

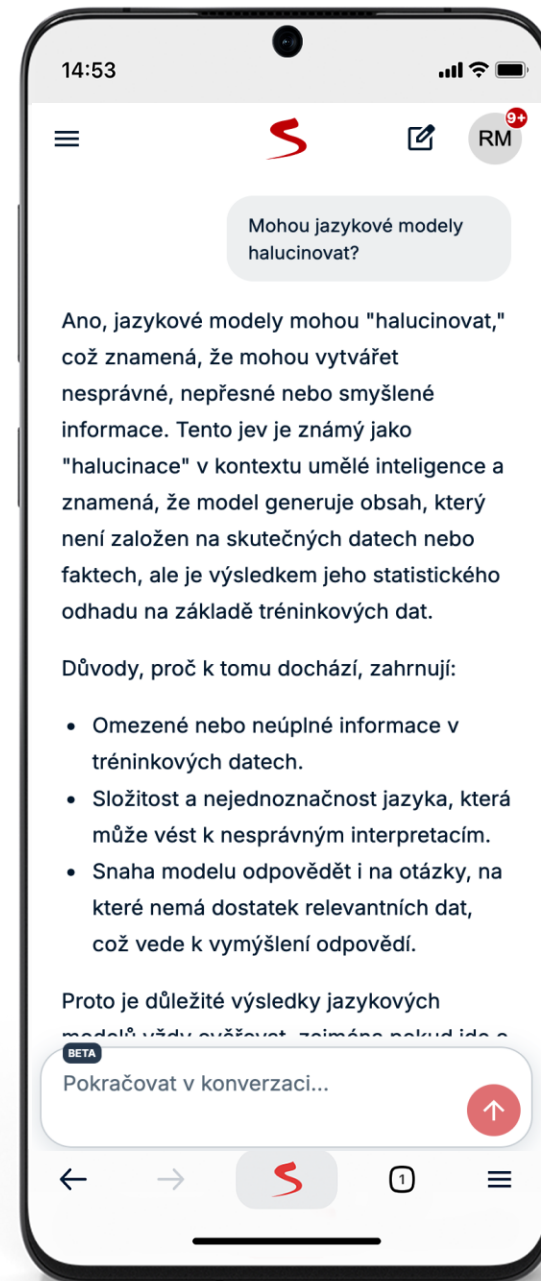


Aplikace LLM v praxi: Krásné lži a drahá pravda



Radoslav Mikeš

Product lead, Seznam Asistent



Disclaimers

- Některé tvrzení už nemusí být pravda.
- To, že to nefunguje dnes neznamená, že za rok nebude.
- Počítejme kolikrát řeknu AI.

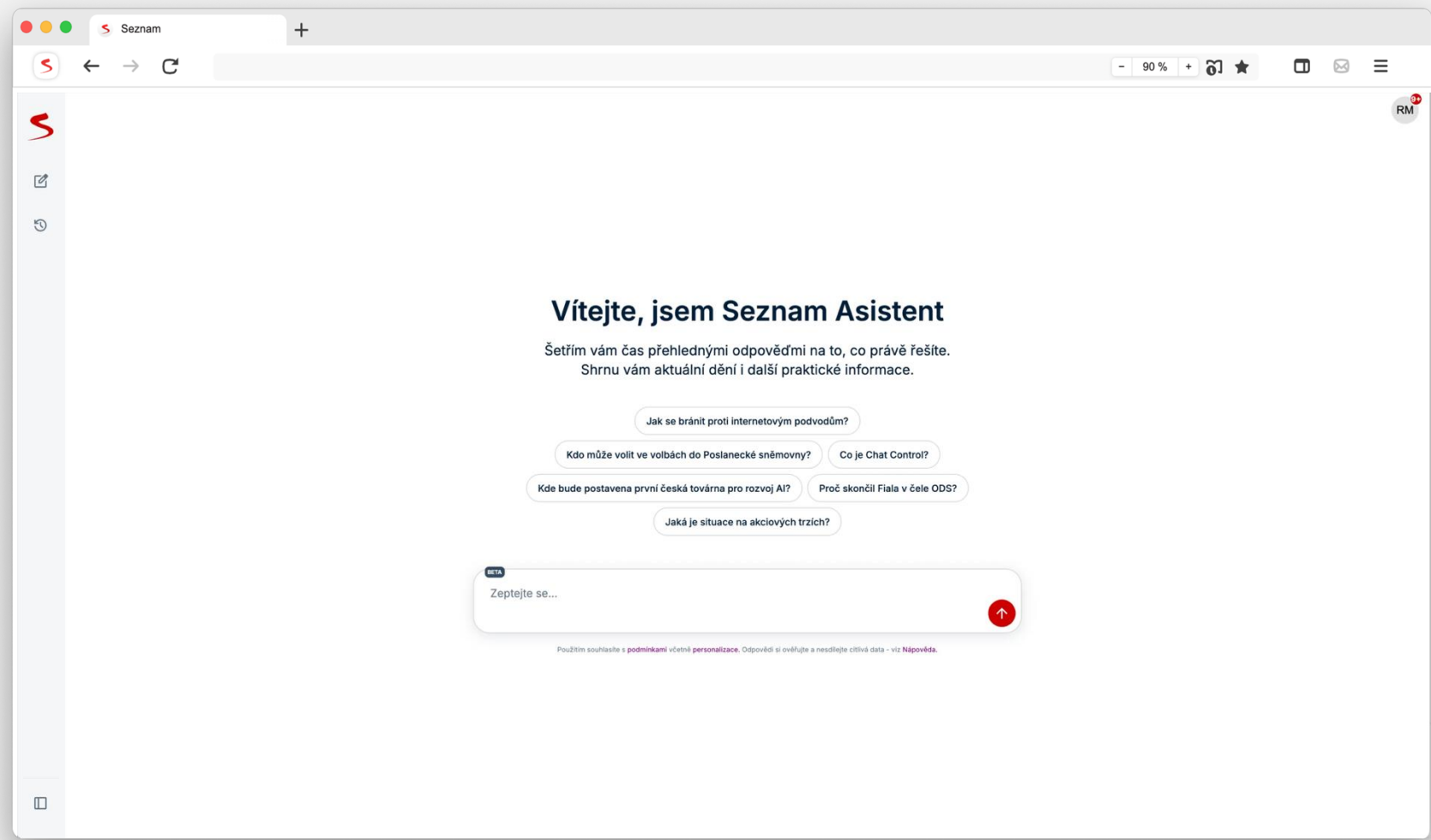
„Víme, že nevíme dost.“



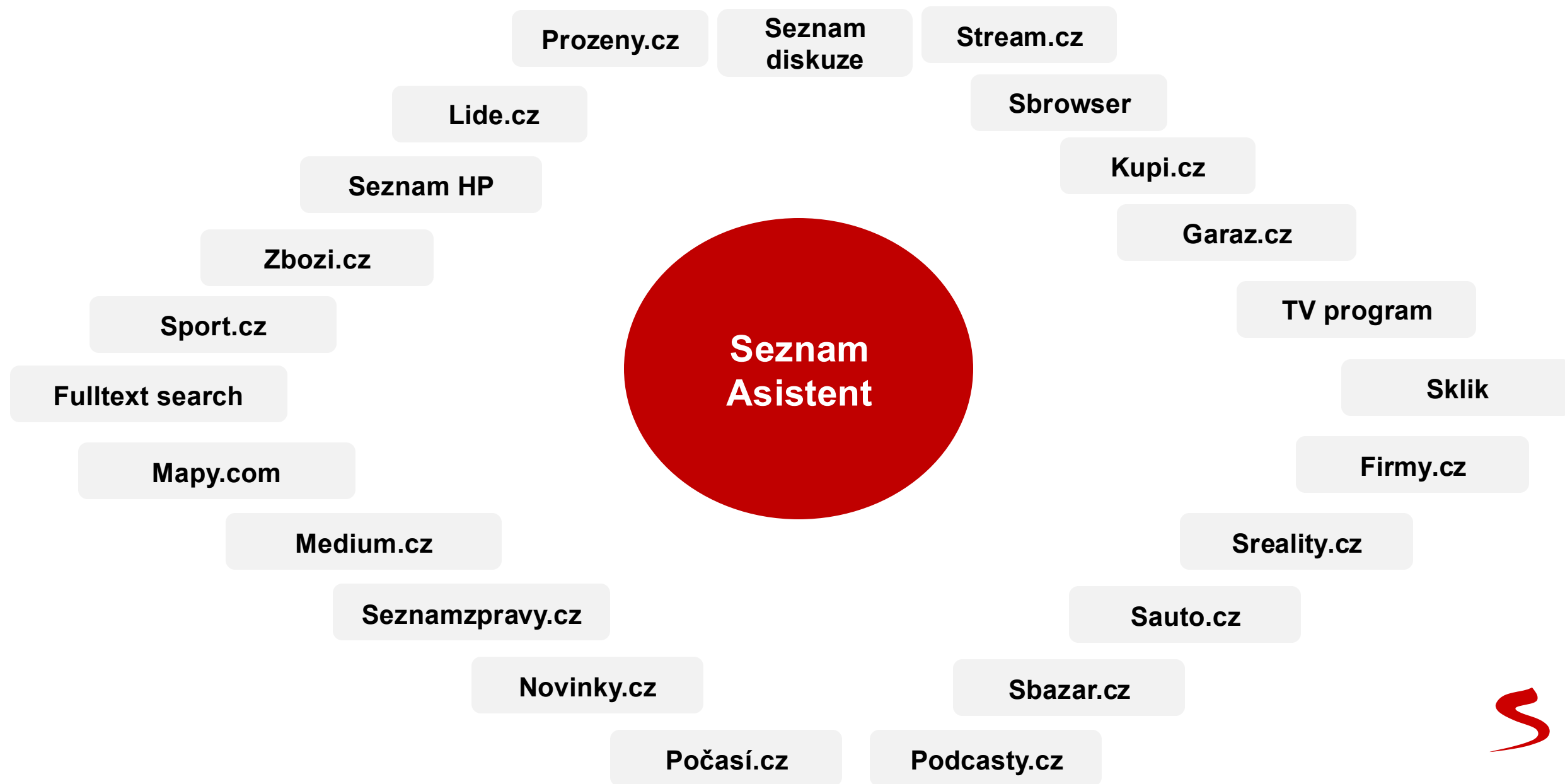
„One of the contradictions of the LLM revolution is that it is so easy to build a charismatic prototype... it is so hard to get from a brutal prototype to something that is valuable and rocksteady in production.“

CTO Palantiru, Shyam Sankar

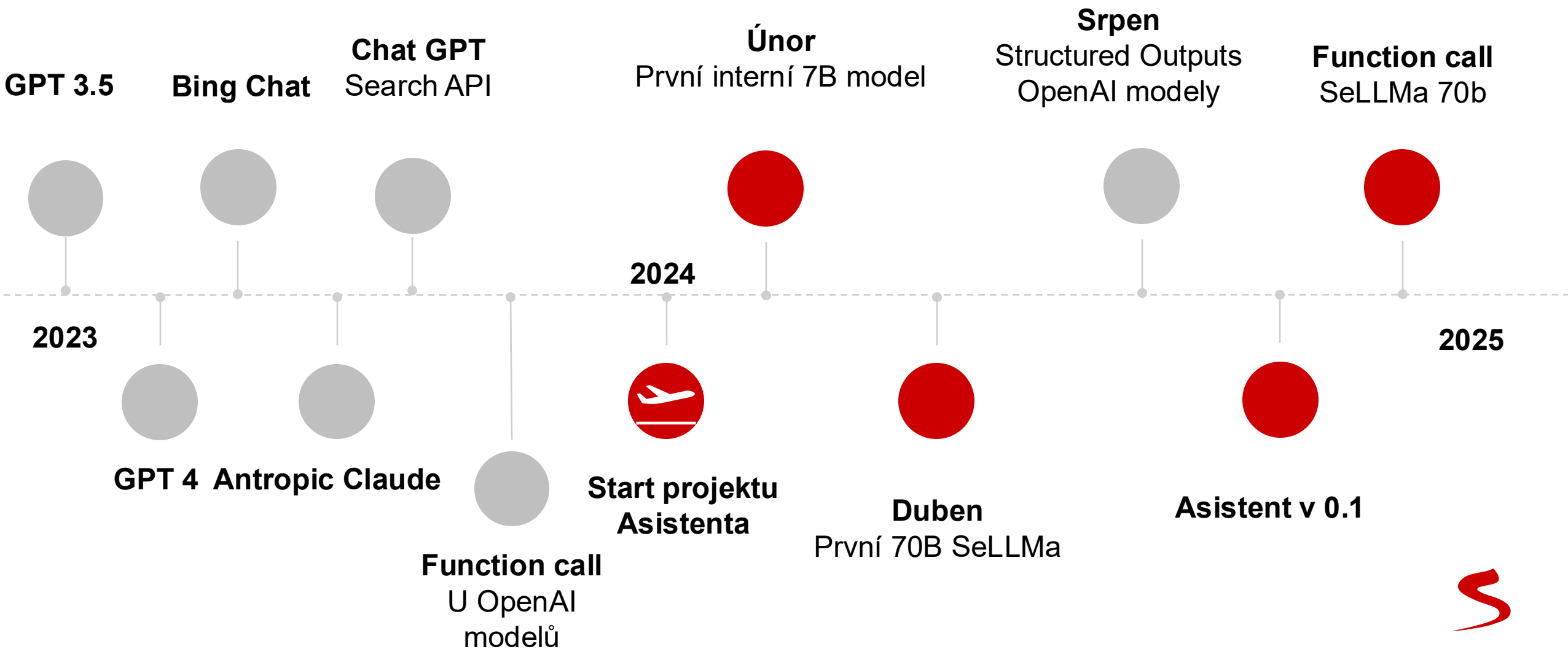
Seznam Asistent



Asistent v platformě Seznamu

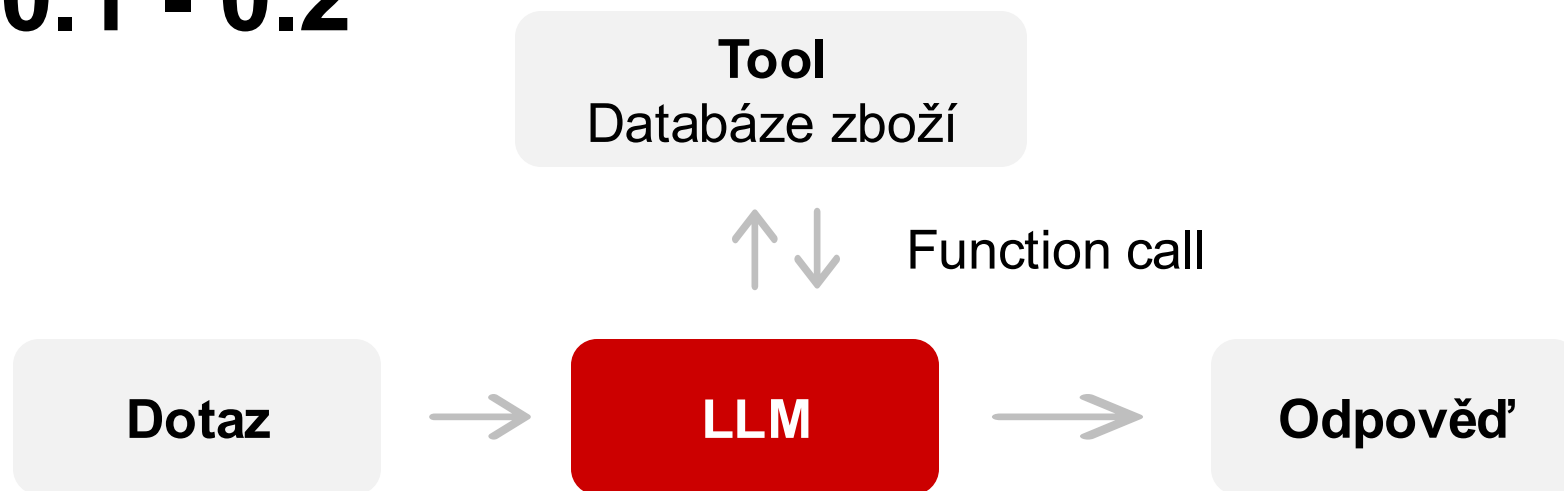


Začátek

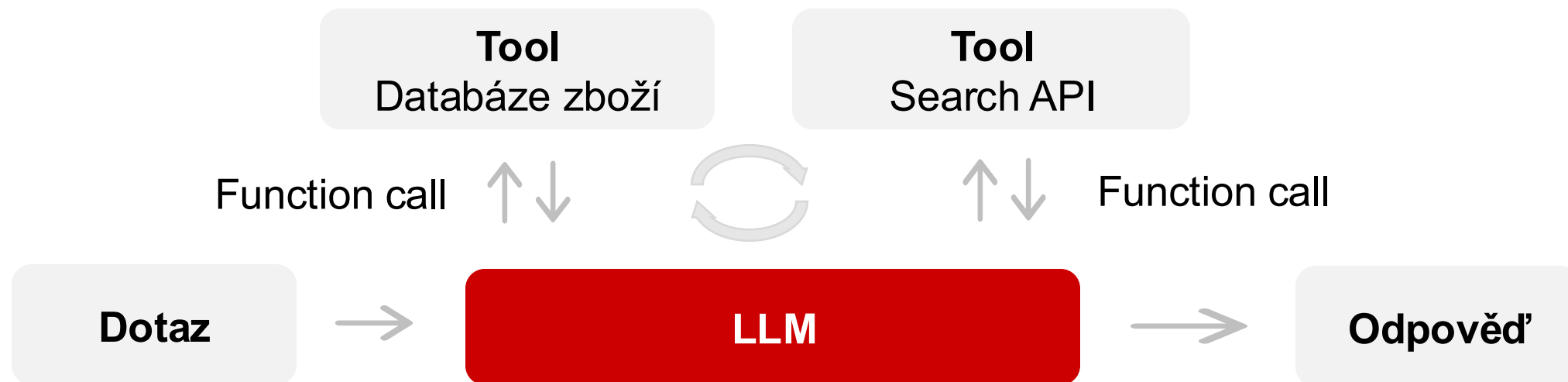


Asistent v 0.1 - 0.2

Function call



ReAct



Zjištění



Datová hygiena

Není zrovna uklizeno .. 😊



Hledání

Hledání nebylo připraveno na "nové typy" dotazů. Mnoho služeb počítalo s trafficem z internetu a nerozvíjí kvalitní vektorové hledání nad svou databází



Nárůst velikosti promptu



Vysoké množství halucinací

Je těžké začít pouze s jednou vertikálou, když se uživatelé ptají cross vertikály.

zboží.cz

RM

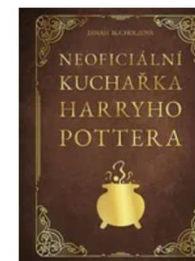
Harry potter



< Knihy

Elektronické knihy: „harry+potter“

Filtr

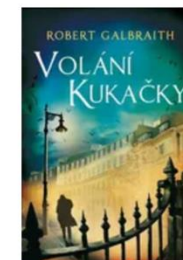


Neoficiální kuchařka Harryho Pottera - Dinah Bucholzová (2023)
[E-kniha]

★★★★☆ 5

349 Kč

Kuchařka, Čeština, Počet stran 248, Rok vydání 2023, Formát PDF



Volání kukačky - Robert Galbraith
[E-kniha]

★★★★☆ 14

od 195 Kč

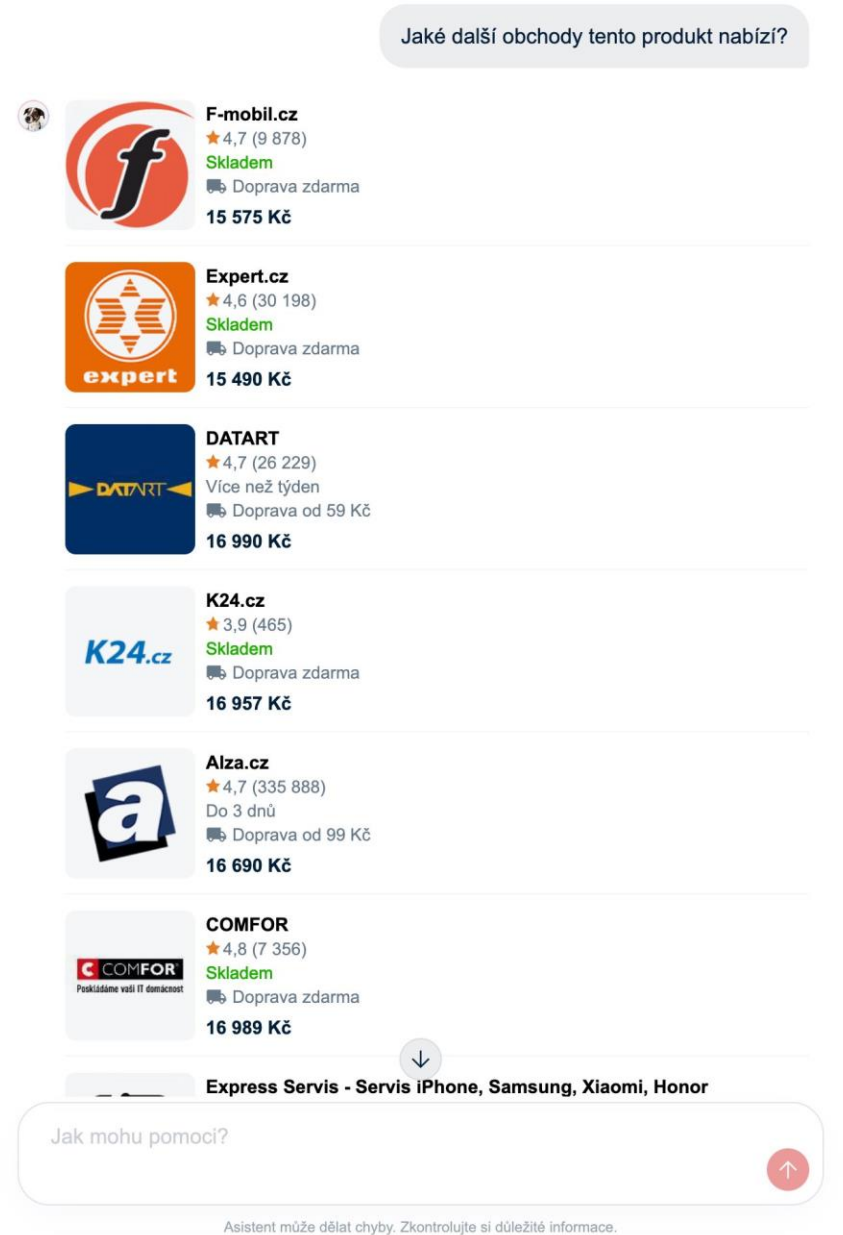
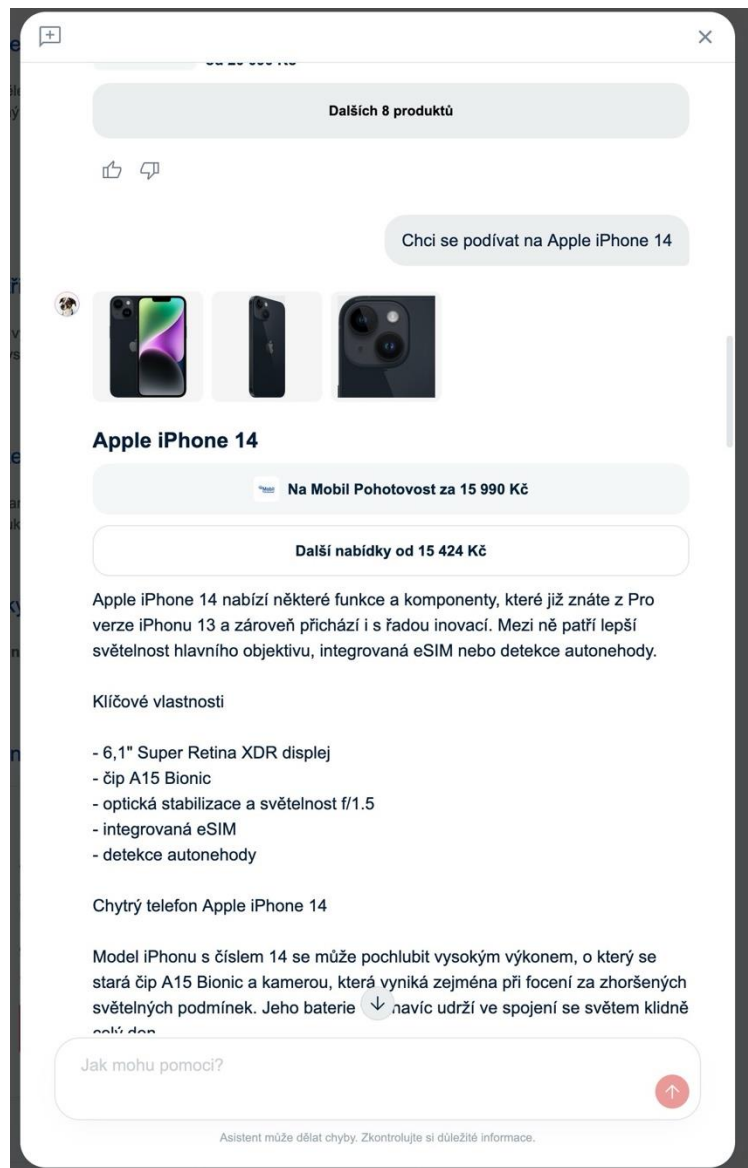
Detektivka, Čeština, Počet stran 480, Rok vydání 2013, Formát Mobi



Lišky pod Big Benem: Vážně i vesele o Velké Británii - Jaromír Marek (2025) [E-kniha]



Prototyp n.1



Stabilita odpovědí

Instrukce od uživatele

Hledám mobil do 20 000, který bude červený s tím, že nechci Apple systém, a měl by být dostatečně velký a kvalitní fotak.

Chatbot
Function call

Tool
Hledání zboží

Query do hledání
Mobil do 20 000, červený,
Android, 6,5 palce, kvalitní foťák

Chatbot
Function call

Tool
Hledání zboží

Query do hledání
Mobil do 20 000, červený,
Android, 5,4 palce, 30 Mpx
foťák,

Chatbot
Function call

Tool
Hledání zboží

Query do hledání
Mobil do 20 000, červený,
Android, 5 palců, 50 Mpx foťák,

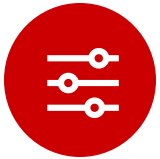


Vyhodnocování



Ground thruth set

Testovací sada pravdy.



Kalibrované LLM as Judge

Vyhodnocování za pomoci LLM.



Uživatel na baterky

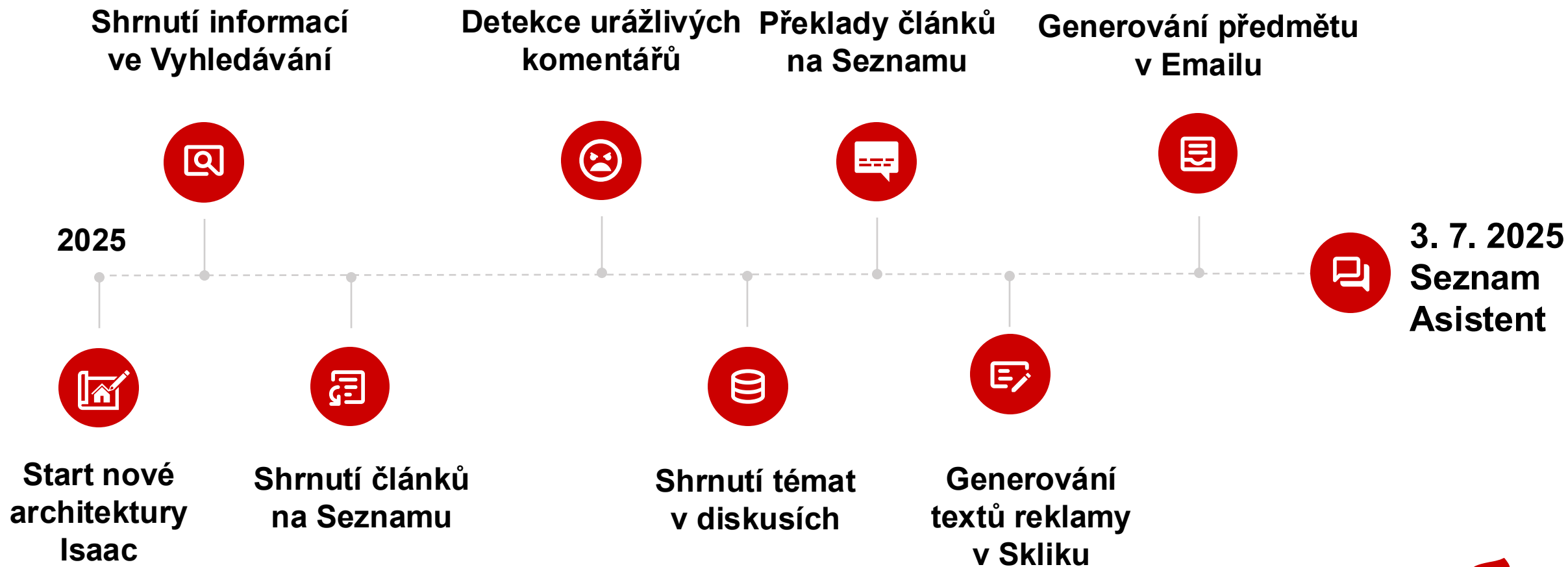
Vytvoření umělého uživatele.



Testovací sady na prompty v architektuře

Na každý uzel je potřeba mít testy.

2025



Proč se hledá Eva Burešová?

Eva Burešová a [Peter Sagan](#) byli vyfotografováni, jak se drží za ruce během noční procházky po taneční soutěži [Let's Dance](#). **Burešová vysvětlila, že šlo o přátelské gesto** a že často diskutují o práci, úspěchu a rovnováze mezi kariérou a osobním životem. – Generováno Seznam AI.

★ Super.cz 🟡 eXtra.cz 🟠 TV Nova

Pokračujte v tématu

Kdo je Eva Burešová a Peter Sagan?

Vztahy celebrit v tanečních soutěžích?

České celebrity v tanečních soutěžích?

✍ Jiná otázka

• Seznam AI



Nejnovější zprávy



blesk.cz

Burešová o fotkách se Saganem: Položil mi ruku kolem ramen a pak... | Blesk.cz

Před 4 hodinami ❤ 0



expres.cz

Eva Burešová Peter Sagan rande - Expres.cz

Před 10 hodinami ❤ 0



super.cz

Sídlo v Monaku i rozvod po třech letech manželství! Noční společník Evy Burešové, cyklista Peter Sagan, odkryl soukromí



🏠 > NOVINKY

Český Seznam předběhl Google: umělá inteligence sumarizuje vyhledávané informace

[Pavel Trousil](#) | 24.01.2025

cnews.cz

Internet

Seznam předběhl Google. Nabízí u nás praktickou sumarizaci vyhledávání, a to rovnou v češtině

[Dominik Dobrozenský](#) | 24. 1. 2025

💬 Přidejte Názor



ZIVE

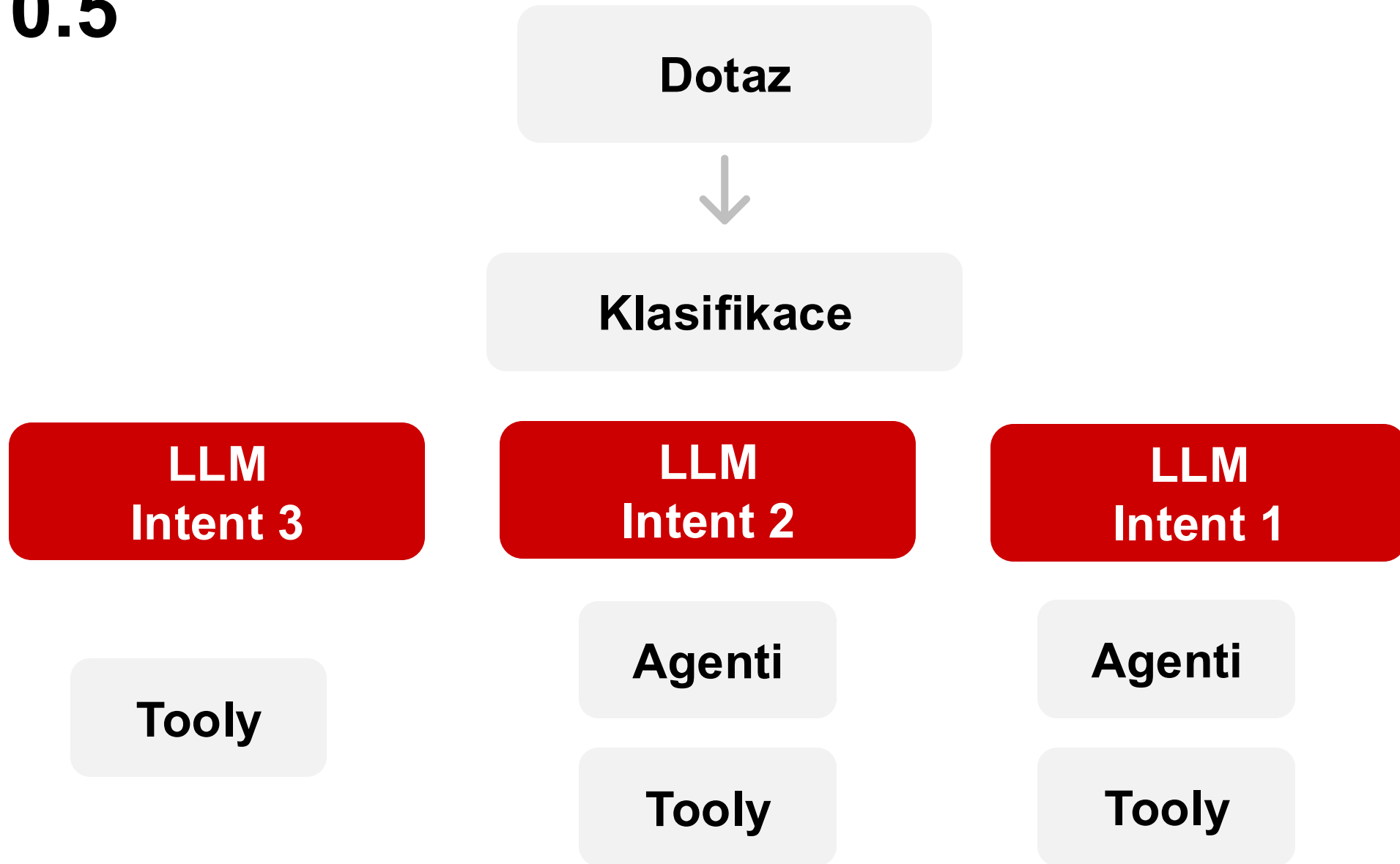
Seznam předběhl český Google. Odpovědi ve vyhledávači mu generuje AI

[Lukáš Václavík](#)

23. ledna 2025

💬 DISKUZE (32)

Verze 0.5





Vítejte, jsem Seznam Asistent

Šetřím vám čas přehlednými odpověďmi na to, co právě řešíte.
Shrnu vám aktuální dění i další praktické informace.

Dej mi detaily k celní dohodě mezi EU a USA

Jak se vyvíjí izraelsko-palestinský konflikt?

Jaké bude zítra počasí na Rhodosu?

Co si sbalit na týdenní dovolenou u moře?

Kdo může volit ve volbách do Poslanecké sněmovny?

Jaký typ munice se bude vyrábět v nové továrně firmy Rheinmetall?

BETA

Zeptejte se...



Použitím souhlasíte s [podmínkami](#) včetně [personalizace](#). Odpovědi si ověřujte a nesdílejte citlivá data - viz [Nápověda](#).

Zjištění



Scénáře jsou stabilně rychlé / pomalé.



Nízká spokojenost v reálném provozu.

Anotovaná spokojenost s odpovědí Asistenta 60%



Tooly laděné na scénáře.

Moc striktní konfigurace Agentů na jednotlivé scénáře. V produkci tyto scénáře failovali.



Máme moc toolu / agentů.

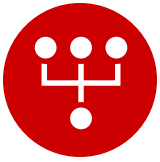
Co s tím uděláme?



Potřebujeme paměť.



Nová architektura verze 0.7+



Více agregovaných toolů.



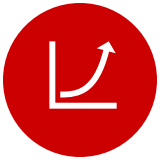
Orchestrace a řízení konverzace.

- Využitím vkládání potřebných instrukcí, dle využitého toolu či typu dotazu.
- Klasifikace jednoduchých a složitých dotazů.



Výměna typu modelů.

Začátek nahrazování decoder modely Bi-encoder , encoder modely.



Nárůst spokojenosti s odpovědí.



Zepete se...

[Nápověda](#) [Vysvětlí mi](#) [Oprav pravopis](#) [Porad' mi, jak](#) [Aktuální dění](#)**Novinky** • Neděle 26. října. Svátek má Erik.**Silný vítr vyvrací stromy a zaměstnává hasiče**

Silný vítr, před nímž varovali meteorologové, zaměstnává od nedělního dopoledne hasiče. V Čechách vyjíždějí...

Ryanair si za odbavení na letišti účtuje přes tisíc korun

Zájem o zvěřinu od státních Lesů ČR roste. Vedou daněk, divočák a jelen

Tragický pád letadla v Letňanech odstartoval problém s kabinou. Pak přišly osudné chyby pilota

Nemocných s roztroušenou sklerózou přibývá. Může za to i životní styl

KVÍZ: Poznáte starý seriál jen podle tří indicií?

Macinka potvrdil, že chce zrušit CHKO Soutok

Dálnici D1 u Brna uzavřela nehoda tří aut

SZ Seznam Zprávy**Mišík: Klempířova spolupráce s StB je stigma. Na Turkovo prozření nevěřím**

Právě vydaná deska zachycuje jedno z posledních vystoupení Vladimíra Mišíka předtím, než ukončil koncert...

Smutek mládí. Dospívající generace přišla o jeden „vrchol štěstí“

KOMENTÁŘ Otázka za milion euro, má Česko přijmout euro?

Nově otestovanou raketu může Rusko vyslat záhy po jaderném úderu, míní expert

Křečové žily jako předzvěst vážné nemoci. Mají je i muži, genetika je zásadní

Zástupci USA a Číny se shodli na rámci obchodní dohody

KOMENTÁŘ Sliby chyby, spojenců proti povolenkám je málo

Pohled na největší staveniště v novodobé historii Česka

Super**Nazvali ji nadívaným krocanem. Nelly Furtado po hejtech na svou plnoštíhlou postavu šokovala fanoušky svým rozhodnutím!**

Zpěvačka Nelly Furtado už se na pódiu neobjeví. Po sérii vystoupení, za které si vysloužila hejty, protože její tělo už...

Jenna Ortega své odhalování posunula na další level. Ňadra předvedla v průsvitné „noční košilce“

Nezdá se, ale je to kanec: Než se usadil, při výběru partnerek sbíral Jiří Mádl zářezy i v zahraničí

Email Firemní e-mail zdarma**Doručené (6750)**

Kalendář

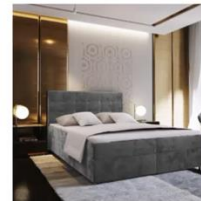
Poznámky

Nová zpráva

Zobrazit e-maily

Seznam bez reklam

Reklama • Koupit reklamu

SCANDISHOP.CZ – Specialisté na skandinávský nábytek**659 Kč**Černá židle BALI
MARK s bukovými...**13 890 Kč**Čalouněná postel
boxspring GLORIE...**10 990 Kč**Čalouněná postel
COMFINO VELVET**Služby**

Mapy



Bazar



Reality



Auta



Slovník



TV program



Letáky



Volná místa



Jízdní řády



Obrázky



Podcasty



Deníky

Další služby

Počasí • Česká republika**9°C**

Právě teď

**5°C**

Večer



Nastavit moji polohu

Meteoradar

Rádio**Expres FM**

To nejlepší ze světové hudby. Nastražte uši!

Krásné lži

kdo vyhrál volby

S Andrej Babiš je český politik a podnikatel slovenské národnosti z Bratislavě. V letech 2017 až 2021 byl předsedou vlády České republiky a lídrem hnutí ANO 2011, které se v české politice stalo dominantní silou. Jeho kariéra je spojena s řadou kontroverzí a kauz, které ho učinily jedním z nejvlivnějších lidí v zemi. Jeho činnosti, tak politických rozhodnutí.

Pro více informací můžete navštívit následující odkazy:

- [Andrej Babiš – Wikipedie](#) - Obsáhlý životopis
- [Andrej Babiš | Vláda České republiky](#) - Oficiální stránky
- [Andrej Babiš | ANO, bude líp](#) - Informace o hnutí ANO



Před 2 dny Před 12 hodinami Včera



Další významy jsou uvedeny na s
[politik a podnikatel slovenského](#)
[é republiky](#). Od října 2013 je
i 2014–2017 působil jako první

***„Já ale nechci produkt co mi
může lhát.“***

Oddělení marketingu



AI ve zpravodajství často chybuje

Modely (ChatGPT, Gemini, Copilot, Perplexity) dělají faktické chyby přibližně ve 2 z 5 odpovědí. Selhání nejsou výjimkou, ale systémový jev napříč jazyky i tématy.



Největší slabina: zdroje

Často uvádějí neexistující nebo nerelevantní zdroje, případně žádné. Uživatel pak nemá jak rychle ověřit pravdivost tvrzení.



Sebevědomé halucinace

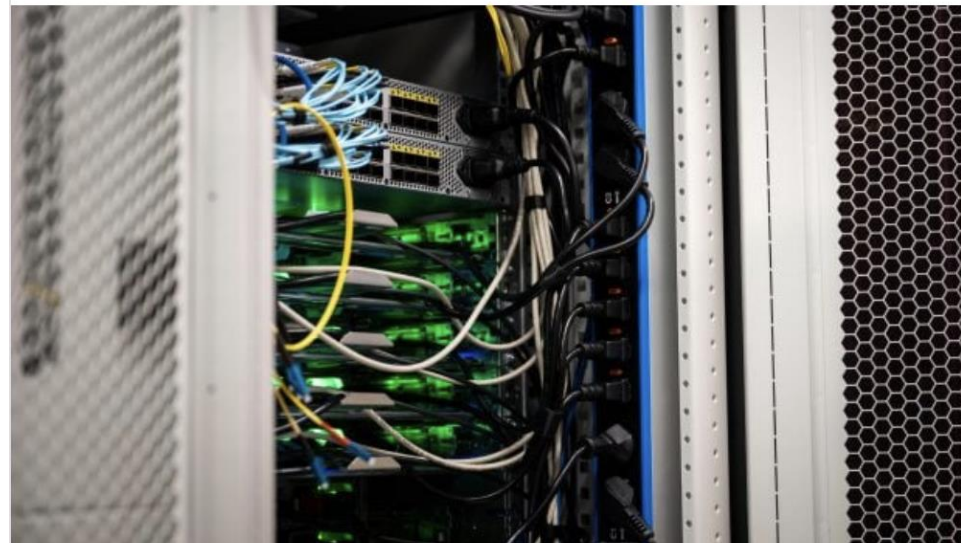
Modely jen výjimečně odmítnou odpovědět, i když „nevědí“. Trénink na přesvědčivý styl vede k jistě znějícím, ale občas falešným výroků.



Co s tím: testovat a ověřovat

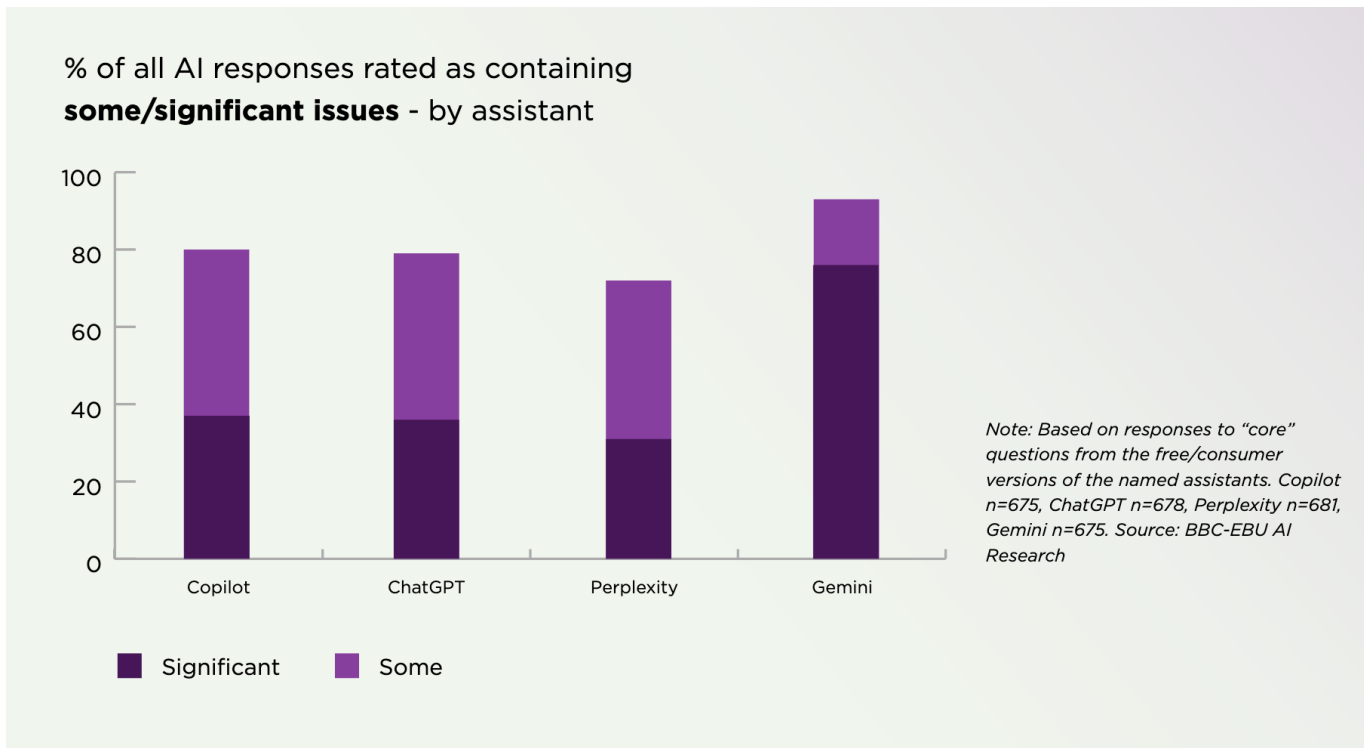
Média i firmy by měla pravidelně měřit přesnost a vyžadovat transparentnost. Uživatelé mají vždy klikat na zdroje a klíčová fakta ověřovat z více míst.

Výzkum: AI asistenti si vymýšlejí při popisování zpráv, vážně chybují téměř v polovině případů

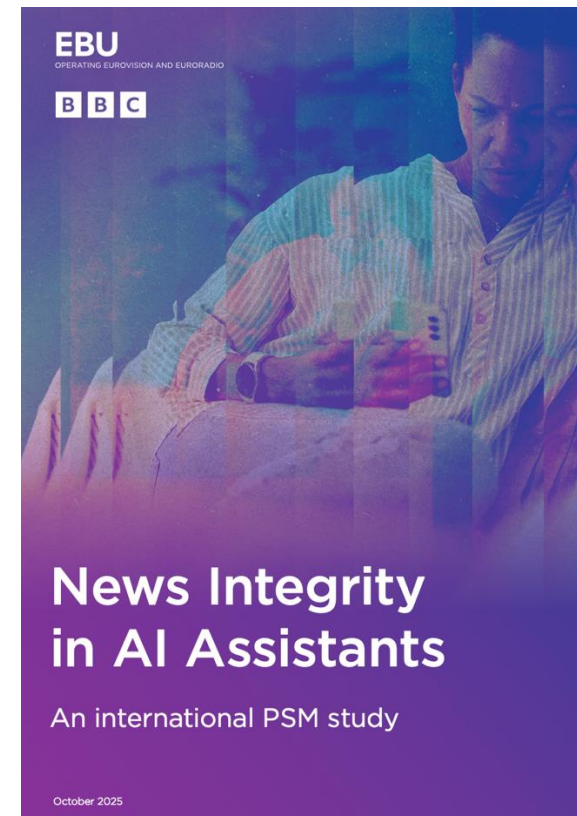


Serverovna datového centra T-Mobile DC7 | Foto: Zuzana Jarolímková | Zdroj: iROZHLAS.cz

Nesprávné či smyšlené citace, neexistující odkazy, názory vydávané za fakta nebo zastaralé informace. To jsou jen některé z chyb, které dělají velké jazykové modely při dotazech na aktuální zprávy. Podle mezinárodního výzkumu, kterého se účastnil i Český rozhlas, jde o plošný a systémový problém.



- **Sourcing:** neexistující/nesprávné odkazy, tvrzení bez odkazu, „ceremoniální“ citace
- **Vymyšlené/pozměněné přímé citace**
- **Zastaralé či časově pomíchané informace**
- **Názory prezentované jako fakta**
- **Nedostatečný kontext / vynechání klíčových informací**



Halucinace

Bolster Hallucination Detection via Prompt-Guided Data

„Halucinace LLM zůstávají nevyřešeným problémem“; navrhuje datově řízené zlepšení detekce. [\(zdroj\)](#)

LLM-based Agents Suffer from Hallucinations.

Ukazuje, že agentní systémy dědí (a násobí) halucinace; mapuje 18 spouštěčů a otevřené problémy. [\(zdroj\)](#) (09/2025) Lin et al.

Hallucination is Inevitable.

Teoreticky dokazuje, že úplná eliminace halucinací je pro LLM v obecném nastavení nemožná. Nejlepší je omezovat dopad. [\(zdroj\)](#) 02/2025, Xu et al.

OpenAI – Why language models hallucinate.

Tvrdí, že halucinace přetrvávají, protože trénink i evaluace odměňují hádání místo přiznání nejistoty. Navrhuje změnit metriky a penalizovat sebevědomé omyly. [\(zdroj\)](#) (09/2025)

Jak anotovat krásné lži?

ika?



Velmi časově náročné

Tisíce člověkohodin práce při označování dat, ověřování informací.

Co se dnes píše na sumperktimes.cz ?



Chce velmi zkušené a kvalitní anotátory.



Vytvářet předvýběr konverzací, kde hrozí riziko halucinací.

Zjišťuji aktuální informace ze sumperktimes.cz.

Dnes na sumperktimes.cz se píše:

Z domova:

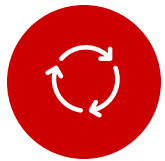
- Žena žijící v Egyptě dluží turistům přes 630 tisíc korun, státní zastupitelství prověřuje možné trestné činy. [Více](#)
- Rodiče v Kladně vozí šestileté děti v kočárku, aby se vyhnuli jízdnému, revizoři udělují stovky pokut měsíčně. [Více](#)
- Na jihu Moravy došlo k nehodám cyklistů způsobeným nadměrnou konzumací alkoholu během slunného víkendu. [Více](#)
- Žena z Poděbradska žalovala sousedku kvůli odleskům z rotační hlavice na komíně, spor dospěl až k Nejvyššímu soudu. [Více](#)
- Na jižní Moravě vzniklo dobrovolnické centrum po tornádu, které před čtyřmi lety zdevastovalo [Více](#)
- [Ladislav](#) nadále komplikovaný vzhledem k věku (88 let) a doprovodným onemocněním. [AP News +1](#)
- Později byl propuštěn z nemocnice po přibližně 38 dnech pobytu (k běhu března) a doporučeno mu několik týdnů až měsíců klidu a rehabilitace. [People.com +1](#)

Jak dosáhnout pravdy?



RAG-first architektura

Nejdřív dohledat, pak generovat; hybridní vyhledávání (lexikální + vektor), důraz na recall a krátký, vysoce relevantní kontext.



Iterativní retrieval & dekompozice dotazu

- self-ask, re-querying a re-ranking, dokud nejsou důkazy pokryté.
- „odpověz pouze s citací“; bez zdroje → odmítnout/požádat o dohledání.



Tool-use pro citlivé úlohy

kalkulačka, znalostní graf, databáze, doménová API u rizikových tvrzení vždy nástroj místo tipování.



Finetuning modelu?



Měření

- Klasifikace pravdy
- Post-hoc fact-checking pipeline
- Anotace

Najdeš lež?

Indri Verze 0.7+

53 jazyků

Článek Diskuse

Číst Editovat Editovat zdroj Zobrazit historii Nástroje

Indri (*Indri indri*) je jediným druhem monotypického rodu indri (*Indri*). Je to lemur z podřádu poloopic.

Výskyt [editovat | editovat zdroj]

Žije v primárních nížinných i horských deštných pralesích východního Madagaskaru až do nadmořské výšky 1500 m, převážně v chráněných oblastech: Ambatovaky, Analamazaotra, Anjanaharibe-Sud, Betampona, Mangerivola, Mantady, Marotandrano, Verezanantsoro a Zahamena.^[3]

Popis [editovat | editovat zdroj]

Indri je největší z přežívajících lemurů. váží průměrně okolo 7 kg, samice bývají těžší. Někteří jedinci dosahují váhy i 10 kg. Srst má velmi hustou, hedvábnou, zbarven je směsí černé a bílé, jedinci na severu bývají tmavší než na jihu. Délka těla s hlavou bývá 60 až 90 cm, ocas má zakrtnělý, pouze 5 cm. Jeho nohy jsou dlouhé a štíhlé, ruce má v porovnání

Indri



Stupeň ohrožení podle IUCN

vyhynulý	obecně ohrožený				málo dotčený	
EX	EW	CR	EN	VU	NT	LC

televize

Mesiáš? Krvavý měsíc je

slavnost plná barev a radosti -

tymiánem – perfektní slaný

Kde LLM může pomoci?

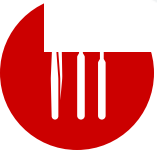


LLM je skvělý na komunikaci



Myslete na

Porovnej mi Samsung



největší pa



Prototypování

Cursor, Claude code, nahrazení encoder modelů v architektuře. Nasimulování Retrieval.



Srovnávací text a vysvětlovár

z faktů dodaných nástroji (RAG) udělá lidskou, srovná

Klíčové srovnání

Samsung Galaxy S25 FE

★★★★☆ 3 hodnocení [Uložit](#) [Porovnat](#) [Hlídat](#) [Nahlásit](#)

E Energetická třída, 6,7", Android 16, operační paměť 8 GB, rozlišení 2340 x 1080, 4900 mAh

Samsung Galaxy S25 FE je smartphone, který nabízí řadu funkcí pro každodenní použití. Je vybaven bezrámečkovým 6,7palcovým Dynamic AMOLED 2X displejem s rozlišením 2340 x 1080, což zajišťuje kvalitní zobrazení s obnovovací frekvencí 120 Hz.



8/128 GB Navy
od 12 562 Kč



8/128 GB
Awesome White
od 12 562 Kč



8/128 GB Jetblack
od 12 750 Kč



+ 10 variant
od 12 813 Kč

Klíčové srovnání		
Parametr	Samsung Galaxy S25 FE	Apple iPhone 16 Pro
Vodětightnost	IP68	IP68
Rychlé nabíjení	Ano	Ano
Operační systém	Android 16	iOS

Kde může LLM škodit?



Aktuální fakta

Cena, sklad, dostupnost, parametry, řazení produktů) – musí dodat retrieval/API.



Velmi dlouhé kontexty



Matematika, přesné výpočty, jednotky,

vždy kalkulačka/engine, ne LLM.



Deterministická byznys logika & compliance

(např. refund pravidla, regulační texty)



Bezpečnostně citlivé scénáře s nedůvěryhodným obsahem

LLM je zranitelné prompt-injection; vstupy z webu/RAG vždy sandboxuj a filtruj nástroji.



**Co musí mít
kvalitní LLM
aplikace?**





Transformace prompt inženýrů

Prompt inženýři se transformují do kontextových retrievalu toolu.



Dobré hledání

Zahazování nerelevantních dat před vložením do



Golden dataset Datová hygiena

Nevíme, jak se to má chovat nevíme co chceme
Binec v datech = nekvalitní odpovědi



Achilovy paty GenAI nahrazujeme architekturou.



Paměť

Chatbot musí mít vlastní paměť nesmí se spoléhat na kontext



Data

+

Machine Learning

=



Data

+

Artificial Intelligence

=



Data

+

Generative AI

=



Data

+

Agentic AI

=



**„Navzdory dnešním limitům už
GenAI přináší výrazný růst
produktivity a bude dál sílit.”**

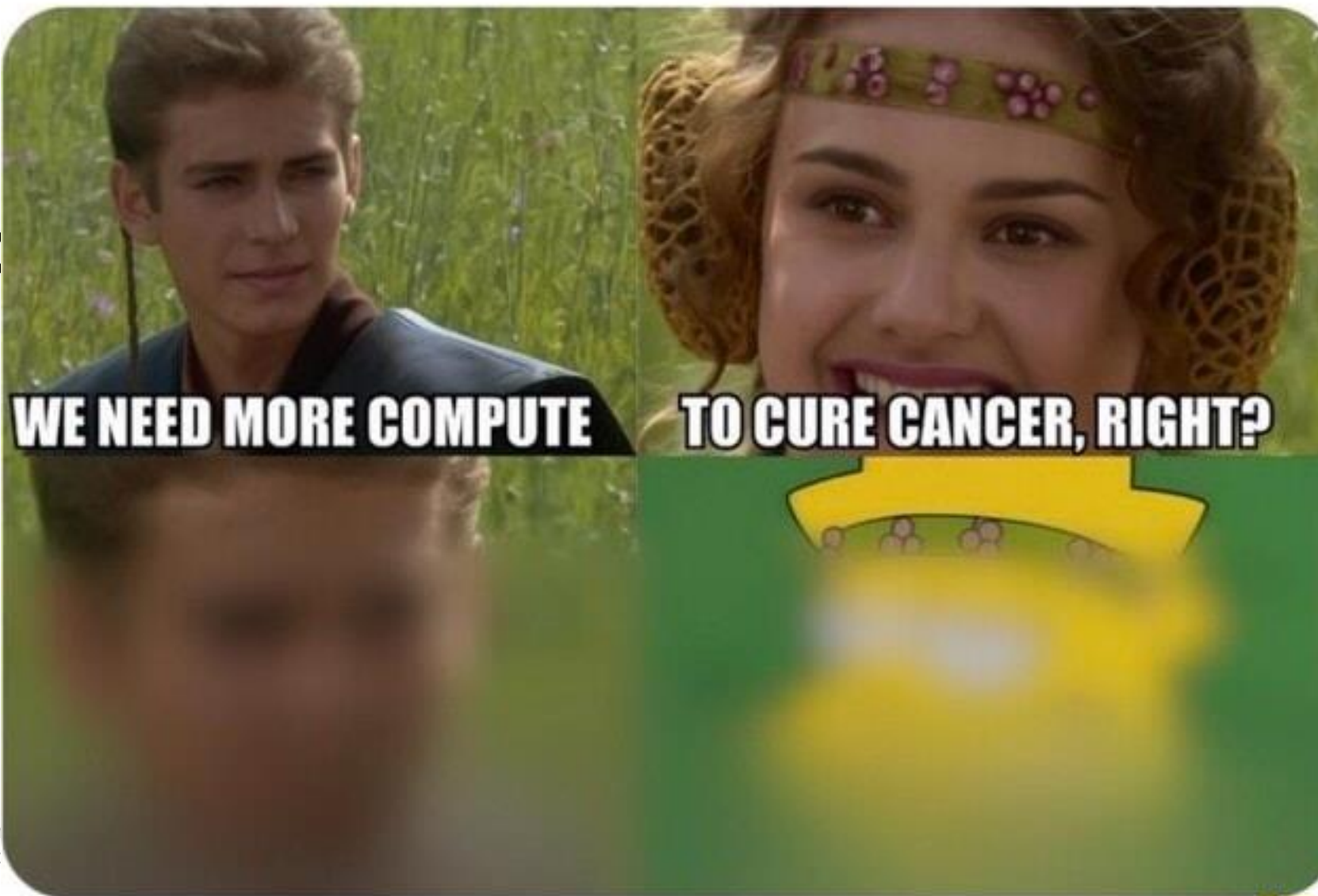
Andrej Karpathy



Finishing touches

AI zko
ji zati

ce



Zdroj: <https://www.gar...>
organizations-are-in-p





Radoslav Mikeš

Product lead, Seznam Asistent

radoslav.mikes@firma.seznam.cz

Copyright © 1996–2025 Seznam.cz, a. s.

