

Benchbook: A Personal Wiki Under Contract

Last modified: 2026-09-13 00:11:34 (PDT)

[benchbook](#) is a plain-markdown, git-versioned personal wiki that an AI agent reads and writes under a written contract. Your AI keeps the wiki; you keep the rules.

The problem it targets is not the code you wrote on a Tuesday, but the reasoning you lost by Thursday: which stack you chose and why, which two approaches you rejected, and which of four config files is live. A README captures what was built, git captures when, and almost nothing captures why. Benchbook captures all three as a side effect of working, because the same agent that helps you build also files the record.

It instantiates Andrej Karpathy's [LLM Wiki](#) pattern (three layers: immutable raw sources, an LLM-owned wiki, and a co-evolved schema file; three operations: ingest, query, lint) and fills in the specifics the gist leaves open after six months of daily use at about 1,563 pages across nine domains.

The repo's value is less in the page count than in the rules that survived contact with real use:

- **update** was removed as a valid log operation after 17 of 26 recent entries were verbose updates duplicating project-page text.
- A central todo file that mirrored project todos was replaced by pointers after 60 of ~76 mirrored items drifted.
- A “split pages over ~500 words” rule was made measurable after a quarter of pages breached it.

Concretely, the contract (`agents-core.md`) is read at every session start, domains are first-class (with packs for `knowledge`, `home`, `projects`, and others), page types (`source`, `entity`, `reference`, `project`) carry required front matter, indexes are per-domain catalogues read before answering (no vector store), and skills package the recurring operations (ingest, query, lint, and more). The agent acts as Librarian, Advisor, and Project Manager against that structure.

The site walks the pattern through three ordinary tasks (ESP32 build, Telegram bot, FPV drone learning path) in [A Day in the Life](#), and is honest about limits and about the anti-bloat discipline that keeps a cheap-to-maintain wiki from growing faster than anyone can read.