There is no way to know from the output.

YouTube and Reddit are the top two cited sources in Google AI Overviews for crypto queries

YouTube was cited 111 times across the AIO prompt set. Reddit was cited 94 times. The third most cited domain was quicknode.com at 26, followed by Wikipedia at 24 and Alchemy at 21. No official protocol page appears in the top 15 cited domains for Google AIO. ethereum.org appears at 11 citations, the closest any protocol's own domain gets.

For a sample of crypto-specific prompts, Google AI Overviews cited YouTube content more than twice as often as it cited official protocol documentation. Reddit at 94 citations reflects the weight Google places on community discussion in this category.

Perplexity's domain profile is different. CoinGecko leads with 6 citations (9% of the Perplexity total), followed by OpenZeppelin and Chainlink's own documentation at 4 each. Exchange learning hubs (BingX and KuCoin) also appear at 4 citations each. Two platforms, two distinct retrieval profiles, neither of which favours official protocol pages.

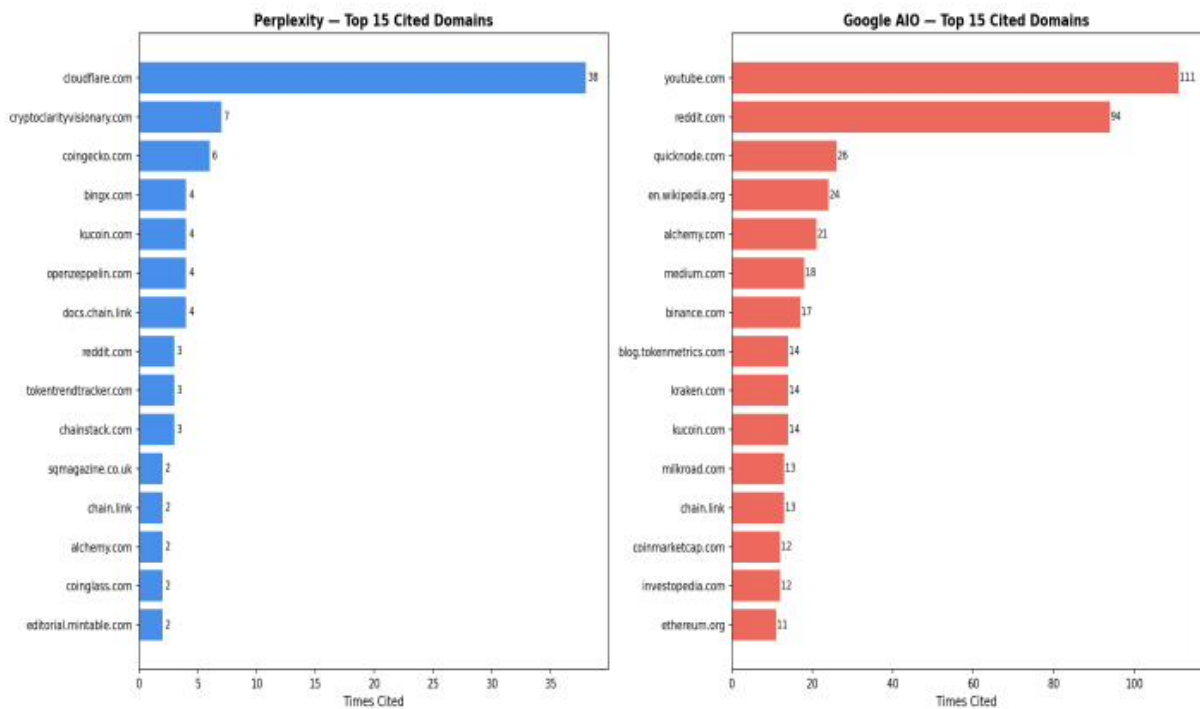


Fig. 7 — Top 15 cited domains: Perplexity (left) and Google AIO (right)

Platform-level differences are significant

The three platforms behave differently enough that cross-platform presence is a more reliable signal than single-platform performance.

ChatGPT produced the broadest mention coverage but provided no citation URLs. **Its mention behaviour appeared to draw heavily on training data, giving older, more established protocols a structural advantage. Protocol teams have no way of auditing which sources ChatGPT uses.**

Perplexity cited sources consistently, drawing primarily from aggregators and developer documentation. Its 69 real citation records produced a clean domain profile: reference sources rather than media or community content.

Google AI Overviews had the lowest mention rate by category across most protocol types. Only 25% of Layer 1 protocols appeared in AIO results. For Infrastructure protocols, the figure was 33%. AIO was more selective and more likely to rely on video and community content when it did surface a protocol.

Safe (Gnosis Safe) produced the sharpest platform divergence in the dataset: zero Perplexity mentions, one ChatGPT prompt appearance, and a Google AIO score of 3/10, placing it 10th overall. Its schema quality (10.5/15) and website readiness (16.2/20) suggest Google's retrieval placed more weight on structured signals for this protocol than Perplexity did.

Category findings: who leads and who lags

Category	ChatGPT rate	Perplexity rate	Google AIO rate
Layer 1	75%	50%	25%
Layer 2	75%	100%	50%
DeFi	60%	70%	40%
Wallets	67%	83%	50%
Staking	50%	100%	50%
Gaming/NFT	67%	100%	33%
Infrastructure	50%	50%	33%

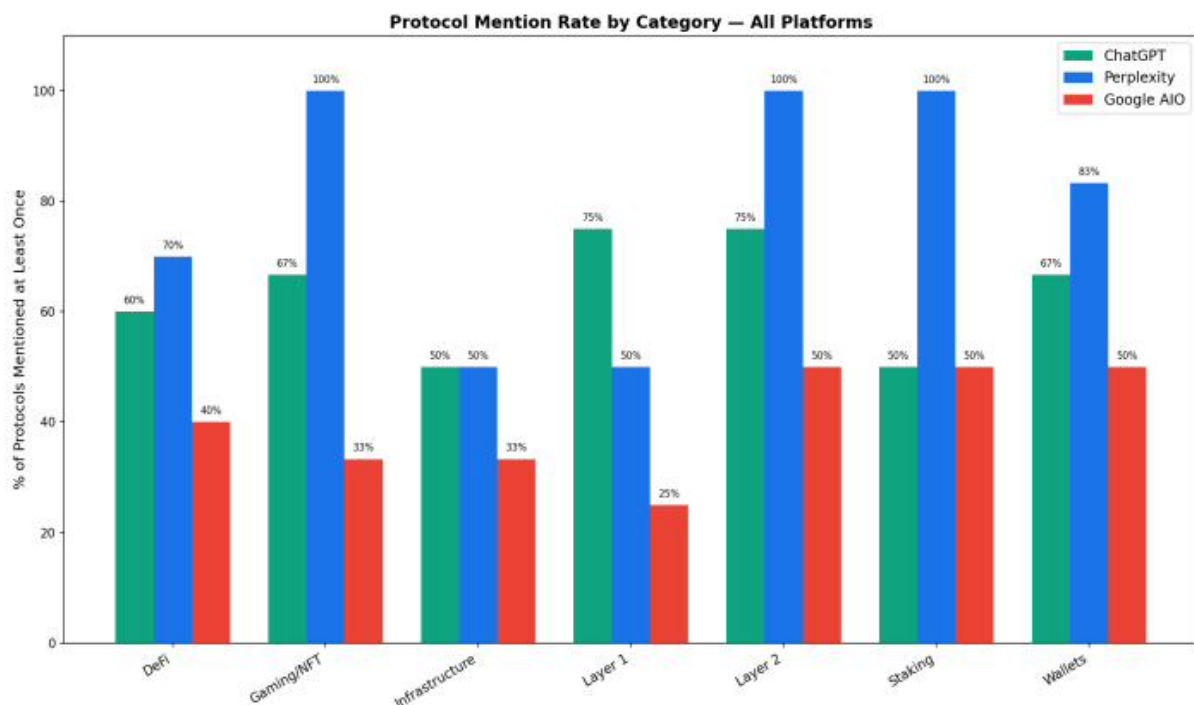


Fig. 8 — Protocol mention rate by category across all platforms

Layer 2 and Staking protocols achieved 100% mention rates on Perplexity. These figures require context: appearing once in a single prompt at position 8 is categorically different from appearing across multiple prompts as a primary recommendation.

Layer 1 protocols had the lowest Google AIO mention rate at 25%, despite having the highest average website readiness score in the study. Ethereum and Solana drove most of the Layer 1 AI presence. Avalanche, Cardano, Polkadot, and Sui collectively contributed very little.

Infrastructure is the weakest category across all three platforms. Chainlink is the only Infrastructure protocol that appears consistently. The Graph, Filecoin, and Arweave registered zero presence across all three platforms despite The Graph scoring 17.5/20 for website readiness.

Strong website readiness, weak AI visibility: the structural gap

Several protocols score highly for website readiness but have minimal AI citation presence. This pattern appears across categories and contradicts the assumption that technical content infrastructure translates directly to AI visibility.

Protocol	Readiness (0-20)	AI presence (0-30)	Overall rank
Cardano	18.8	2	24th
Lido	18.8	10	7th
Aave	18.8	12	2nd
The Graph	17.5	0	33rd
Avalanche	17.5	2	29th
Linea	17.5	1	30th
NEAR Protocol	17.5	3	22nd

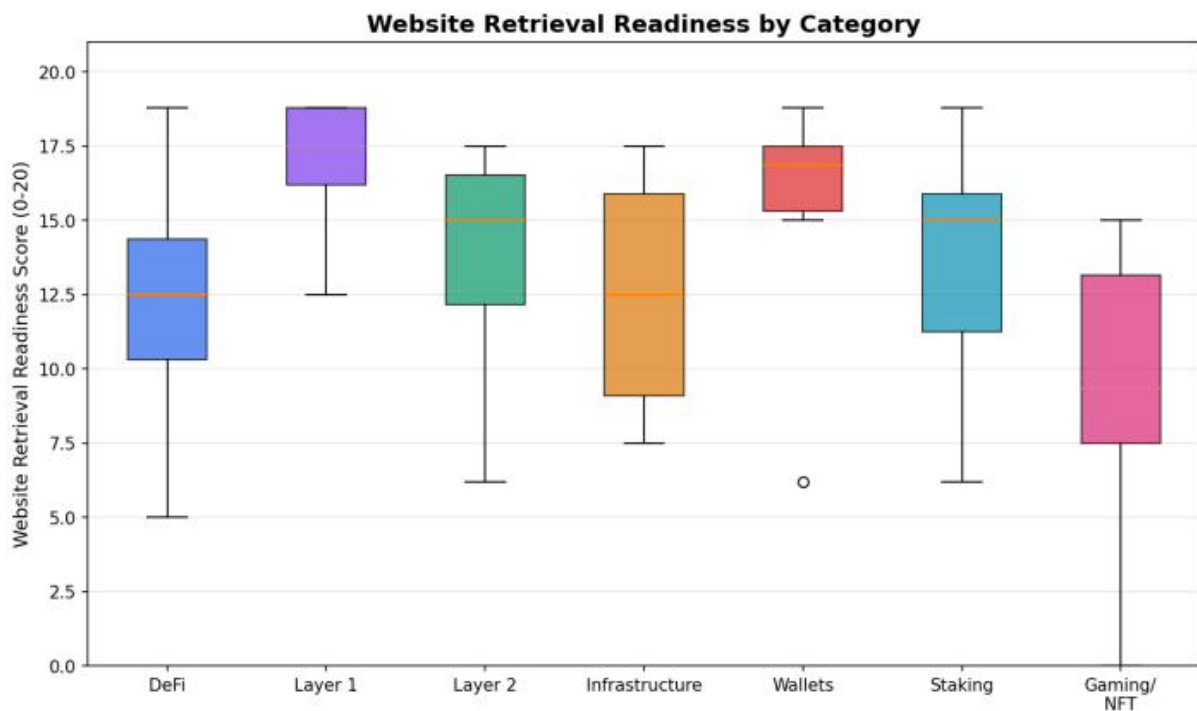


Fig. 9 — Website retrieval readiness by category

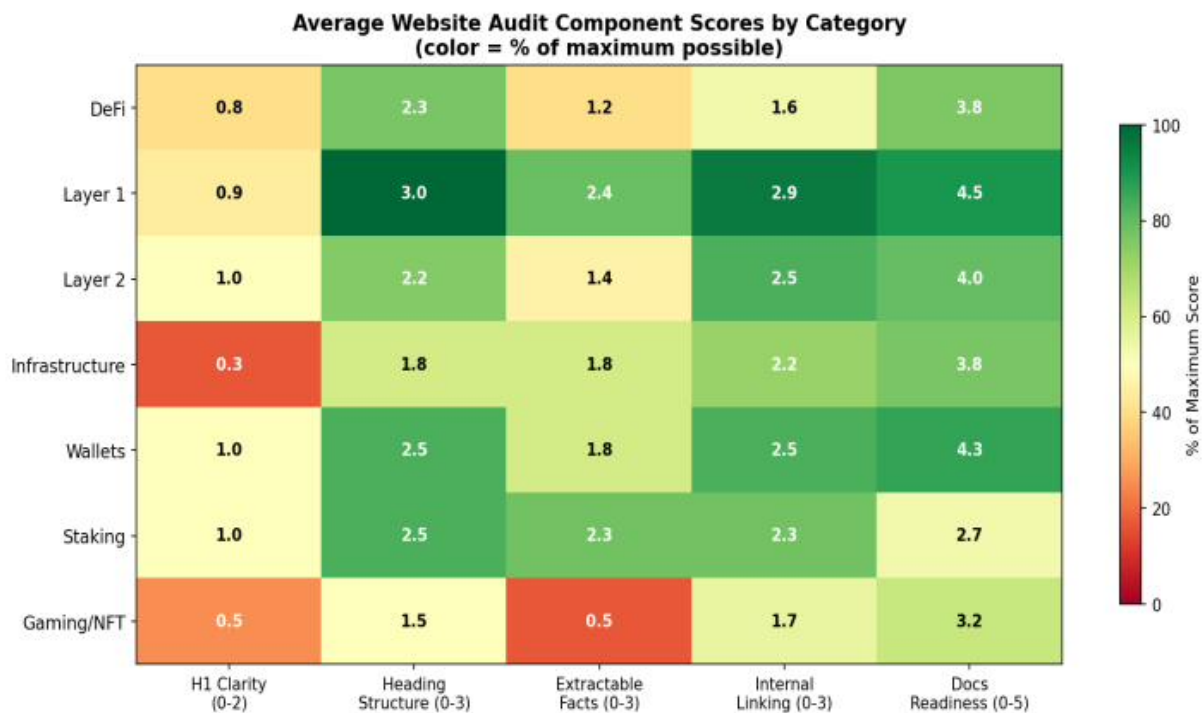


Fig. 10 — Average website audit component scores by category

The Graph scores 17.5/20 for readiness and 7.8/10 for documentation quality. By every structural measure this study applied, it has built a site AI systems should be able to find and read. The problem is not the site. It is the content corpus surrounding the protocol in the wider web.

Full study results: all 50 protocols

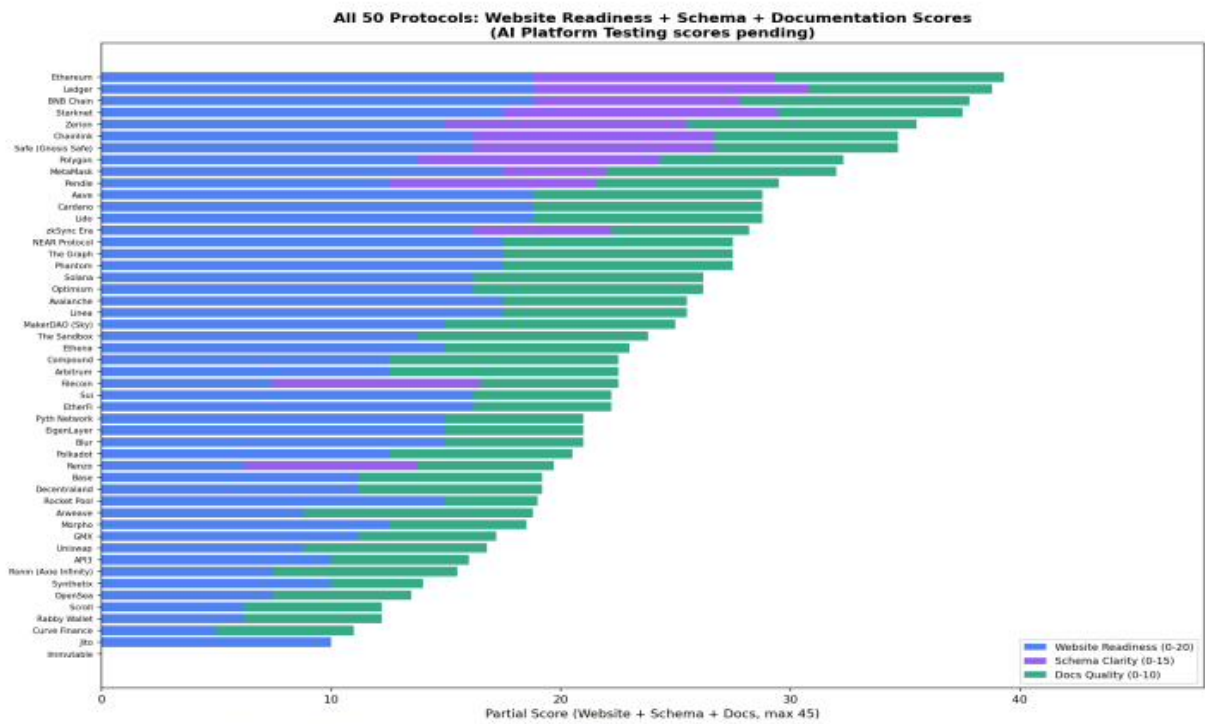


Fig. 11 — All 50 protocols: website readiness, schema, and documentation scores

The six protocols invisible across all three platforms

Protocol	Readiness	Schema	Score	Category
Pendle	12.5	9.0	29.3	DeFi
Ethena	15.0	0	20.6	DeFi
Sui	16.2	0	21.8	Layer 1
The Graph	17.5	0	25.3	Infrastructure
Filecoin	7.5	9.0	23.2	Infrastructure
Arweave	8.8	0	14.4	Infrastructure

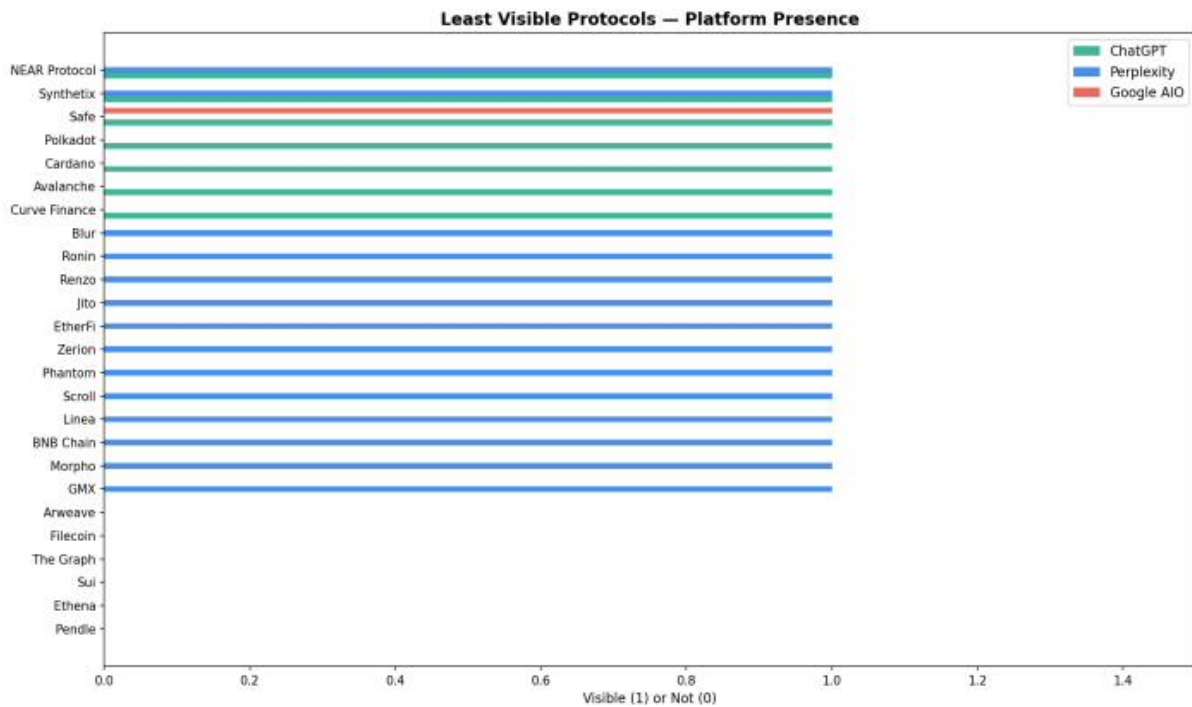


Fig. 12 — Least visible protocols: platform presence across all three AI systems

Sui has a readiness score of 16.2. Ethena scores 15.0. The Graph scores 17.5. None of them appeared in a single AI-generated answer across the full prompt set. Pendle and Filecoin both have schema. Neither appeared anywhere. What these six share is not poor infrastructure but limited third-party content coverage relative to the category-level and comparison queries used in this study.

What the data means

The central finding of this study is not that AI citation visibility is hard to achieve. It is that for most crypto protocols, it is largely unmanaged.

Many protocols in this sample have well-structured sites, indexable documentation, and reasonable schema implementations. What they have not done, in most cases, is build the kind of distributed, third-party content record that AI retrieval systems treat as the primary signal for what a protocol is and why it matters.

The protocols that rank highest share one characteristic more than any other: they have been written about extensively, by people who are not on their payroll, over a long enough period that the coverage has accumulated across formats and platforms. AI systems retrieve from that accumulated record.

What the structural layer actually does

This study does not show schema as a driver of AI citation visibility. A protocol with strong readiness and indexable documentation is easier for AI systems to parse and classify. Beyond that, the data does not show structural scores producing AI visibility in protocols that lack third-party coverage. Cardano scores 18.8 out of 20 for website readiness. Avalanche scores 17.5. Both have near-zero AI presence across all three platforms. The structural layer was not the constraint. The content corpus was.

The third-party source problem is structural, not incidental

Perplexity cited official protocol pages 1% of the time. Google AI Overviews cited them 4% of the time. Protocol teams that assume their official pages are the authoritative source for AI systems are working from an assumption the data does not support. That is not a problem that more schema solves. It requires building the kind of content that third parties cite, share, and link to: original research, developer-focused explainers that third-party tools reference, data that journalists quote.

The 2026 baseline

This is the first year of what will be an annual study. The 2026 findings represent the current state of AI citation visibility in crypto: a sector where a small number of well-established protocols dominate AI-generated answers, where official pages are rarely cited, and where the gap between structural readiness and actual citation is wide enough to suggest that most teams are solving the wrong problem. The study will be repeated in 2027 with the same scoring model and prompt set.

What is AI citation visibility?

AI citation visibility is whether a protocol appears in AI-generated answers in ChatGPT responses, Perplexity citations, or Google AI Overviews when users ask category-level or comparison questions. This study measures it using a repeatable 40-prompt set across three platforms, distinguishing between appearing by name and being cited as a source.

Does schema markup improve AI citations?

The data in this study does not support the claim that schema markup is a reliable path to AI citation. Protocols with the highest schema quality scores are not the most cited. Pendle, the only DeFi protocol with JSON-LD schema, scored zero for AI presence across all three platforms. There is a correlation between schema adoption and higher average mention rates, but it is explained by protocol maturity rather than schema effectiveness. The more defensible claim for schema is entity resolution: helping AI systems confirm what a protocol is and how it relates to other entities.

Can protocols improve their AI citation visibility?

Yes, but the lever is not primarily technical. The protocols that appear most consistently in AI-generated answers have built a distributed content record across third-party platforms: developer documentation that other tools reference, research that journalists cite, explanations that appear on aggregators and in forum discussions. Structural improvements to the site are worth making. They are preparation, not the result.

How will the study be updated?

The study will be repeated annually using the same 50-protocol sample, the same scoring model, and the same 40-prompt set. Annual repetition tracks whether protocol-level investments in content and citation infrastructure produce measurable changes in AI visibility. The 2026 edition establishes the baseline.

CONCLUSION

AI citation visibility in crypto is not a solved problem for most protocols, and it is not going to be solved by the current generation of GEO advice. Structured data, page speed, heading structure, and schema validation are hygiene, not strategy. The protocols that appear in AI-generated answers have earned the source layer that makes citation possible: years of third-party coverage, developer adoption, media attention, and community discussion that AI systems can retrieve and synthesise. That layer is built through content other people find worth citing, not through content the protocol publishes about itself. Most protocol teams have spent more time on the latter. The data shows it.

About this study

Published by David Wood, CryptoContent.dev. ORCID: 0009-0007-3353-7188. Related working paper: Zenodo DOI 10.5281/zenodo.19253709. Full dataset, scoring formulas, and methodology available at cryptocontent.dev/ai-citation-visibility-study-crypto.html.

CryptoContent.dev provides content strategy, GEO, and link-building services for crypto and Web3 projects.