

kapitola 25

Jak modelu předáš své řemeslo?

Soubor, který nikdo nezkompiluje

Podesáté jsem do chatu lepil tentýž odstavec instrukcí.
Čekal jsem API, klíč, trénování — něco těžkého.

Ono to byl jediný textový soubor: **SKILL.md**. Napsán obyčejnou češtinou.
Nic se nezkompilovalo, nic nenainstalovalo.

Řemeslo jsem předal modelu tím, že jsem ho jednou napsal.

Model si tě vybírá sám

Řemeslník vejde do dílny plné zásuvek. Na každé je **štítek** — jedna věta, co je uvnitř. Dovnitř nevidí; vidí jen štítky.

Když je štítek matný nebo lže, zásuvka zůstane zavřená — i kdyby v ní ležel přesně ten nástroj, který úloha potřebuje.

Ten štítek je pole **description**. **Ne kód. Věta.**

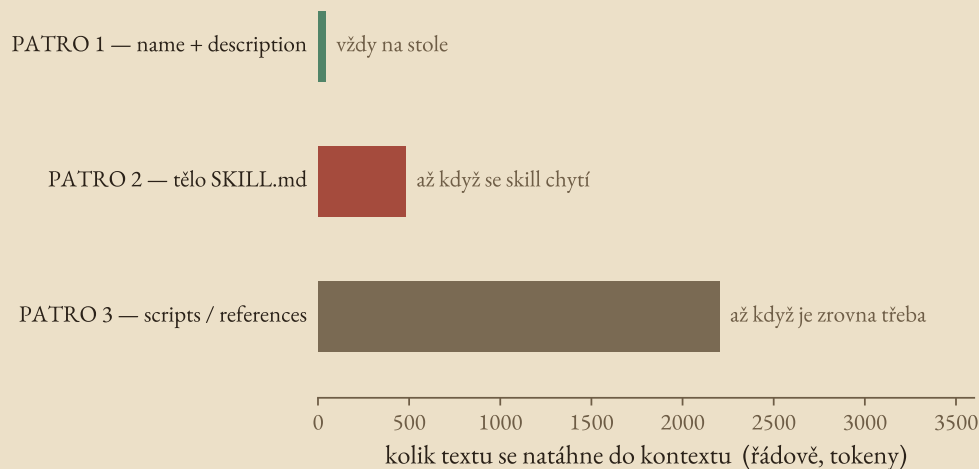
První zákon skillů

- Skill **nikdo nezavolá** — vybere si ho model sám.
- Rozhodne se podle jediné věci, kterou o něm ví: podle **popisu**.
- Sebelepší řemeslo uvnitř leží ve tmě, když je štítek špatný.
- *Skill se nespouští tím, co umí, ale tím, jak je popsáný.*

Tři patra: proč se toho nenačte moc

- Patro 1 — vždy: na stole leží jen name a description. Pár tokenů.
- Patro 2 — když se chytí: natáhne se celé tělo SKILL.md.
- Patro 3 — když je třeba: skripty a reference. Do té doby nic.
- Sto skillů unese stůl líp než deset upovídaných MCP serverů.

Progresivní (líné) načítání



Model se rozhoduje z tenkého proužku nahoře; těžké patro čeká, dokud není třeba.

Spuštění je vyhledávání

Model odhaduje $P(\text{spuštění} = s \mid q, d_s)$ — jak dobře popis d_s sedí na úlohu q .

Popis je index, dotaz je query. Špatný **description** je špatně **zaindexovaný záznam**: existuje, ale nikdo ho nenajde.

A jako každý model se dá **přeučit** — lad' na trénovacích promptech, sud' podle testovacích.

„Jak bych poznal, že se pletu?„

Napiš pět promptů, na které skill *má* naskočit, a pět, na které *nesmí*.
Pusť je. Když chyba není v těle, ale v popisu — první zákon pořád platí.