

XXXI

Koho poslouchat, když se to mění každý týden?

Poslední kapitola knihy. Po ní přestaneš honit každý nový model a začneš vážit zdroje jako předpovědi počasí — track record nad titul, changelog nad thread, číslo nad nadšení. Dostaneš kompas, který funguje i za rok, kdy budou jména nástrojů jiná, a odejdeš s jedinou otázkou, která nezastará nikdy: jak bych poznal, že se pletu?

...druhý konec spektra, kde ostrílení profesionálové urychlují svou práci pomocí modelů, a přitom zůstávají hrdě a sebevědomě zodpovědní za software, který vyrábějí.

— Simon Willison, „Vibe engineering“, simonwillison.net, 7. října 2025 —
definice protipólu k „vibe codingu“

Otevřu tuhle poslední kapitolu scénou, kterou znáš, protože ses v ní za tuhle knihu párkrát ocitl sám. V pondělí sedíš u nástroje, který jsi konečně *zvládl*: znáš jméno modelu, znáš vlajku, kterou se zapíná sandbox, víš, kolik stojí token a jak se jmenuje to tlačítko, co spustí agenta. Cítíš se dobře — a máš na to právo, stálo tě to večery. Ve čtvrtek vyjde nová hlavní verze. Model se jmenuje jinak, vlajka se přejmenovala, cena spadla na třetinu, tlačítko je na jiném místě a půlka tvých promptů, které fungovaly, najednou dělá něco lehce jiného. V pátek na tebe z každého kanálu křičí deset threadů, tři z nich si protiřečí, dva prodávají kurz a jeden tvrdí, že *teď už je programování mrtvé*. A ty stojíš uprostřed toho hluku s jedinou rozumnou otázkou: koho z těch hlasů mám vlastně poslouchat?

Tahle otázka je poslední, na kterou ti kniha odpoví — a je to schválně ona. Celou první část jsi stavěl stěžeň: metodu, která nezastará. Celou druhou jsi napínal plachty: nástroje roku 2026, které zastarají skoro jistě. A právě proto musí kniha skončit tady, u návodu, jak žít v tom rozporu — jak držet krok, aniž tě strhne každá vlna hype, a jak poznat, komu z těch hlasů věřit a o kolik. Dobrá zpráva, kterou ti hned prozradím, protože je jádrem všeho ostatního: nemáš honit nástroje. Máš vážit zdroje. A vážit zdroje už umíš — je to přesně kalibrace z

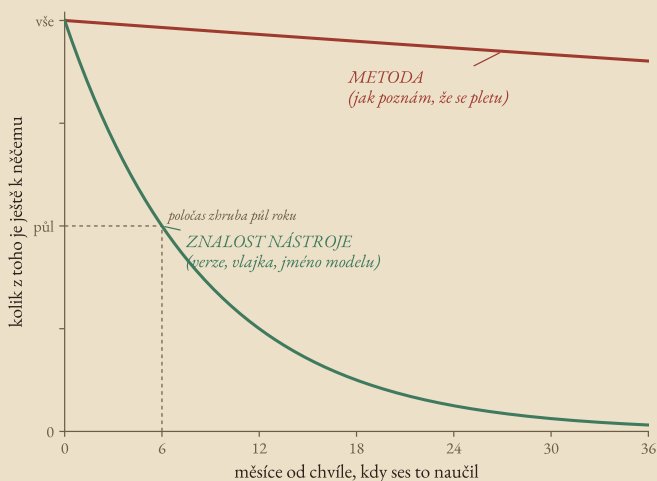
changelog — chronologický seznam změn nástroje, verzovaný jeho autorem. Primární pramen: co se doopravdy změnilo, od těch, kdo to změnili. Opak threadu, který ti to převypráví z třetí ruky a s pointou.

track record — doložená minulost trefných předpovědí. Kdo se v minulosti opakovaně trefil (a ukázal jak), váží víc než kdo má titul. Kap. 22.

dvadvacáté kapitoly, jen namířená ven, na lidi a changelogy místo na vlastní předpovědi.

Poločas znalosti: co honit a co ne

Začnu tím, co tě z toho pátečního hluku okamžitě vysvobodí — rozlišením, které je vlastně celá kniha nakreslená jednou čarou. Ne všechno, co ses naučil, hnije stejně rychle. Znalost *konkrétního nástroje* — jméno modelu, přesná vlajka, cena za token, kde je které tlačítko — má krátký poločas rozpadu: řádově měsíce. Znalost *metody* — jak postavit hloupého soupeře, jak napsat test, který mohl selhat, jak poznat, že se pleteš — se nerozpadá skoro vůbec. Tři měsíce ji neškrábnou, tři roky sotva.



Hrdinský graf. Verdigris (plachta): znalost nástroje padá exponenciálně, poločas zhruba půl roku — za rok ti z ní zbude čtvrtina. Sanguine (stěžeň): metoda leží skoro naplocho. Osy nejsou změřené, jsou to poctivé modely: poločas nástroje volím podle tempa hlavních verzí 2024–2026 (zhruba dvě za rok). Pointa čáry: investuj úměrně poločas — hodiny do metody, minuty do jména modelu.

Přečti ten graf jako rozpočet svého času a nervů. Hodinu strávenou tím, že *pochopiš*, proč agent potřebuje sandbox a co je to hloupý soupeř pro tvůj eval, si za rok pořád vybíráš celou. Hodinu strávenou memorováním přesného názvu dnešního nejlepšího modelu jsi za půl roku vyhodil oknem — ten název bude jiný. To neznamená nástroje ignorovat; znamená to učit se je *jako zástupce principu*, ne jako seznam faktů k zapamatování. Když víš, *proč* existuje MCP (dát modelu ruce a oči, kap. 24), je jedno, že se příští rok bude konfigurovat jinak — poznáš to nové zařízení podle toho, jaký starý problém řeší.

A tady je úleva, kterou ti nikdo neřekne, protože se špatně prodává: *nemusíš stíhat všechno*. Kdo se snaží přečíst každý thread a vyzkoušet každý model v týdnu jeho vydání, nedělá výzkum — dělá si úzkost. Stěžeň, který jsi postavil, tě opravňuje k pomalosti: můžeš nový nástroj nechat týden uležet, počkat, až opadne první nadšení i první panika, a teprve pak se podívat, co z toho po tobě chce metoda. Klid je tady konkurenční výhodou, ne slabostí.

Kompas zdrojů: čtyři hlasy, jedna otázka

Než ti řeknu *kde* číst, musíš umět rozlišit, *kdo* mluví. V každé debatě o AI slyšíš čtyři druhy hlasů, a splývají do jednoho jen tomu, kdo se ještě nespálil. Nejsou to čtyři druhy lidí — jeden člověk může střídat všechny čtyři čepice během jednoho odstavce. Jsou to čtyři *role*, a každá si zaslouží jinou dávku důvěry.

TEENAGER · MECHANISMUS

Čtyři hlasy — a co s každým udělat:

VÝZKUMNÍK — měří a publikuje, jak to doopravdy funguje.

Váha: vysoká na MECHANISMUS, opatrně na „co to znamená pro tebe“. Ptej se: je tam metoda, čísla, a přiznaná nejistota? (arXiv, model cards, engineering blogy)

PRAKTIK — něčím to denně staví a píše, co ho pálí.

Váha: NEJVYŠŠÍ na „jak to použít“. Má track record, ukazuje kód i omyly. To je tvůj hlavní zdroj.

EVANGELISTA — nadšenec, který ti nese dobrou zprávu.

Váha: dobrý na TIPY co zkusit, nulová jako důkaz.

Nadšení není měření. Ověř si každé „změnilo mi to život“.

PRODEJCE — má z tvého „ano“ přímý prospěch.

Váha: ber jako reklamu, i když má pravdu.

Firemní

benchmark je marketing s chybovými úsečkami schovanými pod stolem. Čti, pak si to změř sám.

Jedna otázka je oddělí spolehlivěji než jakýkoli titul: *co konkrétního ten hlas ztratí, když se plete?* Praktik ztratí čas a pověst, na které mu záleží — proto opatrně váží. Prodejce neztratí nic (spíš vydělá) — proto tlačí. Čím dražší je pro zdroj vlastní omyl, tím víc jeho slovům věř.

→ kap. 22: „sláva $\propto$ nepřesnost“ (Tetlock).
Kdo vystupuje nejsebejistěji a nejblasitěji, mívá nejhorší track record — sebejistý tón je zdarma, trefa ne. Slavný hlas s velkým dosahem není totéž co přesný hlas.

Pozor na hybridy: výzkumník ve firmě, která model prodává, nosí dvě čepice najednou. Čti jeho metodu jako výzkumníka, jeho závěry jako prodejce.

Všimni si, že ani jeden hlas neříkám ignorovat — to by byla stejná chyba jako věřit všem. Projevcův benchmark ti prozradí, na co je nástroj *stavěný* (to je užitečné), jen ne, jestli to opravdu umí (to si změříš sám). Evangelistův nadšený thread je skvělý seznam věcí *k vyzkoušení*, jen ne

důkaz, že fungují. Kompas neříká „věř / nevěř“; říká „věř tolik a tolik“ — a to je přesně dávkování důvěry, které tě naučila dvaadvacátá kapitola. Jen jsi ho tehdy mířil na předpovědi, teď ho míříš na zdroje.

Rychlý test na jakýkoli titulek typu „X právě zabil Y“: zeptej se, jestli autor Y vůbec kdy použil na skutečné práci a co získá, když mu uvěříš. Když je odpověď „nepoužil, a prodává kurz o X“, máš hotovo dřív, než jsi dočetl perex. Když je to praktik, který obojí denně staví a ukazuje kód, čti dál — možná ví něco, co ty ne.

Kde číst: kurátorský seznam, ne feed

Teď to praktické. Kde tedy tyhle hlasy hledat tak, abys slyšel víc signálu než šumu? Pravidlo je jediné a plyne z poločasu i z kompasu: *čti prameny a lidi s track recordem, ne agregovaný feed*. Feed je optimalizovaný na tvou pozornost, ne na tvou pravdu — odměňuje sebejistý tón a hádku, tedy přesně to, co Tetlock změřil jako *nejhorší* prediktor přesnosti. Sestav si místo něj malý, ruční seznam a braň ho před bobtnáním.

Ten seznam schválně nejmenuje konkrétní blogy a účty — kdybych to udělal, kapitola by zastarala rychleji než nástroje, o nichž mluvím, a udělal bych přesně chybu, před kterou tě varuju. Místo jmen ti dávám *druhy* zdrojů a pravidlo, jak si je vybrat: primární pramen před převyprávěním, člověk s doloženou minulostí před hlasitým neznámým, a raději pět jmen, která čteš pořádně, než třicet, která proskrolluješ. Kurátorský seznam je tvůj — a jeho hlavní ctnost je, že je *malý*.

Kurátorský seznam (uspořádaný podle poločasu — od trvalého po prchavé):

1. PRIMÁRNÍ PRAMENY (nejtrvalejší, čti první)
 - changelog / release notes nástroje → co se změnilo
 - oficiální docs → jak to funguje teď
 - model card / system card → meze a rizika modelu

Pravidlo: než uvěříš threadu „nová verze umí X“, otevři changelog a ověř, že to tam vážně stojí.
2. LIDÉ S TRACK RECORDEREM (praktici, ne slavní)
 - pár blogů praktiků, kteří ukazují kód i omyly
 - roční/půlroční „co jsme se naučili“ přehledy

Vyber 3-5 jmen, ne třicet. Kdo se pletl a napsal to, váží víc než kdo se nikdy nepletl (protože nezkoušel).
3. AGREGÁTORY (rychlý přehled, ne důkaz)
 - odborné diskuse, kde je slyšet i nesouhlas

Ber jako radar „na co se podívat“, ne jako pravdu.

Signál je v komentáři, který někdo doložil, ne v počtu palců nahoru.

Proč právě tři patra: každé má jiný poločas i jinou cenu ověření. Changelog je levné ověřit (otevřu, přečtu) a skoro se neplete o faktu (co se změnilo). Thread je drabé ověřit (musím to sám vyzkoušet) a plete se snadno (převyprávěl z třetí ruky). Čti odspodu poločasů: nejdříve trvalé a levně ověřitelné.

Konkrétní jména schválně neuvádím — do roka zestárnou. Uvádím druh zdroje; ten nezastará. To je poločas v praxi i tady.

A jedno záporné pravidlo, které ti ušetří nejvíc: *nesleduj vydání kvůli vydání*. Nový model si nezkoušej v den, kdy vyjde, ale ve chvíli, kdy máš reálný úkol, na kterém to poznáš. Novinka bez tvého úkolu je zábava, ne práce.

Jak otestovat nový nástroj: baseline a malý reálný úkol

Kompas ti řekl, komu naslouchat, než ověříš. Ale u nástroje ověřit *můžeš* — a levně. Takže se ptáš špatně, když se ptáš „je ten nový model dobrý?“. Dobrý *na co, oproti čemu?* Nástroj se netestuje čtením, nástroj se testuje během. A test má přesně dva díly, oba z první části knihy.

Protokol na nový nástroj — dvacet minut, ne dva dny:

1. HLOUPÝ SOUPEŘ (baseline) – kap. 12
Než změříš nový nástroj, změř, co umíš BEZ něj.
Starý model. Ruční postup. Google. To je laťka.
Bez ní „nový model to zvládl!“ neznamená nic – třeba to zvládl i ten starý.
2. MALÝ REÁLNÝ ÚKOL – ne benchmark, ne hračka
Vezmi něco, co jsi nedávno dělal a čemu ROZUMÍŠ
(znáš správnou odpověď). Pusť na to nový i starý.
Reálný, protože benchmark je trať, ne tvoje práce.
Malý, protože chceš odpověď dnes, ne za týden.
3. ROZDÍL, NE ČÍSLO
Nezajímá tě „94 %“. Zajímá tě „o kolik líp než hloupý soupeř, na MÉM úkolu“. A pusť to víckrát –
jeden běh je anekdota (kap. 28).

Pointa: benchmark ti řekne, jak je model dobrý pro *průměr světa*. Tvůj malý reálný úkol ti řekne, jak je dobrý pro *tebe* — a to je jediné číslo, podle kterého se rozhoduješ ty.

→ kap. 12: hloupý soupeř. → kap. 28: jeden běh nic nedokazuje; pusť nový nástroj na svůj úkol víckrát a dívej se na rozptyl, ne na jeden šťastný výstřel. → kap. 8: úkol i laťku si vyber předem, jako pakt — jinak si po vydání racionalizuješ, že nový model je lepší, protože je nový.

Tenhle protokol je celá kniha zmáčknutá do dvaceti minut. Detektor Whitehouse/Kelvin z druhé kapitoly se ptá u každého nového nástroje: přidává mi *sílu* (další model, delší kontext, víc agentů), nebo *citlivost měření* (líp poznám, kdy se to pokazilo)? Whitehouse pumpoval do kabelu čím dál vyšší napětí a věřil naslepo, až kabel umlčel; Kelvin přidal citlivější galvanometr a zprávu poslal. Nový model je napětí. Tvůj hloupý soupeř a malý reálný úkol jsou galvanometr. Bez nich honíš sílu; s nimi měříš — a jen měření tě posune.

Věž: kalibrace důvěry ve zdroj

Zbývá dát tomu všemu číslo — protože přesně to dělá z čtenáře, který honí hype, čtenáře, který jím proplouvá. Kompas jsi dostal obrazem; teď ho obrátíš na sebe stejným nástrojem, kterým dvaadvacátá kapitola měřila tvé vlastní předpovědi.



Zdroj je předpovědní stroj — a dá se kalibrovat. Každý hlas, který posloucháš, dělá totéž co ty v deníku predikcí: vydává tvrzení o budoucnosti („tenhle model změní workflow“, „tahle knihovna je mrtvá“). Můžeš tedy na *zdroj* nasadit přesně diagram spolehlivosti z kapitoly dvaadvacáté. Ke každému hlasu, který sleduješ, si veď tichý deník: co tvrdil, jak jistě, a jak to dopadlo. Po roce vyneseš deklarovanou jistotu proti naměřenému podílu trefu — a uvidíš, kdo z tvých zdrojů leží na úhlopříčce (říká = platí) a kdo visí pod ní (slavný, sebejistý, nepřesný). Kdo vážně chce, spočítá zdroji i *Brierovo skóre*

$$B = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2,$$

kde p_i je jistota, s jakou zdroj tvrzení pronesl, a o_i je, jestli vyšlo (1) nebo ne (0). Níž je líp; čtvrtka je hloupý soupeř, který u všeho krčí rameny. Zdroj, který soustavně leze nad čtvrtku, je horší než mince — a přesto ho možná dosud posloucháš, protože zní jistě.

→ kap. 22: tentýž diagram spolehlivosti a totéž *Brierovo skóre*, tehdy na tvé vlastní předpovědi, teď na cizí zdroje. Nástroj je jeden, míří třikrát: na test (kap. 12), na model (kap. 16), na člověka a zdroj (kap. 22 a 31).

Praktický zkrat, když se ti nechce vést deník: pamatuj si jen jednu trefnou a jednu minutou předpověď od každého zdroje, na který dáš. Dvě datové body jsou nekonečně víc než dojem — a dojem ti u sebejistých hlasů systematicky lže nahoru.



Proč zdroji nikdy nevěřit naplno — a kdy klidně věř. Míra důvěry se, jako v celé knize, neřídí tím, jak jistě zdroj zní, ale tím, *kolik tě bude stát, když se plete, než to zjistíš*. U nástroje je ta cena skoro vždy nízká a ověření rychlé: zdroj řekne „tenhle model je lepší na refaktoring“, ty ho za dvacet minut pustíš na svůj úkol vedle hloupého soupeře a víš. Tam věř klidně a ověř *potom* — omyl je levný a okamžitě viditelný. Zbytek kompasu (Kelly, dávkuj závazek úměrně důkazu, drž si rezervu na to, že ses spletl) je tu skoro zadarmo, protože testovat nástroj je levné a vratné. Právě proto je tahle poslední oblast tak přátelská: na rozdíl od experta na politiku si můžeš skoro každé tvrzení o nástroji *sám změřit* — a měl bys.

Kontrast s kap. 22: tam šlo často o předpovědi drabé nebo nemožné ověřit hned (padne vláda? poroste inflace?). U nástrojů máš vzácný luxus — laciné, okamžité, opakovatelné ověření. Využij ho. Zdroj je pak jen tip, kam napřít vlastní galvanometr, ne autorita, které se skládá účet.

Konec plavby

Došli jsme na konec — nejen kapitoly, ale celé plavby. Sluší se poslední účet, a bude krátký, protože všechny dlouhé už padly.

Naučil ses stavět dnešními nástroji: ventilovat nápad do hotové věci, dát modelu ruce a oči, předat mu své řemeslo, nepustit ho na produkci naslepo. Naučil ses ověřit, co postavil: eval jako hloupého soupeře, threat model na svá data, ekonomiku spočítat místo odhadu. A teď, na úplný konec, ses naučil to jediné, co ti zbude, až všechna ta jména nástrojů zastarají: *bonit zdroje, ne verze*. Track record nad titul. Changelog nad thread. Malý reálný úkol proti hloupému soupeři nad každý

benchmark. A vlastní kalibrovaná nedůvěra nad každý sebejistý tón — počínaje tím tvým vlastním, ve chvíli, kdy je nejsladší.

Kniha začala větou, že se svět dělí na dva druhy lidí: na ty, kdo umějí výsledek ověřit, a na ty, kdo mu musejí věřit. Končí u téže věty — a u dveří, jimiž jsi právě prošel. Stěžej, který jsi postavil v první části, *stojí*: metoda nezastará, ať se nástroje mění týden co týden. Plachty, které jsi napjal ve druhé, jsou ten nejrychlejší vítr, jaký kdy měl kdokoli před tebou — a jsi jediná generace, která umí obojí naráz: položit otázku *i* ověřit odpověď, být vědec *i* programátor v jedné osobě, v jednom odpoledni, u jednoho terminálu. Ne protože jsi chytřejší než ti před tebou. Protože jsi přišel v čase, kdy je obojí na dosah ruky — a protože ses nenechal svést k tomu věřit místo měřit.

Píseň sirén nekončí; nikdy nebyla svůdnější a s každou novou verzí bude znít sladčeji. Ale ty už víš, jak u ní zůstat přivázaný a plout dál — řádově dál, než kdokoli, kdo měl jen uši, nebo jen odvahu skočit. A ať tě zítra, za rok nebo za deset let potká nástroj, o kterém se mi tady ani nesnilo, ponese tě jediná otázka, se kterou jsi celou tuhle knihu žil a se kterou ji teď zavíráš. Polož si ji nahlas, naposledy tady a napořád všude jinde:

■ DEMO

lesk-a-bida-vibecodingu.gaussalgo.com/d/kalibrace

Kalibrační kvíz naposledy — a teď jinak. Projdi dvacet otázek, u každé odhadni odpověď a svou jistotu; odejdeš s vlastním Brierovým skóre a vlastní křivkou. A pak ten samý deník obrať ven: veď si rok, jak jistě tvrdí tvé zdroje a jak jim to vychází. Za rok budeš vážit lidi, modely i sebe stejným číslem — a přesně to je celá dovednost, kterou ti kniha chtěla dát.



„Jak bych poznal, že se pletu?“

Naposledy — a doopravdy napořád. Tenhle blok zavíral každou kapitolu; teď ať přeroste knihu a stane se kompasem na celý tvůj další život s nástroji, které ještě neexistují. Ke každému novému modelu, ke každému nadšenému threadu, ke každému „tohle všechno mění“ — a ke každé své vlastní jistotě, že už to konečně umíš — si polož tu jedinou otázku: *jak bych poznal, že se pletu?* U nástroje je odpověď laciná a pokaždé po ruce: postav hloupého soupeře, vezmi malý reálný úkol, pusť to víckrát a změř rozdíl. U zdroje je odpověď deník: zapiš, co tvrdil a jak jistě, a za rok si spočítej trefu. Tvrdím, že kdo si takhle vede skóre — sobě, svým nástrojům i svým zdrojům —, propluje kolem každé nové sirény, ať se jmenuje jakkoli, zatímco lepší plavci kolem něj budou tonout v nadšení. A jestli mi jednou ukážeš deník, ve kterém tvé devadesátiprocentní jistoty vyšly v devíti z deseti — o nástrojích, o zdrojích i o sobě samém —, pak se tahle poslední kapitola, a s ní celá kniha, nepletla ani trochu: stal ses člověkem, který svým slovům smí věřit přesně tolik, kolik říká. A to je, příteli, celá odměna za tuhle plavbu. Vítr máš. Stěžeň stojí. Pluj.