codebase-memory-mcp: Map a Whole Codebase So AI Agents Use 99% Fewer Tokens

Kabui, Charles

2026-06-29

Read at ToKnow.ai

DeusData released codebase-memory-mcp, an open-source tool that indexes a whole repository into a knowledge graph: a map of how functions, files, and packages connect, so the agent queries that map instead of reading files one by one. It plugs into coding assistants through

MCP, the open standard that lets a model call outside tools, and ships as a single binary with no Docker, no API key, and no setup. It reads 158 languages, indexed the 28-million-line Linux kernel in 3 minutes, and answers most lookups in under a millisecond. In a paper testing 31 real repositories, it reached 83% answer quality against 92% for an agent reading files, at 10 times fewer tokens and 2.1 times fewer tool calls.

Tokens are the hidden bill on every agent task, and crawling a repo is where most get spent. One graph query replaces dozens of search-and-read cycles: on a five-question test, that cut usage from 412,000 tokens to 3,400, a 99% drop. The tool runs fully on your machine, stores nothing, and installs into 11 agents like Claude Code and Codex with one command. The trade is honest, a little less accuracy for a fraction of the cost.

Agents are shifting from reading code like prose to querying it like a database. We previously covered CodeGraph, which trimmed tokens 47% the same way; this pushes the idea into one small, language-agnostic binary that any tool can share.

Sources:

- DeusData/codebase-memory-mcp on GitHub
- Codebase-Memory: Tree-Sitter-Based Knowledge Graphs for LLM Code Exploration via MCP (arXiv)
- Paper HTML version
- Model Context Protocol
