Liščí sylogismus

PREMISA 1

Slepice je v kurníku.

PREMISA 2

Kurník je v zahradě.

ZÁVĚR

Slepice je v zahradě.



Cíl: ulovit slepici. Zvolený postup: vydat se do zahrady.

Delší řetězec usuzování

Ilustrativní odpověď LLM propojuje polohu cíle s cestou k němu

Zahrada je na ostrově.
Liška začíná na pevnině.
Spojuje je průchozí most.



NAVRŽENÝ PLÁN

„Přejdi přes most, abys došla do zahrady.“

Předpoklad: most je jedinou přípustnou cestou.

Zdroje: [2]

Více ostrovů, více možných cest

Ostrovy

Místa jako vrcholy grafu

Mosty

Proveditelné přechody

Trasa

Posloupnost platných kroků



Zdroje: [3]

Co rozhoduje o spolehlivosti odpovědi?

Pravdivé premisy	Je slepice stále v kurníku?
Platné odvození	Vyplývá každý závěr z předpokladů?
Proveditelné jednání	Lze každý potřebný most použít?
Vhodně zvolený cíl	Splňuje trasa skutečné zadání?

Plynulost ani délka odpovědi samy o sobě nezaručují správnost.

Zdroje: [1] [3]

Podnikatel a odborník

Příspěvek zkoumá využití UI v praxi formou fiktivního dialogu.

PODNIKATEL

„Co může UI přinést mé organizaci?“

ODBORNÍK

„Pro jaký úkol, na základě jakých podkladů a na čí odpovědnost?“



Zdroje: [1]

Jazykové modely, RAG a agenti

Funkce	Co přináší	V příkladu s ostrovy
Jazykový model	Tvoří odpověď z naučených vzorců a poskytnutého kontextu	Navrhuje trasu podle popisu
RAG	Vyhledává externí podklady jako kontext pro odpověď	Nachází mapu a zprávu o stavu mostu
Agent	Volí kroky a používá povolené nástroje či akce	Ověřuje trasu a provádí povolené kroky

Vyhledávání dodává podklady. Delegování svěřuje pravomoci.

Zdroje: [1] [2] [4] [5]

RAG a kvalita podkladů

Generování rozšířené o vyhledávání

1	**Vyhledání**	Relevantní podklady, k nimž má uživatel oprávněný přístup
2	**Kontext**	Předání vybraných podkladů jazykovému modelu
3	**Generování**	Formulace odpovědi s dohledatelnými zdroji
4	**Ověření**	Kontrola aktuálnosti zdrojů, interpretace a citací

RAG může omezit nepodložené odpovědi. Spolehlivost je nutné ověřit.

Zdroje: [1] [4]

Delegování úkolů agentovi UI

ÚKOL

„Ověř trasu a navrhni plán.“

Provedení akce mění důsledky případné chyby.

OPRÁVNĚNÍ

Číst schválené podklady
Použít nástroj pro ověření trasy
Předat závažná rozhodnutí člověku
Zaznamenat kroky a nejistoty

Oprávnění vymezují rozsah delegování.

Zdroje: [1] [5]

Znalostní asistent v průmyslu

**„Stroj Y hlásí chybu 42.
Co mám dělat?“**

Určit model stroje
Vyhledat správný manuál
Uvést příslušný oddíl
Při chybějících podkladech požádat o pomoc

„Potvrďte prosím model stroje a verzi manuálu, abych mohl vyhledat schválený postup.“

Ilustrativní scénář. Odpověď závisí na skutečné schválené dokumentaci.

Zdroje: [1] [4]

Aplikace a podmínky jejich využití

Oblast	Užitečný úkol	Podmínka spolehlivého využití
Obchod	Vysvětlení dostupnosti zboží a podmínek vrácení	Aktuální údaje a pravidla, možnost předání člověku
Klinická administrativa	Shrnutí záznamů a příprava vysvětlení	Odborná kontrola a ochrana údajů pacientů
Průmyslový provoz	Vyhledání a výklad technické dokumentace	Správná verze a schválené podmínky provozu

Přínosy je třeba ověřit v konkrétním pracovním procesu.

Zdroje: [1]

Lidský a organizační kontext rizik

Neúmyslná selhání

Zastaralé podklady
Nesprávná syntéza
Cíle odměňující nežádoucí výsledek

Úmyslné zneužití

Manipulace
Zneužití osobních údajů
Záměrné překročení oprávnění

Účinný dohled vyžaduje podklady a pravomoc výsledek změnit nebo postup zastavit.

Zdroje: [1]

Různé druhy výpočetních limitů

Otázka	Druh omezení
Existuje algoritmus, který vyřeší každý případ?	Nerozhodnutelnost Příklad: obecný problém zastavení
Rozhodne pevně daná formální teorie každou aritmetickou větu?	Neúplnost Za přesně vymezených předpokladů
Lze řešitelnou úlohu vyřešit efektivně?	Výpočetní složitost Otázka P versus NP zůstává otevřená
Řeší tento model daný úkol spolehlivě?	Empirická úspěšnost Vyžaduje reprezentativní testy

Teoretické výsledky a testování modelů odpovídají na různé otázky.

Zdroje: [1] [6] [7] [8]

Kvantové počítání a výpočetní náročnost

Shor

Faktorizace a diskrétní logaritmus
Polynomiální čas v kvantovém modelu

Grover

Nestrukturované vyhledávání
O(√N) dotazů na orákulum

Praktický přínos závisí na hardwaru, korekci chyb a způsobu kódování úlohy.

NIST v roce 2024 schválil ML-KEM, ML-DSA a SLH-DSA.

Zdroje: [1] [9] [10] [11]

UI závisí na fyzické infrastruktuře

Celosvětová spotřeba elektřiny všech datových center

415 TWh

Odhad pro rok 2024

945 TWh

Projekce pro rok 2030
Základní scénář



IEA, Energy and AI (2025). Údaje zahrnují všechna datová centra, nikoli pouze UI.

Zdroje: [1] [12]

Regulace a odpovědnost

Právní rámec	Hlavní oblast
AI Act (EU)	Povinnosti podle použití, rizika a role organizace
GDPR	Zákonné zpracování a ochrana osobních údajů
Nařízení o digitálních službách	Odpovědnost online zprostředkovatelských služeb a platforem
Odvětvové právo USA	Odpovědnost za škodlivé či diskriminační dopady

Odpovědná osoba musí mít možnost systém přezkoumat, opravit a zastavit.

Zdroje: [1] [13] [14] [15] [16]

Pilotní ověření před nasazením

Krok pilotního ověření	Příklad s ostrovy a mosty
Vymezit úkol a výchozí stav	Dojít do určené zahrady po platné trase
Testovat běžné i změněné podmínky	Uzavřít most, přesunout cíl nebo odebrat podklady
Měřit celý pracovní proces	Správnost, chyby, nároky na kontrolu, čas a náklady
Určit odpovědnost a podmínky pokračování	Schválit rozšíření, systém upravit nebo jej zastavit

Rozšíření vyžaduje doložený přínos a přiměřené postupy pro zvládání chyb.

Zdroje: [1] [3]

Most musí existovat

Podložené premisy. Platné kroky.
Proveditelné akce. Odpovědná rozhodnutí.

„Aby byla technologie úspěšná, musí mít realita přednost před vztahy s veřejností, protože přírodu nelze oklamat.“

Richard P. Feynman, komise pro vyšetření havárie Challengeru, 1986 (český překlad)

Jaké podklady by odůvodnily svěření vašeho úkolu UI?

Zdroje: [1] [17]

Vybrané zdroje: usuzování a podklady

Zdroje pro diskusi po dvacetiminutové přednášce

[1] Snášel. Umělá inteligence a algoritmy v praxi

[2] Wei a kol. Chain-of-Thought Prompting. NeurIPS, 2022

[3] Valmeekam a kol. PlanBench. NeurIPS, 2023

[4] Lewis a kol. Retrieval-Augmented Generation. NeurIPS, 2020

[5] Yao a kol. ReAct. ICLR, 2023

Vybrané zdroje: limity a infrastruktura

Zdroje pro diskusi po dvacetiminutové přednášce

[6] Turing. On Computable Numbers. 1936–1937

[7] Gödel. O formálně nerozhodnutelných větách. 1931

[8] Clay Mathematics Institute. P versus NP

[9] Shor. Kvantová faktorizace a diskrétní logaritmy. 1997

[10] Grover. A fast quantum mechanical algorithm for database search. 1996

[11] NIST. První tři postkvantové standardy FIPS. 13. srpna 2024

[12] Mezinárodní energetická agentura. Energy and AI. 2025

Vybrané zdroje: regulace a odpovědnost

Zdroje pro diskusi po dvacetiminutové přednášce

[13] Evropská komise. Akt o umělé inteligenci

[14] Evropská komise. Vysvětlení ochrany osobních údajů

[15] Evropská komise. Nařízení o digitálních službách

[16] Ministerstvo spravedlnosti USA. Meta: dohoda o reklamě na bydlení. 2023

[17] Feynman. Personal Observations on Reliability of Shuttle. 1986