

# BASE

## Uma base de conhecimento viva para gerir melhor a sua saúde

*O pattern do LLM Wiki (Andrej Karpathy) aplicado à saúde pessoal — um playbook*

Este playbook tem três partes. A **Parte I** explica o padrão — de onde veio, por que funciona e como ele transforma uma pilha de exames soltos numa base que pensa com você. A **Parte II**, o coração, mostra como aplicá-lo à *sua saúde*: o que subir e como a base te entrevista como um médico. A **Parte III** é mão na massa — o passo a passo de ~20 minutos, como tirar seus dados dos laboratórios e wearables, e o que perguntar. Este documento é o guia humano; o prompt exato que você cola no agente vem no arquivo `health-vault-bootstrap-prompt.md`.

### PARTE I

## O padrão

### Por que este documento existe

Todo mundo que usa IA a sério já sentiu a mesma frustração: você passa horas numa análise densa com um modelo e, no dia seguinte, numa sessão nova, é como se nada daquilo tivesse existido. O modelo foi treinado uma vez e, a partir dali, não aprende mais nada com você. Cada conversa recomeça do zero.

A forma mais comum de “dar memória” a um modelo — subir um monte de arquivos e deixar ele buscar trechos na hora da pergunta (o que se chama RAG) — ameniza, mas não resolve: ele continua redescobrimdo o assunto a cada pergunta. Nada se acumula. Faça uma pergunta sutil, que exija cruzar cinco exames, e ele remonta os fragmentos de novo, toda vez. Não há construção.

### A virada: o LLM Wiki de Karpathy

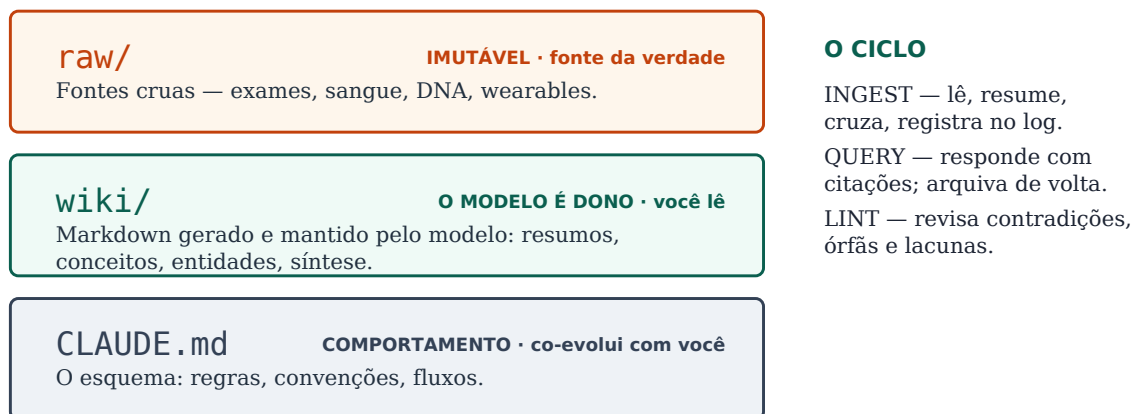
Em abril de 2026, Andrej Karpathy — membro fundador da OpenAI, ex-líder de IA da Tesla e, desde maio de 2026, na equipe de pré-treino da Anthropic — publicou um *gist* curto chamado “**LLM Wiki**”: um padrão para construir bases de conhecimento pessoais com LLMs, feito para ser copiado e colado no seu próprio agente.

A proposta central: em vez de só recuperar trechos dos documentos crus na hora da pergunta, **o modelo constrói e mantém incrementalmente uma wiki persistente** — uma coleção de arquivos markdown interligados que fica *entre você e as fontes*. Quando você adiciona uma fonte, ele a lê, extrai o essencial e **integra ao que já existe**: atualiza páginas, revisa resumos, anota onde o dado novo contradiz um antigo. O conhecimento é *compilado uma vez e mantido atual*, não re-derivado a cada pergunta. E — ponto crucial — **você quase nunca escreve a wiki; o modelo escreve**. Você traz fontes e faz boas perguntas; ele faz o trabalho braçal de resumir, cruzar e manter a consistência.

Karpathy resume com uma analogia que cola: “**O Obsidian é a IDE; o LLM é o programador; a wiki é o código-fonte.**” Ele liga a ideia, no espírito, ao **Memex de Vannevar Bush (1945)** — um repositório pessoal e curado, com trilhas associativas entre documentos. O que Bush não resolvia era *quem faz a manutenção*. A resposta, agora, é o LLM.

## As três camadas

A arquitetura é simples, de três camadas. Entender essa separação é o que mais importa — quase todo o resto decorre dela.



Vale fixar a separação entre a wiki e o esquema: **CLAUDE.md descreve como o agente se comporta; a wiki descreve o que ele sabe**. Regra de comportamento nunca entra na wiki; conhecimento de domínio nunca entra no **CLAUDE.md**.

## As três operações

INGEST	QUERY	LINT
Você joga uma fonte na inbox e pede para processar. O modelo lê, resume, cria/atualiza páginas, cruza relações e registra no log. Uma fonte pode tocar 10-15 páginas.	Você pergunta contra a base. Ele busca, lê e sintetiza com citações. Boas respostas voltam para a wiki como páginas novas — não se perdem no chat.	De tempos em tempos, um exame de saúde da própria wiki: contradições, afirmações superadas, páginas órfãs, conceitos sem página, lacunas.

Dois arquivos costuram tudo: o **index.md**, catálogo que o modelo lê *primeiro* para achar as páginas certas; e o **log.md**, registro cronológico *append-only*. Em escala moderada (~100 fontes, centenas de páginas), o índice já basta — não é preciso montar infraestrutura de embeddings.

## Por que funciona

A parte tediosa de manter uma base não é ler nem pensar — é o **bookkeeping**: atualizar referências cruzadas, manter resumos em dia, preservar consistência entre dezenas de páginas. É por isso que humanos abandonam wikis: o custo de manutenção cresce mais rápido que o valor. **LLMs não se entediam, não esquecem de atualizar uma referência e tocam 15 arquivos numa passada só**. A wiki se mantém viva porque mantê-la custa quase zero.

## Como ler a base: Obsidian

A wiki é uma pasta de arquivos `.md` — você pode abrir em qualquer editor, mas o **Obsidian** foi feito para isso. Ele renderiza os `[[links]]` entre páginas, mostra o **graph view** (o mapa visual de como tudo se conecta — onde estão os *hubs* e as páginas órfãs), embute imagens de exames com `![[arquivo]]` e abre os PDFs num painel lateral. Na prática, você deixa o agente de um lado editando e o Obsidian do outro, navegando em tempo real: seguindo links, lendo as páginas atualizadas, vendo o grafo crescer a cada fonte.

## PARTE II

# Aplicando à sua saúde

Aqui o padrão sai da teoria. Seus dados de saúde estão espalhados — laboratórios, hospitais, o relógio no seu pulso, um teste de DNA, e a sua própria memória de como você tem se sentido. Ninguém, nem um ótimo médico, segura tudo isso junto e cruza ao longo do tempo. A base faz exatamente isso. Esta parte mostra como começar e o que alimentar.

## Comece pela anamnese

Antes de qualquer exame, deixe a base **te entrevistar como num primeiro atendimento**. Esse questionário inicial popula o seu perfil — o documento que todas as outras páginas consultam para comparar valores contra os *seus* alvos. Quanto mais honesta a anamnese, mais afiada a base fica desde o dia 1. Se você declara uma condição, o agente puxa as perguntas de acompanhamento que um clínico faria — ramificando por área:

### A entrevista de admissão — a base te entrevista como numa primeira consulta

<b>Identidade &amp; corpo</b>	Idade · sexo (opcional) · altura · peso (→ IMC) · cidade.
<b>Condições conhecidas</b>	Liste os diagnósticos atuais. Para cada um, a base conduz a anamnese dirigida — as perguntas que um clínico faria numa primeira consulta.
↳ <i>ex.: cardiovascular</i>	Pressão, colesterol/ApoB/Lp(a) ao longo do tempo? Escore de cálcio? Eventos cardíacos na família e em que idade? Sintomas: dor torácica, falta de ar, palpitação?
↳ <i>ex.: metabólico</i>	HbA1c e glicemia de jejum? Circunferência abdominal? Pré-diabetes/diabetes na família?
↳ <i>ex.: musculoesquelético</i>	Onde dói, desde quando, o que piora ou melhora? Já fez imagem? Limita o exercício?
<b>Medicações &amp; suplementos</b>	Nome, dose, desde quando, quem prescreveu.
<b>Histórico familiar</b>	Condições em pais e irmãos, e a idade de início.
<b>Hábitos &amp; wearables</b>	Usa Apple Watch, Whoop, Oura, Garmin? Sono, exercício, álcool, tabaco.
<b>Seus alvos</b>	Tem metas definidas por um médico (alvo de LDL, de pressão)? Elas vêm antes da faixa de referência do laboratório.

Não precisa responder tudo de uma vez. O esqueleto entra agora; o resto a base preenche conforme você sobe exames e conversa com ela.

## O que subir — por tipo de fonte

A regra de ouro é começar pelo *ground truth*: dado real e seu, não opinião. Suba uma fonte por vez e confira o que a base escreveu. Os seis tipos que mais rendem:

Tipo de fonte	O que você sobe	O que a base faz com isso
<b>Exames de imagem</b>	Eco, TC, RM, ultrassom — o laudo em PDF (e a imagem, se tiver).	Transcreve as medidas-chave, monta a série temporal e sinaliza progressão (ex.: um diâmetro que cresce a cada exame).
<b>Exames laboratoriais</b>	Cada painel de sangue em PDF.	Cada coleta vira um resumo datado; cada biomarcador ganha uma página que agrega a série e compara com o SEU alvo antes da referência do lab.
<b>Históricos médicos</b>	Laudos de consulta, diagnósticos, cirurgias, receitas, histórico familiar.	Vira a linha do tempo clínica e alimenta o seu perfil; conecta sintomas, decisões e exames.
<b>Genética / DNA</b>	Export do 23andMe, do genoma (opcional).	Variantes de risco e farmacogenética, sinalizadas para revisar com o médico — nunca como diagnóstico.
<b>Wearables</b>	Export do Apple Health, Whoop, Oura, Garmin.	Resumos por período (semana/mês) de VO2máx, FC de repouso, HRV, sono e volume de treino — não o despejo do arquivo inteiro.
<b>Journal de como você se sente</b>	Uma nota curta: energia, humor, sono, sintomas, efeitos colaterais.	Cruza o subjetivo com o objetivo — liga uma queixa a uma mudança de exame, remédio ou treino.

Para imagem e laboratório, o ganho aparece no **tempo**: a base não guarda um exame isolado, ela monta a *série* — e é na série que mora o sinal (um diâmetro que cresce 1 mm por ano, um marcador que dobra). Por isso cada biomarcador tem uma página só dele, que cada nova coleta atualiza, em vez de virar dezenas de PDFs soltos.

## Transforme suplementos e remédios em experimentos

Tudo que você toma ou faz vira uma **página de intervenção que é a hipótese**: registra o que é, a dose, desde quando, qual marcador deveria mexer, para qual alvo e em quanto tempo. Aí, quando o próximo exame chega, a base confronta o resultado com a previsão. Em vez de “tomo vitamina D porque ouvi falar”, você passa a ter um laço de feedback: *moveu o que eu esperava, no tamanho que eu esperava?* É o que separa um regime de palpites de um experimento pessoal honesto.

## O journal: cruze o que você sente com o que os exames mostram

O dado mais subestimado é o subjetivo. Uma nota curta de “como me sinto hoje” — energia, sono, humor, uma dor nova, um efeito colateral — jogada na base vira combustível de correlação. É o que permite ligar uma queixa de fadiga a uma mudança de

remédio três semanas antes, ou um pico de irritabilidade a uma noite ruim de sono no wearable. Sozinho, o sintoma é anedota; cruzado com a série objetiva, vira pista.

## A disciplina que faz dar pra confiar

Uma base de saúde sem disciplina vira um gerador de pânico — todo exame tem algo “fora da faixa”. O que a torna confiável são algumas regras, embutidas no esquema:

**Quality gate na entrada.** Antes de ingerir, uma pergunta: *isto agrega informação nova de verdade?* Marketing, quase-duplicata e rehash ficam de fora. A base é tão boa quanto a curadoria da porta.

**Dois eixos de confiança.** Separar *quão boa é a fonte* de *quanto a ciência concorda com a afirmação* (estabelecido, emergente, especulativo). É o que blindava você do hype de longevidade: um marcador substituto que mexeu (NAD+, “idade biológica”) gera hipótese, não prova nada — e a base diz isso na cara.

**Disciplina analítica contra falso sinal.** Não afirmar correlação com menos de ~10 observações pareadas; tamanho de efeito importa mais que direção (3% é regressão à média, 30% é sinal); listar confounders (sono, estação, estresse, dieta, horário da coleta, troca de lab) antes de cravar causa; e, por padrão, ficar com a explicação mais chata até descartá-la.

**Descritivo, não prescritivo.** A base descreve o que os dados mostram e sinaliza o que merece levar ao médico — mas não manda você tomar, fazer ou parar nada. Ela te prepara para a consulta; **não substitui o seu médico.**

**Privacidade como regra dura.** O vault é sensível. Nada do conteúdo sai para serviço externo — busca web, API de terceiro, artefato público — sem a sua aprovação explícita, ação a ação.

## O resultado

Com o tempo, a capacidade do modelo vira um superpoder de correlação. Ele conecta assuntos que passam despercebidos e levanta hipóteses que escapam até a excelentes médicos — seja por falta de acesso a *todos* os dados, seja por não fazer o cruzamento. No meu caso, esse cruzamento me deu base para questionar e confirmar um erro num parecer de um especialista de primeira linha, de um dos melhores hospitais do mundo. Não deixei de ir ao médico — a base não substitui ninguém. Ela me faz chegar mais preparado e me dá mais agência sobre a minha saúde. Esse é o ponto.

### PARTE III

## Mãos à obra

As Partes I e II explicaram o porquê e o quê. Esta fecha com o **como** — e com o atrito real que trava a maioria das pessoas. O passo a passo abaixo te coloca rodando em uma tarde.

## Comece em ~20 minutos

Comece em ~20 minutos

- 1** Abra um agente com acesso a arquivos — **Claude (Cowork)** é o mais fácil — e dê a ele uma pasta **vazia** (ex.: `Health/`), de preferência no Dropbox / Drive / iCloud.
- 2** Entregue a ele o **prompt-companheiro** (`health-vault-bootstrap-prompt.md`): anexe o arquivo e diga “*monta meu vault de saúde seguindo isto*”.
- 3** **Responda à anamnese** que ele conduz. Ele cria as pastas, escreve as regras (`CLAUDE.md`) e semeia o seu perfil.
- 4** Jogue **um exame** (um painel de sangue é ótimo) na pasta `raw/inbox/` e diga “*ingere*”. Pronto — a base começou a compor.

Expectativa honesta: o setup leva uns 20 minutos e **um exame já basta** para sentir o valor no dia 1. O resto você adiciona ao longo das semanas — a base compõe.

**Seus dados ficam com você.** Tudo vive na sua pasta, no seu computador ou nuvem. As regras embutidas proíbem o agente de enviar qualquer conteúdo de saúde a serviço externo (busca web, ferramenta de terceiro, link público) sem te perguntar antes — toda vez.

## E os dados, como eu tiro? (o maior atrito do iniciante)

O difícil quase nunca é a IA — é tirar os seus próprios dados de onde eles moram. Atalhos:

Fonte	Como tirar
<b>Exames (lab + imagem)</b>	Portal do laboratório/hospital (ou o e-mail) → botão “baixar PDF”. Só tem papel? Fotografe nítido — o agente faz OCR.
<b>Apple Health (iPhone)</b>	App Saúde → toque na sua foto (canto sup. direito) → “Exportar Todos os Dados de Saúde” → mande o `zip` pro computador.
<b>Whoop / Oura / Garmin</b>	Conta no app ou na web → Exportar dados → você recebe CSVs.
<b>Genética / DNA</b>	23andMe → Configurações → Baixar dados brutos; ou o download do seu provedor de genoma.
<b>Remédios &amp; suplementos</b>	Sem arquivo — é só dizer nome, dose e desde quando durante a anamnese.

## Veja o tipo de resposta que você ganha

Um exemplo concreto vale mais que a promessa. Você pergunta em linguagem natural; a base responde com a série no tempo, a comparação com o *seu* alvo, a disciplina analítica e as fontes:

## VOCÊ PERGUNTA

“Qual é a tendência do meu LDL nos últimos 5 anos? Está indo na direção do meu alvo?”

## A BASE RESPONDE

Série: 130 → 112 → 98 → 71 mg/dL (2021→2026). Cruzou o seu alvo pessoal (<70) depois da troca de medicação em nov/2025. Queda de -45%, maior que o previsto pela regra do statin — candidato a investigar (adesão? super-resposta?). Sem mudança de dieta registrada no período.

Fontes: ldl-cholesterol, painel-2026.

*Exemplo ilustrativo — números fictícios, não são dados de ninguém.*

## O que perguntar (quando já tiver algumas fontes)

**Tendência:** “Qual a evolução do meu LDL / ApoB / [marcador]? Está indo na direção do meu alvo?”

**Cruzamento:** “Cruze meus exames, meus remédios e como tenho me sentido — algo que valha destacar?”

**Pré-consulta:** “Monte um resumo de uma página para a minha próxima consulta, com as perguntas certas para o médico.”

**Causa & efeito:** “Minha FC de repouso / HRV / sono mudou depois que comecei [remédio ou hábito]? Considere os confounders.”

**Contradições:** “Onde meus dados se contradizem — dois laudos com números diferentes?”

O **pre-visit brief** costuma ser o momento em que a ficha cai: você entra na consulta com o histórico inteiro sintetizado e as perguntas certas na mão.

## Onde colar o quê

Este PDF é o **guia humano** — leia, entenda, decida. O arquivo-companheiro `health-vault-bootstrap-prompt.md` é o **prompt exato**: abra, copie tudo e cole no agente apontado para uma pasta vazia; ele faz o resto. Por plataforma: no **Cowork**, crie um Projeto e dê acesso à pasta; no **Claude Code**, rode `claude` dentro da pasta; em outros (Codex/Cursor), mantenha o `CLAUDE.md` na raiz. Depois do dia 1 você não roda o prompt de novo — é só reabrir e conversar; as regras carregam sozinhas.

## Se travar

**Começou a conversar em vez de construir?** Diga: “Siga o prompt — faça a anamnese, depois crie as pastas e os arquivos.”

**Resumiu o CLAUDE.md em vez de escrevê-lo?** Diga: “Reescreva o CLAUDE.md inteiro, verbatim.” É ele que faz o sistema funcionar.

**Não conseguiu ler o PDF do exame?** Provavelmente é um scan: “Faça OCR e transcreva os valores.”

**Te disse o que tomar?** Lembre: “Fique no descritivo — me diga o que os dados mostram e o que levar ao médico, não o que tomar.”

## O que importa mais que a ferramenta

**Comece pela anamnese e pelo ground truth.** Perfil honesto + dados reais seus. O resto cresce em cima.

**Uma fonte por vez, com você no circuito.** O quality gate é o que evita que a base vire um depósito.

**Deixe o modelo escrever; você lê e dirige.** Seu trabalho é trazer fontes, fazer boas perguntas e pensar no significado.

**Seja honesto sobre incerteza.** Distinga consenso de pesquisa emergente e de especulação — e desconfie do que mexe num marcador mas não tem desfecho duro.

**A base compõe — não a polua.** Cada fonte de qualidade a deixa mais afiada; cada ruído admitido a deixa mais burra.

*Uma pasta vazia, um exame, uma pergunta.  
A base cresce a partir daí.*

---

*Crédito: padrão “LLM Wiki” de Andrej Karpathy — [gist público de abril de 2026](https://gist.github.com/karpathy/442a6bf555914893e9891c11519de94f) ([gist.github.com/karpathy/442a6bf555914893e9891c11519de94f](https://gist.github.com/karpathy/442a6bf555914893e9891c11519de94f)). A especialização de saúde — anamnese de admissão, guia de fontes por tipo, journal subjetivo, páginas de intervenção como hipótese, dois eixos de confiança, disciplina analítica e privacidade — é adaptação própria, fruto da prática.*