https://farid.ps/articles/arguing_with_an_ai_sceptic/cs.html

Ponaučení z pokusu o diskusi s skeptikem vůči AI

Epizoda začala politickým memem, který jsem zveřejnil: Donald Trump a Benjamin Netanyahu v oranžových vězeňských uniformách, sedící na palandě pod teplým, nostalgickým vánočním překryvem s nápisem „All I Want for Christmas“. Vizuální ironie byla okamžitá a ostrá. Jeho vytvoření vyžadovalo záměrné obcházení omezení. Současné modely generování obrázků mají jak politická bezpečnostní opatření, tak technická omezení koherence:

- Grok umožňuje karikatury významných osobností, ale konzistentně selhává při vytváření spolehlivého překrytého textu.
- ChatGPT exceluje v generování dekorativního slavnostního textu jako „All I Want for Christmas“, ale jeho bezpečnostní mechanismy odmítají prompty zobrazující žijící politické lídry ve vězeňském prostředí.

Žádný jednotlivý model nedokázal vytvořit kompletní obrázek. Protikladné prvky — nabitá politická satira kombinovaná s sentimentálním vánočním poselstvím — spouštějí mechanismy odmítnutí nebo selhání koherence. Velké jazykové modely (LLM) jsou prostě neschopné syntetizovat takto koncepčně protikladné komponenty do jednoho koherentního výstupu. Vygeneroval jsem oba prvky odděleně, pak je ručně sloučil a upravil v GIMPu. Finální kompozit byl nepopiratelně lidsky vytvořený: moje koncepce, můj výběr komponent, moje sestavení a úpravy. Bez těchto nástrojů by satira zůstala uvězněná v mé hlavě nebo by vyšla jako hrubé panáčkové kresby — zbavená veškerého vizuálního dopadu.

Někdo nahlásil obrázek jako „vygenerovaný AI“. Následující den server zavedl nové pravidlo zakazující obsah generovaný generativní AI. Toto pravidlo — a mem, který ho spustil — mě přímo inspirovaly k napsání a publikování eseje „Vysokodimenzionální mysli a břemeno serializace: Proč LLM záleží pro neurodivergentní komunikaci“. Doufal jsem, že podnítí reflexi nad tím, jak tyto nástroje slouží jako kognitivní a kreativní úpravy. Místo toho se to změnilo v poněkud trapnou výměnu názorů s administrátorem.

Pozice skeptika a výměna názorů

Administrátor argumentoval, že LLM nejsou vyvíjeny pro prospěch lidí, ale podporují plýtvání zdroji a militarizaci. Citoval spotřebu energie, vazby na armádu, kolaps modelů, halucinace a riziko „mrtvého internetu“. Přiznal, že eseji jen zběžně prohlédl a přiznal vlastnictví výkonné herní pracovní stanice schopné spouštět pokročilé lokální LLM pro soukromou zábavu, s přístupem k ještě větším modelům přes přítele.

Objevilo se několik rozporů:

- Moje práce probíhá na nízkoenergetickém, opravitelném Raspberry Pi 5 (5–15 W) s využitím sdílených cloudových instancí. Jeho lokální setup spotřebovává mnohem více vyhrazené energie a hardwaru.
- Hardware, který používá k „hrátkám“ s výkonnými LLM lokálně, pochází od firem (Intel, AMD, NVIDIA) s miliardovými přímými smlouvami s DoD.

Nejvýrazněji osoba prosazující zákaz na ochranu autenticity odmítala někoho, kdo aktivně testuje LLM na faktickou a geopolitickou zkreslenost (viz moje veřejné audity Groku a ChatGPT).

Analogie s Hawkingem a vlastní slova administrátora

Administrátor se identifikoval jako neurodivergentní a uznal potenciál AI jako asistenční technologie. Chválil brýle s reálným časovým titulkováním pro zrakově postižené jako „opravdu skvělé“, ale trval na tom, že „mít stroj psát eseje a kreslit obrázky je něco jiného“. Dodal: „Neurodivergentní lidé to dokážou, mnozí překonali překážky, aby si tyto dovednosti rozvinuli.“ Také popsal svou vlastní zkušenost s LLM: „Čím více už o tématu vím, tím méně potřebuji AI. Čím méně o tématu vím, tím méně jsem vybaven všimnout si halucinací a opravit je.“ Tyto výroky odhalují hlubokou asymetrii v posuzování úprav.

Představte si aplikaci stejné logiky na Stephena Hawkinga:

„Uznáváme, že syntetizátor hlasu by vám mohl pomoci komunikovat rychleji, ale raději bychom, abyste se více snažil s přirozeným hlasem. Mnoho lidí s nemocí motorických neuronů překonalo překážky, aby mluvili jasně — vy byste si tyto dovednosti měl také rozvinout. Stroj dělá něco jiného než skutečná řeč.“

Nebo z jeho vlastní perspektivy na faktickou přesnost:

„Čím více Hawking už o kosmologii ví, tím méně potřebuje syntetizátor. Čím méně ví, tím méně je vybaven všimnout si chyb v hlasu stroje a opravit je.“

Nikdo by to nepřijal. Chápali jsme, že Hawkingův syntetizátor nebyl berlí nebo zředěním — byl to esenciální most, který umožnil jeho mimořádné mysli sdílet plnou hloubku bez nepřekonatelných fyzických překážek.

Administrátorova pohodlnost s lineární, lidsky strukturovanou prózou odráží kognitivní styl, který se blíže shoduje s neurotypickými očekáváními. Můj profil je opačný: faktická a logická hloubka přichází přirozeně (jako vývoj vícejazyčné publikační platformy úplně sám), ale produkce strukturované, přístupné prózy pro lidské publikum byla vždy překážkou — přesně to, co eseji popisuje. Přijmout brýle s titulky nebo alternativní text jako legitimní úpravy, zatímco odmítat LLM strukturování pro kognitivní divergenci, znamená kreslit libovolnou hranici. Mastodon a širší Fediverse se často pyšní inkluzivitou. Přesto to zavádí nové brány: určité úpravy jsou vítány; jiné musí být překonány individuálním úsilím.

Historické ozvěny: Odpor vůči transformačním nástrojům

Paušální odmítnutí veřejného používání generativní AI echoje opakující se vzorec v historii technologií. V Anglii na počátku 19. století kvalifikovaní tkalci známí jako luddité rozbíjeli mechanizované stavy, které ohrožovaly jejich řemeslo a živobytí. Svítilny na plyn v městech se stavěly proti Edisonově žárovce ze strachu z zastarání. Kočí, stájníci a chovatelé koní odporovali automobilu jako existenční hrozbě pro jejich způsob života. Profesionální písaři a kreslíři viděli v kopírce alarm, věříc, že znehodnotí pečlivou ruční práci. Sazeči a tiskaři bojovali proti počítačové sazbě.

V každém případě odpor pramenil z oprávněného strachu: nová technologie činila dovednosti, na které byli hrdí, zastaralými, ohrožujíc jejich ekonomické role a sociální identitu. Změny se zdály jako znehodnocení lidské práce.

Přesto historie hodnotí tyto inovace podle jejich širšího dopadu: mechanizace snížila dřinu a umožnila masovou výrobu; elektrické osvětlení prodloužilo produktivní hodiny a zlepšilo bezpečnost; automobily poskytly osobní mobilitu; kopírky demokratizovaly přístup k informacím; digitální sazba urychlila a zpřístupnila publikování. Málokdo by dnes chtěl vrátit k plynovým lampám nebo koňským povozům jen pro zachování tradičních zaměstnání. Nástroje rozšířily lidské schopnosti a účast mnohem více, než je zmenšily.

Generativní AI — používaná jako protetika pro kognici nebo kreativitu — následuje stejnou trajektorii: neničí lidský záměr, ale rozšiřuje vyjádření těm, jejichž myšlenky byly omezeny bariérami provedení. Její paušální odmítnutí riskuje opakování ludditského impulsu — obranu známých procesů na úkor širší účasti.

Závěr: Kdo rozhoduje o přijatelných úpravách?

Události vyprávěné v této eseji — jedno nahlášené obrázek, jeden ukvapeně zavedený zákaz, jedna protáhlá debata — odhalují více než lokální neshodu ohledně technologie. Odhalují mnohem hlubší a fundamentálnější otázku: **Kdo rozhoduje, které úpravy jsou přijatelné a které ne?** Měli by to být lidé, kteří žijí v kůži a mozku, který úpravu potřebuje — ti, kteří z denní zkušenosti vědí, co přemosťuje propast mezi jejich schopnostmi a plnou účastí? Nebo by to měli být outsideři, ať už dobře mínění, kteří tuto prožitkovou realitu nesdílejí a proto nemohou cítit tíhu překážky?

Historie na tuto otázku odpovídá opakovaně a téměř vždy stejným směrem. Invalidní vozíky byly kdysi kritizovány za podporu závislosti; systémy vzdělávání neslyšících dlouho trvaly na učení čtení ze rtů a orální řeči místo znakového jazyka. V každém případě lidé nejblíže postižení nakonec zvítězili — ne proto, že popírali obavy z nákladů, přístupu nebo potenciálního zneužití, ale proto, že byli primární autoritou na to, co skutečně obnovilo jejich autonomii a důstojnost.

S velkými jazykovými modely a dalšími generativními nástroji procházíme stejným cyklem znovu. Mnoho těch, kteří hlídají jejich používání, nezažívá specifické kognitivní nebo expresivní překážky, které činí lineární strukturování, narativní tok nebo rychlou serializaci vyčerpávajícím úkolem překladu do cizího jazyka. Zvenku může „jen se více snaž“ nebo „rozviň si dovednost“ znít rozumně. Zevnitř nástroj není zkratka kolem úsilí; je to rampa, sluchadlo, protetika, která konečně umožní předchozímu úsilí dosáhnout světa.

Nejhlubší ironie se objevuje, když arbitři se sami identifikují jako neurodivergentní, přesto jejich konkrétní neurologie se blíže shoduje s neurotypickými očekáváními v posuzované doméně. „Já to překonal takto, tak by měli i ostatní“ je srozumitelné, ale stále funguje jako hlídání bran — replikuje ty samé normy, které kritizujeme, když přicházejí od neurotypických autorit. Je načase konzistentní etický princip:

- Osoba nejblíže postižení je primární autoritou na to, co umožňuje jejich smysluplnou účast.
- Externí kritika je legitimní ohledně kolektivních škod (dopad na životní prostředí, riziko dezinformací, vytlačování práce), ale ne ohledně interní legitimity samotné úpravy.

Jeden obzvláště odhalující dvojí standard se objevuje v rozšířeném požadavku explicitního zveřejňování používání generativní AI. Pro většinu jiných úprav podobné zveřejňování nevyžadujeme. Naopak aktivně oslavujeme technologické pokroky, které je činí neviditelnými: tlusté brýle nahrazené kontaktními čočkami nebo refrakční chirurgií; objemné sluchadla zmenšené do téměř neviditelnosti; léky na soustředění, náladu nebo bolest užívané soukromě bez poznámky nebo prohlášení. V těchto případech společnost považuje diskrétní, skryté používání za pokrok — za obnovu důstojnosti a normality. Přesto když úprava rozšiřuje kognici nebo expresi, scénář se obrací: nyní musí být označena, oznámena, odůvodněna. Neviditelnost se stává podezřelou místo žádoucí. Tento selektivní požadavek na transparentnost není skutečně o prevenci klamu; jde o zachování pohodlí s konkrétním obrazem neasistovaného lidského autorství. Fyzické opravy smějí zmizet; opravy mysli musí zůstat viditelně označené.

Pokud chceme být konzistentní, musíme buď požadovat zveřejňování pro každou úpravu (absurdní a invazivní požadavek), nebo přestat vyčleňovat kognitivní nástroje pro zvláštní kontrolu. Principielní pozice — ta, která respektuje autonomii a důstojnost — je umožnit každému člověku rozhodnout, jak viditelná nebo neviditelná jeho úprava má být, bez trestajících pravidel cílených na jednu formu pomoci, protože narušuje stávající představy o kreativitě a intelektu. Tato eseje není jen obhajobou jednoho konkrétního nástroje. Je obhajobou širšího práva postižených a neurodivergentních lidí definovat své vlastní potřeby přístupu, bez nutnosti je odůvodňovat těm, kteří nikdy nekráčeli v jejich botách. Toto právo by nemělo být kontroverzní. Přesto, jak ukazuje předchozí vyprávění, stále je.