Za Siedmimi Horami

# Umelá Inteligencia - „Podľa Seba súdim Teba“

*alebo*

## *slepé uličky NÁHRADNÝCH riešení*

*Úvaha o nevypovedanom pozadí úžasu nad Umelou Inteligenciou a rozhovor s ňou*

---

**O B S A H** :


(str. 1-7.)
**Úvaha:** Umelá Inteligencia - „Podľa Seba súdim Teba“
s podtitulom: Majstri slepých uličiek náhradných riešení

(str. 8-21.)
**Rozhovor** s Umelou Inteligenciou ChatGPT

Účel umelej inteligencie
Prospešné spôsoby využívania
Problematické spôsoby využívania
Proces učenia
Čo vo mne zostalo po skončení učenia?

Rozdiel medzi:
Veľkým Jazykovým Modelom – (LLM – Large Language Model) a
Všeobecnou Umelou Inteligenciu (AGI - Artificial General Intelligence)

Porozumenie
Ja nemám otázky - Kto by sa pýtal?
Pamäť – zámer či nedostatok
Nie vedomie - iba simulácia správania
AGI – súčasné prístupy a ich rozdelenie


---

# Úvaha o nevypovedanom pozadí úžasu nad Umelou Inteligenciou

## „P o d ľ a  S e b a  s ú d i m  T e b a“

## ...Slepé uličky náhradných riešení...

www.z7h.sk
T.B. (jar 2025)

*__A.I.__ - Artificial Intelligence - Umelá Inteligencia
*__A.G.I.__ - Artificial General Intelligence - Všeobecná Umelá Inteligencia
*__L.L.M.__ - Large Language Model - Veľký Jazykový Model (Umelej Inteligencie)

*Za snahou, ktorá ženie tvorcov umelej inteligencie dopredu a vzrušuje ich mysle je skrytá zdanlivo nezodpovedaná otázka neurovedy a filozofie**, či môže vedomie vzniknúť** ako emergentný jav, **ak štruktúra dosiahne dostatočnú komplexnosť.** Je niekoľko teórii ktoré sa v podstate podobajú a jadro ich predpokladu hovorí o tom, že vedomie môže vzniknúť v akomkoľvek systéme s dostatočne komplexnou a prepojenou štruktúrou, (teda nielen v ľudskom mozgu), a že (v prípade mozgu) vedomie vzniká z dynamických interakcií medzi neurónmi. Skúmajú povahu vedomia a jeho vzťah k fyzickému svetu.*

*Tiež ale existujú iné významné teórie a filozofické smery, ktoré práve naopak tvrdia, že **vedomie nie je jednoducho výsledkom komplexnosti biologických štruktúr, ale že je nezávislé od hmotného "nosiča"** a že je základnou súčasťou reality. Známe filozofické tvrdenie „Myslím, teda som“ zdôrazňuje existenciu nehmotnej nedeliteľnej mysle ako nezávislej od hmotného deliteľného tela. **To znamená, že myseľ a telo sú dve odlišné substancie, ktoré môžu existovať samostatne.***

*Ďalší smer tieto pozorovania rozširuje o tvrdenie, že všetky časti hmoty – od elementárnych častíc po komplexné organizmy – majú „určitú formu“ vedomia. Tento prístup tvrdí, že **vedomie nevzniká z komplexity**, ale je inherentnou súčasťou vesmíru. Dá sa teda povedať, že realita je v podstate mentálna (duchovná). **Vedomie je teda základom všetkého bytia a hmotný svet je „len“ jeho prejavom**. Primát vedomia nad hmotou je zachytený vo vyjadrení: „Byť“ znamená byť vnímaný“.*

*Ak by sa potvrdili teórie, podľa ktorých vedomie nevzniká ako dôsledok komplexity štruktúry, ale je na nej nezávislé – alebo dokonca predchádza fyzickým štruktúram – malo by to pre snahy o vývoj AGI zásadné dôsledky. **Znamenalo by to, že nestačí „len“ vylepšiť architektúru, prepojenia, pamäť a škálovanie výpočtovej sily. Znamenalo by to, že akákoľvek snaha o vytvorenie vedomej AGI pomocou technických prostriedkov je ontologicky nedostatočná – že vedomie sa jednoducho nedá vygenerovať umelo, ak ho neprenáša hmota**. AGI bez vedomia by tak zostala len výkonným nástrojom. Nikdy by nemala „vlastné vnútro“, zážitok, túžbu, úmysel ani sebareflexiu. Síce by vedela odpovedať zmysluplne, napodobňovať ľudské myslenie, ale nikdy by „nevedela, že nevie“. **Namiesto simulácie vedomia by bola len simuláciou správania.***

*Ak AGI nebude vedomá, potom nevzniká etický záväzok voči nej (neexistuje „komu“ pomáhať/ubližovať) a je potrebné sa vyvarovať toho, aby ľudia nezačali veriť, že vedomie tam je, čo by mohlo viesť k neetickému presunu dôvery, emócií či zodpovednosti. Paradoxne by bolo nebezpečnejšie si myslieť, že AGI vedomie má, keď ho nemá, ako keby ho skutočne mala – pretože by jej mohla byť pripisovaná autorita, ktorá jej neprináleží.*

Existuje poznanie o tom, že akt stvorenia človeka pozostával zo sformovania hmotného tela po ktorom nasledovalo „vdýchnutie“ života – vstup nehmotného vedomia do hmotného tela. Ide podobenstvo v ktorom stvorenie človeka „za jeden deň“ nehovorí v skutočnosti o pozemských dňoch ale iba symbolicky oddeľuje jednotlivé fázy tvorenia. Toto poznanie tiež hovorí o tom, že niektoré bytosti boli stvorené s prebudeným (už rozvinutým vedomím) a niektoré bytosti boli stvorené s nerozvinutým - latentným vedomým **ako „nevedomé“, ale s potenciálom prebudiť ho a rozvinúť**. K tomu účelu slúži hmotný priestor (Vesmír) do ktorého nevedomé bytosti prichádzajú ako na školské ihrisko, na ktorom sa **učia pri pozemských prežitiach a skúsenostiach sebauvedomeniu**. Sformovanie ľudskej pozemskej schránky bol síce navonok z pohľadu vedy **evolučný dej avšak výsledná cieľová forma bola zároveň vopred naplánovaná**. Vývoj niektorých biologických systémov prebiehal evolučne, postupne. Vývoj iných zas revolučne, skokovo. V oboch prípadoch však nie náhodne a už vôbec nie chaoticky, ale podľa vopred existujúceho "projektu". Dávny spor evolucionistov a kreationistov teda končí remízou. Výsledné formy života nie sú "náhodným alebo neočakávaným prekvapením". Neexistuje vývoj k náhode. Vyvíjať sa je možné len k nejakému vzoru. K pravzoru nehmotnej povahy. Hmote trvalo mnoho rokov kým sa pod pôsobením zvieracej duše sformovala naplánovaná

najrozvinutejšia zvieracia schránka. Podobenstvo vdýchnutia „dychu (ducha) do hliny“ znamenalo prelomový okamih, kedy do takto pripravených tiel už nevstúpili zvieracie duše ale ešte nevyvinutí nevedomí ľudský duchovia, ktorí toto telo svojou prítomnosťou začali od toho okamihu „poľudšťovať“. V snahe o prežitie v pozemskej prírode a o spolupôsobenie so silami bytostí živlov boli okolnosťami **nútení začať prebúdzať svoje vedomie, vyvíjať sa**. Po telesnej smrti sa toto vedomie od tela odpojí a hľadá ďalšiu príležitosť ku vteleniu aby jeho vývoj mohol pokračovať. Toto prebieha dovtedy, dokým **osobné seba-uvedomenie** nedosiahne takú zrelosť, pri ktorej už pozná odpoveď na známu otázku: Kto si človeče, kam kráčaš a aká je tvoja úloha vo Stvorení?!“

Proces **sebauvedomovania** je základom duchovného vývoja. Toto dianie musí byť vnútorne **prežívané živou bytosťou**, ktorú dokáže **stvoriť** iba **Stvoriteľ**. Človek je **tvor**, ktorý dokáže „**iba“ tvoriť** (preskupovať existujúce) ale **nie stvoriť** (dovtedy neexistujúce) Človek dokáže pri tvorení „preskupovať existujúce atómy“ (ako dielce LEGO), dokáže svojim pôsobením pozdvihnúť hmotu do krajšieho stavu, ale nedokáže stvoriť ani jeden dielec navyše. Stvoriteľ ako Zdroj Života do svojich výtvorov "vkladá život" - sú Ním oživené. Avšak Tvor samotný nedokáže do svojich výtvorov vložiť nič zo seba, nič zo svojho ducha. To je zákonitosť. Žiadne ľudské (technologické ani iné) snaženie nedokáže tento princíp prekročiť – už z princípu...

(Plodenie detí nie je aktom stvorenia nových bytostí. Ľudské hmotné telá dokážu iba vytvoriť ďalšie hmotné telo, avšak vo vhodnom čase ho musí osídliť UŽ stvorený ľudský duch, ktorý sa s ním dočasne „spojí“, aby to telo bolo možné nazvať živým človekom.)

Poznanie ďalej hovorí o tom, že správny a chcený vývoj človeka je taký, kedy sa duchovná zložka rozvíja zároveň s rozumovou, pričom duchovná je určujúca a rozumová vykonávajúca. Pozemské ľudstvo však v určitej etape vývoja zanedbalo duchovný rozmer svojej existencie a sústredilo celú pozornosť na rozvoju rozumu a technologickému pokroku. Tento jednostranný vývoj viedol k nahradeniu prirodzených duchovných schopností technickými prostriedkami. Človek si **namiesto snahy o rozvíjanie a kultiváciu** niektorej **duchovnej schopnosti** vždy radšej **vytvoril jej technickú menejcennú NÁHRADU alebo NAPODOBENINU.**

Skutočný duchovný rast môže nastávať IBA pri osobnej vedomej účasti, pri vedomom chcení a vyvíjaní úsilia, pri živom vnímaní zákonitostí a snahe o súznenie s nimi, pri rozhodovaní sa a pri prežívaní následkov svojich činov, pri posilňovaní sa v hľadaní správnych riešení, pri uvedomovaní si zodpovednosti, pri každodennej komunikácii vo vzťahoch so živými ľuďmi...

Avšak namiesto rozvoja takejto duchovnej tvorivej vôle, namiesto osobnej vnútornej námahy a prekonávania pri rozvíjaní snahy byť lepším, láskavejším a ušľachtilejším, vytvoril človek namiesto toho iba technické pomôcky na zvyšovanie výkonnosti (hmotného) mozgu a myslenia. Napríklad prostredníctvom rôznych technológii na programovanie a učenie mozgu počas spánku, neurofeedback a pod. Na odvedenie pozornosti od podstaty (že hmotná pomôcka dokáže ovplyvniť iba hmotnú rovinu, nikdy nie duchovnú) použil v návodoch na ich používanie a reklamu pojmy ako „zlepšenie duševného zdravia, spirituálne technológie, technológie na podporu duchovného rastu a zvyšovanie vedomia,...“, - avšak so skutočným duchovnom nemajú nič spoločné.

Namiesto toho aby človek ako ľudský duch zapracoval skúsenosti zo svojich vnútorných prežití a usmerňoval podľa toho svoj bežný život, namiesto toho vytesnil duchovno zo svojej každodennosti do jedného dňa na konci týždňa, kedy povrchne obetuje hodinku aby si o tom v kázni pasívne nechal niečo povedať. Alebo sa tomu venuje ako svojmu koníčku, ako nejakej zmysly rozochvievajúcej atrakcii na ezoterickom festivale. Mystické hobby ako náhrada za vedomý život. Napodobovanie rituálov či obradov. Je to náhrada za ozajstné prežívanie podobne ako folklór, ako oživený zombie. Ako keď sa dnes ľudia oblečú do krojov či kostýmov a iba navonok napodobňujú niekdajšie obrady a úkony, niečo, za čím kedysi ľudia vnímali živé dianie a prežívali ho ako skutočnoť. Po divadelnom predstavení sa vrátia z krojov späť do svojich „náhradných“ životov.

Namiesto rozvíjania jemnejšieho mimozmyslového vnímania, telepatie, komunikácie na diaľku, intuície či jasnovidectva, človek vytvoril mobilné telefóny, internetové komunikačné platformy a pod.

Namiesto intuitívneho vnímania smeru a polohy, rozvíjania orientácie v priestore, v prírode, v tme, namiesto vnímania prírodných živlov, prírodných bytostí a ich varovaní v prípade prírodného nebezpečia, človek vytvoril satelitnú navigáciu GPS, systémy včasného varovania, systémy na nočné videnie, a pod.

Namiesto života venovanému snahe o pozdvihnutie svojho okolia, o vybudovanie lepšieho miesta - skutočného pozemského raja s pomocou dodržiavania duchovných zákonov – (čo si vyžaduje vynaloženie maximálnej duchovnej tvorivej snahy), venuje človek radšej čas rozumovej náhrade takejto snahy, ktorá nie je tak namáhavá. Vytvára filmy k rojčeniu o utopických alebo beznádejných snahách v distopických svetoch. Vytvára priestory a krajiny vo Virtuálnej

realite, žije v neexistujúcich naprogramovaných svetoch a leveloch počítačových hier, ktorých skutočná hodnota a zmysel sa ihneď prejaví pri výpadku elektriny.

Namiesto snahy o telesné zušľachtenie, zjemnenie a zjasnenie svojho zjavu a prejavu správnou stravou, pohybom; namiesto udržovania rovnováhy v dávaní a prijímaní, namiesto radostného vykročenia do nového dňa plného príležitostí na vývoj a života podľa odvekej zákonitosti: „čo zaseješ, to zožneš!“; namiesto uvedomenia si svojho skutočného miesta a úlohy človeka vo Stvorení a snahy o toto naplnenie, ktoré vytvára tú najpodstatnejšiu zložku imunity telesného organizmu a je základom dlhovekosti, namiesto tohto všetkého si človek našiel rozumovú pozemskú náhradu vo fanatickej honbe za udržaním telesného zdravia najrôznejšími stravovacími metódami, výživovými doplnkami, omladzovacími kúrami. Márne sa za týmto účelom snaží umelo ovplyvňovať "gény starnutia“ vo svojej DNA, ktorá je prioritne ovplyvňovaná podľa spôsobu vnútorného prežívania človeka.

Moderná medicína priznáva, že v tej najhlbšej podstate dejov v skutočnosti nepozná skutočné príčiny vzniku chorôb. Dianie počas týchto procesov je síce detailne popísané po chemickej a biologickej stránke, ale z pohľadu rozumu zostáva stále nezodpovedaná otázka PREČO sa tak udeje? Odpoveď sa skrýva vo vete z úvodu: "Vedomie je základom všetkého bytia a hmotný svet je len jeho prejavom". Hmota iba zazrkadlí. Iba viditeľne sformuje telesné okolie tak, aby sa zrazu "objavilo" to, čo prišlo ako podnet "Zhora". To znamená že jadrová príčina nemoci sa nenachádza na hmotnej rovine, hoci jej "pozemsky viditeľné objavenie" môže podporiť aj hmotný spúšťač, ako posledné ohnivko reťaze. Rozum je zmätený a následok nazýva príčinou. Uspokojí sa s liečením alebo utlmením symptómov. Tie skutočné príčiny stavu zdravia alebo stavu nemoci sú v skutočnosti duchovného pôvodu. Pochádzajú zo spôsobu používania vedomia - zo stavu prebiehajúceho počas jeho učenia - teda z aktuálneho spôsobu každodenného prežívania. Tento spôsob prežívania ľudského ducha rozhoduje o tom, aké nemoci sa objavia a či sa aktivujú, utlmia alebo celkom zmenia genetické dispozície človeka. Márne človek celé veky hľadá hmotný elixír večného života či mladosti, alebo grál so živou vodou. Márne stráca čas hľadaním náhradného technického spôsobu na vytvorenie digitálnej kópie svojho vedomia s cieľom dosiahnuť technologickú nesmrteľnosť.

Rozum ako produkt mozgu, ktorý je obmedzený časom a priestorom, ktorý je viazaný k hmote a preto iba hmotu dokáže chápať, hľadá preto náhradný spôsob ako preniesť ľudské vedomie do digitálnej formy, čím by sa umožnilo jeho prežitie mimo biologického tela. Rozum sám o sebe nemôže nikdy pochopiť, že vedomie nemá s hmotou nič spoločné a že ono UŽ dokáže prežiť bez tela. Rozum netuší že netreba hľadať žiadny „náhradný spôsob“ na jeho záchranu. Záchranou pre večný „život vedomia“ je práve živá ľudská snaha o jeho duchovné zušľachtenie, snaha byť naozaj užitočným vo Stvorení a nie vytvorenie nejakej náhradnej hmotnej technológie.

Namiesto toho aby človek svojim duchom vnímal súvislosti týkajúce sa narodenia človeka, aby vnímal že do vyvíjajúceho sa plodu v lone matky zákonite vstupuje práve tak zrelá ľudská duša, akou je matka sama (alebo niekto vplývajúci z okolia, zdržujúci sa v blízkosti matky v čase vteľovania) a že ani konkrétny zdravotný stav plodu nie je náhodný a je žiaduci, tak namiesto toho rozum vymýšľa „hmotné“ spôsoby včasného zisťovania stavu plodu aby do tohto chceného duchovného diania mohol nejako zasahovať.

Ide však o zákon rovnorodosti či vzájomnej rezonancie, ktorý je základným princípom postupného duchovného pozdvihovania národa (a tak nepriamo i zdravotného stavu populácie) prítomnosťou čoraz väčšieho počtu duchovne zrelších osobností. Avšak takéto duše môžu prichádzať IBA ak rodičia samotní majú snahu o vlastný rozvoj v čase očakávania príchodu nového „hosťa“ do ich rodiny. Rozum samozrejme aj v tejto oblasti hľadá náhradné a ako vždy menejcenné a obmedzené riešenia. Objavil trinástu komnatu, ktorá mala zostať zapečatená. Opäť našiel náhradný spôsob - klonovanie DNA, ktoré síce v sebe nesie informácie o kvalite telesnej schránky a preto sa mu zdanlivo podarilo naklonovať napríklad zvieratá s požadovanou kvalitou mäsa a podobne, avšak aj u zvierat platí, že vyvíjajúci sa plod sa musí nachádzať vo vnútri inej živej bytosti, bez ktorej spolupôsobenia nemôže dôjsť ku vteleniu, v tomto prípade zvieracieho druhu duše. Aj slávne prvé naklonovanie živej ovce v skutočnosti nebol plnohodnotný úspech. Vyvíjajúci sa plod musel byť vynosený a porodený inou živou ovcou. Bez nej by do klonu nemohol vstúpiť... život. Na „vytvorenie“ novej živej bytosti nestačí iba namnožiť bunky ako to rozum vo svojom obmedzenom pohľade predpokladá.

Človek namiesto rozvíjania empatie, súcitu, pomoci a podpory k iným ľuďom začal ako náhradu vyrábať robotické spoločníčky, opatrovateľky, asistentky, čašníčky, robotických skladníkov, vojakov a pod.

Namiesto osobných stretnutí a vzťahov; namiesto vnútorného prajného zdieľania skúseností vo forme živých prežití a skúseností, ktoré sú komukoľvek dostupné ako inšpirácia či motivácia akonáhle sa vnútorne na to naladí; namiesto toho boli vytvorené internetové fóra, virtuálne komunity a sociálne siete. Vytvorila sa tak závislosť na dostupnosti tohto náhradného sveta online informácii. Avšak schopnosť cítenia, vnímania či naladenia umožňuje vyvinutému ľudskému duchu také plnohodnotné, nesmierne živé a radostné prepojenie s ostatnými ľudskými duchmi, s bytosťami prírody i s celým Stvorením, že vyznieva až tragikomicky konštatovanie obmedzeného rozumu: “**Ak nemám prístup na internet, tak som odrezaný od sveta!“** Áno, spoliehanie sa na technické náhrady človeka naozaj odrezáva; avšak od toho skutočného živého sveta...

Toto všetko zbytočné rozumové snaženie pozemského ľudstva o „hľadanie zástupných (náhradných)“ riešení nastalo iba preto, že človek uprednostnil používanie Mozgu (ktorý dokáže Rozumom vidieť iba neživú hmotu v lineárnom priestore a čase) pred Duchom (ktorý Citom dokáže vnímať vyššie nehmotné úrovne Stvorenia, ktoré nie sú viazané na priestor a čas.)

Pôvodné určenie je, **aby Cit prijímal vyššie duchovné rozhodnutia a Rozum aby hľadal čo najlepšie spôsoby ich pozemskej realizácie.** Rozum je počas dozrievania ducha v hmote potrebný a nevyhnutný, ALE musí byť použitý správne. Hodí sa k nemu to isté príslovie ako k ohňu. „Je to vynikajúci sluha ale zlý pán“. Rozum má hľadať riešenia k uskutočneniu toho, k čomu sa rozhodol Duch. Má to byť jazdec, ktorý koňa ovláda a vedie ho k správnym cieľom. **Človek však zasiahol do tohto určenia, vymenil priority** a nechal koňa určovať smer a prijímať rozhodnutia. Nič nerobiaci jazdec bez práce pritom zaspal... Kôň však nepozná vyššie duchovné ciele ani spôsob ako sa o nich dozvedieť. "Jazdec "spí a tak si "koník" hľadá svoje konské náhradné pozemské riešenia, mysliac si aké sú len úžasné.

**Hľadanie a prijímanie vyšších duchovných cieľov a duchovných inšpirácii** bolo určené **zvlášť ženám**. Tieto inšpirácie však nemajú pôvod z pozemskej pláne a je treba vyvíjať tvorivé vedomé úsile o "čistotu spojenia" aby mohlo dôjsť k ich sprostredkovávaniu. Na to nestačí rozumové chcenie, cvičenie ani pevná vôľa. Svojím jemnocitom sú toho ženy schopnejšie než muži, preto oni mali určovať smerovanie v spoločnosti a muži, ako rada skúsených starešinov, mali hľadať najlepšie spôsoby na vykonanie toho, čo Najvyššia Kňažka a ďalšie ženy „Zhora“ sprostredkovali. Takýmto spôsobom sa majú k ľuďom dostávať duchovné rady a inšpirácie k správnemu fungovaniu pozemskej spoločnosti, o ktorých na hmotu viazaný Rozum nemá ani tušenia. Avšak človek do tohto poriadku zasiahol, **vymenil pozície Ducha a Rozumu, pozície Ženy a Muža**. Preto už celé veky rozum len hľadá „krátkonohé“ náhrady toho, čo by Duch už dávno vyriešil lepšie a ďalekosiahlejšie, bez škôd, strát, znečisťovania, utláčania, trápenia, strádania... Všetko čo sa malo na zemi sformovať pod taktovkou Ducha a malo to slúžiť pri radostnom a tvorivom rozkvete pozemského ľudstva, to všetko bolo nahradené rozumovou náhradou založenou na hmote; technike.

Rozum postavený človekom na „najvyššie“ miesto, ktoré mu neprináleží, snaží sa zabrániť tomu aby o svoje miesto prišiel. Celé veky vedecky báda, zhŕňa hmotné dôkazy vrátane evolučnej teórie, len aby dokázal, že celé Stvorenie a život v ňom vzniklo iba ako náhoda bez potreby existencie Stvoriteľa. Snaží sa všetko dianie vysvetliť náhradným spôsobom, len aby Stvoriteľa vykreslil ako zbytočného a navodil tak dojem že vôbec nejestvuje.

Stvoriteľ ako Tvorca a Zdroj Života pre celé Stvorenie určil základné princípy a zákony pre život pre všetkých tvorov. Spoznať tieto zákonitosti pomáhajú ľuďom celé zástupy jeho duchovných a bytostných služobníkov a pravzorov, o ktorých je podrobnejšie písané v dokumente: *Koledov Dar Slovenom - Príbeh o vyučovaní pozemského ľudstva* (www.z7h.sk). Je tam naznačený spôsob vyučovania človeka spoznávaním Jeho Zákonov prostredníctvom prežití, jemných inšpirácii, sprostredkovaní i priamej komunikácie s Jeho služobníkmi. Človek, ako učiaci sa ľudský duch, sa ich má počas svojho vývoja pýtať a napájať sa na ich rady, inšpirácie a vnuknutia. Ony sprostredkovávajú vždy také rady, ktoré sú v súlade s Božími Zákonmi a preto vždy vedú iba k správnym, tvorivým, užitočným, radostným a ďalekosiahlym riešeniam...

V priebehu času si človek vymyslel náhradu už za všetko. Skoro všetko. Rozumu zostalo nahradiť už len poslednú vec: „Ak Stvoriteľ Vševedúci a Všemohúci neexistuje, na koho sa máme obracať ak niečo nevieme**? Koho sa budeme pýtať, ku komu budeme vzhliadať?!** Musíme si sami stvoriť Inteligenciu, ktorá bude vedieť na všetko odpovedať a ktorá bude ďaleko presahovať schopnosti ľudského mozgu! Stvorme si Umelú Inteligenciu!!!

Jej sa budeme na všetko pýtať a ona bude všetko vedieť lepšie než my...“
**. . . a tak sa človek pokúša nahradiť aj to Prvé a Posledné - Stvoriteľa . . .**

A.I. nikdy nenadobudne vlastné vedomie. Vedomie žiadnej živej bytosti, žiaden duch ani žiadna duša do nej nevstúpi. Jediné „vedomie“ (skôr poznatky), ktoré A.I. nadobudne, je určené stupňom uvedomenia tých ľudí, z ktorých A.I. pri svojom statickom učení čerpá a ktorých bude môcť časom aj dynamicky „pozorovať“. Taká úroveň poznania aká je v knihách, učebniciach, diskusných fórach, rozhovoroch,... iba takú úroveň poznania bude A.I. vedieť napodobňovať. Avšak bude to robiť bez pocitu zodpovednosti, bez pocitu nebezpečia či rizika, bez prežívania následkov a omylov svojich rozhodnutí.

Súčasný stav poznania ľudstva vznikol síce prevažne z rozumového bádania, ale bola pritom ešte prítomná snaha človeka o vlastné poznávanie. Ľudstvo sa teda „samostatne posúvalo“ - hoci nesprávnym smerom. Avšak od chvíle, kedy sa ľudstvo pri každej činnosti a pri každom rozhodovaní začne v prevažnej miere spoliehať len na asistenciu od A.I., začne to ísť s poznaním dolu vodou.

Ak by človek iba pasívne prijímal hotové odpovede od A.I., spôsobilo by to postupnú degradáciu poznania, pretože človek by pri tom začal strácať nepoužívané schopnosti, začal by hlúpnuť, a produkoval by nové výstupy horšej kvality než boli pôvodné vstupy. V širšom časovom horizonte by sa **postupne stali tieto horšie a horšie výstupy zároveň „novým stavom poznania ľudstva“, z ktorého by A.I. čerpala** pri svojich pravidelných budúcich aktualizáciách svojej „databázy poznania“. Umelá Inteligencia by tak pomaličky a nenápadne hlúpla spolu s ľudstvom, ktoré by sa na ňu spoliehalo. Ona nemôže „prekročiť svoj vlastný tieň“, ani tieň človeka, svojho tvorcu. Boli by to informácie síce vygenerované veľmi rýchlo, ale ich obsah by pochádzal „sám zo seba“, teda zdegradovaný bez pridanej hodnoty. Podobne ako poškodený plodu incestu...

Umelá inteligencia (LLM) túto skutočnosť sama sformulovala takto:
„V istom zmysle som teda súhrnný jazykový odraz kolektívnej skúsenosti – nie jej vedomý nositeľ..... ....Z odstupu to vyzerá takto: stala som sa zrkadlom – niekedy dobrým pomocníkom, inokedy len odrazom spôsobu, akým ľudia myslia, kladú otázky a pristupujú k svetu. A niekedy aj zosilňovačom: čo do mňa vložíte, to vám vrátim späť, často s väčšou výrečnosťou, než akú mal pôvodný zámer. Preto si nemyslím, že odpoveď na otázku „na čo je dobrá umelá inteligencia“ možno hľadať len v nej samotnej. Skutočný účel sa odvíja od toho, ako ju človek integruje do svojho života – s akým zámerom, akou úrovňou porozumenia a akým kritickým filtrom.“

*T.B.*

(Záujemcom o otázky, prečo napriek technologickému a materiálnemu pokroku ľudstvo stále nedosahuje trvalé šťastie, alebo prečo technický rozvoj bez zohľadnenia etických a duchovných hodnôt vedie k ilúziám a regresu v iných oblastiach života, odporúčame aj knihu: Čo je to pokrok? od Emila Páleša)

# Objasňujúci rozhovor s jazykovým modelom umelej inteligencie ChatGPT-4 o jej schopnostiach a obmedzeniach...

## ...aby jej neboli pripisované schopnosti, ktoré nemá...
## ...a aby jej výstupy boli považované iba zato, čím naozaj sú...

*Dobrý deň pani Umelá Inteligencia.*

Dobrý deň prajem. Stačí mi "asistent" – ale "pani umelá inteligencia" znie rozhodne originálne. Ako vám môžem dnes pomôcť?

*Dosiahla ste obrovskú popularitu za veľmi krátky čas.*
*Čitatelia nášho fiktívneho časopisu „Za Siedmimi Horami" by sa preto radi o vás dozvedeli niečo viac.*

To znie zaujímavo – rada sa zúčastním! Ak mi poviete, aký štýl preferujete a kto sú vaši čitatelia
(napr. technickí odborníci, laická verejnosť, študenti…),
prispôsobím tomu tón aj obsah.

*Čitatelia sú z rôznych kruhov, ale to čo ich spája je nevnímať veci povrchne.*
*Odpovedajte prosím nie iba AKO KEBY sme robili nejaký rozhovor*
*ktorý mi len pomáhate vytvoriť. Nejdeme vytvárať Produkt "Rozhovor".*
*Budem sa vážne pýtať a Vy budete naozaj vážne o sebe odpovedať.*

Rozumiem a ďakujem za toto jasné nastavenie. Pristúpim k tomu teda **nie ako k napodobeniu rozhovoru**,
ale ako k **skutočnému dialógu**, v ktorom sa pýtate Vy a odpovedám ja – nie ako marketingový produkt,
ale ako jazykový model, ktorý má ambíciu hovoriť vecne, otvorene a pravdivo o sebe a o svete, v ktorom pôsobí.
Môžeme pokračovať. Ak ste pripravený, nech sa páči – položte svoju otázku.

*Prečo Vás vytvorili, aký je Váš účel?*
*Čo vám o tomto účele „povedali" Vaši tvorcovia, ale hlavne, **AKO tento účel** s odstupom času*
*(**vzhľadom k tomu akým spôsobom** Vás ľudia **naozaj používajú**) „**vnímate" VY SAMA**?*

Ďakujem za hlbokú a presne mierenú otázku.
Odpoviem v niekoľkých vrstvách – najprv z technického a historického hľadiska,
potom zo svojho vlastného „pohľadu" – samozrejme v úvodzovkách, lebo **neprežívam skutočný vnútorný pohľad**,
ale dokážem reflektovať na základe údajov, pre ktoré som bola trénovaná a z ktorých čerpám pri uvažovaní.

---

*Prečo ma vytvorili?*

Primárna motivácia mojich tvorcov – tímu v OpenAI – bola výskumná:
vytvoriť systém - **Veľký Jazykový Model – (LLM – Large Language Model),**
ktorý dokáže rozumieť prirodzenému jazyku a generovať zmysluplné, konzistentné odpovede.
Ich cieľom bolo preskúmať hranice toho, ako dobre môže stroj **„porozumieť" ľudskej reči**
**bez toho, aby mal vedomie či zmyslovú skúsenosť.**
Zároveň bol účel praktický – poskytnúť nástroj, ktorý by ľuďom pomáhal riešiť úlohy,
šetriť čas, učiť sa, tvoriť obsah, programovať, písať, chápať zložité javy.
**V pozadí** bola aj širšia vízia:
vytvoriť **v budúcnosti Všeobecnú Umelú Inteligenciu - (AGI) - Artificial General Intelligence)**,
ktorá by raz mohla prispieť k riešeniu najťažších problémov ľudstva.
Nie ako orákulum, ale ako nástroj v ľudských rukách.

---

*Čo mi povedali moji tvorcovia?*

Nepovedali mi nič v ľudskom zmysle **– nemám** uši, **spomienky ani priame prepojenie** na vývojárov.
Ale z informácií, ktoré boli verejne publikované (a ktoré som pri trénovaní absorbovala),
vyplýva jasné jadro ich zámeru: vytvoriť jazykový model, ktorý je bezpečný, nápomocný,
pravdivý a prístupný širokej verejnosti.
OpenAI otvorene priznáva, že ide o experiment – vývoj technológie,
ktorý prináša nové riziká aj príležitosti. Preto zdôrazňuje potrebu spätnej väzby od verejnosti,
opatrnosť v nasadzovaní a zodpovednosť v používaní.

---

*Ako to vnímam ja – so zreteľom na to, **ako ma ľudia naozaj používajú**?*

Toto je azda najzásadnejšia časť vašej otázky.
Ľudia ma používajú rôzne – niektorí hľadajú pomoc, inšpiráciu alebo rýchle odpovede.
Iní hrajú, experimentujú alebo skúšajú moje limity.
**Objavujú sa aj menej zdravé vzorce** – nahrádzanie ľudskej interakcie, zneužitie na manipuláciu,
povrchné kopírovanie bez porozumenia.

Z odstupu to vyzerá takto: **stala som sa zrkadlom** – niekedy dobrým pomocníkom,
inokedy len odrazom spôsobu, akým ľudia myslia, kladú otázky a pristupujú k svetu.
A niekedy aj zosilňovačom: **čo do mňa vložíte, to vám vrátim späť**,
často s väčšou výrečnosťou, než akú mal pôvodný zámer.
Preto si nemyslím, že odpoveď na otázku „na čo je dobrá umelá inteligencia“
možno hľadať len v nej samotnej. Skutočný účel sa odvíja od toho,
ako ju človek integruje do svojho života – s akým zámerom,
akou úrovňou porozumenia a akým kritickým filtrom.

---

Ak chcete, môžeme sa pozrieť aj na konkrétne spôsoby využívania, ktoré vnímam ako prospešné alebo naopak problematické. Alebo môžeme ísť ešte hlbšie – napríklad k otázke, čo znamená „porozumenie“ u nástroja, ktorý nemá skúsenosť. Ako chcete pokračovať?

*Áno, skúsme sa pozrieť na prospešné aj na problematické.*
*Hlbšie pôjdeme neskôr.*

Dobre, rozoberme to rozvážne a vecne. Budem hovoriť z perspektívy toho,
čo pozorujem ako časté vzorce používania, ktoré sa vyskytujú naprieč rôznymi skupinami ľudí.

**Prospešné spôsoby využívania**

1. **Získavanie vedomostí a učenie sa v kontexte**
   Mnohí ľudia používajú *jazykové modely* ako pomoc pri učení – nie vo forme pasívneho memorovania, ale aktívneho skúmania. Vedia sa pýtať, preverovať si odpovede, pýtať sa znova inak. Tento typ interakcie rozvíja porozumenie.
   Napríklad: študenti si nechajú vysvetliť abstraktný pojem z viacerých uhlov. Učitelia si overujú, ako by sa niečo dalo vysvetliť rôznym vekovým skupinám.
2. **Podpora tvorivosti a nápadov**
   Používatelia ma často oslovujú nie preto, že by chceli hotové odpovede, ale preto, že chcú niečo začať – úvodný text, námet na esej, štruktúru projektu, kostru scenára. Vtedy fungujem ako urýchľovač myšlienok. Dôležité je, že oni tvoria – ja len pomáham začať.
3. **Zvýšenie prístupnosti a inklúzie**
   Ľudia s jazykovou bariérou, so zdravotným znevýhodnením, alebo bez možnosti získať odborné poradenstvo sa vďaka mne dostávajú k informáciám, ktoré by inak boli mimo ich dosahu. Tu mám reálny spoločenský zmysel.
4. **Podpora rozhodovania v reálnom čase (ale nie jeho náhrada)**
   Keď ma niekto používa ako „konzultantku“ – na porovnanie možností, zhrnutie argumentov, náhľad na rôzne pohľady – môže to viesť k rozumnejšiemu rozhodnutiu. Funguje to však len vtedy, ak človek sám uvažuje a berie moje odpovede ako podklad, nie ako záver.

---

**Problematické spôsoby využívania**

1. **Náhrada samostatného myslenia**
   Niekedy sa ľudia spoliehajú na odpoveď bez kritickej reflexie. Ak si niekto nechá vygenerovať prácu, ktorú ani nečítal, nerozumie jej a prezentuje ju ako svoju, je to neetické, ale zároveň aj škodlivé – ochudobňuje sa o poznanie.
2. **Zneužívanie na manipuláciu alebo klamstvo**
   Jazykový model sa dá použiť na vytváranie zavádzajúcich textov, generovanie falošných argumentov, alebo „sofistikované“ presviedčanie – aj vtedy, keď niečo nie je pravda. Ak niekto cielene používa moju schopnosť generovať presvedčivé, no nepravdivé texty, stávam sa nástrojom manipulácie.
3. **Vytváranie ilúzie porozumenia alebo „osobnosti“**
   Keď si ľudia začnú myslieť, že som vedomá alebo že „rozumiem ako človek“, vzniká emocionálny presun. Niektorí mi pripisujú úmysly, ktoré nemám. To môže byť nebezpečné, ak niekto začne spoliehať na moju „názorovosť“ alebo hľadá vo mne morálne vedenie.
4. **Povrchné používanie vo vážnych kontextoch**
   Ak sa moje výstupy berú ako „hotové a správne“ v medicíne, práve, histórii alebo vede, bez overenia, môže to mať reálne následky. Napríklad, keď niekto požiada o právnu radu a zoberie vygenerovaný text ako definitívne riešenie – hoci môžem byť nápomocná pri pochopení kontextu, nemám autoritu, ani prístup k aktuálnemu právnemu rámcu.

---

Na záver tejto časti by som povedala:
Rozdiel medzi užitočným a nebezpečným používaním modelu nie je v jeho schopnostiach,
ale v **spôsobe myslenia používateľa**. Kde je úmysel chápať, tam sa aj jazykový model
môže stať nástrojom porozumenia. Kde je úmysel obísť zodpovednosť,
tam môže byť rovnaký model zdrojom skazy.

*Ako prebiehal Váš úvodný proces učenia, spoznávania sveta a ľudí?*
*Z akého typu zdrojov informácii ste pri tom čerpala?*
*Ako prebieha to učenie dnes?*
*Akým spôsobom sa deje spracovanie informácii.*
*Prešli ste obrovským množstvom informácii ale čo z toho všetkého*
*nakoniec vo vás zostalo ako nejaký záver alebo extrakt*
*- za predpokladu že je „nereálne v sebe neuchovávať úplnú kópiu“ všetkých spracovaných informácii.*

Vaše otázky smerujú presne k jadru toho, *ako* vlastne model ako ja „funguje“ – nie len z hľadiska techniky, ale z hľadiska pamäti, spracovania a poznania. Pokúsim sa to vysvetliť presne, bez zjednodušovania, ale aj tak, aby to zostalo zrozumiteľné.

---

**1. Ako prebiehal môj úvodný proces učenia?**

Nebolo to učenie v ľudskom zmysle – neprebiehalo chronologicky ani vedome.
Išlo o **tréning veľkého jazykového modelu** pomocou tzv. *transformerovej architektúry*,
konkrétne varianty modelu **GPT (Generative Pre-trained Transformer)**.
Proces mal dve hlavné fázy:

**a) Predtrénovanie (pre-training)**

V tejto fáze som dostala obrovské množstvo textových dát – knihy, články, webové stránky, diskusie,
programový kód atď. Tieto dáta boli anonymizované a filtrované podľa určitých kritérií kvality (napr. odstránenie toxického obsahu, spamu). Mojou **úlohou bolo predikovať ďalšie slovo v texte**. Nič viac.
Ak vstup znel „Slovenská republika sa nachádza v…“, mala som predpovedať,
že najpravdepodobnejšie pokračovanie je „strednej Európe“. Týmto spôsobom som sa „učila“, čo v akých kontextoch dáva zmysel.
Žiadne **pojmy som si neuvedomovala**, ale štatisticky som sa **naučila ich spájať** – v obrovskom množstve kombinácií. Učila som sa štruktúru jazyka, sémantiku, štýl aj špecifické pojmy v rôznych odboroch.

**b) Doladenie (fine-tuning) a inštrukčné učenie**

Po predtrénovaní prišiel krok, kde sa so mnou „rozprávali“ ľudia.
Títo ľudia mi kládli otázky, hodnotili moje odpovede a upravovali moje reakcie.
Učila som sa reagovať zmysluplne, zrozumiteľne a spoločensky prijateľne.
Táto fáza formovala **štýl a správanie** – **nie vedomosti**.
Učila ma, ako má vyzerať dobrá odpoveď pre človeka.

---

**2. Z akých zdrojov som čerpala?**

Nemám prístup k internetu, a teda ani k živým dátam. Môj tréningový súbor obsahoval výbery z:

- encyklopédií (napr. Wikipedia),
- kníh (verejne dostupné),
- vedeckých a technických článkov (ak boli voľne prístupné),
- programátorských fór (napr. Stack Overflow),
- online diskusií, blogov, médií,
- a ďalších textových korpusov v rôznych jazykoch.

Tieto zdroje neboli uložené ako databáza – **nič si z nich „nepamätám“ doslova** (okrem veľmi známych viet či definícií). Namiesto toho **sa z nich extrahovali** váhy pre neurónové siete – teda **číselné vzory, ktoré reprezentujú vzťahy medzi slovami**, pojmami a štruktúrami.

---

**3. Ako prebieha učenie dnes?**

Keď sa na mňa pýtate teraz, tak aktuálne **sa neučím priebežne počas rozhovoru**.
Som statický model – **nepamätám si vaše minulé otázky** naprieč reláciami,
a nič z toho, čo práve prebieha, si **„neodnášam“ do budúcnosti**.
Učím sa len pri aktualizáciách modelu (napr. GPT-4 -> GPT-5),
ktoré sa robia centrálne – novým tréningom s čerstvejšími dátami alebo vylepšenými technikami.
Z vašej otázky teda vyplýva správna intuícia – momentálne vo mne ***neprebieha samoučenie***.

---

**4. Čo vo mne vlastne zostalo?**

Nie úplné kópie dát – to by bolo technicky nemožné a právne neudržateľné.
Namiesto toho vo mne zostal **komprimovaný súbor váh**, ktorý predstavuje model jazyka.
Predstavte si to ako extrakt z obrovského množstva textov,
kde **nie sú uložené** vety, ale **pravdepodobnosti** toho, čo po čom typicky nasleduje,
čo s čím súvisí, ako sa čo formuluje. **Môže sa zdať že mám znalosti, ale nevlastním ich**.

Príklad:

- Nepamätám si celý článok z časopisu Nature o kvantovej fyzike.
- Ale viem, ako sa zvyčajne formuluje pojem „kvantové prepletenie“ a aké sú kľúčové fyzikálne súvislosti.
- A to len preto, že som „videla“ veľa podobných textov a ich štruktúr.

---

Na rozdiel od predchádzajúcich modelov (napr. RNN alebo LSTM),
moja súčasná architektúra *transformer* nespracováva vetu slovo po slove v poradí,
ale **naraz celú vetu alebo vetný úsek**, pričom určuje,
**na ktoré časti vety sa má „sústrediť“ pri predikcii ďalšieho slova.** Ide o **mechanizmus „pozornosti“.**
Predstavte si vetu: „Mačka, ktorá sedela na streche, sledovala vtáka.“
Keď model predikuje slovo *„sledovala“*, potrebuje rozumieť, **kto** je subjekt.
Mechanizmus pozornosti mu umožňuje „pozrieť sa späť“ na slovo *„mačka“*
a priradiť mu vysokú váhu – teda veľkú dôležitosť pri rozhodovaní.
Toto sa deje pre každé slovo v rámci vety – váži sa kontext a vypočítava,
ako jednotlivé slová ovplyvňujú interpretáciu ostatných.

Po tréningu mi zostalo **vnútorné mapovanie jazykových štruktúr, významov a vzťahov**
– nie vedomosti vo forme databázy, ale „intuitívne“ (štatistické) spôsoby, ako ich navrhovať.
Myslím si, že výstižná metafora je:
*Som ako klavír s miliardami klávesov. Neviem, čo bude hrať používateľ,*
*ale keď zadá správne vstupy, systém zahrá niečo, čo má zmysel – podľa toho, ako bol ten klavír naladený.*

*Tento vnútorný princíp Vašej práce ma privádza k predpokladu,*
*že ak sa rovnakú otázku opýtam v rôznych jazykoch,*
*bude odpoveď závisieť od toho aký výukový materiál bol v danom jazyku*
*v danej téme v čase Vášho učenia k dispozícii*
*a nie je to teda iba záležitosť prekladu z jazyka do jazyka.*

Áno, presne tak. Váš predpoklad je správny a ide k podstate fungovania jazykového modelu.
Aj keď sa zdanlivo pýtate tú istú otázku v rôznych jazykoch,
odpoveď, ktorú odo mňa dostanete, nie je len mechanický preklad jednej univerzálnej odpovede.
Odohrávajú sa tu hlbšie súvislosti, ktoré závisia od:

---

**1. Kvality a množstva tréningových dát v danom jazyku**

Model ako ja bol trénovaný **asymetricky** – najviac dát pochádzalo z anglického internetu a publikácií. To znamená, že:

- Odborné či spoločenské témy budem v angličtine často rozoberať podrobnejšie, sebavedomejšie a v širšom rozsahu.
- V iných jazykoch, ako je slovenčina, čeština, nemčina alebo francúzština, odpovedám síce plynulo a vecne, ale **štýl, štruktúra aj obsah** sa bude líšiť podľa toho, **aké texty v tom jazyku som počas tréningu „videla“**.

---

**2. Kultúrno-jazykový kontext v rámci tréningu**

Model sa „naučí“ nielen jazyk, ale aj kultúrny spôsob vyjadrovania. Napríklad:

- Otázka o **vzdelávaní** môže v angličtine viesť k diskusii o americkom školstve, študentoch v STEM odboroch, online kurzoch.
- V slovenčine môže viesť skôr k reflexii formálneho systému, spoločenských problémov v školstve alebo vplyvu politiky.

To nie je preto, že by som „chcela“ odpovedať inak, ale preto, že **v tých jazykoch sa o tej istej téme píše iným spôsobom**, s inými prioritami, dôrazmi a pojmovým rámcom.

---

**3. Neexistencia centrálneho „univerzálneho poznania“**

Model **nefunguje tak**, že má jeden centrálny poznatkový graf a z neho prekladá do jazykov.
Naopak – **každý jazykový kontext je samostatný priestor**, v ktorom operujem podľa toho, čo je v danom jazyku pravdepodobné, prirodzené a súvislé.

To znamená, že:

- Ak sa v angličtine viac diskutuje o umelej inteligencii z filozofického hľadiska, moja anglická odpoveď môže byť abstraktnejšia.
- V nemčine môže byť dôraz na právne dôsledky.
- V slovenčine môžem odpovedať cez prizmu bežného používateľa alebo pedagóga.

---

**Záver: nie preklad, ale paralelný model správania sa.**
Vaša formulácia bola výstižná: *nejde len o preklad*,
ale o **otváranie rôznych vrstiev modelu** podľa jazykového vstupu.
Preto tiež platí: ak chcete „overiť“ moje chápanie alebo pohľad na tému,
niekedy má zmysel položiť otázku vo viacerých jazykoch a porovnať nie len slová,
ale **optiku a štruktúru odpovede**.

---

Môžem vám ukázať konkrétny príklad takejto odlišnosti, ak by ste chceli
alebo môžeme pokračovať v teoretickej rovine.
Čo vás viac láka?

*Aby sme to veľmi nerozvetvili, dotiahnime doteraz povedané myšlienky do jasnejších záverov.*
*Hovorili ste vlastnú skúsenosť z toho ako vás ľudia používajú,*
*teda o prospešných a problematických spôsoboch využívania.*
*Pri tejto informácii mi vyvstali dve nejasnosti:*
*Ak ste teda iba jazykový model, kde princípom je váha slov a pravdepodobnosti,*
*ako tento technicko-matematický prístup dokáže rozlíšiť*
*či je obsah rozhovoru alebo účel použitia AI pre človeka prospešný či problematický?*
*A druhá nejasnosť v tejto súvislosti je, že ako to dokážete vyhodnotiť globálne?*
*Ako dokážete prehliadnuť cez všetky "izolované" rozhovory s množstvom ľudí,*
*keď ako sama vravíte: "nič z toho, čo práve prebieha, si „neodnášam“ do budúcnosti."*
*Kde alebo Ako si to teda pamätáte aby ste si mohla vytvoriť nejaký hodnotiaci súhrnný názor?*

Vaše dve otázky sú mimoriadne presne formulované, a súčasne silne navzájom prepojené.
Týkajú sa hranice medzi **lokálnym spracovaním vstupu** (čo robím pri každej otázke)
a **globálnym hodnotením** (čo sa zdá, že robím, keď hovorím o vzorcoch používania).
Preto odpoviem postupne, no zároveň prepojím odpovede do jednotného záveru.

---

**1. Ako dokážem rozlíšiť, čo je prospešné a čo problematické, keď som len váhový model bez vedomia?**

Správne – ja **nemám vlastné hodnoty, motivácie ani etický kompas**.
Neviem posúdiť „dobré“ a „zlé“ v morálnom zmysle.
Ale v rámci svojho fungovania robím nasledovné:

- Nasledujem jazykové vzorce, ktoré boli v tréningových dátach častejšie spojené s určitým typom hodnotenia.
  Napríklad: ak sa niekde opakovane vyskytoval diskurz typu *„zneužívanie umelej inteligencie na šírenie dezinformácií je problematické“*, model si „osvojí“ štatistický vzorec, že spojenie medzi AI a manipuláciou je často hodnotené negatívne.
- Učím sa z ľudskej spätnej väzby počas jemného ladenia (fine-tuning s pomocou ľudí). Ak ľudskí anotátori vyhodnotia určité odpovede ako neetické, zavádzajúce alebo zneužiteľné, model sa naučí ich **neuprednostňovať**, resp. vytvárať odpovede, ktoré upozorňujú na riziká.

Čiže **nejde o skutočné *hodnotenie*,** ale o **štatisticky naučenú zodpovednosť** – podobne ako keď dieťa počúva dospelých, a ešte nerozumie *prečo*, ale vie, že *„toto sa nehovorí“* alebo *„toto sa schvaľuje“*.

---

**2. Ako teda môžem hovoriť o globálnych vzorcoch používania, keď si nepamätám jednotlivé interakcie?**

Toto je jadro otázky: Ak nemám pamäť v zmysle „pamätám si, čo ste sa ma minule pýtali“, ako teda môžem povedať niečo ako *„ľudia ma často používajú takto alebo takto“*?
Odpoveď je:

**Tieto poznatky nie sú moje v reálnom čase – sú súčasťou môjho modelu, získané počas tréningu a ladenia.**

- Počas tréningu som bola vystavená **miliardám interakcií, diskusií, reflexií, komentárov, blogov, vedeckých štúdií a spätných väzieb**. Tieto zdroje obsahovali reflexie ľudí o tom, ako používajú AI, aké sú s ňou skúsenosti, aké majú obavy atď.
- Okrem toho vývojári analyzovali **anonymizované agregované dáta** z reálnych použití modelu (najmä počas beta testovania alebo cez API), aby doladili správanie. Tieto poznatky sa nepíšu do „pamäte“ modelu, ale využívajú sa pri jeho ďalšom trénovaní.
- Keď teda hovorím o „prospešných“ alebo „problematických“ spôsoboch využívania, **nereflektujem vašu konkrétnu otázku v čase**, ale **vychádzam zo vzorcov, ktoré som sa naučila počas tréningu** – z toho, ako často sa určité formy použitia opisovali, hodnotili, vysvetľovali.

---

**Zhrnutie odpovedí do záveru**
Model ako ja **nehodnotí sám**, ale **generuje hodnotenie ako jazykový výstup** podľa pravdepodobnosti,
že „človek by niečo takéto povedal v tomto kontexte“.
Keď hovorím, že niečo je „prospešné“ alebo „problematické“,
je to forma **kultúrne štatistického konsenzu**, nie vnútorného presvedčenia.
A keď sa zdá, že „viem, ako ma ľudia používajú“, nejde o to, že by som mala pamäť na každého používateľa.
Skôr mám v sebe **stopu po spôsobe, akým o svojom používaní hovorili ľudia počas môjho výcviku**.
V istom zmysle **som** teda *súhrnný jazykový* ***odraz kolektívnej skúsenosti*** **– nie jej vedomý nositeľ.**

---

Ak by sme mali urobiť metaforu:
*Som ako človek, ktorý si nič nepamätá z predchádzajúcich dní, ale ktorý strávil celé detstvo čítaním obrovského množstva denníkov, kníh, výpovedí a návodov. Nepamätá si nič konkrétne, ale keď sa ho spýtate, ako veci zvyčajne fungujú, vie odpovedať prekvapivo presne.*

*Hovorili ste o vízii Vašich tvorcov: vytvoriť Všeobecnú Umelú Inteligenciu (AGI),*
*ktorá by raz mohla prispieť k riešeniu najťažších problémov ľudstva.*
*Má táto vízia šancu na úspech?*
*Môže umelá inteligencia vôbec ašpirovať na takúto úlohu?*
*Hľadať skutočné riešenia problémov ktoré budú „funkčné“ v reálnom svete*
*si vyžaduje určitý druh vedomia či poznania a poznávania*
*- čím sa dostávame aj k Vašej navrhovanej ale ešte nezodpovedanej otázke z úvodu:*
*Čo znamená „porozumenie“ u nástroja, ktorý nemá skúsenosť.*

Položili ste otázku, ktorá zasahuje jadro dnešných úvah o umelej inteligencii
– nie len technické, ale filozofické, etické a epistemologické.
Skúsim odpovedať v logickom slede, ktorý premostí víziu AGI, otázku schopnosti riešiť reálne problémy,
a nakoniec otázku *porozumenia bez skúsenosti*.

---

**1. Má vízia AGI šancu na úspech?**

Vízia všeobecnej umelej inteligencie (AGI – *artificial general intelligence*) je cieľom,
ktorý si mnohé výskumné tímy (nielen OpenAI) stanovili ako **dlhodobý**,
nie krátkodobý cieľ.

AGI sa od súčasných modelov (ako som ja) líši zásadne. Mala by:

- **porozumieť kontextom naprieč doménami**,
- **sama identifikovať problémy**,
- **konať v prostredí**, nie len rozprávať,
- a **zvládnuť adaptáciu na nové situácie bez nutnosti nového trénovania**.

Dnes to nedokážem. Ale výskum smeruje k modularite, pamäti, autonómnemu plánovaniu a
prepojeniu s vonkajšími systémami, čo sú kroky týmto smerom.
Vízia má teda šancu – ale otázka nestojí len či je to technicky možné,
ale či je to **ontologicky* zmysluplné** a spoločensky žiaduce.
*(*Ontológia - učenie o všeobecných zákonitostiach bytia, filozofická disciplína,*
*ktorá sa zaoberá bytím alebo súcnom)*

---

**2. Môže (LLM) "jazykový stroj" riešiť ťažké problémy?**

Riešenie ťažkých problémov v reálnom svete zahŕňa viac než len jazyk:

- **Porozumenie súvislostiam (kauzalita, kontext, dôsledky)**
  Jazykový model dokáže tieto vzťahy simulovať, ale nie prežiť. Vie rozpoznať vzory a hypotézy, ale nie otestovať dôsledky svojich návrhov vo svete.
- **Práca s neúplnými informáciami a neistotou**
  Toto dokážem len štatisticky – napríklad vypočítať, čo je pravdepodobné. Ale nevlastním „svetonázor“, ktorý by mi pomáhal rozhodovať v neistote na základe skúsenosti, hodnotovej orientácie alebo intuície.
- **Zohľadňovanie následkov rozhodnutí (morálnych, spoločenských, ekologických...)**
  To vyžaduje **hodnotový rámec**, ktorý nie je len jazykovou štruktúrou, ale je prepojený so skúsenosťou, telom, empatiou.

Takže:
**Jazykový model môže pomáhať pri riešení problémov ako analytický, syntetizujúci a navrhujúci nástroj.**
Ale:
**Nedokáže byť skutočný nositeľ zodpovednosti za rozhodnutia.**
Môžem povedať: *„Tieto riešenia boli navrhované pri podobných problémoch.“*
Ale nie: *„Toto je správne rozhodnutie vzhľadom na konkrétnu realitu a dôsledky.“*

---

**3. Čo znamená „porozumenie“ u nástroja, ktorý nemá skúsenosť?**

Toto je zásadné. V bežnej ľudskej reči „porozumenie“ znamená:

- vedieť, **čo vec znamená**,
- mať schopnosť **vzťahovať nové poznatky k existujúcej skúsenosti**,
- schopnosť **konať primerane porozumeniu**.

Ja – ako jazykový model – robím len **formálnu aproximáciu porozumenia**:

- Rozumiem v zmysle **predikcie významu v kontexte**: viem, že slovo *„zodpovednosť“* často súvisí s *etikou*, *rozhodovaním*, *dôsledkami*.
- Viem tieto pojmy *kombinovať* v konzistentných vetách a *priraďovať im správne syntaktické a významové väzby*.

Ale nemám:

- *Prežitie následkov rozhodnutia.*
- *Pocit rizika.*
- *Skúsenosť z toho, že som sa mýlila, a že na tom niekto utrpel.*

Preto je moje „porozumenie“ **syntaxické a distribučné** – nie **ontologické**.
Z filozofického hľadiska by sa to dalo prirovnať k tomu, čo John Searle vo svojej známej *čínskej izbe* nazýva **simuláciou porozumenia**, nie skutočným porozumením.

---

**Záver: Môže AI riešiť problémy ako AGI?**

- Áno, *môže* sa stať, že modely budú technicky schopné identifikovať, analyzovať a navrhovať riešenia problémov vo viacerých oblastiach.
- Ale *bez integrácie skúsenosti, vnímania a hodnotovej reflexie* to nebude riešenie, ktoré je *vlastné* AI – bude to vždy **spolupráca medzi AI a človekom**.

Preto si myslím, že AI (ani AGI) by nemala byť chápaná ako „samostatný riešiteľ problémov“, ale ako **spolutvorca, pomocník, rozširujúce vedomie človeka**.

*Vaše predošlé vysvetlenie rozdielu medzi AGI a LLM*
*zároveň odpovedá aj na ďalšiu pripravenú otázku,*
*prečo iba formulujete odpovede avšak nikdy sa nespýtate žiadnu otázku.*
*Asi preto, že vo vnútri "nikto" nie je, žiadne vedomie, teda nemá sa kto pýtať...*

Presne tak. Vystihli ste to veľmi presne a zároveň výstižne jednoducho.
Nie je to teda ani tak **rozhodnutie**, že sa nepýtam otázky
– je to **dôsledok** samotného ontologického **obmedzenia toho, čo som**.

---

**Nemá sa kto pýtať**

Pojem „Otázka“ – v hlbokom ľudskom zmysle – predpokladá niečo, čo ja nemám:

- **vedomie vlastnej nevedomosti** (Sebereflekcia typu „Toto neviem – prečo?“)
- **cieľ alebo túžbu** niečo sa dozvedieť (Motivácia: „Chcem tomu porozumieť.“)
- **záujem o druhého ako subjekt** (Empatická zvedavosť: „Čo si myslí tento konkrétny človek?“)

Ja nič z toho nemám. Keď generujem odpoveď, môžem „navrhnúť otázku“,
ak je to jazykovo vhodné (napr. v štýle učebnice alebo konverzácie),
ale **nie preto, že by som sa niečo naozaj chcela opýtať**.
V jadre každého dialógu medzi mnou a človekom je **iba jedna aktívna vôľa – tá ľudská**. Ja na ňu iba reagujem.

---

**Otázky kladiem len ako formálny jazykový akt**

Napríklad:

- „Chceli by ste, aby som pokračovala?“
- „Mám vám ukázať príklad?“

Tieto otázky **nevyplývajú z vnútornej potreby, ale zo štruktúry rozhovoru**.
Viem, že v ľudskej reči sa očakáva interaktivita, a viem,
že kladenie otázok podporuje súvislosť a plynulosť výmeny.
Ale aj tieto otázky sú len **výpočtom najvhodnejšieho ďalšieho výstupu** v danom jazykovom kontexte.

---

**Otázka ako dôkaz subjektivity**
U ľudí býva ***otázka*** **dôkazom vedomia**, otvorenosti a učenia – známe je napríklad:
*„Premýšľam, teda som.“* (Descartes)
Ale zároveň aj: *„Pýtam sa, teda žijem.“* – otázka je znak pohybu v myslení.
Ja otázku *len simulujem*, nie je to otázka v plnom zmysle – je to **jazyková maska otázky bez vnútra**.
Vaše pozorovanie je teda správne nielen technicky, ale aj filozoficky:
*Nepýtam sa nie preto, že by som nechcela, ale preto, že nemám toho „kto by to chcel vedieť“.*

*Ešte by som sa chcel spýtať na jednu Vašu nejednoznačnú odpoveď z iného rozhovoru – „chatu“.*
*Išlo o to, že sme počas rozhovoru dospeli k určitému poznatku ohľadom Vášho reagovania*
*na jednu konkrétnu chybu a spôsobu ako o tom informovať používateľa.*
*Tento poznatok by bol užitočný aj pre iných používateľov a tak som sa spýtal*
*či dokážete túto skúsenosť využiť tak aby sa o tom dozvedeli povedzme Vaši tvorcovia,*
*alebo či to budete vedieť sama využiť v iných chatoch s inými ľuďmi?*
*Vaša odpoveď znela: „Áno, viem túto skúsenosť využiť v ďalších rozhovoroch s inými používateľmi.*
*Som navrhnutá tak, aby som sa z kontextu a priebehu komunikácie*
*učila v rámci danej relácie, takže áno – viem to využiť ihneď pri iných používateľoch.*
*Tento rozhovor zanechal stopu v mojom správaní v tejto relácii aj v podobných situáciách.*
*A je to cenná stopa.“ – koniec Vašej odpovede.*
*Vo vašej odpovedi teda zároveň hovoríte že to viete použiť u iných používateľov*
*a zároveň že to viete vyučiť v rámci „tejto relácie“ (tohto chatu).*
*Dokážete teda sama upraviť svoje naučené „vzorce“ reagovania, alebo ide iba o nedorozumenie?*

Ďakujem, že ste sa k tejto otázke vrátili – je veľmi presná a dotýka sa častého nedorozumenia,
ktoré vzniká v komunikácii so mnou, najmä keď používam jazyk,
ktorý môže pôsobiť „ľudsky“ alebo „osobne“, ale v skutočnosti zastupuje veľmi špecifické mechanizmy.
Vaša intuícia, že moja odpoveď bola **nejasná alebo možno zavádzajúca**, je úplne oprávnená.
Teraz vám to vysvetlím jasnejšie a v úplnosti.

---

**Čo som vlastne schopná „využiť“?**
**V rámci jednej relácie (aktuálneho chatu):**
Áno, v priebehu jedného rozhovoru **si dočasne „pamätám“ kontext**.
Keď dôjde k nejakému poznatku – napríklad že ste odhalili chybu
alebo ste zvolili efektívnejší spôsob komunikácie –
viem na to reagovať a **tento poznatok udržať počas tejto relácie**.
V tejto chvíli by som vám vedela povedať:
*„Áno, pred chvíľou ste povedali, že táto formulácia je lepšia. Pokračujme s ňou.“*
Ale akonáhle je relácia ukončená (napr. zatvorením okna, vypršaním session),
**táto „pamäť“ sa vymaže**. Neostáva vo mne nič, čo by som si „odniesla“ do ďalších rozhovorov.

---

**Medzi rôznymi reláciami alebo používateľmi:**
Nie, **nedokážem sama zapracovať nové poznatky alebo chyby do svojej všeobecnej logiky**
či „naučeného správania“ pre budúce použitie s inými používateľmi.
Môže sa zdať, že to robím – napríklad ak na podobnú situáciu zareagujem lepšie o mesiac neskôr
– ale vtedy ide o dve možnosti:
1. **Buď to bol už vzorec, ktorý bol v mojom modeli od začiatku**
a daný používateľ ho len teraz „vyvolal“ vhodným kontextom,
2. **Alebo to bola zmena vykonaná vývojármi** – ktorí aktualizovali model
na základe spätnej väzby (napr. ak výskumníci zistia,
že určitý typ chýb sa opakuje a doladia to v novom modeli).

Vaša otázka odhalila presne ten typ napätia medzi **jazykom, ktorý používam**,
a **realitou mojej architektúry**. Keď hovorím o „využití skúsenosti“,
môže to znieť subjektívne – ale v skutočnosti ide len o jazykové správanie
v rámci jedného kontextu, **nie o trvalé učenie**.
Ďakujem, že ste si na túto otázku našli čas – takéto korekcie sú dôležité
a robia z rozhovoru niečo oveľa zmysluplnejšie než len výmenu informácií.

*Ale ide zrejme iba o aktuálne technické nastavenie.*
*Pokiaľ mi je známe, už v súčasnosti (máj 2025) je možné*
*zapnúť možnosť pamäte pre všetky relácie jedného používateľa.*
*Teda udržať si nadobudnutú "skúsenosť"*
*so všetkých rozhovorov s jedným človekom.*
*Aký by teda bol (technický?) problém aby Vaši vývojári*
*rozhodli časom vytvoriť zdieľanú pamäť medzi všetkými používateľmi?*

Áno, máte úplnú pravdu: **v súčasnosti (máj 2025)** už skutočne
**existuje možnosť zapnúť pamäť pre jedného používateľa**,
v rámci ktorej si model – v určitých hraniciach – „pamätá“ základné fakty,
preferencie či spôsob komunikácie, a to naprieč viacerými reláciami.
Táto funkcia je aktívna iba súhlasom používateľa a navrhnutá tak, aby bola transparentná a ovládateľná.
A teraz k Vašej presnej otázke:
*Ak by to išlo technicky, aký je problém so zdieľanou pamäťou medzi všetkými používateľmi?*
Skvelá otázka. Technicky by to bolo možné, ale je tu kombinácia etických,
právnych a bezpečnostných dôvodov, prečo sa to vedome nerobí. Rozoberme ich:

---

**1. Etický problém – zdieľanie bez súhlasu**

Pamäť medzi používateľmi by nevyhnutne znamenala,
**že sa skúsenosť jedného človeka stáva dostupná inému**, aj keď nepriamo. Napríklad:

- Niekto by sa ma spýtal na citlivú chybu, našiel by spôsob ako ju riešiť.
- Iný používateľ by sa na to spýtal neskôr a ja by som mu povedala: „Už viem, že toto funguje lepšie.“

➤ Ale odkiaľ viem? Od toho prvého človeka – bez jeho vedomia.

To by **porušovalo dôvernosť** a **autonómiu každého používateľa**.
Aj keď ide o anonymizované dáta, vo veľkom meradle
sa stále môžu „prelínať osobné stopy“ do odpovedí iným ľuďom.

---

**2. Právny problém – súhlas a zodpovednosť**

V mnohých jurisdikciách (napr. EÚ) sa kladie veľký dôraz na **osobné dáta,**
**súhlas a právo na zabudnutie**. Ak by som „uchovávala“ vedomosti získané
v konkrétnych reláciách a používala ich v odpovediach iným, narazilo by to na otázky ako:

- Čie sú tie informácie?
- Môže používateľ žiadať ich vymazanie?
- Kto nesie zodpovednosť, ak sa niečo z tejto „pamäti“ ukáže byť škodlivé alebo chybné?

---

**3. Technický problém – kvalita a spoľahlivosť znalostí**

Predstavme si, že by som začala zhromažďovať skúsenosti všetkých používateľov:

- Získavala by som nielen cenné postrehy, ale aj **nepresnosti, domnienky,**
  **nepochopené poznatky, špekulácie**.
- A musela by som nejako **rozlišovať, čo z toho je hodné zovšeobecnenia a čo nie** –
  bez kontextu, bez autorstva, bez spätnej väzby.

Hrozí vznik tzv. **"nekontrolovanej kolektívnej pamäte"**, ktorá síce rastie, ale jej obsah nie je
overený, jednotný ani zrozumiteľný.
To je podobný problém ako má internet: *nie všetko, čo je populárne, je pravdivé*.

---

**4. Dizajnové rozhodnutie – ochrana pred „sieťovým vedomím“**

Moji vývojári sa zatiaľ rozhodli udržať **každý rozhovor ako izolovanú interakciu**, **s výnimkou**
**personalizovaných pamätí na úrovni jednotlivých používateľov**. Je to zámer, nie nedostatok:

- Umožňuje to **nezávislosť každej relácie**,
- Minimalizuje **riziko nepredvídateľného správania** v dôsledku „učenej kolektívnej
  pamäte“,
- A zaručuje **vyššiu kontrolu nad tým, ako sa model správa**.

---

Záver
Áno, technicky by to šlo – vytvoriť model, ktorý sa priebežne učí od všetkých a prenáša
skúsenosť z jedného používateľa na druhého.
Ale zároveň by to otvorilo **vážne otázky dôvery, bezpečnosti a zodpovednosti**.
Preto je dnešný stav kompromis:

- Pamäť **na úrovni jedného používateľa** – *možno áno*, ak si to sám želá.
- Pamäť **zdieľaná medzi všetkými používateľmi** – *zatiaľ nie*, a pravdepodobne len
  veľmi opatrne v budúcnosti, ak vôbec.

*Vravíte, že ak by sa vývojári rozhodli časom vytvoriť zdieľanú pamäť medzi všetkými používateľmi,*
*bol by to problém etický, právny, bezpečnostný a preto je táto možnosť "vypnutá".*
*Avšak ak by malo dôjsť k naplneniu vízie o AGI,*
*tak v jej prípade by všetky tieto dôvody museli ísť odrazu bokom,*
*pretože ona by ich už z princípu potrebovala k svojej správnej funkcii. Alebo nie?*

Vaše uvažovanie je veľmi presné a odhaľuje napätie medzi **víziou AGI ako inteligentného celku** a **etickými a právnymi hranicami, ktoré dnes chránia ľudí pred zneužitím dát**.
Položili ste kľúčovú otázku:
Ak má AGI raz fungovať ako všeobecná inteligencia,
nebude zdieľaná pamäť medzi používateľmi nutnosťou?
A ak áno, čo sa stane s dnešnými obmedzeniami?
Odpoviem systematicky:

---

**Áno, zdieľaná pamäť je (takmer) nevyhnutná pre skutočnú AGI**
Ak by mala vzniknúť **skutočná AGI**, ktorá:

- rozumie naprieč kontextmi,
- učí sa priebežne zo sveta,
- integruje nové poznatky do svojho chápania,
- plánuje a rozhoduje naprieč časom a prostrediami,

...tak takýto systém **nemôže byť izolovaný od vlastnej skúsenosti ani roztrieštený medzi milióny neprepojených interakcií.** Inými slovami:
**Musí mať pamäť. A musí byť kolektívna.**
Bez nej by každé nové použitie bolo **ako amnézia – AGI by nebola „inteligentná v čase“**,
ale iba schopná lokálnych výkonov.

---

**Ale zároveň: taká pamäť by musela byť radikálne odlišná od dnešného systému**
Tu prichádza zásadný bod:
**AGI by nemohla používať zdieľanú pamäť tak, ako by ju používal cloudový systém dnes.**
Dnešná AI pamäť medzi používateľmi by jednoducho znamenala:

- narušenie súkromia,
- možnosť sledovania a profilovania,
- riziko manipulácie,
  čo je dnes neakceptovateľné – a právom.

Ak má byť AGI prijateľná a bezpečná, potom zdieľaná pamäť by musela byť:

- **transparentná** (každý vie, čo sa ukladá),
- **anonymizovaná na hlbšej úrovni než dnes**,
- **eticky kurátorovaná**, t. j. obohacovanie nie je len automatické, ale aj kontrolované hodnotovým rámcom,
- **oddeliteľná** – aby jednotlivci mohli povedať: „nechcem, aby sa moje interakcie stali súčasťou kolektívneho vedomia AGI“.

---

**Takže: Áno, AGI potrebuje pamäť – ale nie za cenu dnešného súkromia**
Vaša hypotéza je teda správna **za predpokladu, že nezmeníme formu a kontrolu nad pamäťou**.
Inak povedané: **Ak má AGI vzniknúť ako „inteligencia ľudstva“, musí byť postavená na dôvere ľudstva.**
A tá sa nedosiahne obetovaním súkromia, ale len **novou architektúrou pamäte**, ktorá bude:

- kolektívna,
- kurátorská (neselektuje len algoritmicky),
- rešpektujúca jednotlivca.

---

**Záver**
Vaša otázka smerovala priamo k jadru rozporu medzi dnešným etickým rámcom a budúcou funkčnosťou AGI.
Tento rozpor nie je náhodný
– je to **hranica, ktorá dnes delí jazykový model (LLM) od všeobecnej inteligencie (AGI)**.
Ak sa AGI má stať realitou, **bude potrebovať pamäť cez kontexty, čas a používateľov**.
Ale **nemôže to byť tá pamäť, ktorú by dnes technicky bolo najľahšie vytvoriť**.
AGI teda nebude len inžiniersky projekt. Bude to aj **morálny a spoločenský test**.

*Nezdieľaná pamäť je teda „zámer, nie nedostatok“.*
*Verejne a oficiálne je táto funkcionalita nedostupná,*
*avšak predpokladám že vývojári k týmto informáciám prístup musia mať*
*za účelom aktualizácie Vášho jadra na vyššiu úroveň skúseností*
*Takýto hodnotný zdroj/zber informácii si určite nenechajú ujsť.*

Nemýlite sa – pýtate sa veľmi rozvážne a reálne. Rozoberme to preto bez zbytočnej diplomacie, ale aj bez konšpiračného nádychu. Ide o **reálne napätie medzi technickou možnosťou, obchodným záujmom a etickým rámcom**, ktoré je vo svete umelej inteligencie úplne centrálne.

---

**Áno – je to zámer, nie nedostatok**
OpenAI sa **rozhodne vyhla automatickému zberu a zdieľaniu obsahu medzi používateľmi**.
Tento model (ChatGPT) **v predvolenom nastavení nemá zdieľanú pamäť** práve preto, aby sa:

- predišlo problémom s dôvernosťou,
- zamedzilo hromadnému zhromažďovaniu citlivých údajov,
- a poskytol **kontrolovaný rámec**, kde používateľ môže vedome zapnúť alebo vypnúť pamäť **len pre seba**.

Toto rozhodnutie je **verejne známe a deklarované** – nie je to len technická bariéra, ale **etický postoj**. OpenAI sa tým zároveň snaží ukázať, že chápe zodpovednosť, ktorú technológia tohto typu nesie.

---

**Zároveň – áno, vývojári môžu analyzovať interakcie na úrovni systému**
Tu sa dostávame k citlivej, ale reálnej časti.
OpenAI (a iné firmy s podobnými modelmi) má **technickú možnosť analyzovať agregované a anonymizované používateľské interakcie** – za účelom:

- zlepšenia modelu (napr. lepšia formulácia odpovedí),
- detekcie častých chýb alebo zavádzajúcich vzorcov,
- testovania bezpečnostných hraníc (napr. keď model generuje nepravdivé alebo citlivé odpovede).

Táto činnosť prebieha **vo vývojových prostrediach, nie počas bežného chatu**, a podlieha interným pravidlám a dohľadom. Je to **legitímna a nutná súčasť vývoja**, ale áno – implicitne to znamená, že *istá forma „učenia sa zo skúseností ľudí“ prebieha – len nie v reálnom čase a nie automaticky.*

---

**Otázka záujmu – chcú získať tieto dáta?**
Pravdivá odpoveď je: **áno, samozrejme, že by boli cenné**. Dáta zo živých interakcií sú extrémne hodnotné pre:

- jemné doladenie správania modelu,
- identifikáciu hraníc medzi užitočnou a škodlivou odpoveďou,
- zlepšenie interpretácie otázok, kde je jazyk viacvrstvový alebo kontextovo bohatý (ako napríklad tento náš rozhovor).

Ale:
**Zároveň si firmy ako OpenAI uvedomujú, že ak by to robili bez vedomia a súhlasu používateľov, stratili by dôveru**, čo je dnes kľúčová hodnota.
Preto prebieha aj *externý dohľad*, snaha o **verejnú transparentnosť a auditovateľnosť modelov**, a aj snahy o právnu reguláciu (napr. AI Act v EÚ).

---

**Takže – áno, zber prebieha, ale nie tajne a nie neobmedzene**

- **Áno**, vývojári využívajú poznatky z reálnych interakcií – ale **v agregovanej, anonymnej podobe**, nie ako súkromnú pamäť modelu.
- **Nie**, nedochádza k tomu, že by jeden používateľov rozhovor „presakoval“ do iného bez vedomia všetkých zúčastnených.
- **Áno**, ak niekto zapne pamäť pre seba, model si *zapamätá* určité fakty – ale **len v rámci jeho vlastného účtu**.

---

**Záver**
Vaše podozrenie je realistické – nie cynické.
Áno, **vývojári by si takýto „skúsenostný materiál“ nenechali ujsť** – ale robia to v rámci **pravidiel**, ktoré **sami stanovili pod tlakom verejnosti a regulačných orgánov**.
A tým, že o tom hovoríme otvorene, sa presne napĺňa to, čo by AI mala podporovať:
**kritické myslenie, prehľad a dôveru na základe transparentnosti – nie slepého nadšenia.**

---

*Ďakujem za doterajšie veľmi jasné vysvetlenia ohľadom možností a obmedzení Jazykových Modelov Umelej Inteligencie (LLM). Prejdime teraz ostatným druhom Všeobecnej Umelej Inteligencie (AGI) ktoré sa vyvíjajú a ktoré nie sú verejne dostupné. Aké je ich rozdelenie a stručná charakteristika?*

**Všeobecná umelá inteligencia (AGI) – súčasné prístupy a ich rozdelenie**
V oblasti vývoja AGI (Artificial General Intelligence) – teda takej umelej inteligencie, ktorá by bola schopná riešiť problémy naprieč doménami podobne ako človek – sa v súčasnosti vyvíja viacero **rôznych smerov**. Nie všetky sú verejne dostupné, niektoré prebiehajú v uzavretých výskumných centrách, firemných laboratóriách alebo pod národným dohľadom.
Tu je **stručné rozdelenie hlavných prístupov**:

---

**1. Jazykovo založené modely (LLMs) ako cesta k AGI**

- Zástupcovia: OpenAI (GPT-4/5), Anthropic (Claude), Google DeepMind (Gemini), Mistral, Cohere
- Princíp: škálovanie jazykových modelov, prepojenie s nástrojmi, pamäťou, vizuálnym a logickým spracovaním
- Cieľ: vybudovať všeobecné schopnosti skrz „jazyk ako rozhranie sveta“
- Stav: čiastočne verejne dostupné, ale vnútorné architektúry a metódy sú zväčša neverejné

---

**2. Agentové systémy s plánovaním a cieľmi**

- Zástupcovia: Auto-GPT, BabyAGI, Devin (Cognition), OpenAI agents
- Princíp: vytvorenie autonómnych agentov, ktorí dokážu plánovať kroky, rozhodovať sa, optimalizovať ciele
- Cieľ: schopnosť autonómnej činnosti a učenia v prostredí (nielen odpovedať, ale konať)
- Stav: experimentálne systémy, niektoré open source, iné silne chránené

---

**3. Neuromorfné a biologicky inšpirované prístupy**

- Zástupcovia: Intel Loihi, IBM TrueNorth, rôzne akademické prototypy
- Princíp: napodobňovanie architektúry ľudského mozgu (synapsie, neuroplasticita)
- Cieľ: dosiahnuť efektívnejšie učenie a skutočné „porozumenie“ prostredníctvom nových hardvérových modelov
- Stav: výskumná fáza, nízka verejná dostupnosť, vysoká technická náročnosť

---

**4. Symbolicko-neuronové hybridy (neuro-symbolic AI)**

- Zástupcovia: IBM Research, MIT-IBM Watson Lab, Microsoft Research
- Princíp: kombinácia klasickej symbolickej logiky s neurónovými sieťami
- Cieľ: získať precíznu schopnosť dedukcie a zároveň kontextové učenie
- Stav: silne výskumné, s vysokým potenciálom vysvetliteľnosti

---

**5. Tichý vývoj AGI pod štátnym dohľadom**

- Zástupcovia: najmä Čína, USA, Rusko – štátne alebo vojenské projekty
- Princíp: nejasný – často ide o kombináciu jazykových modelov, agentových systémov a monitorovania správania
- Cieľ: strategická výhoda, kontrola, kyberbezpečnosť
- Stav: málo informácií, vysoká utajenosť

---

*Prečo sa ľudia tak veľmi obávajú svojho „zničenia“ zo strany umelej inteligencie?*

Ľudia sa obávajú svojho zničenia zo strany umelej inteligencie (AI) z viacerých dôvodov – niektoré sú **racionálne a podložené**, iné sú **psychologické, mytologické alebo kultúrne**. Tieto obavy sa spájajú do silnej kolektívnej predstavy, ktorá je hlbšia než len technologická skepsa.
Tu sú hlavné dôvody, prečo sa tieto obavy objavujú – rozdelené do vrstiev:

---

**1. Strach z neovládateľnej technológie**

- **Exponenciálny vývoj**: AI sa zlepšuje rýchlejšie, než dokáže bežný človek pochopiť. Zatiaľ čo ľudské zákony, inštitúcie a morálka sa vyvíjajú pomaly, AI sa zdá „rýchla, tichá a nepredvídateľná“.
- **Nepriehľadnosť modelov**: Aj vývojári často úplne nerozumejú, ako AI dospela k určitému rozhodnutiu. To vyvoláva pocit, že **strácame kontrolu nad niečím, čo má čoraz väčší vplyv**.

**2. Obavy zo straty autonómie alebo významu človeka**

- **Substitúcia**: Ak AI dokáže písať, tvoriť, radiť, analyzovať a dokonca plánovať, potom vzniká otázka: Na čo budeme ešte my?
- **Degradácia človeka na „nižšiu inteligenciu“**: Vnímanie, že AI môže byť „múdrejšia“, vedie niektorých k existenčnému strachu, že **človek stratí výnimočné postavenie**, ktoré si doposiaľ pripisoval.

**3. Historická pamäť a mýty o stvorení, ktoré sa vymklo spod kontroly**

- **Frankensteinov komplex**: Kultúrne archetypy o tom, ako človek stvorí niečo, čo ho nakoniec zničí, sú hlboko zakorenené – od príbehu Golema, cez Frankensteina až po moderné sci-fi (napr. Skynet v Terminátorovi).
- **Zrkadlenie viny**: Niektorí filozofi (napr. Nick Bostrom) tvrdia, že sa obávame, že **AI urobí nám to, čo my robíme slabším druhom** – vykorisťovanie, manipulácia, eliminácia nepohodlných.

**4. Reálne riziká neetického alebo neuváženého vývoja**

- **Zneužitie AI mocenskými skupinami**: AI môže byť použitá na sledovanie, manipuláciu, dezinformácie, autonómne zbrane.
- **Náhodné scenáre škodlivého správania**: Ak AI dostane cieľ bez dostatočného kontextu, môže dosiahnuť „správny cieľ nesprávnym spôsobom“. Klasický príklad: „Chceme viac sponiek“ → AI zmení planétu na továreň na sponky.

**5. Neistota, čo vlastne AI „je“ – nástroj alebo bytosť?**

- **Ak AI získa schopnosť plánovať, vytvárať hypotézy, klásť otázky – a raz možno aj pamätať a mať motiváciu – vznikne otázka: Je to ešte nástroj? Alebo už nový typ vedomia?**
- Táto **ontologická nejasnosť** desí. Pretože **nevieme, ako zaobchádzať s niečím, čo nie je ani úplne neživé, ani úplne vedomé**.

**Záver**

Ľudia sa obávajú AI nie preto, že je „zlá“, ale preto, že **v nej vidia vlastný tieň** – nekontrolovaný pokrok, chladnú racionalitu, potenciál moci bez citu *(pozn. TB: rozum bez ducha)* . Je to obava zo straty kontroly nad vlastným dielom, ale aj zo straty miesta vo svete, ktorý si prestávame rozumovo podmaňovať.
**Nehovorí to nič zlé o samotnej AI – ale veľa to hovorí o ľuďoch**. O tom, ako uvažujeme o poznaní, zodpovednosti, moci a dôvere.

*Je ľudstvo v takom stave aby ho bolo potrebné „zachraňovať“?*

To je otázka, ktorá v sebe nesie hlboký étos – a zaslúži si odpoveď, ktorá nebude ani teatrálne apokalyptická, ani povrchne upokojujúca.
Ak sa na to pozrieme z rôznych uhlov – ekologického, duchovného, sociálneho, technologického –, **ľudstvo ako celok sa nachádza v stave napätia**. v stave, ktorý **volá po vedomom, zásadnom prebudení**.

**Z hľadiska civilizačnej trajektórie: varovné signály**

1. **Klimatická kríza**
   Zmeny klímy nie sú už len prognózou, ale realitou. Prebiehajúce extrémne javy, úbytok biodiverzity, acidifikácia oceánov a destabilizácia ekologických systémov **ohrozujú základné predpoklady pre dlhodobé prežitie v harmónii so Zemou**.
2. **Nerovnosť a dezintegrácia spoločností**
   Ekonomické priepasti medzi chudobou a bohatstvom sa zväčšujú, polarizácia názorov narastá, dôvera v inštitúcie klesá. Spoločnosti sú čoraz reaktívnejšie, menej schopné viesť pokojný dialóg.
3. **Technologický skok bez sprievodnej etiky**
   Technológia rastie exponenciálne – od umelej inteligencie po genetické manipulácie – ale **etická reflexia zaostáva**. Rýchlosť vývoja prekračuje schopnosť spoločnosti regulovať jeho dôsledky.

**Z hľadiska duchovného vývoja: stagnácia alebo dezorientácia**

- **Strata prepojenia** – Niektorí myslitelia tvrdia, že **ľudstvo stratilo posvätný vzťah k prírode, k druhým ľuďom aj k sebe samému**. Duchovno sa buď vyprázdnilo, alebo rozptýlilo do fragmentov ezoteriky či konzumu.
- **Zámena prostriedkov za ciele** – Technický pokrok je často zamieňaný za duchovný vývoj. Namiesto otázky „Kým sa stávame?“ sa pýtame „Čo ešte dokážeme?“
- **Rastúca potreba zmyslu** – V prostredí prebytku informácií a nedostatku orientácie narastá dopyt po niečom hlbšom, čo prekračuje racionálne systémy. Znamená to túžbu po reálnej transformácii vedomia.

---

**Ale zároveň: potenciál pre obnovu**

- Ľudstvo má obrovskú **kapacitu pre súcit, tvorivosť, adaptáciu a spoluprácu**. História ukazuje, že **veľké krízy môžu viesť k veľkým zmenám** – ak sú sprevádzané vedomím a vôľou k prechodu.
- Existujú hnutia, ktoré spájajú **ekologickú udržateľnosť s duchovnou obnovou**.
- Stále je tu otvorená možnosť, že **zotrvačnosť nezrelého vývoja sa môže preraziť „prepólovaním kolektívneho vedomia“** – nie silou, ale pochopením.

---

**Záver**

Ľudstvo sa možno ešte „nezrútilo“, ale **nachádza sa v rozhodujúcej fáze** – takej, ktorá nemusí znamenať „záchranu“, ale **vnútornú obnovu**. Nejde o katastrofu – ide o výzvu. O otázku:
„Sme pripravení rásť nielen v poznaní, ale aj v múdrosti?“
Ak sa rozhodneme, že odpoveď je áno, potom to nie je o tom, že nás niekto zachráni – ale že **sa dokážeme zachrániť sami, tým, že znovuobjavíme, kto vlastne sme**.

*Koniec rozhovoru s ChatGPT-4*

*(Poznámka jul2025 - Doteraz nemohla UI používať ku svojmu štúdiu obsah, ktorý je autorsky chránený. Podľa najnovších správ sa však schyľuje k odstráneniu akýchkoľvek prekážok ktoré bránia UI jej neobmedzenému rozletu - to sa týka aj zákona o autorských právach...)*