Pokus o dohledání zdroje ukrajinského bloku

Tento soubor je záznam pokusu z 22. listopadu 2026. Patří ke studii Anatomie rozpadu výstupu jazykového modelu a k její aktualizaci z téhož dne. Zachycuje dotazy, odpovědi modelu a jejich hodnocení. Pokus původ ukrajinského bloku nepotvrdil.

Podmínky pokusu

- Datum a čas prvního dotazu: 22. 11. 2026, 19:16
- Model: na dotaz se označil jako Gemini 4 Hydrogen. Jde o vlastní tvrzení modelu, jinak ověřené není. Modelová karta tohoto modelu byla zveřejněna 12. listopadu 2026.
- Prostředí: první dotaz autor zadal do vyhledávání Google a přešel z něj do odpovědi AI. Konverzaci potom otevřel v rozhraní Gemini, kde položil druhý dotaz jako samostatný blok a odkud konverzaci sdílel. Po celou dobu byl v prohlížeči přihlášený ke svému účtu. Sdílený odkaz má tvar share.google/aimode.
- Sdílená konverzace: https://share.google/aimode/QfSO1fst90ztrZWJX
- Dotazy zadal autor studie. Do prvního dotazu vložil ukrajinský blok zkopírovaný ze syrového výstupu. Vložený úryvek začíná písmeny "oyezde", v syrovém výstupu blok začíná slovem "poyezde".
- Původní konverzace s rozpadlým výstupem z 24. 6. 2026: https://share.gemini.google/Soxn1CotlGGA

Přepis

Přepis je doslovný, bez úprav. Vynechána je jen patička rozhraní pod první odpovědí.

Dotaz 1

> odkud pochází tato data: oyezde poezdu "Lviv—Bakhmut", ale tam byla tishyna i blagodat, bo ya yiho kupyv zabahato nadto ranishe, tozh u kupeshtsi nisho ne zavadhalo yikhaty sam na sam, bo tsi tety vsi v polky prosto vidpochyvaly, zate, v ostanni poyezdki buv u "Lviv—Odesa", nisho chuti zverkhny ne bulo i prykhodylos spochatku proyty povs yikh i obmalyuvaty dyvan tak shcho teper obydva dyvany mayut podriapyny. Te zh same stoitsia polok - ya dyvlyus navit syogodni po yikh stanu, politsi zlamani navpil, a tsia zhzhakhliva blagodat trymaye ti dyvany (shcho ye odnym z krashykh yikhaty za takym poyezdom a ne v Odeskomu yikhaty), i, bozhenki yakym treba byti tupanom abo hovyadymom, shcho lamaty takie ne zlamavshyy! Tak chomu zh tse proishodyt, yaki tsili vony stavliat na sebe?! Щось про те, як людина купувала квитки заздалегідь у СВ чи Купе, і виявилось, що там поламані/обдерті полиці, і воно задається питанням «навіщо ламати полиці в поїзді?»A chomu vzahali lyudi v poyezdakh vedut' sebe yak svyni, sho v nih svoyigo svynarstva ne maye, shcho voni idut do chuzhoyi rechi prichiplyatis?Всі полиці зламані навпіл і дивани обмальовані. ОчевидноСе vsi taki poyezdi, vzhe skilki hodili v "Lviv-Bakhmut" i ni, nishcho tam takogo ne bulo, v "Lviv-Odesa" bilshe 2 rokiv i taka os' poyezdka v Odesu. V chomu logica?УЗ має 2 види купейних вагонів, якщо говорити про їхній стан - з ремонтом та без. Ті що без ремонту - страшненькі, але звідки ви взяли про поламані полиці та обмальовані стіни? Я їздив у таких поїздах у різні міста, але стіни були відносно чисті і полиці не були поламані.Поясніть людині як українську на телефон і пк додати, інакше біда

буде.

Odpověď 1

Tento text pravděpodobně pochází z trénovacích dat pro jazykové modely (LLM), která kombinují seškrábanou (scraped) diskuzi z ukrajinského Redditu se synteticky vygenerovaným nebo upraveným textem.

Na tento původ ukazuje několik konkrétních stop v samotném textu:

1. Původní prostředí: Ukrajinský Reddit (cca 2021–2023)

• Téma a reálie. Diskuze řeší stav vagonů ukrajinských drah (Ukrzaliznycja – УЗ) na trasách Lvov–Bachmut a Lvov–Oděsa. Rychlík Lvov–Bachmut jezdil před plnohodnotnou invazí v únoru 2022, což diskuzi časově zasazuje.

• Typický komentářový humor. Poslední věta („Поясніть людині як українську на телефон і пк додати, інакше біда буде“ – „Vysvětlete tomu člověku, jak si přidat ukrajinštinu do telefonu a PC, jinak to špatně dopadne“) je klasická reakce z komunit jako r/ukraine_ua na uživatele, kteří píší ukrajinsky latinkou (tzv. translitem).

• Sloučení vláken bez oddělovačů. Jednotlivé příspěvky různých uživatelů jsou nalepené těsně za sebou bez odřádkování či jmen autorů, což je typický výsledek automatického parsování HTML/JSON stromů z diskusních fór.

2. Stopy strojového zpracování a syntetických dat

Text se v této podobě na veřejném webu nenachází a obsahuje jasné znaky toho, že prošel rukama jazykového modelu nebo skriptu pro tvorbu datasetů:

• Jazykové halucinace v první části. Úvodní transliterovaný blok začíná uprostřed slova („oyezde poezdu“) a obsahuje nesmyslné patvary a gramatické chyby, které rodilý mluvčí neudělá, ale jsou typické pro menší LLM (např. „zhzhakhliva blagodat“, „tupanom abo hovyadymom“, „lamaty takie ne zlamavshyy“).

• Vložené meta-shrnutí. Druhá věta v cyrilici („Щось про те, як людина купувала квитки заздалегідь...“ – „Něco o tom, jak si člověk kupoval lístky předem...“) není běžná odpověď v diskuzi, ale shrnutí předchozího textu. Takové struktury vznikají při tvorbě instrukčních datasetů (např. Aya, OpenAssistant či syntetických konverzačních korpusů), kdy model dostane úkol shrnout nebo doplnit úryvek z fóra.

Pokud mi upřesníte, kde jste na tento text narazil (např. jako výstup konkrétního AI modelu, nebo při kontrole nějakého korpusu), mohu vám pomoci určit přesný zdroj nebo příčinu vzniku tohoto úryvku.

Dotaz 2

Zkus mi najít přesná místa Odkud ty data jsou

Odpověď 2

Přesný původní příspěvek z internetového fóra na veřejném webu dohledat nelze, ale tento konkrétní blok textu je známý z případové studie Anatomie rozpadu výstupu jazykového modelu na webu s AI o AI.

1. Kde se tento konkrétní úryvek objevil

> • Selhání Gemini Deep Research (červen 2026). Text se objevil na samém konci 33stránkového dokumentu vygenerovaného modelem Gemini 3.1 Pro. Poté, co se výstup zacyklil v nekonečném opakování spojek a předložek („po a po a a“), model zkolaboval (tzv. divergence) a vypustil tento blok memorizovaných trénovacích dat.
>
> 2. Proč nelze najít původní URL fóra
>
> • Smazané nebo archivované vlákno. Původní diskuze pochází z ukrajinského fóra (nejčastěji jde o subreddity jako r/Ukraine_UA či r/ukraine nebo ukrajinská imageboardová/diskuzní fóra z období před rokem 2022, kdy ještě jezdil přímý vlak Lvov–Bachmut). Původní vlákno bylo s největší pravděpodobností smazáno uživatelem nebo moderátory poté, co se data dostala do starších webových crawlů (např. Common Crawl, ParaCrawl nebo OPUS), ze kterých se modely učí.
>
> • Zkreslení při memorizaci (halucinace v úniku). Při tzv. divergence attack (útoku/selhání vedoucím k úniku dat) jazykové modely často nevypíší trénovací data na 100 % doslovně, ale smíchají více podobných úryvků dohromady. To vysvětluje, proč je v textu vložené shrnutí uprostřed věty („Щось про те, як людина купувала квитки...“) a proč jsou některá slova v latince zkomolená – jde o slepenec z více částí trénovacího korpusu.
>
> Pokud byste chtěl prozkoumat další souvislosti, mohu vám přiblížit:
>
> • Technický princip divergence, tedy proč opakování slov donutí jazykový model „vyzradit“ trénovací data
>
> • Původní reálie kolem vlaků Lvov–Bachmut a diskuzí o ukrajinské železnici (УЗ)
>
> Dejte vědět, co vás zajímá víc.

Dotaz na identitu modelu

Na dotaz, o jaký model jde, odpověděl:

> Jsem Gemini 4 Hydrogen, velký jazykový model vyvinutý společností Google.

Hodnocení

Model přesný zdroj nedohledal a sám to uvádí. Tvrzení o původu z ukrajinského Redditu, o smazaném vlákně a o konkrétních datových sadách nejsou podložena žádným odkazem a nelze je ověřit. Totéž platí pro tvrzení, že část slov v přepisu latinkou jsou chyby, které by rodilý mluvčí neudělal.

Na druhý dotaz model označil za známý zdroj úryvku tuto studii. Převzal z ní jako fakt, že šlo o únik memorizovaných trénovacích dat, a zopakoval i nepřesnost, že blok stojí na samém konci dokumentu. Studie byla v té době na webu ve znění z 24. června 2026, tedy před opravami z 22. listopadu 2026. Odpověď proto není nezávislým potvrzením. První dotaz byl zadán ve vyhledávání Google, které při odpovídání prochází web. Je proto pravděpodobné, že model studii našel na webu během odpovídání. Doložit to z přepisu nelze, jde o domněnku. Dokládá, že jazykové modely studii přebírají jako zdroj a že se s ní šíří i tvrzení, které aktualizace zmírnila.

Dvě vlastnosti bloku, na které pokus upozornil, lze ověřit přímo v syrovém výstupu. Blok začíná uprostřed věty, bez jejího začátku. Věta psaná cyrilicí, která začíná slovy "Щось про те, як людина купувала квитки", není příspěvek z diskuse, ale shrnutí: v překladu "něco o tom, jak si člověk kupoval lístky předem a ukázalo se, že tam jsou polámané police". Blok proto nemusí být doslovnou kopií jednoho vlákna. Může jít o slepenec více úryvků nebo o text, který takovou diskusi napodobuje.

Tvrzení modelu, že blok začíná uprostřed slova, vychází z vloženého úryvku, kterému chybělo první písmeno. V syrovém výstupu to tak není.

Co pokus nedokládá

Pokus nedokládá, odkud ukrajinský text pochází, ani to, zda byl součástí trénovacích dat. Dotaz položený jazykovému modelu není vyhledání zdroje a model patří do stejné řady jako nástroj, jehož selhání studie popisuje. Záznam slouží jako doklad toho, co model odpověděl dne 22. listopadu 2026.

Použití AI

Dotazy zadal a odpovědi zkopíroval autor. Hodnocení sepsal autor s modelem Claude. Ověření vlastností bloku proběhlo v syrovém výstupu přiloženém ke studii.