MCP SERVER | NO CODE | CLOUD HOSTED

Semantic Scholar MCP

Map Research Influence Across Millions of Papers

Semantic Scholar MCP connects your AI agent to a knowledge graph containing over 200 million academic papers. Instantly search across all STEM fields, getting single-sentence summaries of complex research. Track how ideas evolve by finding influential citations and building detailed researcher profiles with metrics like the h-index.

A+ Quality Score 100/100

academic-research | citation-analysis | knowledge-graph | literature-review | scientific-data | research-profiles

MAINTAINED BY VINKIUS CLOUD FOR AI AGENTS · JUNE 2026
vinkius.com/mcp/semantic-scholar

SCAN TO VIEW

The connectivity layer between AI and the world's software.

Vinkius sits between AI and every application. All communication passes through Vinkius Cloud via the Model Context Protocol (MCP) — with governance, observability, and security at every layer.

Your AI Connections Run Through Vinkius Cloud

The world's largest managed MCP catalog

Vinkius is the connectivity layer where AI connects to the software your business already runs. We handle the hosting, the security, the credentials, the uptime — you get agents that actually do things.

We operate the world's largest managed MCP catalog. Major SaaS platforms, CRMs, databases, and cloud providers — running, monitored, production-ready. This MCP server is hosted and maintained by the Vinkius Cloud for AI Agents.

> *The agent doesn't manage credentials, doesn't manage uptime, doesn't manage security. Vinkius does.*
>
> — Architecture principle

Four Pillars of the Vinkius Runtime

01 — Security by design

Credentials stay encrypted at rest via AES-256. The AI agent never touches raw keys — they're injected into a sandboxed V8 isolate at runtime. Actions are logged, and connections have an emergency kill switch.

02 — Built on MCP Fusion

This MCP server was built with **MCP Fusion**, the open-source framework (Apache 2.0) that powers the entire Vinkius catalog. Schema-as-firewall strips undeclared fields, compiled PII redaction runs at zero overhead, and cryptographic lockfiles produce git-diffable audit trails.

03 — Deterministic observability

Eight immutable metrics per endpoint: request volume, p95 latency, error rate, active connections, cost attribution. A live payload feed logs every tool call with mutation detection.

04 — Autonomous operations

Servers are deployed, monitored, and patched autonomously. New capabilities and security patches ship weekly. Zero-downtime deployments ensure continuous availability across all managed MCP servers.

AES-256	Ed25519	24h TTL	V8 Isolate
Encryption at rest	PKI vault signatures	Ephemeral session keys	Sandboxed execution

One Token. Instant Access.

Every MCP server on Vinkius is accessed through a **Connection Token**. Tokens are generated in the cloud dashboard and produce a unique MCP endpoint URL. Paste this URL into any MCP-compatible client — no SDK required.

A single token can serve **multiple AI clients simultaneously**, or you can issue separate tokens per client for granular access control. Each token tracks its own request count, last activity timestamp, and can be individually enabled or revoked.

MCP ENDPOINT

```
https://edge.vinkius.com/{token}/mcp
```

Claude · Cursor · VS Code · Windsurf · Grok · Gemini

Security Is the Architecture

Security in Vinkius is not a feature — it's the foundation of the runtime. The gateway enforces multiple independent protection layers between AI agents and third-party APIs.

01 — Ed25519 PKI Vault

Every workspace has an Ed25519 Master Key. Session keys are generated ephemerally (24h TTL) and signed by the Master Key. Credentials never leave the vault boundary.

02 — V8 Isolate Sandboxing

Tool code runs inside isolated-vm V8 isolates with 64 MB memory caps and per-request timeouts. No filesystem access, no network access except through the SSRF-guarded fetch bridge.

03 — SSRF Guard

All outbound HTTP requests are DNS-resolved and validated before execution. Private IP ranges (10.x, 172.16-31.x, 192.168.x, AWS metadata 169.254.x) are blocked at the network layer.

04 — DLP & PII Redaction

A ResponseGuard pipeline intercepts every tool response. Configurable redaction patterns strip sensitive fields (emails, SSNs, card numbers) before data reaches the AI agent.

05 — Cryptographic Audit Trail

Every request is signed into a SHA-256 hash chain with Ed25519 signatures. Events form a tamper-proof, SIEM-exportable forensic record.

06 — Honeytoken Trap System

Phantom credentials are injected into isolated environments. If a honeytoken is used outside Vinkius infrastructure, the server is quarantined instantly.

Emergency Kill Switch

EU AI Act Art. 14(1) Compliant

The kill switch is an **emergency halt** mechanism — not a simple toggle. When triggered, it executes three actions atomically:

01 — Server deactivated

The MCP server is immediately taken offline across the entire cluster.

02 — All tokens revoked

Every connection token is invalidated. Total lockout — reconnection blocked until new tokens are issued.

03 — WebSocket connections killed

Active connections terminated via Redis pubsub broadcast. Propagates to every runtime node in the cluster.

Full Visibility. Zero Guesswork.

The Vinkius cloud dashboard includes a full MCP Governance suite — real-time analytics and security controls for production AI operations.

Control Plane

KPI dashboard with request volume, latency, success rate, token consumption, and AI-generated operational briefings.

Agent Activity

Which AI clients are connecting, how often, and what they're doing — real-time session tracking.

FinOps

Cost tracking per tool, payload compression savings, budget optimization signals, and consumption trends.

Tool Health

Slowest and most error-prone tools, with actionable root-cause insights and performance baselines.

Firewall & DLP

PII redaction activity, sensitive data protection counters, and security event timeline.

Incident Log

Error trends, failure rates, status-code breakdowns, and forensic audit trail access.

> Get started at **cloud.vinkius.com** — connect your AI agent in under 60 seconds.

Semantic Scholar MCP

4 tools available
Cloud-hosted on Vinkius

Need to dive into deep academic literature? This MCP gives your AI client access to the world's largest knowledge graph of scholarly papers. You can search across 200M+ works, getting instant, single-sentence summaries that distill the core insight from any paper—a massive time saver compared to reading abstracts.

It goes beyond simple searches. You track true influence by seeing which citations meaningfully build upon a work, not just how many people mentioned it. Whether you're analyzing career trajectories or building a literature review, you can look up papers using their DOI, ArXiv ID, or PubMed ID. The Vinkius catalog makes this data available to any compatible client, letting your agent treat academic research like structured data. You find researchers by name and immediately pull up metrics, including paper counts and the h-index, giving you a clear picture of who's leading in their field.

Core Capabilities

01 — Search vast academic databases

Find papers across 200M+ works using AI-generated summaries and filtering by fields like Computer Science or Medicine.

02 — Analyze research influence chains

Determine which other papers cite a specific work, allowing you to trace how an academic idea evolved over time.

03 — Retrieve full paper details by ID

Get complete metadata for any article using its Semantic Scholar ID, DOI, ArXiv ID, or PMID.

04 — Build researcher profiles and metrics

Look up academics to see their total publication count, citation history, and h-index score.

One Click on Vinkius — From Prompt to Execution

Available at **vinkius.com/mcp/semantic-scholar** — connect your AI agent in three steps.

01 Connect your AI client to this MCP in the Vinkius Marketplace.

02 Instruct your agent what you're looking for: a specific topic, an author, or an existing paper ID.

03 Your agent executes the necessary tool call and returns structured data, including summaries and related metrics.

The bottom line is that this MCP turns massive, unstructured academic archives into actionable, queryable datasets.

Built For

This connector serves researchers who spend their days sifting through citation lists and trying to map the evolution of an idea. It's for the PhD student stuck in literature review hell or the R&D scientist needing to quickly assess a competitor's academic standing.

AI/ML Researcher

Needs instant, single-sentence summaries of papers on new architectures (like LoRA) to immediately judge if it's relevant before downloading the PDF.

Graduate Student

Uses the citation graph to map out a complex literature review, showing how one core idea branched into three different sub-fields over two decades.

Research Director

Evaluates potential collaborators or employees by running metrics checks on their academic profiles, looking at h-index and influential citation counts.

What Changes When You Connect

01 Skip the abstract reading. Instead of wasting time skimming lengthy paper abstracts, use the main search function to get an AI-generated TLDR summary for every result immediately.

02 Track true academic impact. Don't just count citations; analyze influential citations to see which works truly advanced a field, giving you a deeper understanding of research importance.

03 Eliminate ID guesswork. Never manually look up details again. You can fetch any paper using its DOI, ArXiv ID, or PMID with the get_semantic_paper tool.

04 Build talent profiles instantly. Instead of piecing together career data from different university sites, use search_semantic_author to pull a comprehensive h-index and citation count for any scholar.

05 Map idea evolution. Use the get_semantic_citations tool to build a precise map of how a core concept was cited and refined by subsequent works.

Real-World Applications

Mapping an evolving technology field

A researcher wants to see how 'Transformer' architectures evolved. They ask their agent to use get_semantic_citations on the foundational paper, allowing them to trace every subsequent development and key breakthrough in the field.

Evaluating a potential collaborator

A hiring manager needs to vet a candidate's research depth. They ask their agent to run search_semantic_author on the candidate's name, instantly retrieving their h-index and total citation count for comparison.

Deep literature review prep

A PhD student needs papers on quantum computing using various identifiers. They use get_semantic_paper, feeding it a mix of DOI and ArXiv IDs to ensure they capture every relevant article in the initial search.

Finding niche scientific breakthroughs

An R&D team needs papers on rare diseases. They run search_semantic_scholar, filtering by Medicine or Biology, getting AI summaries that quickly surface only the most promising, relevant studies.

Patterns to Avoid

Treating it like a general web search

✗ AVOID

Asking your agent to 'Find me papers on deep learning' without specifying criteria results in hundreds of irrelevant links and generic abstracts.

✓ INSTEAD

Instead, use the main search function to narrow down by field (Computer Science) and then immediately filter those results using AI-generated summaries for quick relevance checks.

Forgetting paper identifiers

✗ AVOID

Trying to manually cross-reference papers found on Google Scholar with a specific DOI or ArXiv ID.

✓ INSTEAD

Always use get_semantic_paper. Feed it the DOI, PMID, or ArXiv ID directly; this bypasses manual lookup and guarantees the correct metadata.

Stopping after finding one paper

✗ AVOID

Finding a highly cited paper but stopping there without understanding its impact on later work.

✓ INSTEAD

Immediately run get_semantic_citations. This tells you who built upon that core research, giving you the complete academic context.

The Right Fit

Use this MCP if your workflow requires analyzing structured knowledge from academia: mapping citation networks, comparing researcher metrics (h-index), or summarizing highly technical content across 200M+ papers. If you're doing that, the tools like get_semantic_citations and search_semantic_author are essential. Don't use this if your goal is general fact retrieval—if you just need to know 'what is inflation,' a standard knowledge base tool works better. Also, don't rely on it for pre-print data only available outside the indexed IDs; always prioritize using get_semantic_paper with all available identifiers (DOI, ArXiv, PMID) to maximize coverage.

The Pain of Manual Literature Review

Today, researching a topic means opening Google Scholar, clicking through dozens of abstract pages, and copying key details into a spreadsheet. You spend hours downloading PDFs just to read the summary sections, manually tracking who cited what, and piecing together author metrics from multiple academic websites.

With this MCP, you ask your agent for a research overview. It pulls data across 200M+ papers, instantly summarizing key findings into single sentences. You get structured citation maps and full researcher profiles without ever touching an abstract page.

Semantic Scholar MCP: Structured Academic Data

The biggest time drain is the manual cross-referencing of identifiers, or trying to map a concept's development across different databases. You waste minutes just confirming if an author’s work cited by one source was also tracked under a different ID.

This MCP handles all that complexity behind the scenes. By using tools like get_semantic_paper and search_semantic_author, your agent treats every paper and researcher profile as clean, structured data—ready for immediate analysis.

Semantic Scholar: 4 Tools for Deep Research

These tools allow your agent to search millions of academic records, retrieve full paper metadata by ID, profile researchers, and map out complex citation relationships.

#	TOOL	DESCRIPTION
01	`get_semantic_citations`	Finds all the papers that cite a specific work for impact analysis.
02	`get_semantic_paper`	Retrieves full paper details using its Semantic Scholar ID, DOI, ArXiv ID, or PMID.
03	`search_semantic_author`	Returns a researcher's metrics, including their total papers and h-index score.
04	`search_semantic_scholar`	Searches 200M+ academic papers for summaries, citation counts, and fields of study.

See It in Action

Real prompts you can use once this MCP is connected to your AI agent through Vinkius Cloud.

U What are the most cited papers on transformer architecture in deep learning?

Top results include: 'Attention Is All You Need' by Vaswani et al. (2017) — 🤖 TLDR: The Transformer, a new network architecture based solely on attention mechanisms, achieves superior quality on machine translation tasks. Cited 120,000+ times (25,000+ influential). Fields: Computer Science, Linguistics.

U Get the full details of the LoRA paper using its ArXiv ID arXiv:2106.09685.

Found: 'LoRA: Low-Rank Adaptation of Large Language Models' by Edward Hu et al. (2021). 🤖 TLDR: LoRA freezes the pre-trained model weights and injects trainable rank decomposition matrices, reducing trainable parameters by 10,000x and GPU memory by 3x. Citations: 8,500+ (2,100 influential). Fields: CS, Mathematics. Available at arxiv.org/abs/2106.09685.

U Find the researcher Yann LeCun and show me his publication metrics.

Found researcher Yann LeCun — Author at Meta AI / NYU. Papers: 950+, Total Citations: 380,000+, h-index: 162. Known for pioneering work in convolutional neural networks, self-supervised learning, and modern AI architectures. Profile link and full publication list available.

Frequently Asked Questions

01 **How do I use semantic-scholar to find a specific paper?**

You can use get_semantic_paper. You don't need the title; just provide any unique identifier like its DOI, ArXiv ID (e.g., arXiv:2106.09685), or PMID.

02 **Can semantic-scholar track which papers are most influential?**

Yes. The main search function provides 'influential citation counts.' This is better than a simple count because it measures how meaningfully a paper builds on prior work.

03 **What data does the search_semantic_author tool provide?**

It gives you a clear profile of an academic, including their total number of papers published, their overall citation count, and their career impact metric (h-index).

04 **Is semantic-scholar useful for general topics or just AI/ML?**

It covers all STEM fields—Computer Science, Medicine, Biology, Physics, and more. While it's strong in AI/ML, its scope is vast.

05 **Can I find out what papers cited an old classic paper using get_semantic_citations?**

Absolutely. You provide the ID of the 'classic paper,' and the tool returns a list of all newer works that have built upon it, forming an influence chain.

Go Live in 60 Seconds

Get your connection token from **cloud.vinkius.com**, then paste the endpoint URL into any MCP-compatible client.

YOUR MCP ENDPOINT

```
https://edge.vinkius.com/[TOKEN]/mcp
```

CLIENT	WHERE TO CONFIGURE
Claude AI	Profile → Customize → Connectors → "+" → Add custom connector → Paste endpoint
Cursor	Settings → Features → MCP Servers → "+ Add New MCP Server" → Type: SSE → Paste endpoint
VS Code	Ctrl/Cmd+Shift+P → "MCP: Add Server" → add `"semantic-scholar": { "url": "…" }`
Windsurf	MCP Settings → `mcp_settings.json` → Add endpoint URL
ChatGPT	Settings → Tools & plugins → Add MCP server → Paste endpoint
Gemini	Extensions → Add MCP Server → Paste endpoint URL

ASK AN AI ABOUT THIS

Let your preferred AI explain this MCP server

- Ask ChatGPT
- Ask Claude
- Ask Perplexity
- Ask Gemini
- Ask Grok

READY TO CONNECT

Semantic Scholar is live on Vinkius Cloud.

Get your connection token, paste it into your AI agent, and start building. No SDK. No deployment. Just results.

Start at cloud.vinkius.com →

vinkius.com · support@vinkius.com

INDEPENDENT PLATFORM DISCLAIMER

Vinkius is an independent platform and is not affiliated with, endorsed by, sponsored by, verified by, or otherwise authorized by Semantic Scholar. All third-party trademarks, logos, and brand names are the property of their respective owners. Their use in this document is strictly for informational purposes to identify service compatibility and interoperability.

DOCUMENT INFORMATION

Generated	June 2026
MCP Server	Semantic Scholar MCP
Server ID	`019d7605-903c-7097-8d7c-42da50e68b9c`
Platform	Vinkius Cloud for AI Agents
Endpoint	`https://edge.vinkius.com/{token}/mcp`

LICENSE & USAGE

This document is generated automatically by the Vinkius PDF Engine. Content reflects the MCP server configuration at the time of generation and may change as updates are deployed. For the most current information, visit vinkius.com/mcp/semantic-scholar.